Πανεπιστήμιο Πατρών
Σχολή Θετικών Επιστημών
Τμήμα Μαθηματικών

Μαθηματικές Μέθοδοι Βελτιστοποίησης Προβλημάτων Μεγάλης Κλίμακας

Διδακτορική Διατριβή

Μαριάννα Σ. Αποστολοπούλου

Μάρτιος, 2011

© 2011 Μαριάννα Σ. Αποστολοπούλου
Πανεπιστήμιο Πατρών, Τμήμα Μαθηματικών
Πανεπιστημιούπολη, Ρίο 265 04, Πάτρα

Στο σύζυγό μου Δημήτρη, στους γονείς
μου Σωτήρη & Γωγώ, με αγάπη και
εκτίμηση

Ευχαριστίες

Ολοκληρώνοντας τη διδακτορική μου διατριβή θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν κατά την περίοδο των μεταπτυχιακών μου σπουδών, καθηγητές, συναδέλφους, συγγενείς και φίλους. Ιδιαίτερες ευχαριστίες οφείλω στον επιβλέποντα καθηγητή μου, κ. Παναγιώτη Πιντέλα, για την ειλικρινή και ανιδιοτελή υποστήριξη του και για την αμέριστη συμπαράστασή του καθόλη τη διάρκεια των ερευνητικών μου δραστηριοτήτων.

Επίσης, ευχαριστώ θερμά τον κ. Χαράλαμπο Μπότσαρη, Καθηγητή του Τμήματος Περιφερειακής Οικονομικής Ανάπτυξης του Πανεπιστημίου Στερεάς Ελλάδας και μέλος της Τριμελούς Συμβουλευτικής Επιτροπής, του οποίου η καθοδήγηση και οι συμβουλές υπήρξαν πολύτιμες και ουσιαστικές. Η ενθάρρυνση και η υποστήριξη του κ. Μπότσαρη στα πρώτα μου βήματα με όπλισαν με αυτοπεποίθηση και κουράγιο να συνεχίσω.

Σημαντική βοήθεια και συμπαράσταση είχα από την κα Θεοδούλα Γράψα, Επίκουρη Καθηγήτρια, την οποία ευχαριστώ θερμά. Η κα Γράψα, μέλος της Τριμελούς Συμβουλευτικής Επιτροπής, με στήριξε ηθικά τόσο κατά τη διάρκεια της έρευνάς μου, όσο και στη διάρκεια συγγραφής της διατριβής μου. Οι συμβουλές της υπήρξαν πολύτιμες για τη λογική ροή και τη δομή της διατριβής και με βοήθησαν τα μέγιστα για την τελική παρουσίαση των ερευνητικών αποτελεσμάτων.

Επίσης, θα ήθελα να ευχαριστήσω τον συνεργάτη μου κ. Ιωάννη Λιβιέρη, υποψήφιο διδάκτορα του Τμήματος Μαθηματικών, για την πραγματικά ευχάριστη και επιτυχή συνεργασία που είχαμε.

Τέλος, θερμά ευχαριστώ τον σύζυγο και συνεργάτη μου Δρ. Δημήτρη Σωτηρόπουλο, ο οποίος με στήριξε τα μέγιστα στην ερευνητική μου πορεία και στα δύσκολα χρόνια των μεταπτυχιακών μου σπουδών.

*Μαριάννα Αποστολοπούλου
Πάτρα, Μάρτιος 2011*

Περίληψη

Στην παρούσα διατριβή μελετάμε το πρόβλημα της βελτιστοποίησης μη γραμμικών συναρτήσεων πολλών μεταβλητών,

$$\min_{x \in \mathcal{D}} f(x),$$

όπου η αντικειμενική συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο $\mathcal{D} \subseteq \mathbb{R}^n$. Αναπτύσσουμε μαθηματικές μεθόδους βελτιστοποίησης και αποσκοπούμε στην επίλυση προβλημάτων μεγάλης κλίμακας, δηλαδή προβλημάτων των οποίων οι μεταβλητές είναι πολλές χιλιάδες, ακόμα και εκατομμύρια. Η βελτιστοποίηση προβλημάτων μεγάλης κλίμακας, όπως για παράδειγμα η αποκατάσταση εικόνας, ο διαχωρισμός γραφημάτων ή η πρόβλεψη ασθενειών, αποτελεί σήμερα περιοχή με μεγάλη ερευνητική δραστηριότητα, όχι μόνο λόγω των νέων αλγοριθμικών εξελίξεων αλλά και της ενδογενούς πολυπλοκότητας των προβλημάτων της σύγχρονης εποχής.

Η βασική ιδέα των μεθόδων που αναπτύσσουμε, προκειμένου να επιλύσουμε προβλήματα μεγάλης κλίμακας, έγκειται στη θεωρητική μελέτη των χαρακτηριστικών ιδιοτήτων των Quasi-Newton ενημερώσεων ελάχιστης και μικρής μνήμης. Τα νέα θεωρητικά αποτελέσματα, οδηγούν σε δραστική μείωση της χωρικής πολυπλοκότητας γνωστών αλγορίθμων του μη γραμμικού προγραμματισμού. Ως αποτέλεσμα, οι νέοι αλγόριθμοι που μπορούν να προκύψουν, έχουν χωρική πολυπλοκότητα $\Theta(n)$, όπου n η διάσταση του προβλήματος και μπορούν πλέον να εφαρμοστούν και για μεγάλης κλίμακας προβλήματα.

Η διατριβή αποτελείται από οκτώ κεφάλαια. Τα πρώτα δύο είναι εισαγωγικού περιεχομένου, ενώ τα υπόλοιπα έξι περιλαμβάνουν τα ερευνητικά αποτελέσματα της παρούσας διατριβής. Πιο συγκεκριμένα, στο Κεφάλαιο 1 παραθέτονται βασικές έννοιες των μαθηματικών και της βελτιστοποίησης, οι οποίες είναι απαραίτητες για τη μελέτη και παρουσίαση

των μεθόδων που διαπραγματεύεται η διατριβή. Αναφέρονται οι συνθήκες βελτιστότητας, τόσο στη βελτιστοποίηση χωρίς περιορισμούς, όσο και στη βελτιστοποίηση με περιορισμούς.

Το Κεφάλαιο 2 αναφέρεται στη βασική ιδέα των αλγορίθμων βελτιστοποίησης. Παραθέτονται οι πιο γνωστές μέθοδοι βελτιστοποίησης, όπως η μέθοδος Newton, οι περιορισμένης μνήμης Quasi-Newton μέθοδοι, οι μέθοδοι γραμμικής και καμπυλόγραμμης αναζήτησης. Τέλος αναφέρονται οι μέθοδοι περιοχής εμπιστοσύνης και το υποπρόβλημα περιοχής εμπιστοσύνης. Ιδιαίτερη μνεία γίνεται στις Quasi-Newton ενημερώσεις περιορισμένης μνήμης, στην καμπυλόγραμμη αναζήτηση και στο υποπρόβλημα περιοχής εμπιστοσύνης, καθώς οι συγκεκριμένες μέθοδοι αποτελούν τα ερευνητικά ενδιαφέροντα της παρούσας διατριβής.

Το Κεφάλαιο 3 αποτελεί το πρώτο από τα έξι κεφάλαια στα οποία καταγράφονται τα ερευνητικά αποτελέσματα. Μελετώνται οι χαρακτηριστικές ιδιότητες των ελάχιστης μνήμης Quasi-Newton ενημερώσεων, δηλαδή των SR1, BFGS, DFP και PSB. Διατυπώνονται θεωρήματα αναφορικά με το χαρακτηριστικό πολυώνυμο, των αριθμό των διακριτών ιδιοτιμών και των αντίστοιχων ιδιοδιανυσμάτων της γενικής PSB ενημέρωσης ελάχιστης μνήμης. Μελετώνται διεξοδικά καθεμιά από τις τέσσερις προαναφερθείσες ενημερώσεις και εξάγονται κλειστοί τύποι για τον υπολογισμό των ανωτέρω ποσοτήτων, αποφεύγοντας πλήρως τόσο την αποθήκευση, όσο και την παραγοντοποίηση πινάκων.

Το Κεφάλαιο 4 αναφέρεται στις ελάχιστης και μικρής μνήμης BFGS ενημερώσεις, λόγω της σπουδαιότητας που παρουσιάζουν στους αλγορίθμους βελτιστοποίησης. Μελετάται η περίπτωση όπου ο αρχικός πίνακας είναι ο μοναδιαίος, κλιμακούμενος με μια συγκεκριμένη παράμετρο. Αποδεικνύεται ότι η τιμή της συγκεκριμένης αυτής παραμέτρου μας επιτρέπει να ξέρουμε εκ των προτέρων το πρόσημο όλων των ιδιοτιμών του πίνακα. Επίσης, τεκμηριώνονται ικανές συνθήκες για τον ακριβή αριθμό των διακριτών ιδιοτιμών της BFGS ενημέρωσης μικρής μνήμης. Από αυτές τις συνθήκες, υπολογίζεται το αντίστοιχο ιδιοδιάνυσμα χρησιμοποιώντας είτε κλειστό τύπο ή απλές πράξεις.

Το Κεφάλαιο 5 αναφέρεται στην εφαρμογή των χαρακτηριστικών ιδιοτήτων των Quasi-Newton ενημερώσεων ελάχιστης και μικρής μνήμης, για την επίλυση του υποπροβλήματος περιοχής εμπιστοσύνης μεγάλης κλίμακας, το οποίο αποτελεί ένα πρόβλημα βελτιστοποίησης με περιορισμό.

Προτείνεται μια μέθοδος η οποία είναι ανέξοδη υπολογιστικά και μπορεί να εφαρμοστεί για την επίλυση υποπροβλημάτων πολλών μεταβλητών. Ο προτεινόμενος αλγόριθμος επίλυσης του υποπροβλήματος, εκμεταλλεύεται πλήρως τις ιδιότητες που μελετήθηκαν στα Κεφάλαια 3 και 4. Ο συγκεκριμένος αλγόριθμος αντιμετωπίζει και τις δύο περιπτώσεις που ανακύπτουν στο υποπρόβλημα, την τυπική και τη δύσκολη περίπτωση, χρησιμοποιώντας μόνο εσωτερικά γινόμενα και διανυσματικά αθροίσματα. Τέλος, δίνεται το θεώρημα σύγκλισης του προτεινόμενου αλγορίθμου, όταν αυτός εντάσσεται σε ένα πλαίσιο μεθόδων περιοχής εμπιστοσύνης.

Στο Κεφάλαιο 6 παρουσιάζονται τα αριθμητικά αποτελέσματα του αλγορίθμου που προτάθηκε στο Κεφάλαιο 5. Η προτεινόμενη μέθοδος συγκρίνεται με άλλες τρεις μεθόδους, οι οποίες θεωρούνται οι καλύτερες στη διεθνή βιβλιογραφία για την επίλυση του υποπροβλήματος περιοχής εμπιστοσύνης. Για να αξιολογηθεί η αποδοτικότητα του προτεινόμενου αλγορίθμου, χρησιμοποιούνται στατιστικά μέτρα και προφίλ απόδοσης, από τα οποία γίνεται σαφής η υπεροχή της προτεινόμενης μεθόδου έναντι των υπολοίπων.

Το Κεφάλαιο 7 αναφέρεται στην καμπυλόγραμμη αναζήτηση, στην οποία ο εσσιανός πίνακας προσεγγίζεται από την BFGS ενημέρωση ελάχιστης μνήμης. Οι χαρακτηριστικές ιδιότητες της συγκεκριμένης ενημέρωσης, όπως αυτές μελετήθηκαν στο Κεφάλαιο 3, χρησιμοποιούνται σε έναν αλγόριθμο καμπυλόγραμμης αναζήτησης. Ο προτεινόμενος αλγόριθμος εκμεταλλεύεται την πληροφορία από την καμπυλότητα του προσεγγιστικού εσσιανού πίνακα της αντικειμενικής συνάρτησης. Τεκμηριώνεται ότι ο συγκεκριμένος αλγόριθμος διατηρεί τις καλές ιδιότητες σύγκλισης τόσο της γραμμικής, όσο και της καμπυλόγραμμης αναζήτησης, ενώ η μνήμη που απαιτείται για τον υπολογισμό της πιο αρνητικής ιδιοτιμής και του αντίστοιχου ιδιοδιανύσματος είναι αμελητέα λόγω των αναλυτικών εκφράσεων που χρησιμοποιούνται.

Στο Κεφάλαιο 8 παρουσιάζονται τα αριθμητικά αποτελέσματα του προτεινόμενου καμπυλόγραμμου αλγορίθμου, στην επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης χωρίς περιορισμούς. Τα αποτελέσματα αυτά δείχνουν ότι ο προτεινόμενος αλγόριθμος είναι κατάλληλος για προβλήματα με πολλές μεταβλητές. Για το λόγο αυτό, ο καμπυλόγραμμος αλγόριθμος εφαρμόζεται στην εκπαίδευση νευρωνικών δικτύων και στην κατηγοριοποίηση προβλημάτων βιοϊατρικής μεγάλης κλίμακας. Τα α-

ριθμητικά αποτελέσματα δείχνουν ότι ο προτεινόμενος αλγόριθμος είναι γρήγορος και πολύ αποτελεσματικός.

Τέλος, παρατίθεται ένα ευρετήριο των πιο σημαντικών όρων που χρησιμοποιούνται στην παρούσα Διατριβή.

Περιεχόμενα

Περίληψη	iii
Περιεχόμενα	vii
1 Εισαγωγή	1
1.1 Το πρόβλημα της βελτιστοποίησης	1
1.2 Μαθηματικό υπόβαθρο	3
1.2.1 Διανύσματα και πίνακες	4
1.2.1.1 Ακολουθίες και ρυθμοί σύγκλισης	7
1.2.1.2 Υπόχωροι	8
1.2.1.3 Ο τύπος των Sherman-Morrison-Woodbury	8
1.2.1.4 Η κατά Cholesky παραγοντοποίηση πινάκων	9
1.2.2 Ιδιοτιμές και ιδιοδιανύσματα	10
1.2.2.1 Η μέθοδος της αντίστροφης δύναμης	14
1.2.2.2 Η μέθοδος Lanczos	15
1.2.3 Στοιχεία τοπολογίας	16
1.2.4 Παράγωγοι και η κατά Taylor προσέγγιση	17
1.2.5 Κυρτότητα	19
1.3 Συνθήκες βελτιστότητας	20
1.3.1 Βελτιστοποίηση χωρίς περιορισμούς	21
1.3.2 Βελτιστοποίηση με περιορισμούς	24
1.3.2.1 Συνθήκες βελτιστότητας πρώτης τάξης	25
1.3.2.2 Συνθήκες βελτιστότητας δεύτερης τάξης	28
2 Βασικές μέθοδοι μη γραμμικής βελτιστοποίησης	31
2.1 Η μέθοδος Newton	31
2.2 Οι Quasi-Newton μέθοδοι	33
2.3 Η γενική Powell-Symmetric-Broyden (PSB) ενημέρωση	36
2.3.1 Η SR1 συμμετρική ενημέρωση πρώτης τάξης	38

2.3.2	Η DFP συμμετρική ενημέρωση δεύτερης τάξης . . .	39
2.3.3	Η BFGS συμμετρική ενημέρωση δεύτερης τάξης . .	41
2.3.4	Η PSB ενημέρωση δεύτερης τάξης	42
2.3.5	Η Broyden κλάση	43
2.4	Οι Quasi-Newton ενημερώσεις περιορισμένης μνήμης . . .	44
2.4.1	Οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης . .	45
2.5	Οι μέθοδοι γραμμικής αναζήτησης	46
2.6	Οι μέθοδοι καμπυλόγραμμης αναζήτησης	49
2.6.1	Συνθήκες βήματος δεύτερης τάξης	51
2.6.1.1	Η μέθοδος McCormick	53
2.6.1.2	Η μέθοδος Goldfarb	54
2.6.1.3	Η μέθοδος των Moré & Sorensen	55
2.7	Οι μέθοδοι περιοχής εμπιστοσύνης	57
2.8	Το Υποπρόβλημα Περιοχής Εμπιστοσύνης	60
2.8.1	Η τυπική περίπτωση	63
2.8.1.1	Η μέθοδος Newton στην τυπική περίπτωση	67
2.8.2	Η δύσκολη περίπτωση	72
2.8.3	Ο Αλγόριθμος επίλυσης του ΥΠΕ	75
3	Οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης	79
3.1	Εισαγωγή	79
3.2	Το χαρακτηριστικό πολυώνυμο	83
3.3	Ο αντίστροφος του πίνακα $B + \mu I$	92
3.4	Τα ιδιοδιανύσματα των διακριτών ιδιοτιμών	95
3.4.1	Γραμμικώς εξαρτημένα διανύσματα ενημέρωσης . .	95
3.4.2	Γραμμικώς ανεξάρτητα διανύσματα ενημέρωσης . . .	96
3.5	Η SR1 ενημέρωση ελάχιστης μνήμης	104
3.6	Η BFGS ενημέρωση ελάχιστης μνήμης	107
3.7	Η DFP ενημέρωση ελάχιστης μνήμης	112
3.8	Η PSB ενημέρωση ελάχιστης μνήμης	118
3.9	Συμπεράσματα	123
4	Η BFGS ενημέρωση μικρής μνήμης	125
4.1	Οι 1BFGS και 2BFGS ενημερώσεις	125
4.2	Η 1BFGS ενημέρωση	127
4.2.1	Το χαρακτηριστικό πολυώνυμο της 1BFGS ενημέρω- σης	129

4.2.2	Ο πίνακας $(B_{k+1}^{(1)} + \mu I)^{-1}$	133
4.2.3	Τα ιδιοδιανύσματα των ακραίων ιδιοτιμών	135
4.3	Η 2BFGS ενημέρωση	137
4.3.1	Οι ιδιοτιμές της 2BFGS ενημέρωσης	138
4.3.2	Ο πίνακας $(B_{k+1}^{(2)} + \mu I)^{-1}$	145
4.3.3	Ιδιοδιανύσματα διακριτών ιδιοτιμών	148
4.4	Συμπεράσματα	150
5	Επιλύοντας το Υποπρόβλημα Περιοχής Εμπιστοσύνης	153
5.1	Η προτεινόμενη μέθοδος	153
5.2	Η τυπική περίπτωση	155
5.2.1	Εφαρμογή της γενικής PSB ενημέρωσης ελάχιστης μνήμης	156
5.2.1.1	Η SR1 ενημέρωση ελάχιστης μνήμης	160
5.2.1.2	Η BFGS ενημέρωση ελάχιστης μνήμης	161
5.2.1.3	Η DFP ενημέρωση ελάχιστης μνήμης	162
5.2.1.4	Η PSB ενημέρωση ελάχιστης μνήμης	163
5.2.2	Η 1BFGS ενημέρωση στην τυπική περίπτωση	164
5.2.3	Η συνάρτηση βήματος στην 2BFGS ενημέρωση	166
5.3	Η δύσκολη περίπτωση	170
5.3.1	Η γενική PSB ενημέρωση ελάχιστης μνήμης	171
5.3.1.1	Η SR1 ενημέρωση ελάχιστης μνήμης	173
5.3.1.2	Η BFGS ενημέρωση ελάχιστης μνήμης	174
5.3.1.3	Η DFP ενημέρωση ελάχιστης μνήμης	175
5.3.1.4	Η PSB ενημέρωση ελάχιστης μνήμης	176
5.3.2	Η 1BFGS ενημέρωση στη δύσκολη περίπτωση	177
5.3.3	Η 2BFGS ενημέρωση στη δύσκολη περίπτωση	178
5.4	Ο προτεινόμενος αλγόριθμος επίλυσης του ΥΠΕ	179
5.5	Συμπεράσματα	184
6	Αριθμητικά αποτελέσματα του ΥΠΕ	187
6.1	Επίλυση τυχαίων στιγμιотύπων του ΥΠΕ	187
6.1.1	Χρησιμοποιώντας τις Quasi-Newton ενημερώσεις ε- λάχιστης μνήμης	188
6.1.1.1	Η τυπική περίπτωση	191
6.1.1.2	Η δύσκολη περίπτωση	200

6.1.2	Συγκριτικά αποτελέσματα για την BFGS ενημέρωση ελάχιστης μνήμης	209
6.1.2.1	Τυχαία στιγμιότυπα στην τυπική περίπτωση	211
6.1.2.2	Τυχαία στιγμιότυπα στη δύσκολη περίπτωση	220
6.1.3	Συγκριτικά αποτελέσματα για τις 1BFGS και 2BFGS ενημερώσεις	223
6.2	Ο προτεινόμενος αλγόριθμος βελτιστοποίησης	229
6.3	Συμπεράσματα	239
7	Μια καμπυλόγραμμη αναζήτηση ελάχιστης μνήμης	243
7.1	Το ζεύγος κατευθύνσεων μείωσης	243
7.2	Η περίπτωση του θετικά ορισμένου πίνακα	245
7.3	Η περίπτωση του αόριστου πίνακα	249
7.4	Ο προτεινόμενος αλγόριθμος	255
7.5	Συμπεράσματα	262
8	Αποτίμηση του αλγορίθμου CMBFGS	265
8.1	Βελτιστοποίηση συναρτήσεων της βιβλιοθήκης CUTER	265
8.2	Ο CMBFGS στην εκπαίδευση νευρωνικών δικτύων	273
8.2.1	Το πρόβλημα του καρκίνου του μαστού	274
8.2.2	Το πρόβλημα του διαβήτη	276
8.2.3	Σύνολο SPECT δεδομένων καρδιάς	277
8.2.4	Το πρόβλημα της έγκρισης πίστωσης στην Αυστραλία	279
8.2.5	Το πρόβλημα κολοβακτηριδίων	281
8.2.6	Το πρόβλημα της μαγιάς	282
8.3	Εφαρμογή σε βιοϊατρικά δεδομένα	284
8.4	Συμπεράσματα	288
	Βιβλιογραφία	289
	Ευρετήριο	298
	Κατάλογος Δημοσιεύσεων	301

Κατάλογος Σχημάτων

1.1	Παραδείγματα τοπικού και ολικού ελαχιστοποιητή σε μία διάσταση	22
2.1	Συνθήκες βήματος δεύτερης τάξης	56
2.2	Λύση του ΥΠΕ για διαφορετικές ακτίνες $\Delta_1, \Delta_2, \Delta_3$	63
2.3	$\ d(\lambda)\ $ ως συνάρτηση του λ	65
2.4	Γράφημα της $\ d(\lambda)\ ^2$ για θετικά ορισμένο πίνακα	66
2.5	Γράφημα της $\ d(\lambda)\ ^2$ για μη θετικά ορισμένο πίνακα	67
2.6	Η συνάρτηση $1/\ d(\lambda)\ $	68
2.7	Η δύσκολη περίπτωση: $\ d(\lambda)\ < \Delta$ για $\lambda \in (-\lambda_1, \infty)$	73
2.8	Γράφημα της $\ d(\lambda)\ ^2$ στη δύσκολη περίπτωση	75
6.1	Τυπική περίπτωση: μέσος αριθμός επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα	196
6.2	Τυπική περίπτωση: μέση ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	198
6.3	Τυπική περίπτωση: μέση σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	200
6.4	Δύσκολη περίπτωση: μέσος αριθμός επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα	204
6.5	Δύσκολη περίπτωση: μέση ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	206
6.6	Δύσκολη περίπτωση: μέση σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	208
6.7	Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS, GQTPAR και GLTR, σε σύνολο 12000 στιγμιοτύπων μεσαίας διάστασης ($n = 100, 500$ και 1000).	216

6.8	Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS και GLTR, σε σύνολο 12000 στιγμιοτύπων μεγάλης κλίμακας ($n = 10^4, 10^5$ και 10^6).	217
6.9	Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS και GLTR, σε σύνολο 15000 στιγμιοτύπων πολύ μεγάλης κλίμακας ($n = 1 \cdot 10^6, 2 \cdot 10^6, 3 \cdot 10^6, \dots, 15 \cdot 10^6$).	219
6.10	Δύσκολη περίπτωση: Προφίλ απόδοσης για τους MBFGS και GQTPAR, σε σύνολο 9000 στιγμιοτύπων μεσαίας κλίμακας ($n = 100, 500$ και 1000).	223
6.11	Μέσος όρος επαναλήψεων για μεσαίας διάστασης προβλήματα.	232
6.12	Μέσος όρος επαναλήψεων για μεγάλης διάστασης προβλήματα.	237
6.13	Ολικός CPU χρόνος για μεγάλης διάστασης προβλήματα.	238
8.1	Προφίλ απόδοσης των CMBFGS και LBFGS, σε 1000 διαστάσεις	270
8.2	Προφίλ απόδοσης των CMBFGS και LBFGS, σε 10 000 διαστάσεις	271
8.3	Προφίλ απόδοσης των CMBFGS και LBFGS, σε 100 000 διαστάσεις	272
8.4	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα του καρκίνου του μαστού.	275
8.5	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα του διαβήτη.	277
8.6	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το SPECT πρόβλημα καρδιάς.	278
8.7	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα της αυστραλιανής έγκρισης πίστωσης.	280
8.8	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα των κολοβακτηριδίων.	282
8.9	Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα της μαγιάς.	283

Κατάλογος Αλγορίθμων

1.1	Ο αλγόριθμος της κατά Cholesky παραγοντοποίησης πινάκων	10
2.1	Η μέθοδος Newton	32
2.2	Μέθοδος γραμμικής αναζήτησης	48
2.3	Ο αλγόριθμος καμπυλόγραμμης αναζήτησης του Goldfarb .	55
2.4	Γενικός Αλγόριθμος περιοχής εμπιστοσύνης	59
2.5	Η μέθοδος Newton στην τυπική περίπτωση	69
2.6	Ο αλγόριθμος των Moré & Sorensen για την επίλυση του ΥΠΕ	77
5.1	Υπολογισμός του δοκιμαστικού βήματος	180
7.1	Ο αλγόριθμος CMBFGS	257

Κατάλογος Πινάκων

3.1	Απαιτούμενα εσωτερικά γινόμενα	104
3.2	Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στη BFGS ενημέρωση ελάχιστης μνήμης	111
3.3	Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στη DFP ενημέρωση ελάχιστης μνήμης	118
3.4	Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στην PSB ενημέρωση ελάχιστης μνήμης	123
6.1	Τυπική περίπτωση: επιτυχώς λυμένα στιγμιότυπα	191
6.2	Τυπική περίπτωση: λυμένα στιγμιότυπα στην LSTRS	192
6.3	Τυπική περίπτωση: λυμένα στιγμιότυπα στην GQTPAR	193
6.4	Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα	194
6.5	Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα	195
6.6	Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	196
6.7	Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	197
6.8	Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	198
6.9	Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	199
6.10	Δύσκολη περίπτωση: επιτυχώς λυμένα στιγμιότυπα	201

6.11 Δύσκολη περίπτωση: επιτυχώς λυμένα στιγμιότυπα στην LSTRS	202
6.12 Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιό- τυπα	203
6.13 Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα	203
6.14 Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα . . .	204
6.15 Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	205
6.16 Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα	207
6.17 Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα .	207
6.18 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 100$ σε 4000 στιγμιότυπα	212
6.19 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 500$ σε 4000 στιγμιότυπα	212
6.20 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 1\,000$ σε 4000 στιγμιότυπα	212
6.21 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 10\,000$ σε 4000 στιγμιότυπα	213
6.22 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 100\,000$ σε 4000 στιγμιότυπα	213
6.23 Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 1\,000\,000$ σε 4000 στιγμιότυπα	214
6.24 Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n =$ 100 σε 3000 στιγμιότυπα	221
6.25 Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n =$ 500 σε 3000 στιγμιότυπα	221
6.26 Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n =$ 1\,000 σε 3000 στιγμιότυπα	222
6.27 Συγκριτικά αποτελέσματα για την τυπική περίπτωση ($m =$ 1), 100 στιγμιότυπα ανά διάσταση	226

6.28	Συγκριτικά αποτελέσματα για τη δύσκολη περίπτωση ($m = 1$), 100 στιγμιότυπα ανά διάσταση	226
6.29	Συγκριτικά αποτελέσματα για την τυπική περίπτωση ($m = 2$), 100 στιγμιότυπα ανά διάσταση	227
6.30	Συγκριτικά αποτελέσματα για τη δύσκολη περίπτωση ($m = 2$), 100 στιγμιότυπα ανά διάσταση	227
6.31	Συγκριτικά αποτελέσματα για μεγάλης διάστασης ΥΠΕ ($m = 1$), 100 στιγμιότυπα ανά διάσταση	228
6.32	Συγκριτικά αποτελέσματα για μεγάλης διάστασης ΥΠΕ ($m = 2$), 100 στιγμιότυπα ανά διάσταση	228
6.33	Προβλήματα βελτιστοποίησης	230
6.34	Συγκριτικά αποτελέσματα για μεσαίας διάστασης προβλήματα	231
6.35	Συγκριτικά αποτελέσματα για $n = 1000$	234
6.36	Συγκριτικά αποτελέσματα για $n = 10\,000$	235
6.37	Συγκριτικά αποτελέσματα για $n = 100\,000$	236
7.1	Ποσότητες υπολογισμού της p_{k+1} όταν τα s_k, y_k είναι γραμμικά ανεξάρτητα	247
7.2	Ποσότητες υπολογισμού της p_{k+1} όταν τα s_k, y_k είναι γραμμικά εξαρτημένα	248
7.3	Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της p	248
7.4	Ποσότητες υπολογισμού της d_{k+1} όταν τα s_k, y_k είναι γραμμικά ανεξάρτητα	252
7.5	Ποσότητες υπολογισμού της d_{k+1} όταν τα s_k, y_k είναι γραμμικά εξαρτημένα	253
7.6	Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της d όταν $y_k \neq \kappa s_k$	254
7.7	Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της d_{k+1} όταν $y_k = \kappa s_k$	254
8.1	Προβλήματα	266
8.2	Αποτυχίες για $n = 1\,000$	267
8.3	Αποτυχίες για $n = 10\,000$	268
8.4	Αποτυχίες για $n = 100\,000$	268

8.5	Ολικός CPU χρόνος για τα επιτυχώς λυμένα προβλήματα και συνολικές αποτυχίες σε όλες τις διαστάσεις	269
8.6	Συνολικός αριθμός υπολογιζόμενων κλίσεων αρνητικής καμπυλότητας στα επιτυχώς λυμένα προβλήματα	269
8.7	Αρχιτεκτονικές νευρωνικών δικτύων	286
8.8	Απόδοση γενίκευσης (%) νευρωνικών δικτύων χρησιμοποιώντας τη CMBFGS μέθοδο.	286
8.9	Απόδοση γενίκευσης (%) νευρωνικών δικτύων χρησιμοποιώντας την LBFGS μέθοδο.	287

Εισαγωγή

Στο εισαγωγικό αυτό κεφάλαιο δίνουμε βασικές έννοιες της βελτιστοποίησης μη γραμμικών συναρτήσεων, καθώς και στοιχεία γραμμικής άλγεβρας, τα οποία είναι απαραίτητα για την παρουσίαση των μεθόδων βελτιστοποίησης που διαπραγματεύεται η παρούσα διατριβή. Για περισσότερη θεωρητική μελέτη ο αναγνώστης μπορεί να ανατρέξει σε κλασσικά συγγράμματα [14, 29, 68, 71, 84] που μελετούν ενδελεχώς την εν λόγω περιοχή.

1.1 Το πρόβλημα της βελτιστοποίησης

Η βελτιστοποίηση είναι ένας νέος κλάδος των Εφαρμοσμένων Μαθηματικών, των Υπολογιστικών Μαθηματικών και της Επιχειρησιακής Έρευνας. Προβλήματα βελτιστοποίησης απαντώνται σε πολλά επιστημονικά πεδία, όπως στις Φυσικές επιστήμες, στην Οικονομία και τη Μηχανική, καθώς και σε θέματα που αφορούν στρατιωτικές και διαστημικές τεχνολογίες.

Αν και η βελτιστοποίηση χρονολογείται από πολύ παλιά, δεν αποτελούσε ανεξάρτητο πεδίο έρευνας μέχρι τα τέλη του 1940, όπου ο G.B. Dantzig [26] παρουσίασε τον γνωστό αλγόριθμο Simplex για προβλήματα γραμμικού προγραμματισμού. Στα τέλη της δεκαετίας του 1950, η θεμελίωση και εφαρμογή των μεθόδων συζυγών κλίσεων καθώς και των Quasi-Newton μεθόδων, αποτέλεσε εφελκτήριο ανάπτυξης του μη γραμμικού προγραμματισμού, ως ανεξάρτητο πεδίο έρευνας. Στις μέρες μας, υπάρχουν μέθοδοι βελτιστοποίησης που μπορούν να λύσουν δύσκολα και μεγάλης κλίμακας προβλήματα, με αποτέλεσμα η βελτιστοποίηση να αποτελεί απαραίτητο εργαλείο σε διάφορους τομείς της επιστήμης.

Η θεωρία και οι μέθοδοι βελτιστοποίησης ασχολούνται με την εύρεση της βέλτιστης λύσης προβλημάτων που αναπαρίστανται μέσω ενός μαθηματικού μοντέλου (αιτιοκρατικό ή στοχαστικό). Η ανάπτυξη του μοντέλου

αποσκοπεί στην μετατροπή όλων των βασικών συνιστωσών του προβλήματος σε μαθηματικές ή/και λογικές σχέσεις. Οι πιο βασικές συνιστώσες είναι η *αντικειμενική συνάρτηση* και οι *άγνωστοι* ή *μεταβλητές*, ενώ πολλές φορές υπάρχει κι ένα σύνολο *περιορισμών*. Οι συμβολισμοί που χρησιμοποιούνται σε ένα μαθηματικό πρόβλημα είναι οι ακόλουθοι:

- x είναι το διάνυσμα των μεταβλητών ή αγνώστων.
- $f(x)$ είναι η αντικειμενική συνάρτηση.
- $c(x)$ είναι η διανυσματική συνάρτηση των περιορισμών που πρέπει να ικανοποιούν οι μεταβλητές.

Δεδομένου ότι ένα πρόβλημα μεγιστοποίησης είναι ισοδύναμο με ένα πρόβλημα ελαχιστοποίησης ($\max_x f(x) = \min_x -f(x)$), στη συνέχεια με τον όρο βελτιστοποίηση θα εννοούμε την ελαχιστοποίηση της αντικειμενικής συνάρτησης f . Συνεπώς, ένα πρόβλημα βελτιστοποίησης έχει τη γενική μορφή

$$\min_{x \in \mathcal{D}} f(x), \quad (1.1)$$

όπου $x \in \mathbb{R}^n$ είναι το διάνυσμα των μεταβλητών, $f(x)$ η αντικειμενική συνάρτηση και $\mathcal{D} \subseteq \mathbb{R}^n$ η εφικτή περιοχή που προκύπτει από το σύνολο περιορισμών του προβλήματος. Στην περίπτωση όπου η εφικτή περιοχή είναι όλο το $\mathbb{R}^n$, δηλαδή $\mathcal{D} = \mathbb{R}^n$, το πρόβλημα βελτιστοποίησης (1.1) ονομάζεται *πρόβλημα βελτιστοποίησης χωρίς περιορισμούς* και έχει τη γενική μορφή

$$\min_{x \in \mathbb{R}^n} f(x).$$

Στην αντίθετη περίπτωση, όπου $\mathcal{D} \subset \mathbb{R}^n$, η γενική μορφή ενός *προβλήματος βελτιστοποίησης με περιορισμούς*, είναι η ακόλουθη:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{υπό τους περιορισμούς} \quad \begin{cases} c_i(x) = 0, & i \in \mathcal{E}; \\ c_i(x) \geq 0, & i \in \mathcal{I}, \end{cases}$$

όπου $\mathcal{E}$ και $\mathcal{I}$ είναι τα σύνολα δεικτών των περιορισμών ισότητας (equalities) και ανισότητας (inequalities), αντίστοιχα και όπου $c_i(x)$, $i \in \mathcal{E} \cup \mathcal{I}$, είναι οι συναρτήσεις των περιορισμών (constraints). Όταν τόσο η αντικειμενική συνάρτηση $f(x)$, όσο και οι περιορισμοί $c_i(x)$ είναι γραμμικές συναρτήσεις, τότε το πρόβλημα βελτιστοποίησης ονομάζεται γραμμικό. Σε αντίθετη

περίπτωση, το πρόβλημα καλείται μη γραμμικό.

Εν γένει, για την επίλυση του προβλήματος (1.1) εφαρμόζονται επαναληπτικές διαδικασίες (αλγόριθμοι) οι οποίες ξεκινούν με μια αρχική εκτίμηση της λύσης και δημιουργούν μια ακολουθία από συνεχώς βελτιωμένες εκτιμήσεις της, έως ότου φτάσουν στη βέλτιστη λύση. Κάθε τέτοιος αλγόριθμος θα πρέπει να εγγυάται ότι η ακολουθία των επαναλήψεων που δημιουργούνται από αυτόν, συγκλίνει στη βέλτιστη λύση του εκάστοτε προβλήματος. Αυτή η απαίτηση ικανοποιείται μέσω της θεωρίας σύγκλισης της μαθηματικής μεθόδου που υλοποιεί ο αλγόριθμος. Ο έλεγχος βελτιστότητας της παραγόμενης λύσης επιτυγχάνεται εφόσον ικανοποιούνται ικανές και αναγκαίες συνθήκες, γνωστές ως *συνθήκες βελτιστότητας*. Επιπλέον, η μη ικανοποίηση των συνθηκών βελτιστότητας στην τρέχουσα επανάληψη, συνάγει χρήσιμες πληροφορίες για τον εκάστοτε αλγόριθμο προκειμένου να βελτιώσει την τρέχουσα εκτίμηση της λύσης.

Η στρατηγική που χρησιμοποιείται για την αναζήτηση της βέλτιστης λύσης ανά επανάληψη, χαρακτηρίζει το είδος του αλγορίθμου (π.χ., γραμμική αναζήτηση, καμπυλόγραμμη αναζήτηση, κτλ.). Η πλειονότητα των στρατηγικών αναζήτησης χρησιμοποιούν τις τιμές της αντικειμενικής συνάρτησης, των περιορισμών, αν υπάρχουν, και πιθανώς της πρώτης και δεύτερης τάξης παραγώγου, προκειμένου να ορίσουν το επόμενο σημείο της ακολουθίας των επαναλήψεων. Επίσης, πρέπει να σημειωθεί ότι μερικοί αλγόριθμοι χρησιμοποιούν πληροφορία καμπυλότητας της αντικειμενικής συνάρτησης από προηγούμενες επαναλήψεις (π.χ., Quasi-Newton μέθοδοι), προκειμένου να προσδιορίσουν την κατεύθυνση αναζήτησής τους, εντός του εφικτού χωρίου $\mathcal{D}$.

1.2 Μαθηματικό υπόβαθρο

Στη συνέχεια παρατίθενται οι βασικές έννοιες που απαιτούνται για τη μελέτη των μεθόδων βελτιστοποίησης που προτείνουμε στην παρούσα διατριβή, δηλαδή στοιχεία γραμμικής άλγεβρας και πραγματικής ανάλυσης. Οι έννοιες αυτές αποτελούν τα θεμελιώδη εργαλεία για την ανάλυση και την υλοποίηση των αλγορίθμων βελτιστοποίησης. Επίσης, χρησιμοποιούνται τόσο στον έλεγχο βελτιστότητας της λύσης, όσο και στη θεωρία σύγκλισης των αλγορίθμων. Οι αποδείξεις των θεωρημάτων που αναφέρονται μπορούν να βρεθούν στις αναφορές που παρατέθηκαν στην αρχή αυτού του

κεφαλαίου.

1.2.1 Διανύσματα και πίνακες

Στη συνέχεια, το σύμβολο $\mathbb{R}^n$ θα υποδηλώνει τον πραγματικό Ευκλείδειο n -διάστατο χώρο των διανυσμάτων

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$

όπου κάθε x_i , με $i = 1, \dots, n$, είναι η i -συνιστώσα του διανύσματος x . Για κάθε $x \in \mathbb{R}^n$, θα συμβολίζουμε με x^T το *ανάστροφο* διάνυσμα του x , δηλαδή τον πίνακα γραμμή $x^T = (x_1, x_2, \dots, x_n)$.

Ορισμός 1.1. Έστω τα διανύσματα $x, y \in \mathbb{R}^n$. Το εσωτερικό γινόμενο των x και y , ορίζεται ως

$$x^T y = \sum_{i=1}^n x_i y_i.$$

Αν το εσωτερικό γινόμενο είναι μηδέν, δηλαδή $x^T y = 0$, τότε τα διανύσματα x και y καλούνται *ορθογώνια*.

Μία ορθογώνια διάταξη με m γραμμές και n στήλες και πλήθος στοιχείων $m \cdot n$, λέγεται $m \times n$ *πίνακας*. Ένας πραγματικός $m \times n$ πίνακας $A \in \mathbb{R}^{m \times n}$, έχει τη μορφή

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

και δηλώνεται με $A = (a_{ij})$. Ο *ανάστροφος* πίνακας, A^T , είναι ένας $n \times m$ πίνακας στον οποίον έχουν αντιστραφεί οι στήλες με τις γραμμές του A . Ο *αντίστροφος* πίνακας του A (αν υπάρχει) θα συμβολίζεται με A^{-1} . Ο *Moore-Penrose ψευδοαντίστροφος* ή *γενικευμένος αντίστροφος*, του πίνακα

A , συμβολίζεται με $A^\dagger$ και ορίζεται ως

$$A^\dagger = A^T (AA^T)^{-1}.$$

Ο μοναδιαίος πίνακας ορίζεται ως

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Ένας πίνακας $n \times n$ ονομάζεται *τετραγωνικός*. Αν για ένα τετραγωνικό πίνακα ισχύει $A^T = A$, τότε ο A ονομάζεται *συμμετρικός*. Ένας συμμετρικός πίνακας έχει μόνο πραγματικές ιδιοτιμές, ενώ τα ιδιοδιανύσματά του μπορούν να επιλεγούν ορθοκανονικά. Ένας τετραγωνικός πραγματικός πίνακας A ονομάζεται *ορθογώνιος* αν $AA^T = A^T A = I$.

Ένας πίνακας A για τον οποίο ισχύει $a_{ij} = 0$, όταν $i \neq j$ καλείται *διαγώνιος* και συμβολίζεται ως

$$A = \text{diag}(a_{11}, \dots, a_{nn}) = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Ο πίνακας A λέγεται *τριδιαγώνιος* αν ισχύει $a_{ij} = 0$, για $|i - j| > 1$, *άνω τριγωνικός* αν ισχύει ότι $a_{ij} = 0$, για $j < i$ και *κάτω τριγωνικός* αν $a_{ij} = 0$, για $j > i$.

Ορισμός 1.2. Έστω συμμετρικός πίνακας $A \in \mathbb{R}^{n \times n}$. Ο πίνακας A ονομάζεται *θετικά ορισμένος* ($A > 0$) αν $x A^T x > 0$ για κάθε $x \in \mathbb{R}^n$, με $x \neq 0$. Ο πίνακας A λέγεται *θετικά ημιορισμένος* ($A \geq 0$) αν ισχύει $x A^T x \geq 0$ για κάθε $x \in \mathbb{R}^n$. Ο πίνακας A είναι *αρνητικά ορισμένος* ($A < 0$) ή *αρνητικά ημιορισμένος* ($A \leq 0$), αν ο $-A$ είναι θετικά ορισμένος ή θετικά ημιορισμένος, αντίστοιχα. Ο πίνακας A ονομάζεται *αόριστος* αν δεν είναι ούτε θετικά και ούτε αρνητικά ημιορισμένος (έχει και θετικές και αρνητικές ιδιοτιμές).

Ορισμός 1.3. Μια μετρική ή νόρμα διανύσματος $x \in \mathbb{R}^n$ είναι μια πραγματική συνάρτηση από τον $\mathbb{R}^n$ στον $\mathbb{R}$ με τις ακόλουθες ιδιότητες:

1. $\|x\| \geq 0$ για όλα τα $x \in \mathbb{R}^n$ και $\|x\| = 0$ αν και μόνο αν $x = 0$,
2. $\|ax\| = |a|\|x\|$ για όλα τα $a \in \mathbb{R}$ και $x \in \mathbb{R}^n$,
3. $\|x + y\| \leq \|x\| + \|y\|$ για όλα τα $x, y \in \mathbb{R}^n$.

Πολύ γνωστές νόρμες είναι οι ℓ_p -νόρμες, οι οποίες ορίζονται ως

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad 1 \leq p < \infty.$$

Οι συνηθέστερες από αυτές είναι οι ακόλουθες:

1. η ℓ_1 -νόρμα, $\|x\|_1 = \sum_{i=1}^n |x_i|$,
2. η ℓ_2 -νόρμα ή *Ευκλείδεια* νόρμα, $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} = \sqrt{x^T x}$,
3. η άπειρη νόρμα, $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$.

Όλες οι μετρικές στον $\mathbb{R}^n$ είναι ισοδύναμες, με την έννοια ότι κάθε μία είναι άνω και κάτω φραγμένη από ένα πολλαπλάσιο των άλλων. Συγκεκριμένα, για τις τρεις παραπάνω νόρμες, ισχύουν οι παρακάτω σχέσεις:

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1, \quad \|x\|_1 \leq \sqrt{n}\|x\|_2 \quad \text{και} \quad \|x\|_2 \leq \sqrt{n}\|x\|_\infty.$$

Ακόμα, για τη Ευκλείδεια νόρμα ισχύει η ανισότητα Cauchy-Schwarz:

$$|x^T y| \leq \|x\|_2 \|y\|_2, \quad \text{για όλα τα } x, y \in \mathbb{R}^n.$$

Τέλος, για το εσωτερικό γινόμενο δύο μη μηδενικών διανυσμάτων $x, y \in \mathbb{R}^n$, ισχύει ότι

$$x^T y = \|x\| \|y\| \cos \theta,$$

όπου θ η γωνία που σχηματίζουν τα διανύσματα x, y .

Ορισμός 1.4. Μια μετρική ή νόρμα πίνακα ενός $m \times n$ πίνακα A , η οποία συμβολίζεται με $\|A\|$, είναι μια συνάρτηση από τον $\mathbb{R}^{m \times n}$ στον $\mathbb{R}$ με τις ακόλουθες ιδιότητες:

1. $\|A\| \geq 0$ για όλους τους πίνακες $A \in \mathbb{R}^{m \times n}$ και $\|A\| = 0$ αν και μόνο αν $A = 0$,

2. $\|aA\| = |a|\|A\|$ για όλα τα $a \in \mathbb{R}$ και για όλους τους πίνακες $A \in \mathbb{R}^{m \times n}$,
3. $\|A + B\| \leq \|A\| + \|B\|$ για όλους τους πίνακες $A, B \in \mathbb{R}^{m \times n}$.

Οι γνωστές νόρμες πίνακα είναι η Frobenius νόρμα

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$$

και οι ℓ_p -νόρμες

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p},$$

όπου για $p = 1, 2$ και ∞ , έχουμε αντίστοιχα ότι

$$\begin{aligned} \|A\|_1 &= \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|, \\ \|A\|_2 &= \sqrt{\lambda_{\max}(A^T A)}, \text{ όπου } \lambda_{\max}(\cdot) \text{ η μεγαλύτερη ιδιοτιμή του } A, \\ \|A\|_\infty &= \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|. \end{aligned}$$

1.2.1.1 Ακολουθίες και ρυθμοί σύγκλισης. Μια ακολουθία διανυσμάτων $\{x_k\}$ στον $\mathbb{R}^n$ λέγεται ότι *συγκλίνει* στο σημείο $x^* \in \mathbb{R}^n$, αν ισχύει ότι

$$\lim_{k \rightarrow \infty} \|x_k - x^*\| = 0.$$

Λέμε ότι το $x^* \in \mathbb{R}^n$ είναι *σημείο συσσώρευσης* ή *οριακό σημείο* μιας ακολουθίας $\{x_k\}$ αν υπάρχει μια άπειρη υπακολουθία ακεραίων $k_1, k_2, \dots$ τέτοια ώστε

$$\lim_{i \rightarrow \infty} x_{k_i} = x^*.$$

Μια ακολουθία πινάκων $\{A_k\}$ λέμε ότι *συγκλίνει* στον A , αν ισχύει ότι

$$\lim_{k \rightarrow \infty} \|A_k - A\| = 0.$$

Η μετρική που χρησιμοποιείται δεν παίζει ρόλο στη σύγκλιση ακολουθιών, γιατί όλες οι μετρικές που χρησιμοποιούνται σε πεπερασμένης διάστασης χώρους είναι ισοδύναμες.

Ο *ρυθμός σύγκλισης* αποτελεί ένα μέτρο εκτίμησης της ταχύτητα σύγκλισης μιας ακολουθίας. Το μαθηματικό αυτό μέτρο υπολογίζει την υπολογιστική αποδοτικότητα ενός επαναληπτικού αλγορίθμου. Ένας από τους πιο αποτελεσματικούς τρόπους αξιολόγησης του ρυθμού σύγκλισης, είναι να κάνουμε σύγκριση ανάμεσα στην πρόοδο που έχουμε σε κάθε βήμα ενός αλγορίθμου, με την πρόοδο που είχαμε στο προηγούμενο βήμα. Λέμε ότι μια ακολουθία έχει τάξη σύγκλισης r , με $r \geq 1$, στο x^* αν

$$\alpha \leq \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^r} \leq \beta, \quad \text{με } \alpha, \beta > 0.$$

Συγκεκριμένα, όταν $r = 1$ ο ρυθμός σύγκλισης λέγεται *γραμμικός*. Όταν $r = 2$, έχουμε σύγκλιση 2ης τάξεως, η οποία ονομάζεται *τετραγωνική σύγκλιση*. Επίσης, αν ισχύει ότι

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|} = 0,$$

ο ρυθμός σύγκλισης ονομάζεται *υπεργραμμικός*.

1.2.1.2 Υπόχωροι. Δοθέντος του Ευκλείδειου χώρου $\mathbb{R}^n$, το υποσύνολο $\mathcal{S} \subset \mathbb{R}^n$ είναι ένας *υπόχωρος* του $\mathbb{R}^n$ όταν $\alpha x + \beta y \in \mathcal{S}$, για όλα τα $\alpha, \beta \in \mathbb{R}$ και για κάθε $x, y \in \mathcal{S}$. Για κάθε υπόχωρο $\mathcal{S} \subset \mathbb{R}^n$, τα διανύσματα $s_1, s_2, \dots, s_m$ που ανήκουν στον $\mathcal{S}$ καλούνται *γραμμικά ανεξάρτητα* αν δεν υπάρχουν πραγματικοί αριθμοί $\alpha_1, \alpha_2, \dots, \alpha_m$, τέτοιοι ώστε $\alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_m s_m = 0$, εκτός αν $\alpha_1 = \alpha_2 = \dots = 0$. Για κάθε υπόχωρο $\mathcal{S} \subset \mathbb{R}^n$, τα διανύσματα $s_1, s_2, \dots, s_m$ που ανήκουν στον $\mathcal{S}$ ονομάζονται *γραμμικά εξαρτημένα* αν υπάρχουν πραγματικοί αριθμοί $\alpha_1, \alpha_2, \dots, \alpha_m$, όχι όλοι μηδέν, τέτοιοι ώστε $\alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_m s_m = 0$. Ισοδύναμα, ένα από τα διανύσματα s_i , $i = 1, \dots, m$ είναι γραμμικός συνδυασμός των άλλων. Η *τάξη* ενός πίνακα A συμβολίζεται με $\text{rank} A$ και είναι ο μεγαλύτερος αριθμός στηλών του πίνακα που συνιστούν ένα γραμμικά ανεξάρτητο σύνολο.

1.2.1.3 Ο τύπος των Sherman-Morrison-Woodbury. Αν ένας αντιστρέ-

ψιμος τετραγωνικός πίνακας $A \in \mathbb{R}^{n \times n}$ επιδέχεται μια ενημέρωση πρώτης τάξης, της μορφής

$$\bar{A} = A + ab^T,$$

όπου $a, b \in \mathbb{R}^n$, τότε αν ο $\bar{A}$ είναι και αυτός αντιστρέψιμος και ισχύει ότι

$$\bar{A}^{-1} = A^{-1} - \frac{A^{-1}ab^TA^{-1}}{1 + b^TA^{-1}a}.$$

Ο παραπάνω τύπος μπορεί να επεκταθεί και για μεγαλύτερης τάξης ενημερώσεις. Πράγματι, αν U , C και V πίνακες διάστασης $n \times k$, $k \times k$ και $k \times n$, αντίστοιχα. Τότε,

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}V^TA^{-1}. \quad (1.2)$$

Ο τύπος (1.2), είναι γνωστός ως τύπος των Sherman-Morrison-Woodbury.

1.2.1.4 Η κατά Cholesky παραγοντοποίηση πινάκων. Η παραγοντοποίηση πινάκων είναι σημαντική στο σχεδιασμό και την ανάλυση αλγορίθμων. Οι αλγόριθμοι βελτιστοποίησης μη γραμμικών συναρτήσεων σχεδόν πάντα απαιτούν να λυθεί τουλάχιστον ένα γραμμικό σύστημα n εξισώσεων με n αγνώστους, της μορφής

$$Ax = b,$$

όπου $A \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^n$ και $x \in \mathbb{R}^n$ το διάνυσμα των αγνώστων. Στην παραγοντοποίηση πινάκων, η βασική ιδέα για να λύσουμε αριθμητικά το γραμμικό σύστημα, είναι η εξής: αναλύουμε τον πίνακα A σε γινόμενο πινάκων, δηλαδή,

$$A = A_1 \cdot A_2 \cdot \dots \cdot A_m$$

και στη συνέχεια λύνουμε τα επιμέρους γραμμικά συστήματα

$$\begin{aligned} A_1 b_1 &= b \\ A_2 b_2 &= b_1 \\ &\vdots \\ A_m b_m &= b_{m-1}, \end{aligned}$$

οπότε το $x = b_m$ είναι η λύση του αρχικού γραμμικού συστήματος $Ax = b$. Ένας πολύ γνωστός τρόπος παραγοντοποίησης συμμετρικών πινάκων που χρησιμοποιείται για την αριθμητική επίλυση γραμμικών συστημάτων της μορφής $Ax = b$, είναι η κατά Cholesky παραγοντοποίηση. Όταν ο πίνακας $A \in \mathbb{R}^{n \times n}$ είναι συμμετρικός και θετικά ορισμένος, τότε μπορεί να αναλυθεί ως

$$A = LL^T,$$

όπου L είναι ένας κάτω τριγωνικός πίνακας. Πρέπει να σημειωθεί ότι στην περίπτωση όπου ο L έχει θετικά διαγώνια στοιχεία, ορίζεται μονοσήμαντα. Αφού παραγοντοποιήσουμε τον $A = LL^T$, λύνουμε το σύστημα $Ly = b$ ως προς y και τέλος λύνουμε το σύστημα $L^T x = y$, ως προς x . Ο Αλγόριθμος 1.1 αναφέρεται μόνο στα κάτω τριγωνικά στοιχεία του A , αφού λόγω συμμετρίας του πίνακα, αρκεί να αποθηκευτούν μόνο αυτά.

Αλγόριθμος 1.1 Ο αλγόριθμος της κατά Cholesky παραγοντοποίησης πινάκων

Δοθέντος συμμετρικού και θετικά ορισμένου πίνακα $A \in \mathbb{R}^{n \times n}$ κάνε:

για $i = 1, 2, \dots, n$;

 Θέσε l_{ii} το $\sqrt{a_{ii}}$;

 για $j = i + 1, i + 2, \dots, n$

 Θέσε l_{ji} το a_{ji}/l_{ii} ;

 για $k = i + 1, i + 2, \dots, j$

 Θέσε a_{jk} το $a_{jk} - l_{ji}l_{ki}$;

 τέλος

 τέλος

τέλος

1.2.2 Ιδιοτιμές και ιδιοδιανύσματα

Σε αυτή την ενότητα δίνουμε τις βασικές ιδιότητες των ιδιοτιμών και των ιδιοδιανυσμάτων, τις οποίες θα χρησιμοποιήσουμε στη συνέχεια. Οι αποδείξεις των θεωρημάτων που αναφέρονται μπορούν να βρεθούν στις αναφορές που παρατέθηκαν στην αρχή αυτού του κεφαλαίου.

Ορισμός 1.5. Μια βαθμωτή ποσότητα λ ονομάζεται ιδιοτιμή ενός $n \times n$ πίνακα A , αν υπάρχει μη μηδενικό n -διάστατο διάνυσμα u , τέτοιο ώστε

$$Au = \lambda u.$$

Το διάνυσμα u ονομάζεται ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ .

Ο αριθμός λ είναι μία ιδιοτιμή του A , αν και μόνο αν $p(\lambda) \equiv \det(A - \lambda I) = 0$. Το σύνολο όλων των ιδιοτιμών του A καλείται φάσμα του A . Ένας πίνακας είναι *μη ιδιάζον* (ομαλός) αν όλες οι ιδιοτιμές του είναι διάφορες του μηδενός. Αντίθετα, *ιδιάζον* λέγεται όταν τουλάχιστον μία ιδιοτιμή του ισούται με μηδέν. Για τους ιδιάζοντες πίνακες μπορούμε να υπολογίσουμε μόνο τον γενικευμένο ψευδοαντίστροφο $A^\dagger$, γιατί ο A^{-1} δεν ορίζεται. Με άλλα λόγια, ο ιδιάζον πίνακας δεν αντιστρέφεται. Αν (λ, u) είναι ένα ιδιοζεύγος ενός μη ιδιάζον πίνακα A , τότε το $(1/\lambda, u)$ είναι το αντίστοιχο ιδιοζεύγος του A^{-1} . Επιπλέον, αν (λ, u) είναι ένα ιδιοζεύγος του πίνακα A , τότε το $(\lambda + \mu, u)$ είναι ένα ιδιοζεύγος του πίνακα $A + \mu I$, όπου μ βαθμωτό μέγεθος. Στη συνέχεια συνοψίζουμε τις βασικές ιδιότητες που ισχύουν για συμμετρικούς πίνακες και τις οποίες θα χρειαστούμε στη συνέχεια.

Έστω $A \in \mathbb{R}^{n \times n}$ συμμετρικός πίνακας. Τότε:

- όλες οι ιδιοτιμές του είναι πραγματικές,
- είναι θετικά ορισμένος, αν και μόνο αν όλες οι ιδιοτιμές του είναι θετικές,
- είναι θετικά ημιορισμένος, αν και μόνο αν όλες οι ιδιοτιμές του είναι μη αρνητικές,
- είναι αρνητικά ορισμένος, αν και μόνο αν όλες οι ιδιοτιμές του είναι αρνητικές,
- είναι αρνητικά ημιορισμένος, αν και μόνο αν όλες οι ιδιοτιμές του είναι μη θετικές,
- είναι αόριστος, αν και μόνο αν έχει τόσο θετικές, όσο και αρνητικές ιδιοτιμές.

Για συμμετρικό πίνακα A , η φασματική του ανάλυση είναι:

$$A = \sum_{i=1}^n \lambda_i u_i u_i^T,$$

όπου $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ είναι οι n πραγματικές ιδιοτιμές του και $u_1, u_2, \dots, u_n$ τα αντίστοιχα ιδιοδιανύσματα. Αυτή η ανάλυση μπορεί να επαναδιατυπωθεί ορίζοντας τους πίνακες U και Λ , όπου ο U είναι ο πίνακας των ιδιοδιανυσμάτων του A και $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$. Τότε,

$$A = U\Lambda U^T.$$

Για θετικά ορισμένους συμμετρικούς πίνακες έχουμε ότι $\|A\|_2 = \lambda_n$, όπου λ_n η μεγαλύτερη ιδιοτιμή του πίνακα A και $\|A^{-1}\|_2 = 1/\lambda_1$, όπου λ_1 η μικρότερη ιδιοτιμή του A .

Έστω A συμμετρικός πίνακας και $x \in \mathbb{R}^n$ μια καλή προσέγγιση ενός ιδιοδιανύσματος του A . Η ακόλουθη ποσότητα

$$R = \frac{x^T A x}{x^T x}$$

ονομάζεται ποσότητα Rayleigh και είναι μια καλή προσέγγιση της ιδιοτιμής που αντιστοιχεί στο x .

Έστω A τετραγωνικός πίνακας με ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$. Τότε, το *ίχνος* του συμβολίζεται με $\text{tr}(A)$ και ορίζεται ως το άθροισμα των διαγώνιων στοιχείων του, δηλαδή

$$\text{tr}(A) = \sum_{i=1}^n \alpha_{ii}.$$

Επίσης, το *ίχνος* ισούται με το άθροισμα των ιδιοτιμών του A , δηλαδή,

$$\text{tr}(A) = \sum_{i=1}^n \lambda_i.$$

Οι ιδιότητες του *ίχνος* που θα φανούν χρήσιμες στην ανάλυσή μας, είναι

οι ακόλουθες:

$$\operatorname{tr}(A) = \operatorname{tr}(A^T) \quad (1.3)$$

$$\operatorname{tr}(aA + bB) = a \operatorname{tr}(A) + b \operatorname{tr}(B) \quad (1.4)$$

$$\operatorname{tr}(uv^T) = \operatorname{tr}(u^T v), \quad (1.5)$$

όπου A, B τετραγωνικοί πίνακες, a, b βαθμωτά μεγέθη και u, v n -διάστατα διανύσματα.

Η *ορίζουσα* ενός τετραγωνικού πίνακα A συμβολίζεται με $\det(A)$ και ισούται με το γινόμενο των ιδιοτιμών του, δηλαδή

$$\det(A) = \prod_{i=1}^n \lambda_i.$$

Οι ιδιότητες των οριζουσών που θα χρειαστούμε, είναι οι ακόλουθες:

$$\det(A^T) = \det(A) \quad (1.6)$$

$$\det(AB) = \det(A) \det(B) \quad (1.7)$$

$$\det(I + xy^T + uv^T) = (1 + y^T x)(1 + v^T u) - (x^T v)(y^T u), \quad (1.8)$$

όπου A, B τετραγωνικοί πίνακες και x, y, u και v , n -διάστατα διανύσματα.

Τέλος, το πολυώνυμο

$$p(\lambda) = \det(A - \lambda I)$$

ονομάζεται *χαρακτηριστικό πολυώνυμο* του πίνακα A , ενώ η εξίσωση $p(\lambda) = 0$, ονομάζεται *χαρακτηριστική εξίσωση* του A . Μέσω της χαρακτηριστικής εξίσωσης, μπορούμε να υπολογίσουμε τις ιδιοτιμές του A . Το παρακάτω θεώρημα είναι γνωστό ως θεώρημα Cayley-Hamilton και δηλώνει ότι ο πίνακας A ικανοποιεί τη χαρακτηριστική του εξίσωση, δηλαδή ο πίνακας $p(A)$ είναι ο μηδενικός πίνακας.

Θεώρημα 1.1. (Cayley-Hamilton) Ένας τετραγωνικός πίνακας A ικανοποιεί τη χαρακτηριστική του εξίσωση· δηλαδή, αν $A \in \mathbb{R}^{n \times n}$ και $p(\lambda)$ το χαρακτηριστικό του πολυώνυμο, τότε $p(A) = 0$.

Το μικρότερου βαθμού πολυώνυμο $q(\lambda)$ για το οποίο ισχύει $q(A) = 0$, ονομάζεται *ελάχιστο πολυώνυμο* του A .

Θεώρημα 1.2. Το χαρακτηριστικό πολυώνυμο $p(\lambda)$ του A διαιρείται με το ελάχιστο πολυώνυμο $q(\lambda)$. Επίσης, αν λ ιδιοτιμή του A , τότε $q(\lambda) = 0$ και αντίστροφα.

Το επόμενο θεώρημα είναι γνωστό ως θεώρημα του Wilkinson [88] ή θεώρημα συνδιαπλοκής ιδιοτιμών (interlacing theorem). Δηλώνει ότι αν ένας πίνακας A ενημερωθεί από έναν πίνακα πρώτης τάξης, τότε οι ιδιοτιμές του αρχικού και του ενημερωμένου πίνακα συνδιαπλέκονται μεταξύ τους.

Θεώρημα 1.3. (θεώρημα συνδιαπλοκής ιδιοτιμών) Έστω A ένας $n \times n$ συμμετρικός πίνακας με ιδιοτιμές $\lambda_n \geq \lambda_{n-1} \geq \dots \geq \lambda_1$. Επίσης, έστω $v \in \mathbb{R}^n$ διάνυσμα, σ βαθμωτό μέγεθος και $A^* = A + \sigma vv^T$ συμμετρικός πίνακας με ιδιοτιμές $\lambda_n^* \geq \lambda_{n-1}^* \geq \dots \geq \lambda_1^*$, ο οποίος διαφέρει κατά ένα πίνακα πρώτης τάξης από τον A . Τότε οι ιδιοτιμές τους λ_i και λ_i^* , με $i = 1, 2, \dots, n$, αντίστοιχα, συνδιαπλέκονται μεταξύ τους ως εξής:

(1) αν $\sigma > 0$, τότε $\lambda_n^* \geq \lambda_n \geq \lambda_{n-1}^* \geq \lambda_{n-1} \geq \dots \geq \lambda_1^* \geq \lambda_1$;

(2) αν $\sigma < 0$, τότε $\lambda_n \geq \lambda_n^* \geq \lambda_{n-1} \geq \lambda_{n-1}^* \geq \dots \geq \lambda_1 \geq \lambda_1^*$.

Σε πολλές μεθόδους βελτιστοποίησης που χρησιμοποιούνται για την επίλυση μη γραμμικών προβλημάτων, κατά τη διάρκεια της όλης διαδικασίας, είναι απαραίτητο να βρούμε ή έστω να προσεγγίσουμε κάποια ιδιοτιμή ενός πίνακα ή ένα ιδιοδιάνυσμα που αντιστοιχεί σε αυτή. Μια συνηθισμένη τακτική είναι να χρησιμοποιήσουμε επαναληπτικές μεθόδους, όπως είναι η μέθοδος της αντίστροφης δύναμης και η μέθοδος Lanczos, τις οποίες θα περιγράψουμε συνοπτικά στη συνέχεια.

1.2.2.1 Η μέθοδος της αντίστροφης δύναμης. Έστω ότι θέλουμε να υπολογίσουμε μια οποιαδήποτε διακεκριμένη ιδιοτιμή λ_j ενός πίνακα A καθώς και το αντίστοιχο ιδιοδιάνυσμα u_j , δοθέντος ότι γνωρίζουμε μία καλή εκτίμηση λ της λ_j . Τότε μπορούμε να εφαρμόσουμε τη μέθοδο της αντίστροφης δύναμης (inverse power method) ή αλλιώς αντίστροφης επανάληψης (inverse iteration method), η οποία προτάθηκε από τον Wielandt [87] και είναι μια γνωστή επαναληπτική μέθοδος για τον υπολογισμό ιδιοτιμών και ιδιοδιανυσμάτων.

Δοθέντος ενός μη μηδενικού διανύσματος u_0 , η αντίστροφη επανάληψη δημιουργεί μια ακολουθία διανυσμάτων u_i , εφαρμόζοντας επαναληπτικά το σχήμα

$$u_i = \left(A - \lambda I \right)^{-1} \frac{u_{i-1}}{\|u_{i-1}\|}, \quad i \geq 1$$

όπου $\hat{\lambda} = \lambda + \epsilon$, $\epsilon \rightarrow 0$ και λ είναι μια διακεκριμένη ιδιοτιμή του A ή μια καλή προσέγγισή της. Η ακολουθία των επαναλήψεων u_i συγκλίνει σε ένα ιδιοδιάνυσμα που αντιστοιχεί στην λ . Συνήθως, το αρχικό διάνυσμα u_0 επιλέγεται να είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$. Επιπλέον, η λύση u_i κανονικοποιείται και αυτή.

Αν λ είναι μια ακριβής ιδιοτιμή του πίνακα A και όχι μια προσέγγισή της, η μέθοδος συγκλίνει σε μία μόνο επανάληψη [52], παρέχοντας ένα κλειστό τύπο για το αντίστοιχο ιδιοδιάνυσμα. Πράγματι, η λύση $(A - \hat{\lambda}I)u_i = u_{i-1}/\|u_{i-1}\|$ είναι πολύ κοντά στην $(A - \hat{\lambda}I)u_i = 0$, λόγω του ότι ο πίνακας $A - \hat{\lambda}I$ είναι ιδιάζον. Έτσι έχουμε ότι $Au_i = \hat{\lambda}u_i \approx \lambda u_i$, που είναι ακριβώς αυτό που θέλουμε. Επομένως, στη συγκεκριμένη περίπτωση, ο κλειστός τύπος υπολογισμού του ιδιοδιανύσματος που αντιστοιχεί στη διακριτή ιδιοτιμή λ ενός πίνακα $A \in \mathbb{R}^{n \times n}$, είναι ο ακόλουθος:

$$u = (A - \hat{\lambda}I)^{-1} \frac{(1, 1, \dots, 1)^T}{\sqrt{n}}.$$

1.2.2.2 Η μέθοδος Lanczos. Η μέθοδος Lanczos είναι μια επαναληπτική μέθοδος που εισήχθη από τον Lanczos, η οποία μπορεί να χρησιμοποιηθεί για τον υπολογισμό των ιδιοτιμών και ιδιοδιανυσμάτων ενός πίνακα A . Δοθέντος ενός τετραγωνικού συμμετρικού πίνακα A και μοναδιαίου διανύσματος v_1 , η μέθοδος αυτή δημιουργεί ταυτόχρονα έναν τριδιαγώνιο πίνακα T και έναν ορθοκανονικό πίνακα V , τέτοιους ώστε

$$A = V^T T V.$$

Ο αλγόριθμος της μεθόδου μπορεί να συναχθεί εύκολα, ως ακολούθως: Έστω οι πίνακες

$$T = \begin{pmatrix} \alpha_1 & \beta_1 & \dots & 0 \\ \beta_1 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \beta_{n-1} \\ 0 & \dots & \beta_{n-1} & \alpha_n \end{pmatrix}$$

και $V = (v_1, v_2, \dots, v_n)$. Τότε από την εξίσωση $V^T AV = T$ έχουμε ότι

$$Av_j = \alpha_j v_j + \beta_{j-1} v_{j-1} + \beta_j v_{j+1}, \quad j = 1, 2, \dots, n-1, \quad \text{όπου} \quad \beta_0 v_0 = 0.$$

Τα διανύσματα v_j ονομάζονται διανύσματα Lanczos.

1.2.3 Στοιχεία τοπολογίας

Υποθέτουμε ότι $\|\cdot\|$ είναι μια μετρική στον $\mathbb{R}^n$ και έστω σημείο $x \in \mathbb{R}^n$. Θεωρούμε ένα υποσύνολο $\mathcal{S}$ του $\mathbb{R}^n$. Ορίζουμε ως *ανοιχτή σφαίρα* ακτίνας ε και κέντρου $x \in \mathcal{S}$, το σύνολο

$$\mathcal{O}_\varepsilon(x) = \{y \mid \|y - x\| < \varepsilon\}.$$

Το σύνολο $\mathcal{S}$ ονομάζεται *ανοιχτό* στον $\mathbb{R}^n$ αν για κάθε διάνυσμα $x \in \mathcal{S}$ υπάρχει βαθμωτό μέγεθος $\varepsilon(x) > 0$ τέτοιο ώστε $\mathcal{O}_{\varepsilon(x)}(x) \subset \mathcal{S}$. Τόσο το κενό σύνολο όσο και ο $\mathbb{R}^n$ είναι ανοιχτά σύνολα στον $\mathbb{R}^n$. Επίσης η τομή πεπερασμένου αριθμού ανοιχτών συνόλων, όπως επίσης και η ένωση αυθαίρετου αριθμού ανοιχτών συνόλων, είναι και αυτές ανοιχτά σύνολα.

Ένα σύνολο $\mathcal{S}$ ονομάζεται *κλειστό* στον $\mathbb{R}^n$ αν και μόνο αν το συμπλήρωμά του $\mathbb{R}^n \setminus \mathcal{S}$, είναι ανοιχτό. Επομένως, αμφότερα το κενό σύνολο και ο $\mathbb{R}^n$ είναι κλειστά σύνολα στον $\mathbb{R}^n$. Η *κλειστή σφαίρα* ακτίνας ε γύρω από το $x \in \mathcal{S}$,

$$\mathcal{B}_\varepsilon(x) = \{y \mid \|y - x\| \leq \varepsilon\},$$

είναι κλειστό σύνολο, καθώς είναι τα καρτεσιανά γινόμενα κλειστών διαστημάτων της μορφής $[a_i, b_i]$. Η ένωση πεπερασμένου πλήθους κλειστών συνόλων, όπως και η τομή αυθαίρετου αριθμού κλειστών συνόλων, είναι επίσης κλειστά σύνολα.

Ένα υποσύνολο $\mathcal{S}$ του $\mathbb{R}^n$ λέγεται *φραγμένο* αν υπάρχει σταθερά $k > 0$ τέτοια ώστε

$$\|x - y\| \leq k \quad \text{για όλα τα} \quad x, y \in \mathcal{S}.$$

Τέλος, ένα σύνολο $\mathcal{S}$ του $\mathbb{R}^n$ είναι *συμπαγές* αν και μόνο αν είναι κλειστό και φραγμένο. Το σημαντικό πλεονέκτημα των συμπαγών συνόλων είναι ότι ακολουθείς που ανήκουν σε συμπαγή σύνολα έχουν συγκλίνουσες υπακολουθίες και ότι το μεγαλύτερο κάτω φράγμα (infimum) και το μικρότερο άνω φράγμα (supremum) συνεχών συναρτήσεων ανήκουν σε

κλειστά σύνολα.

1.2.4 Παράγωγοι και η κατά Taylor προσέγγιση

Αυτή η υποενότητα αναφέρεται στην θεμελιώδη ιδιότητα της διαφορισμότητας συναρτήσεων, που τις περισσότερες φορές αποτελεί προϋπόθεση για την εφαρμογή των αλγορίθμων του μαθηματικού προγραμματισμού. Έστω $f : \mathbb{R} \rightarrow \mathbb{R}$ μια μονοδιάστατη συνάρτηση. Η πρώτη παράγωγος ορίζεται ως

$$f'(x) = \frac{df}{dx} = \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon) - f(x)}{\epsilon}.$$

Η συνάρτηση f ονομάζεται *παραγωγίσιμη* στο x αν υπάρχει το παραπάνω όριο. Έστω τώρα η n -διάστατη συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Η μερική παράγωγος της f ως προς x_i ορίζεται ως το όριο (αν υπάρχει)

$$\frac{\partial f}{\partial x_i} = \lim_{\epsilon \rightarrow 0} \frac{f(x + \epsilon e_i) - f(x)}{\epsilon},$$

όπου e_i είναι το i μοναδιαίο διάνυσμα. Η συνάρτηση f ονομάζεται *διαφορίσιμη* αν όλες οι μερικές παράγωγοι της f υπάρχουν. Το διάνυσμα

$$\nabla f(x) = g(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix},$$

ονομάζεται *κλίση* της f . Αν η f είναι διαφορίσιμη στο x , τότε είναι και συνεχής σε αυτό. Αν οι παράγωγοι είναι και αυτές συνεχείς συναρτήσεις, τότε η f ονομάζεται *συνεχώς διαφορίσιμη* συνάρτηση. Αν οι *δεύτεροι παράγωγοι* της f ,

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \frac{\partial f(x)}{\partial x_j}$$

υπάρχουν για όλα τα i, j , με $1 \leq i, j \leq n$ και είναι συνεχείς συναρτήσεις του x , τότε η f ονομάζεται *δύο φορές συνεχώς διαφορίσιμη* συνάρτηση. Αυτές οι n^2 παράγωγοι αναπαρίστανται από έναν $n \times n$ συμμετρικό πίνακα,

γνωστό ως *εσσιανό* πίνακα ή απλά *εσσιανή* της f :

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f(x)}{\partial^2 x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial^2 x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial^2 x_n} \end{pmatrix}.$$

Οι συναρτήσεις που έχουν παραγώγους όλων των τάξεων ή μέχρι ενός συγκεκριμένου βαθμού σε κάποιο σύνολο που μας ενδιαφέρει, ονομάζονται *ομαλές*.

Ένα από τα πιο συχνά χρησιμοποιούμενα θεωρήματα της πραγματικής ανάλυσης στη βελτιστοποίηση, είναι το γνωστό **Θεώρημα Μέσης Τιμής**:

Θεώρημα 1.4. (Θεώρημα Μέσης Τιμής) Έστω ότι η συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι συνεχώς παραγωγίσιμη και έστω $p \in \mathbb{R}^n$. Τότε

$$f(x+p) = f(x) + \nabla f(x + \alpha p)^T p,$$

για κάποιο $\alpha \in (0, 1)$. Αν επιπρόσθετα η f είναι δυο φορές συνεχώς παραγωγίσιμη, τότε έχουμε ότι

$$\nabla f(x+p) = \nabla f(x) + \int_0^1 \nabla^2 f(x + \alpha p) p \, d\alpha,$$

και

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x + \alpha p) p,$$

για κάποιο $\alpha \in (0, 1)$.

Το παραπάνω θεώρημα δείχνει ότι αν είναι γνωστές οι πρώτης και δεύτερης τάξης παράγωγοι μιας συνάρτησης σε ένα σημείο x , μπορούμε να φτιάξουμε προσεγγίσεις αυτής της συνάρτησης σε όλα τα σημεία μιας γειτονιάς του x . Συγκεκριμένα, μπορούμε να προσεγγίσουμε την $f(x+p)$ με την κατά *Taylor* πρώτη τάξης προσέγγισή της,

$$f(x+p) \approx f(x) + \nabla f(x)^T p \quad (1.9)$$

ή με τη κατά Taylor δεύτερης τάξης προσέγγισή της,

$$f(x+p) \approx f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x) p. \quad (1.10)$$

Οι εξισώσεις (1.9) και (1.10) ονομάζονται *γραμμικό* και *τετραγωνικό μοντέλο* της συνάρτησης f γύρω από ένα σημείο x , αντίστοιχα. Τα μοντέλα αυτά χρησιμοποιούνται ευρέως στους αλγορίθμους βελτιστοποίησης.

Η κατά Lipschitz συνεχής συνάρτηση. Έστω $\mathcal{X}$ και $\mathcal{Y}$ ανοιχτά σύνολα και συνάρτηση $f : \mathcal{X} \rightarrow \mathcal{Y}$. Υποθέτουμε ότι οι $\|\cdot\|_{\mathcal{X}}$ και $\|\cdot\|_{\mathcal{Y}}$, είναι μετρικές των δύο αυτών συνόλων, αντίστοιχα. Τότε λέμε ότι μια τέτοια συνάρτηση είναι *συνεχής κατά Lipschitz* στο $x \in \mathcal{X}$, αν υπάρχει $L(x)$ τέτοιο ώστε

$$\|f(z) - f(x)\|_{\mathcal{Y}} \leq L(x) \|z - x\|_{\mathcal{X}}, \quad \text{για όλα τα } z \in \mathcal{X}.$$

Η σταθερά Lipschitz $L(x)$ εξαρτάται τόσο από το x , όσο και από τις μετρικές $\|\cdot\|_{\mathcal{X}}$ και $\|\cdot\|_{\mathcal{Y}}$. Η συνάρτηση είναι *τοπικά* συνεχής κατά Lipschitz κοντά στο x , αν το σύνολο $\mathcal{X}$ είναι μια ανοιχτή σφαίρα με κέντρο το x . Είναι συνεχής κατά Lipschitz στο σύνολο $\mathcal{X}$ αν η σταθερά είναι ανεξάρτητη του x , δηλαδή αν υπάρχει L τέτοιο ώστε

$$\|f(z) - f(x)\|_{\mathcal{Y}} \leq L \|z - x\|_{\mathcal{X}}, \quad \text{για όλα τα } x, z \in \mathcal{X}.$$

1.2.5 Κυρτότητα

Ένα σύνολο $\mathcal{S}$ του $\mathbb{R}^n$ ονομάζεται *κυρτό* αν για οποιαδήποτε σημεία $x, y \in \mathcal{S}$ και για κάθε $\lambda \in [0, 1]$, ισχύει ότι

$$\lambda x + (1 - \lambda)y \in \mathcal{S}.$$

Το σημείο $\lambda x + (1 - \lambda)y$ ονομάζεται *κυρτός συνδυασμός* των x και y .

Μια συνάρτηση f , ορισμένη σε ένα κυρτό σύνολο $\mathcal{S}$ είναι *κυρτή*, αν για όλα τα $x, y \in \mathcal{S}$ και για κάθε $\lambda \in [0, 1]$, έχουμε ότι

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Λέμε ότι μια συνάρτηση f είναι *κοίλη*, αν η $(-f)$ είναι κυρτή.

Αν μια συνάρτηση f είναι συνεχώς διαφορίσιμη, τότε ισχύει ότι

$$f(y) \geq f(x) + \nabla_x^T f(x)(y - x).$$

Επίσης, αν είναι δύο φορές συνεχώς διαφορίσιμη, ο εσσιανός πίνακας είναι θετικά ημιορισμένος, για όλα τα $x \in \mathcal{S}$.

Ορίζουμε ως *καμπυλότητα* της f στο $x \in \mathcal{S}$ κατά μήκος μιας κατεύθυνσης d , το κλάσμα

$$\frac{d^T \nabla^2 f(x) d}{\|d\|^2}.$$

Αν η f είναι δύο φορές συνεχώς παραγωγίσιμη και ισχύει η ανίσωση

$$d^T \nabla^2 f(x) d > 0,$$

τότε λέμε ότι η κατεύθυνση d είναι *κατεύθυνση θετικής καμπυλότητας*. Αντίστοιχα, αν ισχύει η ανίσωση

$$d^T \nabla^2 f(x) d < 0,$$

λέμε ότι η κατεύθυνση d είναι *κατεύθυνση αρνητικής καμπυλότητας*. Προφανώς, η καμπυλότητα μιας κυρτής συνάρτησης είναι πάντα μη αρνητικός αριθμός, ενώ ισχύει και το αντίθετο. Δηλαδή, αν ο εσσιανός πίνακας μιας δύο φορές διαφορίσιμης συνάρτησης είναι θετικά ημιορισμένος σε όλο το $\mathcal{S}$, ή η καμπυλότητα της f σε κάθε σημείο $x \in \mathcal{S}$ κατά μήκος κάθε κατεύθυνσης d είναι μη αρνητική, τότε η f είναι κυρτή.

1.3 Συνθήκες βελτιστότητας

Στην παρούσα διατριβή ασχολούμαστε με τη βελτιστοποίηση μη γραμμικών συναρτήσεων. Συγκεκριμένα, θέλουμε να υπολογίσουμε τοπικά ελάχιστα μιας αντικειμενικής συνάρτησης f , n πραγματικών μεταβλητών x , όπου το x ανήκει σε μια κλειστή περιοχή *εφικτών* σημείων. Αυτή η εφικτή περιοχή μπορεί να είναι όλο το $\mathbb{R}^n$, οπότε το πρόβλημα είναι χωρίς περιορισμούς, ή κάποιο υποσύνολο του $\mathbb{R}^n$, οπότε έχουμε να επιλύσουμε ένα πρόβλημα με περιορισμούς. Πιο αναλυτικά, θεωρούμε το γενικό μαθηματικό πρόβλημα

$$\min_{x \in \mathcal{S}} f(x), \quad (1.11)$$

όπου η f είναι μια ομαλή συνάρτηση από τον $\mathbb{R}^n$ στον $\mathbb{R}$ και το εφικτό σύνολο $S \subseteq \mathbb{R}^n$. Για να χαρακτηρίσουμε τη λύση του προβλήματος (1.11) ως βέλτιστη, απαραίτητη προϋπόθεση είναι, το σημείο αυτό να ανήκει στο εφικτό σύνολο S .

Στη συνέχεια θα αναφέρουμε τις συνθήκες βελτιστότητας, οι οποίες είναι τα μαθηματικά εργαλεία που χρησιμοποιούμε ώστε να επαληθεύσουμε αν ένα σημείο είναι ή όχι η βέλτιστη λύση του εκάστοτε προβλήματος που μελετάμε. Αυτές οι συνθήκες ελέγχουν αν ένα σημείο x^* έχει τη μικρότερη συναρτησιακή τιμή στη γειτονιά του. Επιπρόσθετα, η πληροφορία που δίνουν οι έλεγχοι βελτιστότητας χρησιμοποιούνται για τον υπολογισμό του επόμενου δοκιμαστικού σημείου. Οι συνθήκες βελτιστότητας είναι η βάση της ανάπτυξης και της ανάλυσης των επαναληπτικών αλγορίθμων. Στην πράξη, οι αλγόριθμοι τερματίζουν όταν αυτές οι συνθήκες ισχύουν προσεγγιστικά. Οι αποδείξεις των θεωρημάτων που αναφέρονται στη συνέχεια καθώς και περισσότερες λεπτομέρειες για τις συνθήκες βελτιστότητας, μπορούν να βρεθούν στις αναφορές που παρατέθηκαν στην αρχή αυτού του κεφαλαίου.

1.3.1 Βελτιστοποίηση χωρίς περιορισμούς

Τα προβλήματα βελτιστοποίησης χωρίς περιορισμούς μπορούν να εκφραστούν μαθηματικά ως εξής:

$$\min_{x \in \mathbb{R}^n} f(x), \quad (1.12)$$

όπου η αντικειμενική συνάρτηση της μεταβλητής x $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι ομαλή.

Ορισμός 1.6. Ένα σημείο x^* ονομάζεται ασθενής τοπικός ελαχιστοποιητής αν υπάρχει $\delta > 0$ τέτοιο ώστε $f(x^*) \leq f(x)$ για όλα τα $x \in \mathbb{R}^n$, με $\|x - x^*\| < \delta$. Ένα σημείο x^* λέγεται αυστηρά τοπικός ελαχιστοποιητής αν υπάρχει $\delta > 0$ τέτοιο ώστε $f(x^*) < f(x)$ για όλα τα $x \in \mathbb{R}^n$, με $x \neq x^*$ και $\|x - x^*\| < \delta$.

Ορισμός 1.7. Ένα σημείο x^* ονομάζεται ολικός ελαχιστοποιητής της f αν $f(x^*) \leq f(x)$, για όλα τα $x \in \mathbb{R}^n$. Ένα σημείο x^* λέγεται αυστηρά ολικός ελαχιστοποιητής της f αν $f(x^*) < f(x)$, για όλα τα $x \in \mathbb{R}^n$, με $x \neq x^*$.

Το παρακάτω σχήμα απεικονίζει τους τύπους ελαχιστοποιητών.

Σχήμα 1.1: Παραδείγματα τοπικού και ολικού ελαχιστοποιητή σε μία διάσταση

Όταν η αντικειμενική συνάρτηση είναι κυρτή, όλοι οι τοπικοί ελαχιστοποιητές είναι και ολικοί. Ονομάζουμε *ελάχιστο* την τιμή της αντικειμενικής συνάρτησης f στον ελαχιστοποιητή.

Οι συνθήκες βελτιστότητας για το πρόβλημα (1.12) συνάγονται υποθέτοντας πρώτον, ότι ένα σημείο x^* είναι τοπικός ελαχιστοποιητής της συνάρτησης f και δεύτερον, θεσπίζοντας κάποιες ιδιότητες που πρέπει να πληροί το συγκεκριμένο σημείο, αναφορικά με της πρώτης και δεύτερης τάξης παραγώγους της αντικειμενικής συνάρτησης f στο σημείο αυτό.

Ορισμός 1.8. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ διαφορίσιμη στο $x \in \mathbb{R}^n$. Αν υπάρχει διάνυσμα $d \in \mathbb{R}^n$ τέτοιο ώστε $\nabla f(x)^T d < 0$, τότε η d ονομάζεται κατεύθυνση μείωσης της f στο x .

Θεώρημα 1.5. (Πρώτης Τάξης Αναγκαία Συνθήκη) Έστω $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο $\mathcal{D}$. Αν το $x^* \in \mathcal{D}$ είναι τοπικός ελαχιστοποιητής του προβλήματος (1.12), τότε

$$\nabla f(x^*) = 0. \quad (1.13)$$

Η απαίτηση (1.13) είναι γνωστή ως πρώτης τάξης (αναγκαία) συνθήκη. Ένα σημείο x^* που ικανοποιεί τη σχέση (1.13) αναφέρεται ως *πρώτης τάξης κρίσιμο* ή *πρώτης τάξης στάσιμο* σημείο της f . Κάθε τοπικός ελαχιστοποιητής πρέπει να είναι πρώτης τάξης κρίσιμο σημείο.

Ορισμός 1.9. Ένα σημείο $x^* \in \mathbb{R}^n$ ονομάζεται *στάσιμο* ή *κρίσιμο* σημείο της διαφορίσιμης συνάρτησης f , αν $\nabla f(x^*) = 0$.

Θεώρημα 1.6. (Δεύτερης Τάξης Αναγκαία Συνθήκη) Έστω $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ δύο φορές συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο $\mathcal{D}$. Αν το x^* είναι τοπικός ελαχιστοποιητής του προβλήματος (1.12), τότε ισχύει η (1.13) και ο εσσιανός πίνακας της f στο x^* είναι θετικά ημιορισμένος. Δηλαδή,

$$\nabla f(x^*) = 0 \quad \text{και} \quad s^T \nabla^2 f(x^*) s \geq 0, \quad \text{για όλα τα } s \in \mathbb{R}^n. \quad (1.14)$$

Οι απαιτήσεις (1.14) είναι γνωστές ως δεύτερης τάξης (αναγκαίες) συνθήκες, ενώ το x^* που τις ικανοποιεί ονομάζεται *δεύτερης τάξης κρίσιμο σημείο* της f . Προφανώς, ο εσσιανός της f είναι οπωσδήποτε θετικά ημιορισμένος στον ελαχιστοποιητή.

Το ακόλουθο θεώρημα αναφέρει υπό ποιες συνθήκες ένα σημείο x^* είναι τοπικός ελαχιστοποιητής της αντικειμενικής συνάρτησης f .

Θεώρημα 1.7. (Δεύτερης Τάξης Ικανή Συνθήκη) Έστω $f : \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ δύο φορές συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο $\mathcal{D}$. Αν

$$\nabla f(x^*) = 0 \quad \text{και} \quad s^T \nabla^2 f(x^*) s > 0 \quad \text{για όλα τα } s \neq 0 \in \mathbb{R}^n, \quad (1.15)$$

τότε το $x^* \in \mathcal{D}$ είναι αυστηρά τοπικός ελαχιστοποιητής του προβλήματος (1.12).

Το Θεώρημα 1.5 δηλώνει ότι αν το x^* είναι τοπικός ελαχιστοποιητής, τότε είναι πρώτης τάξης κρίσιμο σημείο. Εντούτοις, το αντίστροφο δεν ισχύει. Αν το x^* είναι πρώτης τάξης κρίσιμο σημείο, μπορεί να είναι είτε τοπικός ελαχιστοποιητής ή τοπικός μεγιστοποιητής, αλλά μπορεί να μην είναι τίποτα από τα δύο. Στην τελευταία περίπτωση, το x^* ονομάζεται *σαγματικό σημείο*. Συνεπώς, ένα πρώτης τάξης κρίσιμο σημείο δεν είναι κατ' ανάγκη τοπικός ελαχιστοποιητής. Όμως, αν η αντικειμενική συνάρτηση, η οποία είναι διαφορίσιμη, είναι κυρτή, τότε τα κρίσιμα σημεία είναι τοπικοί και ολικοί ελαχιστοποιητές του προβλήματος (1.12).

Θεώρημα 1.8. Έστω $\mathcal{S} \subset \mathbb{R}^n$ ένα μη κενό κυρτό σύνολο και $f : \mathcal{S} \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Έστω $x^* \in \mathcal{S}$ τοπικός ελαχιστοποιητής του προβλήματος $\min_{x \in \mathcal{S}} f(x)$.

1. Αν η f είναι κυρτή, τότε το x^* είναι και ολικός ελαχιστοποιητής.
2. Αν η f είναι αυστηρά κυρτή, τότε το x^* είναι μοναδικός ολικός ελαχιστοποιητής.

Το παρακάτω θεώρημα λέει ότι για διαφορίσιμη κυρτή συνάρτηση, το κρίσιμο σημείο είναι ολικός ελαχιστοποιητής. Οπότε, οι ακόλουθες συνθήκες βελτιστότητας είναι ικανές και αναγκαίες.

Θεώρημα 1.9. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ συνεχώς διαφορίσιμη και κυρτή. Τότε το x^* είναι ολικός ελαχιστοποιητής του προβλήματος (1.12) αν και μόνο αν $\nabla f(x^*) = 0$.

1.3.2 Βελτιστοποίηση με περιορισμούς

Η γενική μαθηματική έκφραση ενός προβλήματος βελτιστοποίησης με περιορισμούς είναι η ακόλουθη:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{υπό τους περιορισμούς} \quad \left\{ \begin{array}{l} c_i(x) = 0, \quad i = 1, \dots, m_e \\ c_i(x) \geq 0, \quad i = m_e + 1, \dots, m \end{array} \right\},$$

όπου η αντικειμενική συνάρτηση f και οι συναρτήσεις περιορισμών c_i είναι ομαλές, πραγματικές συναρτήσεις στον $\mathbb{R}^n$ και τουλάχιστον μία είναι μη γραμμική. Στα προβλήματα βελτιστοποίησης με περιορισμούς, απαιτούμε η μεταβλητή x να ανήκει σε μια εφικτή περιοχή Ω , η οποία είναι το σύνολο των σημείων x που ικανοποιούν τους περιορισμούς. Αν θέσουμε $\mathcal{E} = \{i = 1, \dots, m_e\}$ και $\mathcal{I} = \{i = m_e + 1, \dots, m\}$, το εφικτό σύνολο μπορεί να οριστεί ως

$$\Omega = \{x \mid c_i(x) = 0, \quad i \in \mathcal{E}, \quad c_i(x) \geq 0, \quad i \in \mathcal{I}\}.$$

Τα σύνολα $\mathcal{E}$ και $\mathcal{I}$ είναι δύο ξένα μεταξύ τους σύνολα δεικτών και $c_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ($i \in \mathcal{E} \cup \mathcal{I}$) είναι οι περιορισμοί ισότητας και ανισότητας. Επομένως, ένα πρόβλημα βελτιστοποίησης με περιορισμούς μπορεί να διατυπωθεί ως εξής:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{υπό τους περιορισμούς} \quad \left\{ \begin{array}{l} c_i(x) = 0, \quad i \in \mathcal{E}; \\ c_i(x) \geq 0, \quad i \in \mathcal{I}, \end{array} \right\}, \quad (1.16)$$

Αν ένα σημείο x ικανοποιεί όλους του περιορισμούς λέγεται εφικτό σημείο. Το σύνολο όλων των εφικτών σημείων ονομάζεται εφικτό σύνολο. Αν το πρόβλημα δεν έχει εφικτά σημεία, τότε ονομάζεται μη εφικτό. Οι διαφορετικού τύπου λύσεις περιορίζονται σε εφικτά σημεία μόνο.

Ορισμός 1.10. Αν $x^* \in \Omega$ και αν $f(x) \geq f(x^*)$, για κάθε $x \in \Omega$, τότε το x^* είναι ολικός ελαχιστοποιητής του προβλήματος (1.16). Αν $x^* \in \Omega$ και $f(x) > f(x^*)$, για κάθε $x \in \Omega$, με $x \neq x^*$, τότε το x^* είναι αυστηρά ολικός ελαχιστοποιητής του προβλήματος (1.16).

Ορισμός 1.11. Αν $x^* \in \Omega$ και αν υπάρχει γειτονιά $\mathcal{B}(x^*, \delta)$ του x^* , τέτοια ώστε $f(x) \geq f(x^*)$, για κάθε $x \in \Omega \cap \mathcal{B}(x^*, \delta)$, τότε το x^* ονομάζεται τοπικός ελαχιστοποιητής του προβλήματος (1.16), όπου $\mathcal{B}(x^*, \delta) = \{x \mid \|x - x^*\|_2 \leq \delta\}$, με $\delta > 0$.

Αν $x^* \in \Omega$ και αν υπάρχει γειτονιά $\mathcal{B}(x^*, \delta)$ του x^* , τέτοια ώστε $f(x) > f(x^*)$, για κάθε $x \in \Omega \cap \mathcal{B}(x^*, \delta)$, με $x \neq x^*$, τότε το x^* ονομάζεται αυστηρά τοπικός ελαχιστοποιητής του προβλήματος (1.16).

Ορισμός 1.12. Αν $x^* \in \Omega$ και αν υπάρχει γειτονιά $\mathcal{B}(x^*, \delta)$ του x^* τέτοια ώστε το x^* να είναι ο μοναδικός ελαχιστοποιητής στο σύνολο $\Omega \cap \mathcal{B}(x^*, \delta)$, τότε το x^* είναι ένας μεμονωμένος τοπικός ελαχιστοποιητής.

Προφανώς, ένας ολικός ελαχιστοποιητής είναι και τοπικός. Πριν δώσουμε τον ορισμό των ενεργών περιορισμών, χρειάζεται να θεωρήσουμε το σύνολο

$$\mathcal{I}(x) = \{i \mid c_i(x) = 0, i \in \mathcal{I}\}.$$

Ορισμός 1.13. Για κάθε $x \in \mathbb{R}^n$ το σύνολο των δεικτών των περιορισμών για τους οποίους ισχύει η ισότητα, δηλαδή

$$\mathcal{A}(x) = \{i \in \mathcal{E} \cup \mathcal{I}(x)\}.$$

ονομάζεται ενεργό σύνολο περιορισμών στο x .

Λέμε ότι οι περιορισμοί των οποίων οι δείκτες ανήκουν στο σύνολο ενεργό σύνολο $\mathcal{A}(x)$, είναι ενεργοί. Αν $j \notin \mathcal{A}(x)$, τότε ο περιορισμός $c_j(x)$ είναι ανενεργός στο x .

1.3.2.1 Συνθήκες βελτιστότητας πρώτης τάξης. Πριν δώσουμε τις συνθήκες βελτιστότητας πρώτης τάξης για προβλήματα με περιορισμό, θα δώσουμε τον ορισμό της εφικτής κατεύθυνσης.

Ορισμός 1.14. (Σύνολο Εφικτών Κατευθύνσεων (EK)) Έστω $x^* \in \Omega$ και $0 \neq d \in \mathbb{R}^n$. Αν υπάρχει $\delta > 0$ τέτοιο ώστε $x^* + td \in \Omega$, για κάθε $t \in [0, \delta]$,

τότε η d ονομάζεται εφικτή κατεύθυνση του Ω στο x^* . Το σύνολο όλων των εφικτών κατευθύνσεων του Ω στο x^* είναι το σύνολο

$$EK(x^*, \Omega) = \{d \mid x^* + td \in \Omega, \forall t \in [0, \delta]\}.$$

Ορισμός 1.15. (Σύνολο Γραμμικοποιημένων Εφικτών Κατευθύνσεων) Έστω $x^* \in \Omega$ και $d \in \mathbb{R}^n$. Αν

$$\begin{aligned} d^T \nabla c_i(x^*) &= 0, & i \in \mathcal{E}, \\ d^T \nabla c_i(x^*) &\geq 0, & i \in \mathcal{I}(x^*), \end{aligned}$$

τότε η d ονομάζεται γραμμικοποιημένη εφικτή κατεύθυνση του Ω στο x^* . Το σύνολο όλων των γραμμικοποιημένων εφικτών κατευθύνσεων του Ω στο x^* είναι το

$$\Sigma GEK(x^*, \Omega) = \left\{ d \mid \begin{array}{l} d^T \nabla c_i(x^*) = 0, \quad i \in \mathcal{E} \\ d^T \nabla c_i(x^*) \geq 0, \quad i \in \mathcal{I}(x^*) \end{array} \right\}.$$

Ορισμός 1.16. (Σύνολο Εφικτών Κατευθύνσεων Ακολουθίας) Έστω $x^* \in \Omega$ και $d \in \mathbb{R}^n$. Αν υπάρχουν ακολουθίες d_k ($k = 1, 2, \dots$) και $\delta_k > 0$, ($k = 1, 2, \dots$) τέτοια ώστε $x^* + \delta_k d_k \in \Omega$, για κάθε k και $d_k \rightarrow d$, $\delta_k \rightarrow 0$, τότε η οριακή κατεύθυνση d ονομάζεται εφικτή κατεύθυνση ακολουθίας του Ω στο x^* . Το σύνολο όλων των εφικτών κατευθύνσεων ακολουθίας του Ω στο x^* είναι το

$$\Sigma EKA(x^*, \Omega) = \left\{ d \mid \begin{array}{l} x^* + \delta_k d_k \in \Omega, \forall k \\ d_k \rightarrow d, \delta_k \rightarrow 0 \end{array} \right\}.$$

Για να ορίσουμε τις συνθήκες βελτιστότητας πρώτης τάξης, χρησιμοποιούμε τη συνάρτηση Lagrange. Η Λαγκρανσιανή ή συνάρτηση Lagrange του προβλήματος (1.16) ορίζεται ως

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x),$$

όπου $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T \in \mathbb{R}^m$ είναι το διάνυσμα των πολλαπλασιαστών Lagrange.

Θεώρημα 1.10. (Πρώτης Τάξης Αναγκαίες Συνθήκες-KKT conditions)

Έστω x^* ένας τοπικός ελαχιστοποιητής του προβλήματος (1.16). Αν ισχύει η ικανότητα περιορισμού

$$\Sigma EKA(x^*, \Omega) = \Sigma GEK(x^*, \Omega), \quad (1.17)$$

τότε υπάρχουν πολλαπλασιαστές Lagrange λ_i^* τέτοιοι ώστε να ικανοποιούνται οι ακόλουθες συνθήκες στο (x^*, λ^*) :

$$\nabla f(x^*) - \sum_i^m \lambda_i^* \nabla c_i(x^*) = 0 \quad (1.18)$$

$$c_i(x^*) = 0, \quad \forall i \in \mathcal{E} \quad (1.19)$$

$$c_i(x^*) \geq 0, \quad \forall i \in \mathcal{I} \quad (1.20)$$

$$\lambda_i^* \geq 0, \quad \forall i \in \mathcal{I} \quad (1.21)$$

$$\lambda_i^* c_i(x^*) = 0, \quad \forall i \in \mathcal{I} \quad (1.22)$$

Λέμε ότι το σημείο x^* που ικανοποιεί τις συνθήκες KKT είναι ένα *KKT σημείο* ή ένα *πρώτης τάξης κρίσιμο σημείο* για το πρόβλημα (1.16). Οι συνθήκες (1.18)-(1.22) είναι γνωστές και ως συνθήκες Karush-Kuhn-Tucker ή πιο απλά, ως συνθήκες KKT. Η συνθήκη (1.18) αναφέρεται συχνά ως *συνθήκη στασιμότητας*, γιατί μπορεί να γραφεί ως

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = \nabla f(x^*) - \sum_i^m \lambda_i^* \nabla c_i(x^*) = 0.$$

Οι συνθήκες (1.19) και (1.20) ονομάζονται *συνθήκες εφικτότητας*. Η συνθήκη (1.21) είναι η συνθήκη μη αρνητικότητας για τους πολλαπλασιαστές Lagrange. Τέλος, η συνθήκη (1.22) είναι η *συνθήκη συμπληρωματικής χαλαρότητας*, γιατί δηλώνει ότι δεν μπορούν ταυτόχρονα τα λ_i^* και $c_i(x^*)$ να είναι μη μηδενικά, ή ισοδύναμα οι πολλαπλασιαστές Lagrange που αντιστοιχούν σε ανενεργούς περιορισμούς είναι μηδέν.

Ορισμός 1.17. (Ικανότητα Γραμμικώς Ανεξάρτητων Περιορισμών (ΙΓΑΠ))
Αν οι κλίσεις των ενεργών περιορισμών $\nabla c_i(x^*)$, $i \in \mathcal{A}(x^*)$ είναι γραμμικά ανεξάρτητες, τότε ισχύει η ικανότητα γραμμικώς ανεξάρτητων περιορισμών (LICQ).

Θεώρημα 1.11. Έστω x^* τοπικός ελαχιστοποιητής του προβλήματος (1.16). Αν ισχύει η ΙΓΑΠ, δηλαδή αν οι $\nabla c_i(x^*)$, $i \in \mathcal{A}(x^*)$ είναι γραμμικά ανεξάρ-

τητες, τότε υπάρχουν πολλαπλασιαστές Lagrange $\lambda_i, i = 1, \dots, m$, τέτοιοι ώστε να ισχύουν οι συνθήκες (1.18)-(1.22).

Αν το πρόβλημά μας είναι η ελαχιστοποίηση μιας κυρτής συνάρτησης f σε ένα κυρτό σύνολο Ω , τότε έχουμε ένα πρόβλημα κυρτού προγραμματισμού,

$$\min f(x), \quad \text{ως προς } x \in \Omega. \quad (1.23)$$

Για αυτού του είδους τα προβλήματα ισχύουν τα παρακάτω θεωρήματα.

Θεώρημα 1.12. Κάθε τοπικός ελαχιστοποιητής ενός προβλήματος κυρτού προγραμματισμού (1.23) είναι και ολικός ελαχιστοποιητής, ενώ το σύνολο $\mathcal{S}$ των ολικών ελαχιστοποιητών είναι κυρτό.

Θεώρημα 1.13. Το σημείο KKT ενός προβλήματος κυρτού προγραμματισμού είναι ο ελαχιστοποιητής του.

Θεώρημα 1.14. Ένα πρόβλημα κυρτού προγραμματισμού με αντικειμενική συνάρτηση f αυστηρά κυρτή, έχει μοναδικό ελαχιστοποιητή.

1.3.2.2 Συνθήκες βελτιστότητας δεύτερης τάξης. Χρησιμοποιώντας πληροφορία παραγώγων δεύτερης τάξης μπορούμε να θεσπίσουμε τις δεύτερης τάξης συνθήκες βελτιστότητας για το πρόβλημα (1.16). Για να εισάγουμε αυτές τις συνθήκες ορίζουμε τα σύνολα

$$\mathcal{I}_+(x^*) = \{i \mid i \in \mathcal{I}(x^*) \text{ με } \lambda_i^* > 0\},$$

$$\mathcal{S}(x^*, \lambda^*) = \left\{ d \mid \begin{array}{l} d^T \in \Sigma \text{EKA}(x^*, \Omega) \\ \sum_{i=1}^m \lambda_i^* c_i(x_k) = 0 \end{array} \right\}$$

και

$$\mathcal{N}(x^*, \lambda^*) = \left\{ d \mid \begin{array}{l} d \neq 0 \\ d^T \nabla c_i(x^*) = 0, i \in \mathcal{E} \cup \mathcal{I}_+(x^*) \\ d^T \nabla c_i(x^*) \lambda \geq 0, i \in \mathcal{I}(x^*) \setminus \mathcal{I}_+(x^*) \end{array} \right\}.$$

Θεώρημα 1.15. (Δεύτερης τάξης αναγκαίες συνθήκες) Έστω x^* η τοπική λύση του προβλήματος με περιορισμούς (1.16). Αν ισχύει η ικανότητα περιορισμού (1.17), τότε έχουμε ότι

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) d \geq 0, \quad \forall d \in \mathcal{S}(x^*, \lambda^*),$$

όπου $\mathcal{L}(x, \lambda)$ η συνάρτηση *Lagrange*. Επιπλέον, αν $\mathcal{S}(x^*, \lambda^*) = \mathcal{N}(x^*, \lambda^*)$, τότε

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) d \geq 0, \forall d \in \mathcal{N}(x^*, \lambda^*).$$

Θεώρημα 1.16. (Δεύτερης τάξης ικανές συνθήκες) Έστω x^* ένα KKT σημείο του προβλήματος (1.16). Αν

$$d^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*) d > 0, \forall d \in \mathcal{N}(x^*, \lambda^*),$$

τότε το x^* είναι μια αυστηρά τοπική λύση του προβλήματος (1.16).

Βασικές μέθοδοι μη γραμμικής βελτιστοποίησης

Στο παρόν κεφάλαιο παρουσιάζουμε τις *τοπικές μεθόδους* και *τεχνικές ολικής σύγκλισης* που μπορούν να εφαρμοστούν για την επίλυση μη γραμμικών προβλημάτων βελτιστοποίησης. Οι μέθοδοι αυτές, αποτελούν τη βάση των μεθόδων βελτιστοποίησης που προτείνουμε στα επόμενα κεφάλαια, αποσκοπώντας πάντα στο να γίνει εφικτή η εφαρμογή τους σε προβλήματα μεγάλης κλίμακας. Οι μέθοδοι που παρουσιάζουμε έχουν μόνο τοπικές ιδιότητες σύγκλισης, ενώ οι τεχνικές ολικής σύγκλισης επιβάλλεται να χρησιμοποιηθούν ώστε να εξασφαλιστεί η σύγκλιση σε ένα τοπικό ελάχιστο, όταν ο αλγόριθμος έχει ξεκινήσει από ένα αυθαίρετο αρχικό σημείο, το οποίο εν γένει δεν ανήκει στη γειτονιά του τοπικού ελαχίστου. Οι τοπικές μέθοδοι που αναφέρουμε είναι η μέθοδος Newton [68], οι Quasi-Newton μέθοδοι [68] και οι περιορισμένης μνήμης Quasi-Newton μέθοδοι [56, 67]. Οι αναφερόμενες τεχνικές ολικής σύγκλισης είναι οι μέθοδοι γραμμικής αναζήτησης [68], καμπυλόγραμμης αναζήτησης [35, 42, 61, 63] και περιοχής εμπιστοσύνης [25].

2.1 Η μέθοδος Newton

Η μέθοδος Newton είναι ίσως η πιο γνωστή από όλες τις επαναληπτικές μεθόδους. Χρησιμοποιείται στην επίλυση του προβλήματος

$$\min_{x \in \mathbb{R}^n} f(x), \quad (2.1)$$

όπου η αντικειμενική συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι δυο φορές συνεχώς διαφορίσιμη. Σε ότι ακολουθεί θα δηλώνουμε την πρώτη παράγωγο της f

με

$$g(x) = \nabla f(x).$$

Η βασική ιδέα της Newton έχει ως εξής: αντί να ελαχιστοποιήσουμε την f , ελαχιστοποιούμε το τετραγωνικό μοντέλο της, το οποίο κατασκευάζεται στην τρέχουσα επανάληψη x_k . Έστω λοιπόν ότι η εσσιανή $\nabla^2 f(x_k)$ της f είναι θετικά ορισμένη στο x_k . Τότε, το παραγόμενο τετραγωνικό μοντέλο της f με κέντρο το σημείο x_k ,

$$m_k(x_k + s) = f(x_k) + g(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s,$$

είναι κυρτό, όπου έχουμε θέσει $s = x - x_k$. Οπότε το μοντέλο έχει μοναδικό ελαχιστοποιητή το σημείο στο οποίο η κλίση του μοντέλου μηδενίζεται. Αυτό το σημείο είναι προφανώς η λύση του συστήματος

$$\nabla m_k(x_k + s) = 0, \quad \text{δηλαδή} \quad g(x_k) + \nabla^2 f(x_k) s = 0.$$

Επομένως, ελαχιστοποιώντας το $m_k(x_k + s)$ στην τρέχουσα επανάληψη, βρίσκουμε ότι $s_k = -[\nabla^2 f(x_k)]^{-1} g(x_k)$ και έτσι προκύπτει το επαναληπτικό σχήμα

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} g(x_k), \quad (2.2)$$

όπου το διάνυσμα $s_k = x_{k+1} - x_k$, είναι γνωστό ως *κατεύθυνση Newton*. Όταν η εσσιανή είναι θετικά ορισμένη για όλα τα x , η κατεύθυνση Newton είναι καλά ορισμένη και μπορούμε να συνοψίσουμε τα βασικά βήματα στον ακόλουθο αλγόριθμο:

Αλγόριθμος 2.1 Η μέθοδος Newton

Βήμα 1: Δοθέντος αρχικού σημείου $x_0 \in \mathbb{R}^n$, $\epsilon > 0$, θέσε $k := 0$.

Βήμα 2: Αν $\|g(x_k)\| \leq \epsilon$, σταμάτα.

Βήμα 3: Υπολόγισε την κατεύθυνση s_k επιλύοντας το γραμμικό σύστημα $\nabla^2 f(x_k) s = -g(x_k)$, ως προς s .

Βήμα 4: Θέσε $x_{k+1} = x_k + s_k$.

Βήμα 5: Θέσε $k := k + 1$ και πήγαινε στο Βήμα 2.

Εύκολα μπορούμε να δείξουμε ότι στην περίπτωση όπου η εσσιανή είναι θετικά ορισμένη, πάντα η παραγόμενη κατεύθυνση είναι μία *κατεύθυνση μείωσης*. Πράγματι, από τη σχέση $\nabla^2 f(x_k)s = -g(x_k)$, συνεπάγεται ότι

$$g(x_k)^T s_k = -s_k^T \nabla^2 f(x_k) s_k \leq -\sigma_k \|s_k\|^2, \quad \text{για κάποιο } \sigma_k > 0.$$

Εντούτοις, μακριά από ένα τοπικό ελάχιστο, η εσσιανή $\nabla^2 f(x)$ μπορεί να μην είναι θετικά ορισμένη, με αποτέλεσμα η κατεύθυνση Newton να μην είναι κατεύθυνση μείωσης.

Θεώρημα 2.1. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ δυο φορές συνεχώς διαφορίσιμη και έστω x_k σημείο κοντά στη βέλτιστη λύση x^* του προβλήματος ελαχιστοποίησης (2.1), με $g(x^*) = 0$. Αν η $\nabla^2 f(x^*)$ είναι θετικά ορισμένη και η $\nabla^2 f(x)$ ικανοποιεί τη συνθήκη Lipschitz

$$|\nabla_{ij}^2 f(x) - \nabla_{ij}^2 f(y)| \leq M \|x - y\|, \quad \text{για κάποιο } M, \forall i, j \in \{1, \dots, n\}$$

τότε η επανάληψη Newton (2.2) είναι καλά ορισμένη και η ακολουθία σημείων $\{x_k\}$ που δημιουργείται, συγκλίνει τετραγωνικά στο x^* .

Η μέθοδος Newton συγκλίνει σε μία μόνο επανάληψη για όλες τις τετραγωνικές αντικειμενικές συναρτήσεις. Όμως, ένα πολύ βασικό της μειονέκτημα είναι ότι δεν έχει ολική σύγκλιση. Αυτό πρακτικά σημαίνει ότι η μέθοδος συγκλίνει μόνο όταν το αρχικό σημείο είναι αρκετά κοντά στο x^* . Επιπλέον, η Newton χρειάζεται δεύτερης τάξης πληροφορία, η οποία μπορεί είτε να μην είναι διαθέσιμη ή ο υπολογισμός της να είναι απαγορευτικός σε μεγάλης κλίμακας προβλήματα, λόγω της χωρικής πολυπλοκότητάς της που είναι $\Theta(n^2)$. Σε αυτές τις περιπτώσεις χρησιμοποιούμε είτε τις Quasi-Newton μεθόδους ή τις περιορισμένης μνήμης Quasi-Newton μεθόδους, αντίστοιχα.

2.2 Οι Quasi-Newton μέθοδοι

Οι Quasi-Newton μέθοδοι είναι μια κλάση μεθόδων, στις οποίες δεν χρειάζεται ο υπολογισμός της εσσιανής, αλλά μιας προσέγγισής της. Στηρίζονται στην ενημέρωση του προσεγγιστικού εσσιανού πίνακα με τρόπο που δεν απαιτεί μεγάλο υπολογιστικό κόστος. Το μεγαλύτερο πλεονέκτημά τους είναι ότι δεν χρειάζονται τον υπολογισμό δεύτερων παραγώγων μιας και

χρησιμοποιούν πληροφορία μόνο από την αντικειμενική συνάρτηση και τις παραγώγους πρώτης τάξης.

Οι Quasi-Newton μέθοδοι μπορούμε να πούμε ότι “γεννήθηκαν” στα μέσα του 1950 από τον W.C. Davidon [27]. Ο αλγόριθμος του Davidon αποτελεί τον πρώτο Quasi-Newton αλγόριθμο στην ιστορία της βελτιστοποίησης. Αρκετά αργότερα αποδείχτηκε από τους Powell [73] και Fletcher [37] ότι ο αλγόριθμος αυτός ήταν πολύ πιο γρήγορος και αξιόπιστος από τους ήδη υπάρχοντες αλγόριθμους μη γραμμικής βελτιστοποίησης. Αξίζει να αναφέρουμε ότι η μέθοδος του Davidon παρέμεινε για πάνω από τριάντα χρόνια τεχνική αναφορά, μέχρις ότου δημοσιεύτηκε επίσημα το 1991 στο έγκριτο περιοδικό SIAM Journal on Optimization [28]. Μέχρι τις μέρες μας, πολλά επιστημονικά άρθρα ασχολούνται με αυτόν τον αλγόριθμο και προτείνονται παραλλαγές του.

Για να προσεγγίσουν οι Quasi-Newton μέθοδοι την εσσιανή της αντικειμενικής συνάρτησης με έναν συμμετρικό πίνακα, χρησιμοποιούν ένα τετραγωνικό μοντέλο της f , γύρω από μια συγκεκριμένη επανάληψη. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ δύο φορές συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο $\mathcal{D} \subset \mathbb{R}^n$. Η τετραγωνική προσέγγιση της f στο x_{k+1} είναι

$$f(x) \approx f(x_{k+1}) + g_{k+1}^T(x - x_{k+1}) + \frac{1}{2}(x - x_{k+1})^T \nabla^2 f(x_{k+1})(x - x_{k+1}),$$

όπου $g_{k+1} = \nabla f(x_{k+1})$. Η παράγωγος της προσέγγισης αυτής είναι

$$g(x) \approx g_{k+1} + \nabla^2 f(x_{k+1})(x - x_{k+1}).$$

Θέτοντας $x = x_k$, $s_k = x_{k+1} - x_k$ και $y_k = g_{k+1} - g_k$, έχουμε ότι $\nabla^2 f(x_{k+1})^{-1}y_k \approx s_k$. Αν η συνάρτηση f είναι κυρτή με εσσιανή την $\nabla^2 f$, ισχύει η ισότητα, δηλαδή ισχύει ότι

$$\nabla^2 f(x_{k+1})^{-1}y_k = s_k \quad \text{ή} \quad \nabla^2 f(x_{k+1})s_k = y_k.$$

Θεωρούμε τώρα το τετραγωνικό μοντέλο της f στο x_{k+1} ,

$$m_{k+1}(x) = f(x_{k+1}) + g_{k+1}^T(x - x_{k+1}) + \frac{1}{2}(x - x_{k+1})^T B_{k+1}(x - x_{k+1}), \quad (2.3)$$

όπου B_{k+1} είναι η προσέγγιση της εσσιανής στο x_{k+1} . Στο τετραγωνικό μοντέλο (2.3), αν θέσουμε όπου $x = x_{k+1}$, τότε ικανοποιούνται οι συνθή-

κες παρεμβολής

$$\begin{aligned} m_{k+1}(x_{k+1}) &= f(x_{k+1}), \\ \nabla m_{k+1}(x_{k+1}) &= g_{k+1}. \end{aligned}$$

Επιπλέον, απαιτούμε να ισχύει η συνθήκη παρεμβολής $\nabla m_{k+1}(x_k) = g_k$, από όπου έπεται η συνθήκη: $g_{k+1} + B_{k+1}(x_k - x_{k+1}) = g_k$. Έτσι, έχουμε τη λεγόμενη Quasi-Newton εξίσωση, ή αλλιώς εξίσωση της τέμνουσας (secant equation),

$$B_{k+1}(x_{k+1} - x_k) = g_{k+1} - g_k.$$

Αφού $s_k = x_{k+1} - x_k$ και $y_k = g_{k+1} - g_k$, η παραπάνω εξίσωση μπορεί να γραφεί ως ακολούθως:

$$B_{k+1}s_k = y_k, \quad (2.4)$$

ή ισοδύναμα,

$$s_k = H_{k+1}y_k, \quad (2.5)$$

όπου H_{k+1} ο αντίστροφος του προσεγγιστικού εσσιανού πίνακα B_{k+1} . Δοθείσης δηλαδή της μετατόπισης s_k και της διαφοράς των κλίσεων y_k , η εξίσωση της τέμνουσας απαιτεί ο συμμετρικός και θετικά ορισμένος πίνακας B_k να απεικονίζει το s_k στο y_k . Αυτό είναι εφικτό, μόνο αν ισχύει η λεγόμενη *συνθήκη καμπυλότητας*

$$s_k^T y_k > 0, \quad (2.6)$$

δηλαδή, αν ο πίνακας B_{k+1} είναι θετικά ορισμένος. Όταν η συνθήκη καμπυλότητας (2.6) ικανοποιείται, η εξίσωση τέμνουσας (2.4) έχει μοναδική λύση B_{k+1} .

Επομένως, στις Quasi-Newton μεθόδους είναι απαραίτητο να παράγουμε τον B_{k+1} ή τον αντίστροφό του H_{k+1} , με κάποιες συμφέρουσες μεθόδους, ώστε αφενός μεν να ισχύει η εξίσωση τέμνουσας (2.4) ή (2.5), αφετέρου δε να μην απαιτείται μεγάλο υπολογιστικό κόστος. Γενικά, ένας τέτοιος πίνακας παράγεται ενημερώνοντας τον B_k (ή τον H_k) με κάποιο τρόπο, ώστε να παραχθεί ο B_{k+1} (αντίστοιχα ο H_{k+1}). Σε όλες τις Quasi-Newton μεθόδους, ο αρχικός πίνακας B_0 (ή H_0) τίθεται να είναι είτε ο μοναδιαίος, δηλαδή $B_0 = I$, ή ένα πολλαπλάσιό του $B_0 = \theta I$, $\theta \in \mathbb{R}^*$. Ο

διαφορετικός τρόπος ενημέρωσης, χαρακτηρίζει τις διαφορετικές Quasi-Newton μεθόδους. Οι πιο γνωστές από αυτές είναι οι SR1, DFP, BFGS και PSB, οι οποίες αποτελούν την κλάση των Quasi-Newton ενημερώσεων, τη γενική PSB ενημέρωση, όπως είναι γνωστή.

2.3 Η γενική Powell-Symmetric-Broyden (PSB) ενημέρωση

Θα μπορούσαμε να πούμε ότι η γενική PSB ενημέρωση, αποτελεί το γεννιότερο των Quasi-Newton ενημερώσεων. Από αυτή την κλάση πινάκων, μπορούν να παραχθούν όλες οι ευρέως γνωστές Quasi-Newton ενημερώσεις. Δοθέντος ενός συμμετρικού πίνακα B , παράγεται μια ακολουθία πινάκων, η οποία συγκλίνει σε έναν συμμετρικό πίνακα, ο οποίος ικανοποιεί την εξίσωση της τέμνουσας (2.4).

Έστω $B \in \mathbb{R}^{n \times n}$ συμμετρικός πίνακας και

$$C_1 = B + \frac{(y - Bs) c^T}{c^T s},$$

όπου $c \in \mathbb{R}^n$ και $c^T s \neq 0$. Η ενημέρωση αυτή είναι γνωστή ως η *γενική ενημέρωση πρώτης τάξης του Broyden*. Εν γένει, ο πίνακας C_1 δεν είναι συμμετρικός. Γι αυτό θεωρούμε τον πίνακα

$$C_2 = \frac{1}{2} (C_1 + C_1^T).$$

Ο C_2 είναι μεν συμμετρικός, αλλά δεν ικανοποιεί πάντα την εξίσωση της τέμνουσας (2.4). Γι αυτό το λόγο, συνεχίζουμε τη διαδικασία και δημιουργούμε μια ακολουθία πινάκων $\{C_k\}$ ως εξής:

$$\begin{aligned} C_{2k+1} &= C_{2k} + \frac{(y - Bs) c^T}{c^T s}, \\ C_{2k+1} &= \frac{1}{2} (C_{2k+1} + C_{2k+1}^T), \quad k = 0, 1, \dots \end{aligned} \quad (2.7)$$

όπου $C_0 = B$. Δηλώνοντας με $Q(y, s) = \{C \in \mathbb{R}^{n \times n} \mid Cs = y\}$ το σύνολο πινάκων που ικανοποιούν την εξίσωση τέμνουσας (2.4), κάθε C_{2k+1} πίνακας είναι ο κοντινότερος πίνακας στον C_{2k} , που ανήκει στο $Q(y, s)$, ενώ κάθε C_{2k+2} είναι ο κοντινότερος συμμετρικός πίνακας στον C_{2k+1} , που

ανήκει στο $Q(y, s)$. Το όριο αυτής της ακολουθίας πινάκων $\{C_k\}$ είναι ο πίνακας [73]

$$\bar{B} = B + \frac{(y - Bs)c^T + c(y - Bs)^T}{s^T c} - \frac{(y - Bs)^T s}{s^T c} \frac{cc^T}{s^T c}, \quad (2.8)$$

ο οποίος είναι συμμετρικός και ικανοποιεί την Quasi-Newton εξίσωση (2.4).

Θεώρημα 2.2. (Powell [73]) Έστω συμμετρικός πίνακας $B \in \mathbb{R}^{n \times n}$, $c, s, y \in \mathbb{R}^n$ και $s^T c \neq 0$. Έστω η ακολουθία πινάκων $\{C_k\}$ που ορίζεται από τη σχέση (2.7), με $C_0 = B$. Τότε η ακολουθία $\{C_k\}$ συγκλίνει στον πίνακα $\bar{B}$, ορισμένο στην (2.8).

Ο πίνακας (2.8) παράγει μια κλάση ενημερώσεων δεύτερης τάξης (rank two) που προκύπτει από συμμετρικοποίηση. Αν ξαναγράψουμε τη σχέση (2.8) θεωρώντας και τους δείκτες της επανάληψης, η σχέση αυτή γίνεται

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) c_k^T + c_k (y_k - B_k s_k)^T}{s_k^T c_k} - \frac{(y_k - B_k s_k)^T s_k}{s_k^T c_k} \frac{s_k s_k^T}{s_k^T c_k}, \quad (2.9)$$

η οποία και αποτελεί τη γενική PSB ενημέρωση. Σε αυτή την κλάση ενημερώσεων, ανάλογα με την επιλογή του διανύσματος c_k , παράγεται και μία συγκεκριμένη Quasi-Newton ενημέρωση. Πιο αναλυτικά, αν θέσουμε στην (2.9)

- $c_k = y_k - B_k s_k$, προκύπτει η SR1 ενημέρωση πρώτης τάξης,
- $c_k = y_k$, έχουμε την DFP ενημέρωση δεύτερης τάξης,
- $c_k = (1/w_{k+1})y_k + (w_k/w_{k+1})B_k s_k$, όπου $w_k = (s_k^T y_k / s_k^T B_k s_k)^{1/2}$, δημιουργείται η BFGS ενημέρωση δεύτερης τάξης, και
- τέλος, για $c_k = s_k$, έχουμε την PSB ενημέρωση δεύτερης τάξης.

2.3.1 Η SR1 συμμετρική ενημέρωση πρώτης τάξης

Η SR1 (Symmetric Rank 1 update) είναι μια Quasi-Newton ενημέρωση, η οποία προκύπτει από τη γενική PSB (2.9), αν θέσουμε

$$c_k = y_k - B_k s_k.$$

Η SR1 παράγει τον B_{k+1} ενημερώνοντας τον B_k μέσω μιας ενημέρωσης της μορφής $B_{k+1} = B_k + E_k$, όπου ο E_k είναι ένας πίνακας μικρότερης τάξης. Έτσι, στην περίπτωση της ενημέρωσης πρώτης τάξης, έχουμε

$$B_{k+1} = B_k + uv^T, \quad (2.10)$$

όπου $u, v \in \mathbb{R}^n$. Από την εξίσωση της τέμνουσας (2.4) έχουμε ότι

$$B_{k+1}s_k = (B_k + uv^T)s_k = y_k,$$

οπότε,

$$(v^T s_k)u = y_k - B_k s_k. \quad (2.11)$$

Η παραπάνω σχέση δείχνει ότι το διάνυσμα u πρέπει να είναι στην κατεύθυνση του $y_k - B_k s_k$. Υποθέτοντας ότι $y_k - B_k s_k \neq 0$ (αλλιώς ο B_k θα ικανοποιούσε την εξίσωση τέμνουσας) και ότι το διάνυσμα u είναι τέτοιο ώστε $u^T s_k \neq 0$, από τις (2.10) και (2.11) έχουμε ότι

$$B_{k+1} = B_k + \frac{1}{v^T s_k} (y_k - B_k s_k) v^T, \quad (2.12)$$

όπου ο πίνακας B_{k+1} ονομάζεται (όπως είδαμε πιο πριν) Broyden γενική ενημέρωση πρώτης τάξης. Επειδή όμως ο B_{k+1} πρέπει να είναι συμμετρικός πίνακας, θέτουμε ως v το διάνυσμα $v = y_k - B_k s_k$. Έτσι, καταλήγουμε στην εξίσωση

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}. \quad (2.13)$$

Ο παραπάνω πίνακας ονομάζεται συμμετρική ενημέρωση πρώτης τάξης ή απλά SR1 ενημέρωση.

Έστω $H_k = B_k^{-1}$ ο αντίστροφος προσεγγιστικός εσσιανός πίνακας στην k επανάληψη. Τότε από τον τύπο των Sherman-Morrison-Woodbury (

σελ. 8), έχουμε ότι ο αντίστροφος της SR1 ενημέρωσης (2.13) είναι ο

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)(s_k - H_k y_k)^T}{(s_k - H_k y_k)^T y_k}. \quad (2.14)$$

Μερικές φορές, ο παρανομαστής $(y_k - B_k s_k)^T s_k$ της SR1 (2.13) είναι πολύ μικρός ή ακόμα και μηδέν, με αποτέλεσμα να υπάρχουν σοβαρές αριθμητικές δυσκολίες ή ακόμα και να καταρρεύσει ο αλγόριθμος. Έχει παρατηρηθεί στην πράξη ότι η SR1 συμπεριφέρεται καλά αν παραλείψουμε την ενημέρωση στην περίπτωση που ο παρανομαστής είναι μικρός. Πιο συγκεκριμένα, η ενημέρωση γίνεται μόνο όταν

$$|s_k^T (y_k - B_k s_k)| \geq r \|s_k\| \|y_k - B_k s_k\|, \quad (2.15)$$

όπου το $r \in (0, 1)$ είναι ένας πολύ μικρός αριθμός, της τάξης του 10^{-8} . Αν η ανίσωση (2.15) δεν ισχύει, απλώς θέτουμε $B_{k+1} = B_k$, δηλαδή χρησιμοποιούμε τον ίδιο προσεγγιστικό πίνακα για την επόμενη επανάληψη του αλγορίθμου.

2.3.2 Η DFP συμμετρική ενημέρωση δεύτερης τάξης

Η DFP ενημέρωση (Davidon, Fletcher, Powell) είναι μια συμμετρική ενημέρωση δεύτερης τάξης, δηλαδή ο B_k ή αντίστοιχα ο αντίστροφος του H_k , ενημερώνεται από την πρόσθεση δύο πινάκων πρώτης τάξης. Η DFP προκύπτει από τη γενική PSB ενημέρωση, αν στον πίνακα (2.9) θέσουμε

$$c_k = y_k.$$

Έστω ο αντίστροφος προσεγγιστικός πίνακας H_k . Η δεύτερης τάξης ενημέρωση θα είναι της μορφής

$$H_{k+1} = H_k + a u u^T + b v v^T, \quad (2.16)$$

όπου τα βαθμωτά μεγέθη $a, b \in \mathbb{R}$, καθώς και τα διανύσματα $u, v \in \mathbb{R}^n$ πρέπει να καθοριστούν. Από την Quasi-Newton εξίσωση (2.5) και τη σχέ-

ση (2.16) έχουμε:

$$\begin{aligned} H_{k+1}y_k &= s_k \\ (H_k + auu^T + bvv^T)y_k &= s_k \\ H_k y_k + ay_k^T uu + bv^T y_k v &= s_k. \end{aligned} \quad (2.17)$$

Προφανώς τα u και v δεν μπορούν να οριστούν μονοσήμαντα, ωστόσο οι πιο απλή επιλογή είναι να θέσουμε $u = s_k$ και $v = H_k y_k$. Αντικαθιστώντας τα διανύσματα u και v στην (2.17) βρίσκουμε ότι

$$\begin{aligned} H_k y_k + as_k^T y_k s_k + by_k^T H_k y_k H_k y_k &= s_k \\ (1 + by_k^T H_k y_k) H_k y_k + (as_k^T y_k - 1) s_k &= 0, \end{aligned}$$

οπότε $a = 1/s_k^T y_k$ και $b = -1/y_k^T H_k y_k$. Έτσι, η σχέση (2.16) γίνεται

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k^T y_k H_k}{y_k^T H_k y_k}. \quad (2.18)$$

Η εξίσωση (2.18) αποτελεί την πρώτη Quasi-Newton ενημέρωση που προτάθηκε από τον Davidon [27] το 1959, ενώ μελετήθηκε περαιτέρω από τους Fletcher & Powell [38], από τους οποίους έλαβε και το όνομά της ως DFP ενημέρωση. Από τον τύπο των Sherman-Morrison-Woodbury (σελ. 8), βρίσκουμε ότι ο προσεγγιστικός εσσιανός είναι ο πίνακας

$$\begin{aligned} B_{k+1} = B_k &+ \frac{(y_k - B_k s_k) y_k^T + y_k (y_k - B_k s_k)^T}{s_k^T y_k} \\ &- \frac{(y_k - B_k s_k)^T s_k}{(s_k^T y_k)^2} y_k y_k^T. \end{aligned} \quad (2.19)$$

Η DFP μέθοδος έχει χρησιμοποιηθεί εκτενώς σε πολλές εφαρμογές και έχει παίξει σημαντικό ρόλο στη θεωρητική ανάλυση των αριθμητικών αλγορίθμων. Ωστόσο, εμπειρισταωμένες αριθμητικές μελέτες της μεθόδου έχουν δείξει ότι είναι υπολογιστικά ασταθής και μερικές φορές παράγει αριθμητικά ιδιάζοντες προσεγγιστικούς πίνακες της εσσιανής. Εν αντιθέσει, η ευρέως γνωστή BFGS ενημέρωση, που θα δούμε στην επόμενη ενότητα, δεν παρουσιάζει αριθμητικές αστάθειες με αποτέλεσμα η απόδοσή της να είναι ανώτερη από την απόδοση της DFP.

2.3.3 Η BFGS συμμετρική ενημέρωση δεύτερης τάξης

Όπως και η DFP, έτσι και η BFGS (Broyden, Fletcher, Goldfarb, Shanno) είναι μια συμμετρική ενημέρωση δεύτερης τάξης, δηλαδή ο B_k ενημερώνεται από την πρόσθεση δύο πινάκων πρώτης τάξης. Προκύπτει από τη γενική PSB ενημέρωση (2.9), αν θέσουμε

$$c_k = \frac{1}{w_{k+1}} y_k + \frac{w_k}{w_{k+1}}, \quad \text{όπου} \quad w_k = \left(\frac{s_k^T y_k}{s_k^T B_k s_k} \right)^{1/2}.$$

Θεωρούμε λοιπόν τη γενική μορφή μιας ενημέρωσης δεύτερης τάξης

$$B_{k+1} = B_k + a u u^T + b v v^T, \quad (2.20)$$

όπου τα βαθμωτά μεγέθη $a, b \in \mathbb{R}$, όπως επίσης και τα διανύσματα $u, v \in \mathbb{R}^n$ θα πρέπει να προσδιοριστούν. Από την Quasi-Newton εξίσωση (2.4) και την (2.20) έχουμε:

$$\begin{aligned} B_{k+1} s_k &= y_k \\ (B_k + a u u^T + b v v^T) s_k &= y_k \\ B_k s_k + a s_k^T u u + b v^T s_k v &= y_k. \end{aligned} \quad (2.21)$$

Για άλλη μια φορά τα u και v δεν μπορούν να οριστούν μονοσήμαντα, αλλά οι προφανείς επιλογές είναι να θέσουμε $u = y_k$ και $v = B_k s_k$. Τότε από την (2.21) έχουμε:

$$\begin{aligned} B_k s_k + a s_k^T y_k y_k + b s_k^T B_k s_k B_k s_k &= y_k \\ (1 + b s_k^T B_k s_k) B_k s_k + (a s_k^T y_k - 1) y_k &= 0, \end{aligned}$$

οπότε $a = 1/s_k^T y_k$ και $b = -1/s_k^T B_k s_k$. Έτσι, η σχέση (2.20) γίνεται

$$B_{k+1} = B_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}, \quad (2.22)$$

που ονομάζεται BFGS ενημέρωση, επειδή ανακαλύφθηκε ξεχωριστά από τους Broyden [24], Fletcher [36], Goldfarb [41] και Shanno [80]. Η BFGS ενημέρωση μπορεί να υπολογιστεί από την DFP ενημέρωση, αντικαθιστώντας απλά τους $H_{k+1} \leftrightarrow B_{k+1}$ και $s_k \leftrightarrow y_k$.

Για να υπολογίσουμε τον αντίστροφο H_{k+1} του B_{k+1} , χρησιμοποιούμε τον τύπο των Sherman-Morrison-Woodbury (σελ. 8), από όπου λαμβάνουμε ότι ο H_{k+1} εκφράζεται από την ακόλουθη σχέση:

$$H_{k+1} = H_k + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{s_k^T y_k}. \quad (2.23)$$

2.3.4 Η PSB ενημέρωση δεύτερης τάξης

Η PSB ενημέρωση, επίσημα γνωστή ως η Powell-συμμετρική-Broyden ενημέρωση (Powell-Symmetric-Broyden) λόγω του Powell [73], είναι και αυτή μια ενημέρωση δεύτερης τάξης. Προκύπτει από τη γενική PSB ενημέρωση (2.9), αν θέσουμε

$$c_k = s_k$$

και είναι ο πίνακας

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T + s_k (y_k - B_k s_k)^T}{s_k^T s_k} - \frac{(y_k - B_k s_k)^T s_k}{s_k^T s_k} \frac{s_k s_k^T}{s_k^T s_k}. \quad (2.24)$$

Για την εύρεση του αντιστρόφου, από τον τύπο των Sherman-Morrison-Woodbury (σελ. 8), λαμβάνουμε την έκφραση

$$H_{k+1} = H_k - \frac{h_k h_k^T (s_k^T H_k s_k) - (h_k s_k^T H_k + H_k s_k h_k^T) (s_k^T H_k y_k)}{(h_k^T y_k) (s_k^T H_k s_k) - (s_k^T H_k y_k)^2} + \frac{H_k s_k s_k^T H_k (h_k^T y_k)}{(h_k^T y_k) (s_k^T H_k s_k) - (s_k^T H_k y_k)^2},$$

όπου $h_k = H_k y_k - s_k$.

2.3.5 Η Broyden κλάση

Αν θεωρήσουμε το συνδυασμό των DFP και BFGS ενημερώσεων, παίρνουμε την κλάση ενημερώσεων

$$B_{k+1} = (1 - \theta)B_{k+1}^{BFGS} + \theta B_{k+1}^{DFP} \quad (2.25)$$

$$= \theta B_{k+1}^{BFGS} + \theta w_k w_k^T \quad (2.26)$$

$$= B_{k+1}^{DFP} + (\theta - 1)w_k w_k^T \quad (2.27)$$

$$= B_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \theta w_k w_k^T, \quad (2.28)$$

όπου θ παράμετρος και

$$w_k = (s_k^T B_k s_k)^{1/2} \left(\frac{y_k}{s_k^T y_k} - \frac{B_k s_k}{s_k^T B_k s_k} \right).$$

Οι σχέσεις (2.26)-(2.28), περιγράφουν την Broyden κλάση. Προφανώς, για $\theta = 0$ έχουμε την BFGS ενημέρωση, ενώ για $\theta = 1$ έχουμε τη DFP ενημέρωση.

Ο αντίστροφος του πίνακα B_{k+1} , είναι ο πίνακας

$$H_{k+1} = \phi H_{k+1}^{BFGS} + (1 - \phi)H_{k+1}^{DFP} \quad (2.29)$$

$$= H_{k+1}^{BFGS} + (\phi - 1)v_k v_k^T \quad (2.30)$$

$$= H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \phi v_k v_k^T, \quad (2.31)$$

όπου ϕ είναι παράμετρος και

$$v_k = (y_k^T H_k y_k)^{1/2} \left(\frac{s_k}{s_k^T y_k} - \frac{H_k y_k}{y_k^T H_k y_k} \right).$$

Αν $0 \leq \phi \leq 1$, οι σχέσεις (2.30)-(2.31), περιγράφουν την κυρτή Broyden κλάση ενημερώσεων. Για $\phi = 0$ έχουμε την DFP ενημέρωση, για $\phi = 1$ τη BFGS και για $\phi = \frac{s_k^T y_k}{y_k^T (s_k - H_k y_k)}$, την SR1 ενημέρωση.

Η σχέση που συνδέει τις παραμέτρους ϕ και θ είναι η εξής:

$$\theta = \frac{\phi - 1}{\phi - 1 - \phi\mu}, \quad \text{όπου} \quad \mu = \frac{y_k^T H_k y_k s_k^T B_k s_k}{(s_k^T y_k)^2}.$$

Προφανώς, επειδή ισχύει ότι $w_k^T s_k = 0$ και $v_k^T y_k = 0$, η Broyden κλάση (2.26)-(2.28) και (2.30)-(2.31), ικανοποιεί την εξίσωση τέμνουσας (2.4) και (2.5), για κάθε παράμετρο ϕ και θ .

2.4 Οι Quasi-Newton ενημερώσεις περιορισμένης μνήμης

Οι Quasi-Newton ενημερώσεις δεν είναι εφαρμόσιμες για μεγάλης κλίμακας προβλήματα βελτιστοποίησης γιατί οι υπολογιστικές απαιτήσεις αυξάνονται αναλογικά με το n^2 . Γι αυτό το λόγο αναπτύχθηκαν οι περιορισμένης μνήμης Quasi-Newton μέθοδοι, γνωστές στη διεθνή βιβλιογραφία ως limited-memory Quasi-Newton methods [67, 68]. Σε προβλήματα βελτιστοποίησης μεγάλης κλίμακας, οι μέθοδοι αυτές είναι οι πιο κατάλληλες για να προσεγγίσουν την εσσιανή της αντικειμενικής συνάρτησης.

Οι περιορισμένης μνήμης Quasi-Newton μέθοδοι χρησιμοποιούν πληροφορία καμπυλότητας μόνο από τις m πιο πρόσφατες επαναλήψεις, με σκοπό να κατασκευάσουν τον προσεγγιστικό εσσιανό πίνακα. Η πληροφορία αυτή αποθηκεύεται σε m διανυσματικά ζεύγη $\{s_i, y_i\}$, όπου $i = k-m, \dots, k-1$. Συνεπώς, αποθηκεύονται μόνο μερικά n -διάστατα διανύσματα, αντί να αποθηκευθούν και να υπολογιστούν οι $n \times n$ προσεγγίσεις της εσσιανής. Αφού υπολογιστεί η νέα επανάληψη, το παλαιότερο ζεύγος από τα $\{s_i, y_i\}$ διαγράφεται και αντικαθίσταται από το καινούργιο ζεύγος $\{s_k, y_k\}$. Η επιλογή του αριθμού m των αποθηκευμένων ζευγών $\{s_i, y_i\}$, εξαρτάται τόσο από τη διάσταση του προβλήματος που έχουμε να λύσουμε, όσο και από τη μνήμη του υπολογιστή.

Η πιο γνωστή από τις Quasi-Newton μεθόδους περιορισμένης μνήμης είναι η περιορισμένης μνήμης BFGS μέθοδος, γνωστή ως L-BFGS [56, 67]. Ο αντίστροφος προσεγγιστικός πίνακας H_{k+1} δίνεται από τον τύπο

$$\begin{aligned} H_{k+1} &= H_k + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{s_k^T y_k} \\ &= \left(I - \frac{s_k y_k^T}{s_k^T y_k}\right) H_k \left(I - \frac{y_k s_k^T}{s_k^T y_k}\right) + \frac{s_k s_k^T}{s_k^T y_k}. \end{aligned}$$

Θέτοντας

$$\rho_k = \frac{1}{s_k^T y_k} \quad \text{και} \quad V_k = I - \rho_k y_k s_k^T,$$

ο πίνακας H_{k+1} μπορεί να γραφεί ως

$$H_{k+1} = V_k^T H_k V_k + \rho_k s_k s_k^T, \quad (2.32)$$

δηλαδή ο H_k ενημερώνεται χρησιμοποιώντας το ζεύγος $\{s_k, y_k\}$. Στην L-BFGS μέθοδο σώζουμε m ζεύγη (s_i, y_i) , $i = k - m, k - m + 1, \dots, k - 1$ για να ανακατασκευάσουμε τον H_k . Στην k επανάληψη, αφού διαλέξουμε έναν αρχικό πίνακα $H_k^{(0)}$, εφαρμόζουμε τον τύπο (2.32) επαναληπτικά m φορές, δηλαδή

$$H_k^{(j+1)} = V_{k-m+j}^T H_k^{(j)} V_{k-m+j} + \rho_{k-m+j} s_{k-m+j} s_{k-m+j}^T, \quad j = 0, \dots, m-1$$

οπότε λαμβάνουμε

$$\begin{aligned} H_k = & (V_{k-1}^T \dots V_{k-m}^T) H_k^{(0)} (V_{k-m} V_{k-m+1} \dots V_{k-1}) \\ & + \rho_{k-m} (V_{k-1}^T \dots V_{k-m+1}^T) s_{k-m} s_{k-m}^T (V_{k-m+1} \dots V_{k-1}) \\ & + \rho_{k-m+1} (V_{k-1}^T \dots V_{k-m+2}^T) s_{k-m+1} s_{k-m+1}^T (V_{k-m+2} \dots V_{k-1}) \\ & + \dots \\ & + \rho_{k-1} s_{k-1} s_{k-1}^T. \end{aligned}$$

Μια επιλογή του αρχικού πίνακα $H_k^{(0)}$ θα μπορούσε να είναι είτε ο μοναδιαίος ή κάποιο πολλαπλάσιό του, όπως για παράδειγμα

$$H_k^{(0)} = \frac{s_k^T y_k}{y_k^T y_k} I \quad \text{ή} \quad H_k^{(0)} = \frac{s_k^T s_k}{s_k^T y_k} I.$$

2.4.1 Οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης

Οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης [68] χρησιμοποιούνται ευρέως σε προβλήματα βελτιστοποίησης μεγάλης κλίμακας [4, 17, 33]. Σε αυτές τις μεθόδους ο προηγούμενος προσεγγιστικός εσσιανός πίνακας, πριν ενημερωθεί, αντικαθίσταται από το μοναδιαίο ή κάποιο πολλαπλάσιό του, ενώ μόνο το πιο πρόσφατο διανυσματικό ζεύγος (s_k, y_k) φυλάσσεται σε κάθε επανάληψη ($m = 1$). Επομένως, έχουμε μια ενημέρωση της

μορφής

$$B_{k+1} = U(B_k^{(0)}, s_k, y_k),$$

όπου ο $B_k^{(0)} = \theta I$, $\theta \in \mathbb{R}^*$, είναι ο πίνακας που αντικαθιστά τον προηγούμενο προσεγγιστικό εσσιανό. Συνεπώς, η κλάση της ελάχιστης μνήμης γενικής PSB ενημέρωσης είναι η ακόλουθη:

$$B_{k+1} = B_k^{(0)} + \frac{(y_k - B_k^{(0)} s_k) c_k^T + c_k (y_k - B_k^{(0)} s_k)^T}{s_k^T c_k} - \frac{(y_k - B_k^{(0)} s_k)^T s_k}{s_k^T c_k} \frac{s_k s_k^T}{s_k^T c_k}.$$

Αν θέσουμε $B_k^{(0)} = \theta I$, τότε η παραπάνω ενημέρωση μπορεί να γραφεί ως εξής:

$$B_{k+1} = \theta I + \frac{(y_k - \theta s_k) c_k^T + c_k (y_k - \theta s_k)^T}{s_k^T c_k} - \frac{(y_k - \theta s_k)^T s_k}{s_k^T c_k} \frac{s_k s_k^T}{s_k^T c_k},$$

όπου για

- $c_k = y_k - \theta s_k$, έχουμε την ελάχιστης μνήμης SR1 ενημέρωση,
- $c_k = y_k$, δημιουργείται η ελάχιστης μνήμης DFP ενημέρωση,
- $c_k = (1/w_{k+1})y_k + (\theta w_k/w_{k+1})s_k$, όπου $w_k = (s_k^T y_k / \theta s_k^T s_k)^{1/2}$, λαμβάνουμε την ελάχιστης μνήμης BFGS ενημέρωση και για
- $c_k = s_k$, προκύπτει η ελάχιστης μνήμης PSB ενημέρωση.

2.5 Οι μέθοδοι γραμμικής αναζήτησης

Οι μέθοδοι γραμμικής αναζήτησης αναφέρονται στη βελτιστοποίηση μιας μονοδιάστατης συνάρτησης και αποτελούν τη βάση της βελτιστοποίησης συναρτήσεων πολλών μεταβλητών. Δοθέντος ενός σημείου x_k , οι μέθοδοι γραμμικής αναζήτησης δημιουργούν μια ακολουθία επαναλήψεων μέσω του σχήματος

$$x_{k+1} = x_k + \alpha_k d_k.$$

Δοθείσης μιας κατεύθυνσης μείωσης d_k , σκοπός είναι να βρούμε ένα κατάλληλο μήκος βήματος α_k , έτσι ώστε στη νέα επανάληψη η αντικειμενική συνάρτηση να έχει μικρότερη τιμή από ότι στην προηγούμενη επανάληψη. Αν θέσουμε

$$\phi(\alpha) = f(x_k + \alpha d_k),$$

το πρόβλημα της αναζήτησης ενός μήκους βήματος στην κατεύθυνση d_k τέτοιου ώστε

$$\phi(\alpha_k) < \phi(0),$$

είναι απλά μια γραμμική αναζήτηση για το α . Η επιλογή που θα έδινε το βέλτιστο α_k είναι θεωρητικά η λύση του ακόλουθου μονοδιάστατου προβλήματος ελαχιστοποίησης

$$\min_{\alpha > 0} \phi(\alpha), \quad \text{όπου} \quad \phi(\alpha) = f(x_k + \alpha d_k). \quad (2.33)$$

Εντούτοις, ο ακριβής ελαχιστοποιητής του προβλήματος (2.33) είναι υπολογιστικά ακριβός και μη απαραίτητος. Επομένως, αντί να λυθεί το πρόβλημα (2.33) επακριβώς και να εκτελεστεί μία *ακριβής γραμμική αναζήτηση*, συνήθως προτιμάται η επίλυση του προβλήματος (2.33) προσεγγιστικά, εκτελώντας μία *μη-ακριβή γραμμική αναζήτηση* (inexact), στην οποία επιβάλλονται κάποιες συνθήκες, ώστε να εξασφαλιστεί μια επαρκής μείωση της f . Οι συνθήκες αυτές χρησιμοποιούν πληροφορία πρώτης τάξης, ενώ οι πιο ευρέως χρησιμοποιούμενες είναι η *συνθήκη Armijo* [13], οι *συνθήκες Goldstein* [44], και οι *συνθήκες Wolfe* [89].

Η συνθήκη Armijo [13] δηλώνει ότι δοθέντων $\beta \in (0, 1)$, $\rho \in (0, \frac{1}{2})$ και $\tau > 0$, υπάρχει τουλάχιστον ένας μη αρνητικός ακέραιος m_k , τέτοιος ώστε

$$f(x_k + \beta^{m_k} \tau d_k) \leq f(x_k) + \rho \beta^{m_k} \tau \nabla f(x_k)^T d_k. \quad (2.34)$$

Ο Goldstein [44] θεώρησε το διάστημα

$$J = \{\alpha > 0 \mid f(x_k + \alpha d_k) < f(x_k)\}.$$

Για να εγγυηθούμε ότι η αντικειμενική συνάρτηση μειώνεται επαρκώς, διαλέγουμε ένα μήκος βήματος α τέτοιο ώστε να είναι μακριά από τα δύο άκρα του διαστήματος J . Οι συνθήκες για να το επιτύχουμε αυτό είναι οι

ακόλουθες:

$$f(x_k) + (1 - \rho)\alpha \nabla f(x_k)^T d_k \leq f(x_k + \alpha d_k) \quad (2.35)$$

$$f(x_k + \alpha d_k) \leq f(x_k) + \rho\alpha \nabla f(x_k)^T d_k, \quad (2.36)$$

όπου $0 < \rho < 1/2$.

Οι πιο ευρέως χρησιμοποιούμενες συνθήκες είναι οι συνθήκες Wolfe [89]:

$$f(x_k + \alpha d_k) \leq f(x_k) + c_1 \alpha \nabla f(x_k)^T d_k, \quad (2.37)$$

$$\nabla f(x_k + \alpha d_k)^T d_k \geq c_2 \nabla f(x_k)^T d_k, \quad (2.38)$$

όπου $0 < c_1 < c_2 < 1$. Η γεωμετρική αναπαράσταση της δεύτερης συνθήκης (2.38), γνωστής ως *συνθήκη καμπυλότητας*, είναι ότι η κλίση $\phi'(\alpha_k)$ στο δεκτό σημείο πρέπει να είναι μεγαλύτερη ή ίση από ένα πολλαπλάσιο c_2 της αρχικής κλίσης.

Ένας τυπικός αλγόριθμος γραμμικής αναζήτησης είναι ο ακόλουθος:

Αλγόριθμος 2.2 Μέθοδος γραμμικής αναζήτησης

Βήμα 1: Δοθέντων x_0 , και $\epsilon \rightarrow 0$, θέσε $k = 0$, και υπολόγισε την $\nabla f(x_0)$;

Βήμα 2: Αν $\|\nabla f(x_k)\| \leq \epsilon$, σταμάτα;

Βήμα 3: Υπολόγισε μια κατεύθυνση μείωσης d_k τέτοια ώστε

$$d_k^T \nabla f(x_k) < 0.$$

Βήμα 4: Βρες $\alpha_k > 0$ τέτοιο ώστε να ικανοποιούνται οι συνθήκες Armijo (2.34), Goldstein (2.35)-(2.36) ή Wolfe (2.37)-(2.38).

Βήμα 4: Θέσε $x_{k+1} = x_k + \alpha_k p_k$.

Βήμα 5: Υπολόγισε την $\nabla f(x_{k+1})$, θέσε $k = k + 1$ και πήγαινε στο Βήμα 2;

Στην πράξη, πολύ συχνά χρησιμοποιούμε μόνο τη συνθήκη (2.37), αν διαλέξουμε ένα κατάλληλο α που δεν είναι πολύ μικρό. Η μέθοδος αυτή

ονομάζεται μέθοδος γραμμικής αναζήτησης υπαναχώρησης (backtracking line search). Η ιδέα είναι, να δοκιμάσουμε πρώτα το μήκος βήματος $\alpha = 1$. Αν το $x_k + \alpha_k d_k$ δεν είναι αποδεκτό, τότε μειώνουμε το α έως ότου το $x_k + \alpha_k d_k$ ικανοποιήσει τη συνθήκη (2.37)

Για να αποδείξουμε την ιδιότητα της μείωσης της f στη μέθοδο της γραμμικής αναζήτησης, προσπαθούμε να αποφύγουμε την περίπτωση όπου οι κατευθύνσεις αναζήτησης $\alpha_k d_k$ είναι σχεδόν κάθετες με την αρνητική κλίση $-\nabla f(x_k)$. Άρα, η γωνία ϑ_k μεταξύ της $\alpha_k d_k$ και της $-\nabla f(x_k)$ πρέπει να είναι ομοιόμορφα φραγμένη μακριά από τις 90° , δηλαδή

$$\vartheta_k \leq \frac{\pi}{2} - \mu, \quad \text{για κάθε } k, \quad (2.39)$$

όπου $\mu > 0$. Η $\vartheta_k \in [0, \frac{\pi}{2}]$ ορίζεται από τη σχέση

$$\cos \vartheta_k = -\frac{\alpha_k d_k^T \nabla f(x_k)}{\|\alpha_k d_k\| \|\nabla f(x_k)\|}.$$

Αλλιώς, το εσωτερικό γινόμενο $\alpha_k d_k^T \nabla f(x_k)$ τείνει στο μηδέν, με αποτέλεσμα η $\alpha_k d_k$ να μην είναι κατεύθυνση μείωσης. Το ακόλουθο θεώρημα αποτελεί το θεώρημα ολικής σύγκλισης του γενικού Αλγόριθμου 2.2 ανακρίβους γραμμικής αναζήτησης:

Θεώρημα 2.3. Έστω ότι το μήκος βήματος α_k στον Αλγόριθμο 2.2, ορίζεται από τις συνθήκες Armijo (2.34), Goldstein (2.35)-(2.36) ή Wolfe (2.37)-(2.38). Έστω ακόμα ότι η d_k ικανοποιεί τη σχέση (2.39). Αν η ∇f υπάρχει και είναι ομοιόμορφα συνεχής στο σύνολο $\{x | f(x) \leq f(x_0)\}$, τότε είτε $\nabla f(x_k) = 0$ για κάποιο k , ή $f(x_k) \rightarrow -\infty$ ή $\nabla f(x_k) \rightarrow 0$.

2.6 Οι μέθοδοι καμπυλόγραμμης αναζήτησης

Θεωρούμε το πρόβλημα βελτιστοποίησης

$$\min f(x), \quad x \in \mathcal{D} \subset \mathbb{R}^n,$$

στο οποίο η αντικειμενική συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι συνεχώς διαφορίσιμη στο ανοιχτό σύνολο $\mathcal{D}$. Για την επίλυση αυτού του προβλήματος, πέρα από τις μεθόδους γραμμικής αναζήτησης, οι οποίες χρησιμοποιούν πρώτης τάξης πληροφορία, χρησιμοποιούνται και οι μέθοδοι καμπυλόγραμ-

μης αναζήτησης, οι οποίες χρησιμοποιούν πληροφορία δεύτερης τάξης. Η δεύτερης τάξης πληροφορία εισάγεται μέσω της κατεύθυνσης αρνητικής καμπυλότητας. Υπενθυμίζουμε εδώ ότι, για μια συνάρτηση η οποία είναι δύο φορές συνεχώς διαφορίσιμη, το διάνυσμα $d \in \mathbb{R}^n$ είναι μια κατεύθυνση αρνητικής καμπυλότητας σε ένα σημείο $x \in \mathbb{R}^n$, αν ισχύει η σχέση

$$d^T \nabla^2 f(x) d < 0,$$

στο x . Αυτή η κατεύθυνση συνδέεται με το ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του εσσιανού πίνακα της f . Στην κατεύθυνση αρνητικής καμπυλότητας, το τετραγωνικό μοντέλο της αντικειμενικής συνάρτησης

$$m_k(x_k + d) = f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T \nabla^2 f(x_k) d,$$

είναι μη φραγμένο από κάτω, με αποτέλεσμα να υπάρχει δυνατότητα μεγάλης μείωσης της τιμής της f , προς την κατεύθυνση αυτή.

Ιστορικά, η χρήση της κατεύθυνσης αρνητικής καμπυλότητας προτάθηκε το 1968 από τους Fiacco & McCormick [35] και το 1976 από τους Botsaris & Jacobson [23]. Το 1977, ο McCormick [61] τροποποίησε τον κανόνα του Armijo (2.34) έτσι ώστε η μέθοδος να μπορεί να χρησιμοποιεί πληροφορία δεύτερης τάξης, ειδικά στη περίπτωση όπου ο εσσιανός πίνακας δεν είναι θετικά ημιορισμένος. Το 1979, οι Moré & Sorensen [63], καθώς και ο Goldfarb [42] το 1980, πρότειναν μια παρόμοια προσέγγιση, στην οποία η αναζήτηση μιας καινούργιας επανάληψης γινόταν μέσω της καμπύλης

$$C = \{x(\alpha) : x(\alpha) = x + \phi_1(\alpha)p + \phi_2(\alpha)d, \alpha \geq 0\}$$

με $\phi_1(0) = \phi_2(0) = 0$, συνδυάζοντας δύο κατευθύνσεις: την κατεύθυνση p , που είναι μια κατεύθυνση τύπου Newton, και την κατεύθυνση αρνητικής καμπυλότητας, d . Επιπλέον, στη δουλειά των Moré & Sorensen [63], χρησιμοποιήθηκε ένας κανόνας τύπου Armijo, ώστε να εξασφαλίσει τη σύγκλιση της μεθόδου σε δεύτερης τάξης κρίσιμα σημεία, όπου ο εσσιανός είναι θετικά ημιορισμένος. Πρόσφατα, στην εργασία των Gould et al. [46], προτάθηκε ένας εναλλακτικός τρόπος χρησιμοποίησης της κατεύθυνσης αρνητικής καμπυλότητας. Η κατεύθυνση αυτή χρησιμοποιήθηκε

μαζί με την κατεύθυνση τύπου Newton σε μια κατάλληλα διαμορφωμένη γραμμική αναζήτηση. Τέλος, στην εργασία των Olivares et al. [70], οι συγγραφείς απέδειξαν συνθήκες που παρέχουν τη δυνατότητα σε κάθε επανάληψη, να μπορεί να επιλεγεί είτε μια γραμμική αναζήτηση (η οποία χρησιμοποιεί μία μόνο από τις κατευθύνσεις) ή μια καμπυλόγραμμη αναζήτηση, στην οποία συνδυάζονται και οι δύο κατευθύνσεις.

Ο υπολογισμός της κατεύθυνσης αρνητικής καμπυλότητας είναι ισοδύναμος με τον υπολογισμό του ιδιοδιανύσματος που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του εσσιανού πίνακα [22, 21, 20]. Οι Moré & Sorensen [63] χρησιμοποίησαν την παραγοντοποίηση των Bunch & Parlett, ενώ οι Ferris et al. [34] ενσωμάτωσαν αυτή την προσέγγιση με ένα μη μονότονο πλαίσιο αναζήτησης. Και στις δύο περιπτώσεις, ο υπολογισμός της κατεύθυνσης αρνητικής καμπυλότητας απαιτεί την παραγοντοποίηση πινάκων καθώς και στην αποθήκευσή τους, κάτι το οποίο είναι απαγορευτικό, όταν η διάσταση του προβλήματος είναι πολύ μεγάλη. Στην εργασία των Lucidi et al. [58], ο υπολογισμός της κατεύθυνσης αρνητικής καμπυλότητας βασίστηκε στη μέθοδο Lanczos. Όμως, επειδή η μέθοδος απαιτεί την αποθήκευση πίνακα, μόνο λίγα διανύσματα της διαδικασίας Lanczos αποθηκεύονται και όχι όλα όσα θα έπρεπε. Άλλες δουλειές που χρησιμοποιούν την κατεύθυνση αρνητικής καμπυλότητας, την υπολογίζουν μέσω της μεθόδου Lanczos, σε συνδυασμό με τη μέθοδο των συζυγών κλίσεων [46, 59]. Σε αυτές τις μεθόδους, είναι απαραίτητο να επαναλαμβάνεται η αναδρομική διαδικασία, έτσι ώστε να ξαναδημιουργούνται τα Lanczos διανύσματα, όπου αυτό είναι απαραίτητο. Πιο πρόσφατα, το 2007, οι Fasano & Roma [31] πρότειναν τη χρήση αλγορίθμων συζυγών κλίσεων για τον υπολογισμό της κατεύθυνσης αρνητικής καμπυλότητας.

2.6.1 Συνθήκες βήματος δεύτερης τάξης

Στη συνέχεια θα παρουσιάσουμε αναλυτικά τις συνθήκες βήματος δεύτερης τάξης, πάνω στις οποίες βασίζονται οι μέθοδοι καμπυλόγραμμης αναζήτησης. Οι συνθήκες αυτές χρησιμοποιούν την κατεύθυνση αρνητικής καμπυλότητας.

Υπόθεση 2.1. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ δύο φορές συνεχώς διαφορίσιμη συνάρτηση σε ένα ανοιχτό σύνολο $\mathcal{D}$ και υποθέτουμε ότι για κάποιο $x_0 \in \mathcal{D}$, το

σύνολο

$$\Omega = \{x \in \mathcal{D} : f(x) \leq f(x_0)\}$$

είναι συμπαγές.

Υπό την Υπόθεση 7.1, μπορούμε να κατασκευάσουμε ένα τοπικό τετραγωνικό μοντέλο της αντικειμενικής συνάρτησης f , από το αντίστοιχο ανάπτυγμα της σειράς Taylor, για όλες τις επαναλήψεις $x_k \in \Omega$, ως ακολούθως:

$$m_k(x_k + d) = f(x_k) + g(x_k)^T p + \frac{1}{2} p^T B_k p,$$

όπου B_k είναι ένας θετικά ορισμένος συμμετρικός πίνακας. Με αυτό τον τρόπο, η Quasi-Newton κατεύθυνση p_k μπορεί να υπολογιστεί από τον τύπο

$$p_k = \begin{cases} -B_k^{-1} g_k, & \text{αν } B_k > 0; \\ -g_k, & \text{διαφορετικά.} \end{cases}$$

Η κατεύθυνση αρνητικής καμπυλότητας d_k , σχετίζεται με το ιδιοδιάνυσμα u_k που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του πίνακα B_k και υπολογίζεται από τον τύπο

$$d_k = \begin{cases} 0, & \text{αν } B_k > 0; \\ -\text{sgn}(u_k^T g_k) u_k, & \text{διαφορετικά.} \end{cases}$$

Και οι δύο κατευθύνσεις p_k και d_k θα πρέπει να είναι κατευθύνσεις μείωσης.

Ορισμός 2.1. Ένα ζεύγος κατευθύνσεων (p, d) ονομάζεται ζεύγος μείωσης σε ένα σημείο x στο οποίο η $f(x)$ είναι δυο φορές συνεχώς διαφορίσιμη, αν ισχύουν οι σχέσεις

$$p^T \nabla f(x) < 0, \quad d^T \nabla f(x) \leq 0, \quad \text{και} \quad d^T \nabla^2 f(x) d = 0,$$

όταν $\nabla^2 f(x) \geq 0$, και

$$p^T \nabla f(x) \leq 0, \quad d^T \nabla f(x) \leq 0, \quad \text{και} \quad d^T \nabla^2 f(x) d < 0,$$

διαφορετικά.

Από μια πιο πρακτική σκοπιά, σε μια δοσμένη επανάληψη x_k , το ζεύγος (p_k, d_k) υποτίθεται ότι είναι ικανό ζεύγος κατευθύνσεων μείωσης, αν οι

$\{p_k\}$ και $\{d_k\}$ είναι φραγμένες και ικανοποιούν τις ακόλουθες συνθήκες [42]:

Συνθήκη 2.1. $g_k^T p_k = 0$ έπεται ότι $g_k = 0$ και $p_k = 0$

Συνθήκη 2.2. $g_k^T p_k \rightarrow 0$ έπεται ότι $g_k \rightarrow 0$ και $p_k \rightarrow 0$

Συνθήκη 2.3. $d_k^T B_k d_k \rightarrow 0$ έπεται ότι $\min(\lambda_1, 0) \rightarrow 0$ και $d_k \rightarrow 0$, όπου λ_1 είναι η μικρότερη ιδιοτιμή του πίνακα B_k .

Οι Συνθήκες 2.1 και 2.2 είναι οι βασικές που πρέπει να ισχύουν για τις κατευθύνσεις τύπου Newton, οι οποίες πρέπει να είναι κατευθύνσης μείωσης. Εξασφαλίζουν ότι οι $p_k \neq 0$ και $g_k \neq 0$ δεν είναι ορθογώνιες και δεν γίνονται τέτοιες γρήγορα. Η Συνθήκη 2.3 εξασφαλίζει ότι η d_k περιέχει πληροφορία που σχετίζεται με την πιο μικρή ιδιοτιμή του εσσιανού ή του προσεγγιστικού εσσιανού πίνακα.

2.6.1.1 Η μέθοδος McCormick. Ο McCormick [61] πρότεινε τη χρήση της κατεύθυνσης αρνητικής καμπυλότητας μέσω της αναζήτησης πάνω στην καμπύλη

$$x(\alpha) = x_k + \phi_1(\alpha)p_k + \phi_2(\alpha)d_k, \quad (2.40)$$

όπου (p_k, d_k) είναι ένα ζεύγος κατευθύνσεων μείωσης στο x_k , ενώ τα $\phi_1(\alpha)$ και $\phi_2(\alpha)$ είναι πολυώνυμα, τέτοια ώστε $\phi_1(0) = \phi_2(0) = 0$. Αν θέσουμε $\Phi(\alpha) = f(x(\alpha))$, τότε για $\rho \in (0, 1)$ υπάρχει $\bar{\alpha} > 0$ τέτοιο ώστε να ισχύει η σχέση

$$\Phi(\alpha) \leq \Phi(0) + \rho \left[\Phi'(0)\alpha + \frac{1}{2}\Phi''(0)\alpha^2 \right],$$

για όλα τα $\alpha \in [0, \bar{\alpha}]$. Η παραπάνω σχέση προϋποθέτει ότι είτε $\Phi'(0) < 0$ ή $\Phi'(0) = 0$ και $\Phi''(0) < 0$. Η πιο απλή επιλογή που θα μπορούσε να γίνει αναφορικά με τα πολυώνυμα $\phi_1(\alpha)$ και $\phi_2(\alpha)$ στη σχέση (2.40), είναι να θέσουμε

$$\phi_1(\alpha) = \alpha^2 \quad \text{και} \quad \phi_2(\alpha) = \alpha,$$

που οδηγεί στην επανάληψη

$$x(\alpha) = x_k + \alpha^2 p_k + \alpha d_k. \quad (2.41)$$

Θέτοντας $\alpha = \gamma^i$, όπου $\gamma \in (0, 1)$, τότε η (2.41) γίνεται

$$x_k(i) = x_k + \gamma^{2i} p_k + \gamma^i d_k \in \mathcal{D}. \quad (2.42)$$

Η δεύτερης τάξης συνθήκη Armijo απαιτεί να βρούμε ένα δείκτη $i(k)$, ο οποίος να είναι ο μικρότερος μη αρνητικός ακέραιος i , τέτοιος ώστε να ισχύει η σχέση

$$f(x_k(i)) \leq f(x_k) + \rho \gamma^{2i} \left(g_k^T p_k + \frac{1}{2} d_k^T B_k d_k \right), \quad (2.43)$$

όπου $\rho \in (0, 1)$ και $x_k(i(k)) = x_{k+1}$. Ο McCormick διάλεξε να θέσει στη σχέση (2.42), $\gamma^2 = \frac{1}{2}$. Υπάρχει ένας πεπερασμένος δείκτης $i(k)$ που να ικανοποιεί την (2.43) προϋποθέτοντας ότι

$$p_k^T g_k < 0, \quad \text{όποτε} \quad g_k \neq 0 \quad (2.44)$$

και

$$d_k^T B_k d_k < 0, \quad \text{όποτε} \quad g_k = 0. \quad (2.45)$$

Μόνο όταν το x_k είναι δεύτερης τάξης κρίσιμο σημείο δεν υπάρχει ζεύγος κατευθύνσεων μείωσης που να ικανοποιεί τις σχέσεις (2.44)-(2.45), και ο αλγόριθμος τερματίζεται. Το ακόλουθο θεώρημα είναι το θεώρημα σύγκλισης της συνθήκης Armijo δεύτερης τάξης:

Θεώρημα 2.4. (McCormick,[61]) Έστω ότι η f ικανοποιεί την Υπόθεση 7.1. Έστω ότι οι $\{\|p_k\|\}$ και $\{\|d_k\|\}$ είναι φραγμένες. Αν η $\{x_k\}$ ικανοποιεί τις (2.42) και (2.43), τότε

$$\lim_{k \rightarrow \infty} g_k^T p_k = 0 \quad \text{και} \quad \lim_{k \rightarrow \infty} d_k^T B_k d_k = 0.$$

2.6.1.2 Η μέθοδος Goldfarb.

Ο Goldfarb [42] θεώρησε ότι η επανάληψη της μορφής

$$x_k(\alpha) = x_k + \alpha^2 p_k + \alpha d_k. \quad (2.46)$$

δεν είναι ιδανική.

Αλγόριθμος 2.3 Ο αλγόριθμος καμπυλόγραμμης αναζήτησης του Goldfarb

Βήμα 1. Δοθέντων γ, ρ , όπου $0 < \gamma, \rho < 1$ και αρχικού σημείου x_0 , όρισε το x_{k+1} , $k = 0, 1, 2, \dots$, ως ακολούθως:

Βήμα 2. Διάλεξε ένα ζεύγος κατευθύνσεων μείωσης (p_k, d_k) στο x_k . Αν δεν υπάρχει τέτοιο ζεύγος, σταμάτα. Αλλιώς, έστω $i(k) + 1$ ο μικρότερος μη αρνητικός ακέραιος τέτοιος ώστε

$$f(x_k(\gamma^i)) - f(x_k) \leq \rho \left(\gamma^i p_k^T g_k + \frac{1}{2} \gamma^{4i} d_k^T B_k d_k \right).$$

Βήμα 3. Θέσε $x_{k+1} = x_k(\gamma^{i(k)+1})$.

Υποστήριξε ότι ο τύπος (2.46) μπορεί να είναι καλός σε μια γειτονιά ενός σαγματικού σημείου, αλλά όχι και μακριά από αυτό. Ο Goldfarb έδωσε μια παρόμοια με την προηγούμενη δεύτερης τάξης συνθήκη Armijo, βασισμένη σε μια επανάληψη της μορφής

$$x_k(\alpha) = x_k + \alpha p_k + \alpha^2 d_k.$$

Θεώρημα 2.5. (Goldfarb, [42]) Έστω ότι η f ικανοποιεί την Υπόθεση 7.1 και έστω ότι μια κατάλληλη ακολουθία ζευγών μείωσης $\{p_k, d_k\}$ χρησιμοποιείται στον Αλγόριθμο 2.3 με

$$\begin{aligned} -p_k^T g_k &\geq c_1 \|p_k\|^2 \\ p_k^T B_k p_k &\leq c_2 \|p_k\|^2, \end{aligned}$$

όπου $0 < c_1, c_2 < \infty$. Τότε $g_k \rightarrow 0$, $p_k \rightarrow 0$, $\lambda_k \rightarrow 0$ και $d_k \rightarrow 0$.

2.6.1.3 Η μέθοδος των Moré & Sorensen.

Θεωρούμε την επανάληψη της μορφής

$$x(\alpha) = x_k + \alpha^2 p_k + \alpha d_k, \quad (2.47)$$

όπου οι (p_k, d_k) είναι ένα ζεύγος μείωσης στο x_k . Αντικαθιστώντας την (2.47) στις συνθήκες Wolfe (2.37)-(2.38) (σελ. 48), ζητάμε από το μήκος

βήματος α να ικανοποιεί τις συνθήκες

$$f(x(\alpha)) \leq f(x) + c_1 \alpha^2 \left[\nabla f(x)^T p + \frac{1}{2} d^T \nabla^2 f(x) d \right], \quad (2.48)$$

$$\nabla f(x(\alpha))^T x'(\alpha) \geq c_2 [\nabla f(x)^T d + 2\alpha \nabla f(x)^T p + \alpha d^T \nabla^2 f(x) d], \quad (2.49)$$

όπου $0 < c_1 \leq c_2 < 1$. Όταν $d = 0$, αυτές οι συνθήκες ανάγονται στις συνθήκες Wolfe (2.37)-(2.38) (σελ. 48). Οι συνθήκες (2.48)-(2.49) ονομάζονται συνθήκες Wolfe δεύτερης τάξης. Η γεωμετρική αναπαράσταση των δεύτερης τάξης συνθηκών Wolfe φαίνεται στο Σχήμα 2.1.

Σχήμα 2.1: Συνθήκες βήματος δεύτερης τάξης

Αν οι (p_k, d_k) είναι ένα ζεύγος μείωσης στο x_k και θέσουμε

$$\Phi_k(\alpha_k) = f(x_k + \alpha^2 p_k + \alpha d_k),$$

τότε οι (2.48)-(2.49) είναι ισοδύναμες με τις συνθήκες

$$\begin{aligned} \Phi_k(\alpha_k) &\leq \Phi_k(0) + \frac{1}{2} c_1 \Phi_k''(0) \alpha_k^2, \\ \Phi_k'(\alpha_k) &\geq c_2 [\Phi_k'(0) + \Phi_k''(0) \alpha_k]. \end{aligned}$$

Θεώρημα 2.6. (Moré & Sorensen, [63]) Έστω ότι η f ικανοποιεί την Υ-πόθεση 7.1. Αν η ακολουθία $\{x_k\}$ ικανοποιεί τις δεύτερης τάξης συνθήκες

Wolfe (2.48)-(2.49), τότε

$$\lim_{k \rightarrow \infty} g_k^T p_k = 0 \quad \text{και} \quad \lim_{k \rightarrow \infty} d_k^T B_k d_k = 0.$$

2.7 Οι μέθοδοι περιοχής εμπιστοσύνης

Οι μέθοδοι περιοχής εμπιστοσύνης είναι επαναληπτικές και παράγουν μια τοπική λύση του προβλήματος που μελετάμε. Ιστορικά θα μπορούσαμε να πούμε ότι έχουν τις ρίζες τους από τον Levenberg [54] το 1944 και αργότερα από τον Marquardt [60]. Από τις εργασίες των Levenberg και Marquardt εμπνεύστηκε ο Powell [73] για να φτιάξει τον πρώτο αλγόριθμο περιοχής εμπιστοσύνης.

Σε κάθε επανάληψη x_k ενός αλγορίθμου περιοχής εμπιστοσύνης, δημιουργείται ένα τετραγωνικό μοντέλο που προσεγγίζει την αντικειμενική συνάρτηση γύρω από μια γειτονιά με κέντρο το x_k . Η γειτονιά αυτή ονομάζεται *περιοχή εμπιστοσύνης* (*trust region*), διότι θεωρούμε ότι το μοντέλο σε αυτή τη γειτονιά θα είναι μια πολύ καλή προσέγγιση της f . Η περιοχή εμπιστοσύνης ορίζεται ως το σύνολο των σημείων τα οποία απέχουν από το x_k , το πολύ μια δοσμένη ακτίνα *περιοχής εμπιστοσύνης* (*trust region radius*). Στη συνέχεια υπολογίζεται ένα βήμα σε ένα δοκιμαστικό σημείο μέσα στην περιοχή εμπιστοσύνης, επιλύοντας το λεγόμενο *υποπρόβλημα περιοχής εμπιστοσύνης* (*trust region subproblem*), στο οποίο υπολογίζεται μια επαρκής μείωση του μοντέλου. Στο δοκιμαστικό αυτό σημείο, υπολογίζεται η τιμή της αντικειμενικής συνάρτησης και συγκρίνεται με τη μείωση της τιμής που προβλέφθηκε από το μοντέλο. Αν το πηλίκο της μείωσης που επιτεύχθηκε από την f προς την προβλεπόμενη μείωση είναι επαρκώς θετικό, τότε η τρέχουσα επανάληψη ορίζεται ως το νέο δοκιμαστικό σημείο και το μέγεθος της περιοχής εμπιστοσύνης, αυξάνεται. Σε αντίθετη περίπτωση, το μέγεθός της μειώνεται και ο αλγόριθμος ξαναπροσπαθεί να βρει νέο βήμα στην καινούργια περιοχή εμπιστοσύνης.

Η περιοχή εμπιστοσύνης είναι η περιοχή γύρω από την τρέχουσα επανάληψη

$$\mathcal{B}_k = \{x \in \mathbb{R}^n \mid \|x - x_k\| \leq \Delta_k\},$$

όπου Δ_k είναι η ακτίνα της περιοχής εμπιστοσύνης. Η επόμενη επανάληψη x_{k+1} υπολογίζεται βρίσκοντας το ελάχιστο του τετραγωνικού μοντέλου

που προσεγγίζει την αντικειμενική συνάρτηση,

$$m_k(x_k + d) = f_k + \nabla f_k^T d + \frac{1}{2} d^T B_k d,$$

όπου $f_k = f(x_k)$, $\nabla f_k = \nabla f(x_k)$ και B_k είναι ένας συμμετρικός πίνακας, θετικά ορισμένος ή μη. Επομένως, ψάχνουμε μια λύση του υποπροβλήματος

$$\min_{d \in \mathbb{R}^n} m_k(x_k + d) \quad \text{υπό τον περιορισμό} \quad \|d\| \leq \Delta_k, \quad (2.50)$$

γνωστό ως υποπρόβλημα περιοχής εμπιστοσύνης (Trust Region Subproblem).

Η επιλογή της ακτίνας Δ_k βασίζεται στη σχέση ανάμεσα στο τετραγωνικό μοντέλο και την αντικειμενική συνάρτηση σε προηγούμενες επαναλήψεις. Δοθέντος ενός βήματος d_k , έστω ότι το x_k είναι η τρέχουσα επανάληψη και το $x_k + d_k$ η λύση του υποπροβλήματος (2.50). Τότε η διαφορά

$$Ared_k = f(x_k) - f(x_k + d_k)$$

είναι η *πραγματική μείωση* της αντικειμενικής συνάρτησης, ενώ η διαφορά

$$Pred_k = m_k(0) - m_k(d_k)$$

είναι η *προβλεπόμενη μείωση* της αντικειμενικής συνάρτησης. Προκειμένου να εκτιμήσουμε πόσο καλά προσεγγίζει το μοντέλο την αντικειμενική συνάρτηση, ορίζουμε το *πηλίκο*

$$\rho_k = \frac{Ared_k}{Pred_k}. \quad (2.51)$$

Επειδή το βήμα d_k υπολογίζεται από την ελαχιστοποίηση του μοντέλου m_k πάνω σε μια περιοχή που περιλαμβάνει και το βήμα $d = 0$, η προβλεπόμενη μείωση είναι πάντα μη αρνητικός αριθμός. Έτσι, αν το πηλίκο έχει τιμή κοντά στο 1, τότε το μοντέλο προσεγγίζει αρκετά καλά την αντικειμενική συνάρτηση. Σε αυτή την περίπτωση η περιοχή εμπιστοσύνης διευρύνεται, δηλαδή εμπιστευόμαστε το μοντέλο και του παρέχουμε τη δυνατότητα να παράγει ένα μεγαλύτερο βήμα. Αν το ρ_k είναι αρνητικό ή πολύ κοντά στο μηδέν, σημαίνει ότι η καινούργια τιμή $f(x_k + d_k)$ της αντικειμενι-

κής συνάρτησης είναι μεγαλύτερη από την προηγούμενη τιμή $f(x_k)$, άρα το βήμα d_k πρέπει να απορριφθεί. Σε αυτή την περίπτωση, η περιοχή εμπιστοσύνης μειώνεται και το υποπρόβλημα (2.50) επιλύεται ξανά. Σε οποιαδήποτε άλλη περίπτωση, η περιοχή εμπιστοσύνης δεν μεταβάλλεται.

Ο παρακάτω αλγόριθμος είναι γενικός αλγόριθμος που συνοψίζει τα βασικά βήματα που ακολουθούμε στη μέθοδο περιοχής εμπιστοσύνης.

Αλγόριθμος 2.4 Γενικός Αλγόριθμος περιοχής εμπιστοσύνης

Βήμα 1. Δίνονται ένα αρχικό σημείο x_0 , ένα άνω φράγμα $\bar{\Delta}$ για την ακτίνα της περιοχής εμπιστοσύνης, μια αρχική ακτίνα $\Delta_0 \in (0, \bar{\Delta})$ και οι σταθερές $\epsilon \geq 0$, $0 < \eta_1 \leq \eta_2 < 1$ και $0 < \gamma_1 < 1 < \gamma_2$. Θέσε $k := 0$.

Βήμα 2. Αν $\|g\| \leq \epsilon$, σταμάτα.

Βήμα 3. Λύσε προσεγγιστικά ως προς d_k , το υποπρόβλημα περιοχών εμπιστοσύνης

$$\begin{aligned} \min_{d \in \mathbb{R}^n} m_k(x_k + d) &= f_k + \nabla f_k^T d + \frac{1}{2} d^T B_k d, \\ \text{υπό τον περιορισμό} &= \|d\| \leq \Delta_k. \end{aligned}$$

Βήμα 4. Υπολόγισε την $f(x_k + d_k)$ και το πηλίκο ρ_k . Θέσε:

$$x_{k+1} = \begin{cases} x_k + d_k, & \text{αν } \rho_k \geq \eta_1; \\ x_k, & \text{διαφορετικά.} \end{cases}$$

Βήμα 5. Θέσε:

$$\Delta_{k+1} \in \begin{cases} (0, \gamma_1 \Delta_k], & \text{αν } \rho_k < \eta_1; \\ [\gamma_1 \Delta_k, \Delta_k], & \text{αν } \rho_k \in [\eta_1, \eta_2); \\ [\Delta_k, \min \gamma_2 \Delta_k, \bar{\Delta}], & \text{αν } \rho_k \geq \eta_2 \text{ και } \|d_k\| = \Delta_k. \end{cases}$$

Αύξησε το k κατά 1 και πήγαινε στο βήμα 2.

Στον Αλγόριθμο 2.4, το $\bar{\Delta}$ είναι ένα άνω φράγμα για όλα τα Δ_k . Οι επαναλήψεις για τις οποίες ισχύει $\rho_k \geq \eta_2$ και άρα $\Delta_{k+1} \geq \Delta_k$, ονομάζονται πολύ επιτυχημένες επαναλήψεις. Οι επαναλήψεις για τις οποίες ισχύει $\rho_k \geq \eta_1$ και άρα $x_{k+1} = x_k + d_k$ ονομάζονται επιτυχημένες. Τέλος, οι επαναλήψεις στις οποίες ισχύει $\rho_k < \eta_1$ και άρα $x_{k+1} = x_k$, λέγονται αποτυχημένες επαναλήψεις.

Υπόθεση 2.1. Θεωρούμε ότι η νόρμα των προσεγγιστικών εσσιανών πινά-

των B_k είναι ομοιόμορφα φραγμένη, δηλαδή υπάρχει θετική σταθερά M τέτοια ώστε $\|B_k\| \leq M$ για κάθε k .

Υπόθεση 2.2. Θεωρούμε ότι η αντικειμενική συνάρτηση f είναι συνεχώς διαφορίσιμη και φραγμένη στο σύνολο $\{x | f(x) \leq f(x_0)\}$.

Υπόθεση 2.3. Θεωρούμε ότι η λύση d_k του υποπρόβληματος (2.50) ικανοποιεί την ανίσωση $\|d_k\| \leq \gamma \Delta_k$, για κάποια σταθερά $\gamma \geq 1$.

Θεώρημα 2.7. Έστω ότι ισχύουν οι υποθέσεις 2.1–2.3. Αν ο Αλγόριθμος 2.4 έχει πεπερασμένες επιτυχημένες επαναλήψεις, τότε συγκλίνει σε πρώτης τάξης στάσιμο σημείο.

Θεώρημα 2.8. Έστω ότι ισχύουν οι υποθέσεις 2.1–2.3. Αν ο Αλγόριθμος 2.4 έχει άπειρες επιτυχημένες επαναλήψεις, τότε ισχύει ότι

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0.$$

Θεώρημα 2.9. Έστω ότι ισχύουν οι υποθέσεις 2.1–2.3. Τότε, $\lim_{k \rightarrow \infty} g_k = 0$.

2.8 Το Υποπρόβλημα Περιοχής Εμπιστοσύνης

Το υποπρόβλημα περιοχών εμπιστοσύνης (ΥΠΕ)

$$\min_{d \in \mathbb{R}^n} \phi(d) = \nabla f^T d + \frac{1}{2} d^T B d \quad \text{υπό τον περιορισμό} \quad \|d\| \leq \Delta, \quad (2.52)$$

όπου B ένας $n \times n$ πραγματικός, συμμετρικός και πιθανόν μη θετικά ορισμένος πίνακας, $g \in \mathbb{R}^n$, Δ θετικό βαθμωτό μέγεθος και d το πραγματικό n -διάστατο άγνωστο διάνυσμα, είναι γνωστό στη διεθνή βιβλιογραφία ως TRS (Trust-Region Subproblem). Αποτελεί από μόνο του ένα πρόβλημα βελτιστοποίησης με περιορισμούς και δεν εμφανίζεται μόνο στις μεθόδους περιοχής εμπιστοσύνης, αλλά συναντάται ευρέως και σε άλλες εφαρμογές, όπως για παράδειγμα σε προβλήματα βελτιστοποίησης με περιορισμούς [25], διαχωρισμού γραφημάτων [49], αποκατάστασης εικόνας [78], μοριακής μοντελοποίησης [18] κ.ο.κ.

Σε έναν αλγόριθμο περιοχής εμπιστοσύνης, το ΥΠΕ (2.52) αποτελεί το σημαντικότερο και δυσκολότερο βήμα του. Στην επανάληψη x_k υπολογίζεται ένα δοκιμαστικό βήμα d_k μέσω της επίλυσης του ΥΠΕ (2.52), όπου η

$\phi_k(d)$ είναι μια τετραγωνική προσέγγιση της αντικειμενικής συνάρτησης f , $g_k = \nabla f(x_k)$, $B_k \in \mathbb{R}^{n \times n}$ είναι είτε ο εσσιανός πίνακας ή μια προσέγγισή του στο x_k , ενώ $\Delta_k > 0$ είναι η ακτίνα περιοχής εμπιστοσύνης.

Το 1966 οι Goldfeld, Quandt και Trotter [43] έλυσαν υποπροβλήματα όπου η αντικειμενική συνάρτηση ήταν μη κυρτή. Το μειονέκτημα των παραπάνω μεθόδων ήταν ότι δεν μπορούσαν να χειριστούν την περίφημη *δύσκολη* περίπτωση (hard case), όπως ονομάζεται στη βιβλιογραφία. Το 1981 ο Gay [39] έδειξε πώς να αναγνωρίζει τη δύσκολη περίπτωση αριθμητικά, ενώ το 1983, οι Moré & Sorensen δημοσίευσαν το πλέον κλασικό άρθρο, Computing a trust region step, για το υποπρόβλημα περιοχής εμπιστοσύνης [64]. Σε αυτό το άρθρο ανέπτυξαν ένα τρόπο για να επιλύουν αποτελεσματικά το πρόβλημα στη λεγόμενη τυπική περίπτωση, γνωστή διεθνώς ως standard case. Αυτό το πέτυχαν εφαρμόζοντας τη μέθοδο Newton σε μια σχεδόν γραμμική συνάρτηση. Στη δύσκολη περίπτωση θεμελίωσαν μία πρόταση η οποία επέτρεπε την αποτελεσματική επίλυση του υποπροβλήματος στην περίπτωση αυτή.

Το επόμενο θεώρημα αποτελεί τις Karush-Kuhn-Tucker πρώτης τάξης αναγκαίες συνθήκες βελτιστότητας για το ΥΠΕ (2.52) και αποτελεί τη βάση για την εύρεση μιας σχεδόν ακριβούς λύσης του.

Θεώρημα 2.10. [64] Το διάνυσμα d^* είναι η λύση του υποπροβλήματος περιοχής εμπιστοσύνης (ΥΠΕ)

$$\min_{d \in \mathbb{R}^n} \phi(d) = \nabla f^T d + \frac{1}{2} d^T B d \quad \text{υπό τον περιορισμό} \quad \|d\| \leq \Delta,$$

αν και μόνο αν υπάρχει πραγματικός αριθμός $\lambda^* \geq 0$ τέτοιος ώστε οι ακόλουθες συνθήκες να ικανοποιούνται:

$$(B + \lambda^* I) d^* = -g, \quad (2.53)$$

$$\|d^*\| \leq \Delta, \quad (2.54)$$

$$\lambda (\Delta - \|d^*\|) = 0, \quad \text{και} \quad (2.55)$$

$$(B + \lambda^* I) \quad \text{θετικά ημιορισμένος}. \quad (2.56)$$

Επιπλέον, αν ο $(B + \lambda^* I)$ είναι θετικά ορισμένος, τότε η λύση d^* είναι μοναδική.

Από το Θεώρημα 2.10 μπορούμε να χαρακτηρίσουμε τις λύσεις του ΥΠΕ

διαχωρίζοντας τις ακόλουθες περιπτώσεις:

1. Αν $\lambda_1 > 0$ (ο πίνακας B είναι θετικά ορισμένος), τότε

α. αν $\|B^{-1}g\| \leq \Delta$, τότε $\lambda^* = 0$ και $d^* = B^{-1}g$.

β. αν $\|B^{-1}g\| > \Delta$, τότε για το μοναδικό $\lambda^* \in (0, \infty)$ τέτοιο ώστε να ισχύει $\|(B + \lambda^*I)^{-1}g\| = \Delta$, η λύση είναι $d^* = -(B + \lambda^*I)^{-1}g$.

2. Αν $\lambda_1 \leq 0$ (ο πίνακας B δεν είναι θετικά ορισμένος) τότε

α. αν υπάρχει $\lambda^* \in (-\lambda_1, \infty)$ τέτοιο ώστε $\|(B + \lambda^*I)^{-1}g\| = \Delta$, τότε η λύση είναι $d^* = -(B + \lambda^*I)^{-1}g$.

β. αλλιώς, $\lambda^* = -\lambda_1$ και $d^* = -(B - \lambda_1 I)^{-\dagger}g + \tau u_1$, όπου ο πραγματικός αριθμός $\tau \in \mathbb{R}$ είναι τέτοιος ώστε $\|-(B - \lambda_1 I)^{-\dagger}g + \tau u_1\| = \Delta$ και u_1 είναι ένα ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 , τέτοιο ώστε $\|u_1\| = 1$.

Οι περιπτώσεις 1α, 1β και 2α είναι γνωστές στη βιβλιογραφία ως η *τυπική περίπτωση* (standard case), ενώ η 2β είναι η περίφημη *δύσκολη περίπτωση* (hard case). Η τελευταία εμφανίζεται όταν ο πίνακας B είναι μη θετικά ορισμένος και το διάνυσμα g είναι κάθετο σε όλα τα ιδιοδιανύσματα που αντιστοιχούν στην πιο αρνητική ιδιοτιμή του πίνακα. Η περίπτωση αυτή παρουσιάζει αριθμητικές δυσκολίες λόγω του ότι ο πίνακας $B - \lambda_1 I$ είναι ιδιάζον. Επιπρόσθετα, χρειάζεται να υπολογιστεί ένα ιδιοδιάνυσμα που αντιστοιχεί στη μικρότερη ιδιοτιμή του B . Συνοψίζοντας, η δύσκολη περίπτωση έχει μελετηθεί διεξοδικά από τον Gay [39] και έχει τεκμηριωθεί επαρκώς στην εργασία των Moré & Sorensen [64].

Το Θεώρημα 2.10 μας υποδεικνύει ότι είναι απαραίτητο να βρούμε τη λύση του γραμμικού συστήματος $(B + \lambda I)d = -g$ και να υπολογίσουμε την απαιτούμενη τιμή του λ . Αφού ο B είναι συμμετρικός, υπάρχει ορθογώνιος πίνακας U και διαγώνιος πίνακας Λ , τέτοιοι ώστε να ισχύει

$$B = U^T \Lambda U,$$

όπου $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, με $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, είναι οι ιδιοτιμές του B . Προφανώς,

$$B + \lambda I = U^T (\Lambda + \lambda I) U.$$

Η τιμή του λ που ψάχνουμε πρέπει να είναι τέτοια ώστε $\lambda \geq -\lambda_1$, καθώς σε αυτή την περίπτωση ο πίνακας $B + \lambda I$ είναι θετικά ημιορισμένος, ενώ αν $\lambda > -\lambda_1$ τότε ο $B + \lambda I$ είναι θετικά ορισμένος, με αποτέλεσμα ο ελαχιστοποιητής του μοντέλου να είναι μοναδικός.

Σχήμα 2.2: Λύση του ΥΠΕ για διαφορετικές ακτίνες $\Delta_1, \Delta_2, \Delta_3$

2.8.1 Η τυπική περίπτωση

Ως τυπική περίπτωση (standard case), χαρακτηρίζεται η περίπτωση όπου για τον πολλαπλασιαστή Lagrange ισχύει $\lambda \in (-\lambda_1, \infty)$. Έστω λοιπόν ότι $\lambda > -\lambda_1$. Τότε ο $B + \lambda I$ είναι θετικά ορισμένος και το γραμμικό σύστημα $(B + \lambda I)d = -g$ έχει μοναδική λύση, την

$$d(\lambda) = -(B + \lambda I)^{-1}g = -U^T(\Lambda + \lambda I)Ug. \quad (2.57)$$

Αν η τιμή $\lambda = 0$ ικανοποιεί τις σχέσεις (2.53) και (2.56) με $\|d(0)\| \leq \Delta$, τότε η λύση του ΥΠΕ είναι η $d(0) = -B^{-1}g$, δηλαδή η λύση βρίσκεται στο εσωτερικό της περιοχής εμπιστοσύνης. Αλλιώς, ψάχνουμε μια τιμή $\lambda > 0$ στο σύνορο της περιοχής εμπιστοσύνης με ακτίνα Δ . Σε διαφορετική περίπτωση ($\lambda \neq 0$), αναζητούμε μία τιμή $\lambda > 0$ τέτοια ώστε να ικανοποιεί

την ισότητα

$$\|d(\lambda)\| = \Delta. \quad (2.58)$$

Με άλλα λόγια, η λύση να βρίσκεται στο σύνορο της περιοχής εμπιστοσύνης. Στην περίπτωση αυτή, το πρόβλημά μας ανάγεται στην επίλυση μίας μονοδιάστατης μη γραμμικής εξίσωσης (2.58), ως προς λ . Για $\lambda \neq \lambda_j$ από την (2.57) έχουμε

$$d(\lambda) = - \sum_{j=1}^n \frac{\gamma_j}{\lambda_j + \lambda} u_j, \quad (2.59)$$

όπου το γ_j συμβολίζει την j συνιστώσα του Ug . Επομένως, από την ορθοκανονικότητα του πίνακα έχουμε ότι

$$\|d(\lambda)\|^2 = \sum_{j=1}^n \frac{(\gamma_j)^2}{(\lambda_j + \lambda)^2}, \quad (2.60)$$

Επειδή $\lambda > -\lambda_1$, έχουμε ότι $\lambda_j + \lambda > 0$ για όλα τα $j = 1, 2, \dots, n$, άρα η $\|d(\lambda)\|$ είναι συνεχής και μη αύξουσα συνάρτηση του λ στο διάστημα $(-\lambda_1, \infty)$. Μάλιστα, έχουμε ότι

$$\lim_{\lambda \rightarrow \infty} \|d(\lambda)\| = 0. \quad (2.61)$$

Επιπρόσθετα, όταν $\gamma_j \neq 0$, ισχύει

$$\lim_{\lambda \rightarrow -\lambda_j} \|d(\lambda)\| = \infty. \quad (2.62)$$

Αυτά τα χαρακτηριστικά φαίνονται στο Σχήμα 2.3. Προφανώς η τιμή του Δ τέμνει το γράφημα της $\|d(\lambda)\|$ σε ακριβώς ένα σημείο, το βέλτιστο λ^* . Όταν ο B είναι θετικά ορισμένος και $\|B^{-1}g\| \leq \Delta$, η τιμή $\lambda = 0$ ικανοποιεί τις σχέσεις (2.53)-(2.56), άρα δεν χρειάζεται να ψάξουμε για άλλο λ . Αν ο B είναι θετικά ορισμένος και $\|B^{-1}g\| > \Delta$, υπάρχει μία ακριβώς τιμή του λ τέτοια ώστε $\|d(\lambda)\| = \Delta$, άρα ψάχνουμε για μια λύση της (2.58) στο διάστημα $(0, \infty)$.

Σχήμα 2.3: $\|d(\lambda)\|$ ως συνάρτηση του λ

Παράδειγμα 2.1. Έστω ότι το πρόβλημα ορίζεται ως

$$g = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

όπου ο πίνακας B είναι θετικά ορισμένος. Το γράφημα της $\|d(\lambda)\|^2$ φαίνεται στο Σχήμα 2.4. Η συνάρτηση είναι πάντα θετική, τείνει στο μηδέν καθώς το λ τείνει στο άπειρο και έχει πόλους στις αντίθετες ιδιοτιμές του πίνακα, αρκεί το $(\gamma_j)^2 \neq 0, j = 1, 2, 3, 4$.

Σύμφωνα με το Θεώρημα 2.10 ενδιαφερόμαστε μόνο για $\lambda > 0$. Αν το Δ είναι αρκετά μεγάλο, και συγκεκριμένα αν $\Delta^2 > 1.5$ τότε παρατηρούμε ότι $\|d(\lambda)\|^2 < \Delta^2$ για όλα τα $\lambda \geq 0$. Αν όμως μειώνουμε συνεχώς το Δ , τότε για $\Delta^2 < 1.5$ πρέπει να αυξήσουμε το λ για να διατηρήσουμε την ισότητα $\|d(\lambda)\| = \Delta$. Παρατηρούμε ότι για κάθε $\Delta^2 < 1.5$ υπάρχει μία και μόνο τιμή του λ έτσι ώστε να διατηρείται αυτή ισότητα. Γεωμετρικά, καθώς ο B είναι θετικά ορισμένος, το μοντέλο της συνάρτησης έχει μοναδικό ελάχιστο χωρίς περιορισμούς. Έτσι, για $\Delta^2 < 1.5$ υπάρχει μοναδικό $\lambda = \lambda^*$ τέτοιο ώστε $\|d(\lambda)\| - \Delta = 0$, το οποίο αν χρησιμοποιηθεί στην επίλυση του γραμμικού συστήματος $d(\lambda) = -(B + \lambda I)^{-1} g$ δίνει το βέλτιστο βήμα $d(\lambda^*)$.

Σχήμα 2.4: Γράφημα της $\|d(\lambda)\|^2$ για θετικά ορισμένο πίνακα

Παράδειγμα 2.2. Έστω ότι το πρόβλημα ορίζεται ως

$$g = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

όπου ο B είναι αόριστος. Το γράφημα της $\|d(\lambda)\|^2$ φαίνεται στο Σχήμα 2.5. Παρατηρούμε πάλι ότι η συνάρτηση είναι παντού θετική, ότι τείνει στο μηδέν καθώς το λ τείνει στο άπειρο και ότι έχει πόλους στις αντίθετες ιδιοτιμές του πίνακα, αρκεί το $(\gamma_1)^2 \neq 0$. Άρα η $\|d(\lambda)\|^2$ παίρνει όποια θετική τιμή θέλουμε για το μοναδικό λ ανάμεσα στο μεγαλύτερο πόλο της και το άπειρο. Στο συγκεκριμένο παράδειγμα ο μεγαλύτερος πόλος είναι το 2, που είναι η αντίθετη τιμή της πιο μικρής ιδιοτιμής του πίνακα, της $\lambda_1 = -2$. Επομένως ο πίνακας $B + \lambda I$ είναι θετικά ορισμένος για όλα τα $\lambda > 2$. Καθώς και το Δ είναι θετικό, συμπεραίνουμε ότι υπάρχει μία και μόνο τιμή του $\lambda > 2$ για την οποία ισχύει $\|d(\lambda)\| - \Delta = 0$. Γεωμετρικά, ο B είναι μη θετικά ορισμένος, άρα το μοντέλο της συνάρτησης δεν έχει κάτω φράγμα. Επομένως, η λύση του υποπροβλήματος θα πρέπει να βρίσκεται στο σύνορο του περιορισμού. Ο ελαχιστοποιητής του μοντέλου βρίσκεται από την επίλυση του γραμμικού συστήματος $d(\lambda) = -(B + \lambda I)^{-1} g$ για κάποια θετική τιμή

του λ , η οποία βρίσκεται από την επίλυση της εξίσωσης $\|d(\lambda)\| - \Delta = 0$.

Σχήμα 2.5: Γράφημα της $\|d(\lambda)\|^2$ για μη θετικά ορισμένο πίνακα

2.8.1.1 Η μέθοδος Newton στην τυπική περίπτωση. Όπως προαναφέραμε, η επίλυση του ΥΠΕ (2.52) στην τυπική περίπτωση, ανάγεται στην επίλυση της μονοδιάστατης μη γραμμικής εξίσωσης

$$\|d(\lambda)\| - \Delta = 0,$$

ως προς λ , και στη συνέχεια, για αυτή τη συγκεκριμένη τιμή του λ , στην επίλυση ενός γραμμικού συστήματος

$$d(\lambda) = -(B + \lambda I)^{-1} g$$

για την εύρεση του βήματος $d(\lambda)$. Επομένως χρειάζεται να βρούμε τη ρίζα της εξίσωσης $\|d(\lambda)\| - \Delta = 0$. Για αυτό το λόγο χρησιμοποιούμε τη μέθοδο Newton. Ως γνωστόν, η μέθοδος Newton βασίζεται σε μια κατά Taylor προσέγγιση της $\|d(\lambda)\|$ και η επιτυχία της στηρίζεται στο να έχουμε καλά συμπεριφερόμενες παραγώγους στην περιοχή αναζήτησης. Καθώς όμως υπάρχει πόλος στο $-\lambda_1$, είναι πολύ πιθανό να παρατηρηθούν μεγάλες διακυμάνσεις των τιμών της παραγώγου στην περιοχή αναζήτησης, η

οποία είναι και η περιοχή που περιμένουμε να βρούμε τη λύση της εξίσωσης. Επομένως, η μέθοδος Newton ενδέχεται να αποβεί αναξιόπιστη ή μη αποτελεσματική. Γι αυτό το λόγο, οι Moré & Sorensen θεώρησαν τη συνάρτηση

$$\psi(\lambda) = \frac{1}{\|d(\lambda)\|} - \frac{1}{\Delta}, \quad (2.63)$$

για να εφαρμόσουν τη μέθοδο Newton. Η εξίσωση

$$\psi(\lambda) = 0 \quad (2.64)$$

είναι γνωστή στη διεθνή βιβλιογραφία ως *secular equation*. Εύκολα αποδεικνύεται ότι η εξίσωση $\psi(\lambda) = 0$ έχει τις ίδιες λύσεις με την $\|d(\lambda)\| - \Delta = 0$, ενώ η συνάρτηση $\psi(\lambda)$ έχει το πλεονέκτημα να είναι σχεδόν γραμμική στην περιοχή αναζήτησης [76]. Αυτό έχει ως αποτέλεσμα η μέθοδος Newton να συμπεριφέρεται εν γένει καλύτερα όταν εφαρμόζεται στην $\psi(\lambda) = 0$ από ότι στην $\|d(\lambda)\| - \Delta = 0$. Γενικά, αν υποθέσουμε ότι $g \neq 0$, η $\psi(\lambda)$ για $\lambda > -\lambda_1$, είναι γνησίως αύξουσα και κυρτή.

Σχήμα 2.6: Η συνάρτηση $1/\|d(\lambda)\|$

Η μέθοδος Newton για την εύρεση ρίζας της εξίσωσης $\psi(\lambda) = 0$ αντι-

καθιστά την εκτίμηση $\lambda_k > -\lambda_1$ με τη βελτιωμένη εκτίμηση

$$\lambda_{k+1} = \lambda_k - \frac{\psi(\lambda_k)}{\psi'(\lambda_k)}, \quad k = 0, 1, 2, \dots \quad (2.65)$$

Για να εφαρμόσουμε τη μέθοδο Newton χρειαζόμαστε τη συναρτησιακή τιμή της $\psi(\lambda_k)$ καθώς και την πρώτη παράγωγό της, την $\psi'(\lambda_k)$.

Αλγόριθμος 2.5 Η μέθοδος Newton στην τυπική περίπτωση

Βήμα 1. Δοθέντων $\lambda > -\lambda_1$ και $\Delta > 0$, για $k = 0, 1, 2, \dots$:

Βήμα 2. Παραγοντοποίησε κατά Cholesky τον πίνακα $B + \lambda I = R^T R$.

Βήμα 3. Λύσε ως προς $d(\lambda_k)$ το σύστημα $R^T R d(\lambda_k) = -g$.

Βήμα 4. Λύσε ως προς $q(\lambda_k)$ το σύστημα $R^T q(\lambda_k) = d(\lambda_k)$.

Βήμα 5. Θέσε $\lambda_{k+1} = \lambda_k + \left(\frac{\|d(\lambda_k)\| - \Delta}{\Delta} \right) \left(\frac{\|d(\lambda_k)\|}{\|q(\lambda_k)\|} \right)^2$.

Η συναρτησιακή τιμή μπορεί να υπολογιστεί από την επίλυση του γραμμικού συστήματος $d(\lambda) = -(B + \lambda I)^{-1} g$, ενώ η $\psi'(\lambda_k)$ μέσω της παραγώγου της $d(\lambda)$. Και οι δύο ποσότητες μπορούν να υπολογιστούν μέσω της επίλυσης γραμμικών συστημάτων που περιλαμβάνουν τον $B + \lambda I$. Στην περιοχή αναζήτησης αυτός ο πίνακας είναι συμμετρικός και θετικά ορισμένος, άρα μπορούμε να χρησιμοποιήσουμε την κατά Cholesky παραγοντοποίηση για να λύσουμε τα συστήματα.

Ας εξετάσουμε λίγο πιο αναλυτικά το επαναληπτικό σχήμα της Newton (5.4). Έχουμε:

$$\begin{aligned} (B + \lambda_k I) &= U (\Lambda + \lambda_k I) U^T = R^T R \\ (R^T)^{-1} U &= RU (\Lambda + \lambda_k I)^{-1}. \end{aligned} \quad (2.66)$$

Επίσης,

$$\begin{aligned} U^T R^T RU &= \Lambda + \lambda_k I \\ \left[RU (\Lambda + \lambda_k I)^{-1/2} \right]^T \left[RU (\Lambda + \lambda_k I)^{-1/2} \right] &= I. \end{aligned} \quad (2.67)$$

Άρα, ο $RU (\Lambda + \lambda_k I)^{-1/2}$ είναι ορθοκανονικός πίνακας. Έχουμε:

$$\begin{aligned}
 q(\lambda_k) &= (R^T)^{-1} d(\lambda_k) \\
 &= - (R^T)^{-1} (R^T R)^{-1} g \\
 &= - (R^T)^{-1} U (\Lambda + \lambda_k I)^{-1} U^T g \\
 &= - (R^T)^{-1} U (\Lambda + \lambda_k I)^{-1} \gamma \\
 &= RU (\Lambda + \lambda_k I)^{-2} \gamma,
 \end{aligned}$$

όπου η τελευταία ισότητα προκύπτει μέσω της (2.66). Επομένως, χρησιμοποιώντας την (2.67) βρίσκουμε ότι

$$\begin{aligned}
 \|q(\lambda_k)\|^2 &= \|RU (\Lambda + \lambda_k I)^{-1/2} (\Lambda + \lambda_k I)^{-3/2} \gamma\|^2 \\
 &= \|(\Lambda + \lambda_k I)^{-3/2} \gamma\|^2 \\
 &= \sum_{j=1}^n \frac{\gamma_j^2}{(\lambda_j + \lambda_k)^3}.
 \end{aligned}$$

Λαμβάνοντας υπόψη ότι $\|d(\lambda)\|^2 = \sum_{j=1}^n \frac{(\gamma_j)^2}{(\lambda_j + \lambda)^2}$, παραγωγίζοντας ως προς λ βρίσκουμε:

$$\begin{aligned}
 \frac{d}{d\lambda} \|d(\lambda)\|^2 \Big|_{\lambda=\lambda_k} &= 2 \sum_{j=1}^n \frac{\gamma_j^2}{(\lambda_j + \lambda_k)^3} \\
 &= 2 \|q(\lambda_k)\|^2.
 \end{aligned}$$

Από τις παραπάνω ισότητες έχουμε

$$\begin{aligned}
 \frac{d}{d\lambda} \|d(\lambda)\| \Big|_{\lambda=\lambda_k} &= \frac{d}{d\lambda} (\|d(\lambda)\|^2)^{1/2} \Big|_{\lambda=\lambda_k} \\
 &= \frac{\frac{d}{d\lambda} \|d(\lambda)\|^2 \Big|_{\lambda=\lambda_k}}{2 \|d(\lambda_k)\|} \\
 &= \frac{\|q(\lambda_k)\|^2}{\|d(\lambda_k)\|}.
 \end{aligned}$$

Επομένως,

$$\begin{aligned}
\psi'(\lambda_k) &= \frac{d}{d\lambda} \psi(\lambda) \Big|_{\lambda=\lambda_k} \\
&= \frac{d}{d\lambda} \left(\frac{1}{d(\lambda)} - \frac{1}{\|d(\lambda)\|} \right) \Big|_{\lambda=\lambda_k} \\
&= \frac{\frac{d}{d\lambda} \|d(\lambda)\| \Big|_{\lambda=\lambda_k}}{\|d(\lambda_k)\|^2} \\
&= \frac{\|q(\lambda_k)\|^2}{\|d(\lambda_k)\|^3},
\end{aligned}$$

και το λ_{k+1} μπορεί να εκφραστεί ως

$$\begin{aligned}
\lambda_k - \frac{\psi(\lambda_k)}{\psi'(\lambda_k)} &= \lambda_k - \left(\frac{1}{\|d(\lambda_k)\|} - \frac{1}{\Delta} \right) \frac{\|d(\lambda_k)\|^3}{\|q(\lambda_k)\|^2} \\
&= \lambda_k + \left(\frac{\|d(\lambda_k)\| - \Delta}{\Delta} \right) \left(\frac{\|d(\lambda_k)\|}{\|q(\lambda_k)\|} \right)^2,
\end{aligned}$$

όπως δίνεται στον Αλγόριθμο 2.5.

Αν βρεθεί λ_0 τέτοιο ώστε $\lambda_0 > -\lambda_1$ και $\psi(\lambda_0) > 0$, τότε ο Αλγόριθμος 2.5 συγκλίνει τετραγωνικά, αφού η $\psi(\lambda)$ είναι κυρτή και γνησίως αύξουσα στο $(-\lambda_1, \infty)$. Επομένως η $\psi(\lambda) = 0$ έχει μοναδική ρίζα και η μέθοδος Newton εγγυάται τετραγωνική σύγκλιση όταν αρχικοποιηθεί με ένα κατάλληλο λ εντός του διαστήματος όπου η συνάρτηση είναι αύξουσα και κυρτή και ικανοποιεί τη σχέση $\psi(\lambda) > 0$.

Ο αλγόριθμος των Moré & Sorensen αξιοποιεί μία διαδικασία διασφάλισης της μεθόδου Newton έτσι ώστε $\lambda > -\lambda_1$ και $\lambda \in [\lambda_L, \lambda_U]$, όπου το $[\lambda_L, \lambda_U]$ είναι ένα διάστημα εμπιστοσύνης του λ . Για να ανήκει το λ στο διάστημα εμπιστοσύνης $[\lambda_L, \lambda_U]$, η πιο απλή επιλογή είναι να διαλέξουμε το μέσο $\lambda = \frac{1}{2}(\lambda_L + \lambda_U)$. Μια καλύτερη λύση όμως (κατόπιν πειραματικής έρευνας) είναι να θέσουμε

$$\lambda = \max \left(\sqrt{\lambda_L \lambda_U}, \lambda_L + 0.01(\lambda_U - \lambda_L) \right)$$

Αρχικές τιμές των λ , λ_L και λ_U , μπορούν να παρέχουν οι ανισότητες της ποσότητας Rayleigh,

$$(\lambda + \lambda_1)^2 \leq \frac{\|(B + \lambda I) d(\lambda)\|^2}{\|d(\lambda)\|^2} = \left(\frac{\|g\|}{\|d(\lambda)\|} \right)^2 \leq (\lambda + \lambda_n)^2, \quad (2.68)$$

από τις οποίες έχουμε ότι

$$\frac{\|g\|}{\Delta} - \lambda_n \leq \lambda \leq \frac{\|g\|}{\Delta} - \lambda_1.$$

Οι Moré & Sorensen έθεσαν

$$\lambda_L := \max \left[0, \max_i (-b_{i,i}), \frac{\|g\|}{\Delta} - \|B\|_1 \right] \quad \text{και} \quad \lambda_U := \frac{\|g\|}{\Delta} + \|B\|_1,$$

όπου $\|B\|_1 = \max_{1 \leq j \leq n} \sum |b_{i,j}|$. Τέλος, αν το λ_L είναι μηδέν, τότε συνήθως διαλέγουμε το $\lambda_0 = 0$. Η τιμή $\lambda_0 = \|g\|/\Delta$ έχει επίσης προταθεί.

2.8.2 Η δύσκολη περίπτωση

Ας θυμηθούμε ότι στα προηγούμενα, υποθέσαμε ότι $\gamma_1 \neq 0$, στην περίπτωση του μη θετικά ορισμένου πίνακα B . Όταν αυτή η συνθήκη δεν ισχύει, τότε το όριο $\lim_{\lambda \rightarrow -\lambda_j} \|d(\lambda)\| \neq \infty$ για $\lambda_j = \lambda_1$ και άρα μπορεί να μην υπάρχει τιμή του $\lambda \in (-\lambda_1, \infty)$ για την οποία να ισχύει $\|d(\lambda)\| = \Delta$ (Σχήμα 2.7). Η περίπτωση αυτή αναφέρεται στη βιβλιογραφία ως δύσκολη περίπτωση (hard case). Συμβαίνει στην περίπτωση του μη θετικά ορισμένου πίνακα B , όταν όλα τα ιδιοδιανύσματα που αντιστοιχούν στη λ_1 είναι κάθετα με το διάνυσμα g . Με μια πρώτη ματιά δεν είναι προφανές το πως θα διαλέξουμε τα λ και $d(\lambda)$ έτσι ώστε να ικανοποιούνται οι συνθήκες του Θεωρήματος 2.10 στη δύσκολη περίπτωση. Η τεχνική της μεθόδου Newton για την εύρεση του βέλτιστου λ προφανώς δεν δουλεύει, αφού δεν υπάρχει $\lambda \in (-\lambda_1, \infty)$. Το Θεώρημα 2.10 όμως μας διαβεβαιώνει ότι υπάρχει βέλτιστο λ στο διάστημα $[-\lambda_1, \infty)$. Άρα, προφανώς $\lambda^* = -\lambda_1$.

Για να βρούμε το βήμα d , δεν αρκεί να διαγράψουμε τους όρους για τους οποίους ισχύει $\lambda_j = \lambda_1$ στη σχέση (2.59) και να θέσουμε

$$d(\lambda) = \sum_{j: \lambda_j \neq \lambda_1} \frac{\gamma_j}{\lambda_j + \lambda_1} u_j.$$

Σε αυτή την περίπτωση όπου $\lambda^* = -\lambda_1$, ο πίνακας $B - \lambda_1 I$ είναι ιδιάζον, άρα υπάρχει διάνυσμα z τέτοιο ώστε $\|z\| = 1$ και $(B - \lambda_1 I)z = 0$.

Σχήμα 2.7: Η δύσκολη περίπτωση: $\|d(\lambda)\| < \Delta$ για $\lambda \in (-\lambda_1, \infty)$

Μάλιστα, το z είναι ένα ιδιοδιάνυσμα του B που αντιστοιχεί στη λ_1 , οπότε από την ορθογωνιότητα του U έχουμε ότι $u_j^T z = 0$ για $\lambda_j = \lambda_1$. Από αυτή την ιδιότητα έπεται ότι αν θέσουμε

$$d(\lambda) = \sum_{j:\lambda_j \neq \lambda_1} \frac{\gamma_j}{\lambda_j + \lambda_1} u_j + \tau z, \quad (2.69)$$

για οποιοδήποτε $\tau \in \mathbb{R}$, θα έχουμε ότι

$$\|d(\lambda)\|^2 = \sum_{j:\lambda_j \neq \lambda_1} \frac{\gamma_j^2}{(\lambda_j + \lambda_1)^2} + \tau^2.$$

Άρα, είναι πάντοτε εφικτό να βρούμε κατάλληλο $\tau \in \mathbb{R}$, τέτοιο ώστε να ισχύει $\|d\| = \Delta$.

Ας δούμε δύο τεχνικές λεπτομέρειες της δύσκολης περίπτωσης. Πρώτον, δοθέντος του R , ένα κανονικοποιημένο διάνυσμα z μπορεί να βρεθεί μέσω μιας τεχνικής του πακέτου LINPACK (τώρα LAPACK [3]). Δεύτερον, ένα βαθμωτό μέγεθος τ τέτοιο ώστε $\|d + \tau z\| = \Delta$, μπορεί να υπολογιστεί

από τον τύπο

$$\tau = \frac{\Delta^2 - \|d\|^2}{d^T z + \operatorname{sgn}(d^T z)^2 \sqrt{(d^T z)^2 + (\Delta^2 - \|d\|^2)}}. \quad (2.70)$$

Η έκφραση αυτή προκύπτει από το γεγονός ότι υπάρχουν δύο τιμές του τ για τις οποίες ισχύει ότι $\|d + \tau z\| = \Delta$. Για κάθε $z \in \mathbb{R}^n$ έχουμε:

$$\begin{aligned} \phi(d + z) &= \frac{1}{2}(d + z)^T B(d + z) + g^T(d + z) \\ &= \frac{1}{2}(d + z)^T (B + \lambda I)(d + z) + g^T(d + z) - \frac{1}{2}\lambda\|d + z\|^2 \\ &= \frac{1}{2}(d + z)^T R^T R(d + z) - d^T R^T R(d + z) - \frac{1}{2}\lambda\|d + z\|^2 \\ &= \frac{1}{2}d^T R^T R d + d^T R^T R z + \frac{1}{2}z^T R^T R z \\ &\quad - d^T R^T R d - d^T R^T R z - \frac{1}{2}\lambda\|d + z\|^2 \\ &= -\frac{1}{2}(\|Rd\|^2 + \lambda\|d + z\|^2) + \frac{1}{2}\|Rz\|^2. \end{aligned} \quad (2.71)$$

Χρησιμοποιώντας τη σχέση (2.71) έχουμε

$$\begin{aligned} \phi(d + \tau z) &= -\frac{1}{2}(\|Rd\|^2 + \lambda\|d + \tau z\|^2) + \frac{1}{2}\|R\tau z\|^2 \\ &= -\frac{1}{2}(\|Rd\|^2 + \lambda\Delta^2) + \frac{1}{2}\tau^2\|Rz\|^2. \end{aligned}$$

Η παραπάνω σχέσεις συνεπάγονται ότι πρέπει να διαλέξουμε τη λύση τ με την μικρότερη απόλυτη τιμή, δηλαδή ανάμεσα σε δύο πιθανά βήματα στο σύνορο του περιορισμού, να διαλέξουμε το πιο κοντινό στο d .

Παράδειγμα 2.3. Έστω ότι το πρόβλημα ορίζεται ως ακολούθως:

$$g = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad B = \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

όπου ο B είναι αόριστος. Οι τιμές $\lambda < 2$ είναι απαγορευμένες λόγω του

Θεωρήματος 2.10. Αν $\Delta > 1.2$ τότε δεν υπάρχει αποδεκτή τιμή του λ . Σε αυτή την περίπτωση πρέπει να υπολογίσουμε ένα ιδιοδιάνυσμα z που αντιστοιχεί στη λ_1 και ένα τ ώστε να εξασφαλίσουμε ότι $\|d + \tau z\| = \Delta$.

Σχήμα 2.8: Γράφημα της $\|d(\lambda)\|^2$ στη δύσκολη περίπτωση

2.8.3 Ο Αλγόριθμος επίλυσης του ΥΠΕ

Στον αλγόριθμο των Moré & Sorensen για την επίλυση του ΥΠΕ (2.52), γίνεται χρήση της τεχνικής πακέτου LINPACK. Η τεχνική αυτή συνοψίζεται στα ακόλουθα. Έστω u_1 ένα ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 , οπότε ισχύει ότι $(B - \lambda_1 I)u_1 = 0$. Προσπαθούμε να βρούμε ένα διάνυσμα u τέτοιο ώστε να ελαχιστοποιεί το $u^T (B - \lambda I) u$, ή ισοδύναμα να μεγιστοποιεί το $w = (B - \lambda I)^{-1} u = R^{-T} R^{-1} u$. Θα επιλέξουμε τη συνιστώσα του u που είναι μεταξύ του -1 και του 1 με σκοπό να κάνουμε το $R^{-1} u$ μεγάλο. Αυτό μπορεί να επιτευχθεί με το να βεβαιώσουμε ότι σε κάθε βήμα της αντικατάστασης $Rw = u$, του πρόσημο του u είναι έτσι επιλεγμένο ώστε να κάνει το w όσο πιο μεγάλο γίνεται. Συγκεκριμένα, έστω ότι έχουμε ορίσει τις πρώτες $k-1$ συνιστώσες του w κατά τη διάρκεια της αντικατάστασης $Rw = u$. Τότε η k συνιστώσα ικανοποιεί τη σχέση $r_{kk} w_k = u_k - \sum_{i=1}^{k-1} r_{ki} w_i$, και διαλέγουμε το u_k να είναι ± 1 , ανάλογα με

το πιο κλάσμα

$$\frac{1 - \sum_{i=1}^{k-1} r_{ki} w_i}{r_{kk}} \quad \text{ή} \quad \frac{-1 - \sum_{i=1}^{k-1} r_{ki} w_i}{r_{kk}}$$

είναι μεγαλύτερο. Έχοντας βρει το w , το u είναι απλά το $R^{-T}w/\|R^{-T}w\|$, το οποίο έχει την ιδιότητα $u^T (B + \lambda I) u \rightarrow 0$ καθώς το $\lambda \rightarrow -\lambda_1$.

Έστω ότι κατά την παραγοντοποίηση κατά Cholesky, έχουμε συναντήσει ένα αρνητικό οδηγό στο k στάδιο της ανάλυσης, για κάποιο $k \leq n$. Τότε η παραγοντοποίηση έχει αποτύχει (αφού ο B είναι μη θετικά ορισμένος). Σε αυτή την περίπτωση είναι δυνατό να προσθέσουμε τον αριθμό $\delta = \sum_{j=1}^{k-1} r_{kj}^2 - b_{kk} \geq 0$ στην k διαγώνιο του $B + \lambda I$ έτσι ώστε ο κύριος $k \times k$ υποπίνακας του $B + \lambda I + \delta e_k e_k^T$ να είναι ιδιάζον. Είναι επίσης εύκολο να βρούμε διάνυσμα u για το οποίο ισχύει $(B + \lambda I + \delta e_k e_k^T) u = 0$, χρησιμοποιώντας τους παράγοντες Cholesky που συσσωρεύονται μέχρι το βήμα k . Θέτοντας $u_j = 0$ για $j > k$, $u_k = 1$ και λύνοντας το

$$u_j = -\frac{\sum_{i=j+1}^k r_{ij} u_i}{r_{jj}} \quad \text{για} \quad j = k-1, k-2, \dots, 1,$$

υπολογίζουμε το απαιτούμενο διάνυσμα. Μπορούμε τώρα να υπολογίσουμε ένα κάτω φράγμα για την $-\lambda_1$, χρησιμοποιώντας την ταυτότητα $e_k^T u = u_k = 1$ και τις Rayleigh ανισότητες (2.68). Έχουμε:

$$0 = \frac{u^T (B + \lambda I) u}{u^T u} + \delta \frac{(e_k^T u)^2}{u^T u} \geq \lambda + \lambda_1 + \frac{\delta}{\|u\|^2},$$

από το οποίο έπεται ότι το φράγμα για τη $-\lambda_1$ είναι $\lambda + \frac{\delta}{\|u\|^2} \leq -\lambda_1$. Στον ολοκληρωμένο αλγόριθμο των Moré & Sorensen για την επίλυση του ΥΠΕ (2.52), οι πιθανές τιμές του λ χωρίζονται σε τρία ξένα μεταξύ τους σύνολα:

$$\begin{aligned} \mathcal{N} &= \{\lambda \mid \lambda \leq \max(0, -\lambda_1)\} \\ \mathcal{L} &= \{\lambda \mid \max(0, -\lambda_1) \leq \lambda \leq \lambda^*\} \\ \mathcal{G} &= \{\lambda \mid \lambda > \lambda^*\} \end{aligned}$$

Αλγόριθμος 2.6 Ο αλγόριθμος των Moré & Sorensen για την επίλυση του ΥΠΕ

Δοθέντων $\Delta > 0$, λ_L , λ_U , λ_0 , ϵ_{easy} , ϵ_{hard} ; για $k = 0, 1, 2, \dots$:

Βήμα 1. Προσπάθησε να παραγοντοποιήσεις κατά Cholesky τον πίνακα $B + \lambda_k I = R^T R$, αν δεν έχει ήδη γίνει προσπάθεια.

1a. Αν η παραγοντοποίηση επιτύχει, τότε $\lambda_k \in \mathcal{F}$. Λύσε το σύστημα $RR^T d = -g$. Αν $\|d\| < \Delta$, τότε $\lambda_k \in \mathcal{G}$; έλεγξε για εσωτερική σύγκλιση. Αλλιώς (αν η παραγοντοποίηση αποτύχει), $\lambda_k \in \mathcal{L}$. Αλλιώς $\lambda_k \in \mathcal{N}$.

Βήμα 2. Αν $\lambda_k \in \mathcal{G}$, θέσε $\lambda_U \leftarrow \lambda_k$. Διαφορετικά, θέσε $\lambda_L \leftarrow \lambda_k$.

Βήμα 3. Αν $\lambda_k \in \mathcal{F}$,

3a. Λύσε το σύστημα $Rq = d$ και θέσε $\lambda_{k+1} = \lambda_k + \left(\frac{\|d\| - \Delta}{\Delta} \right) \left(\frac{\|d\|}{\|q\|} \right)^2$.

3b. Αν $\lambda_k \in \mathcal{G}$,

(i) Χρησιμοποίησε τη LINPACK μέθοδο για να βρεις ένα μοναδιαίο διάνυσμα u , έτσι ώστε να γίνει η ποσότητα $u^T (B + \lambda_k I) u$ μικρή.

(ii) Θέσε το $\lambda_L \leftarrow \max(\lambda_L, \lambda_k - u^T (B + \lambda_k I) u)$.

(iii) Βρες τη λύση τ της εξίσωσης $\|d + \tau u\| = \Delta$, η οποία κάνει την τιμή του μοντέλου $\phi(d + \tau u)$ ελάχιστη και θέσε το $d \leftarrow d + \tau u$.

Αλλιώς, (αν $\lambda_k \notin \mathcal{F}$)

3γ. Χρησιμοποίησε τη μερική παραγοντοποίηση για να βρεις δ και u τέτοια ώστε $[(B + \lambda_k I) + \delta e_k e_k^T] u = 0$.

3δ. Θέσε το $\lambda_L \leftarrow \max(\lambda_L, \lambda_k + \frac{\delta}{\|u\|^2})$.

Βήμα 4. Έλεγξε για τερματισμό:

4a. Αν $\lambda_k \in \mathcal{F}$ και $|\|d\| - \Delta| \leq \epsilon_{easy} \Delta$, σταμάτα, με $d = d(\lambda_k)$.

4b. Αν $\lambda_k = 0 \in \mathcal{G}$, σταμάτα, με $d = d(0)$.

4γ. Αν $\lambda_k \in \mathcal{G}$ και η LINPACK μέθοδος δίνει ένα u και ένα τ τέτοια ώστε $\tau^2 u^T (B + \lambda_k I) u \leq \epsilon_{hard} (d(\lambda_k)^T (B + \lambda_k I) d(\lambda_k) + \lambda_k \Delta^2)$, σταμάτα, με $d = d(\lambda_k) + \tau u$.

Βήμα 5. Αν $\lambda_k \in \mathcal{L}$ και $g \neq 0$ θέσε $\lambda_k \leftarrow \lambda_{k+1}$. Αλλιώς, αν $\lambda_k \in \mathcal{G}$, προσπάθησε να παραγοντοποιήσεις τον $B + \lambda_{k+1} I = R R^T$.

5a. Αν η παραγοντοποίηση είναι επιτυχής, τότε $\lambda_{k+1} \in \mathcal{L}$. Θέσε $\lambda_k \leftarrow \lambda_{k+1}$.

5b. Αλλιώς, $\lambda_{k+1} \in \mathcal{N}$. Θέσε $\lambda_L \leftarrow \max(\lambda_L, \lambda_{k+1})$ και έλεγξε για εσωτερική σύγκλιση; αλλιώς θέσε $\lambda_k \leftarrow \max(\sqrt{\lambda_L \lambda_U}, \lambda_L + 0.01(\lambda_U - \lambda_L))$. Αλλιώς, θέσε $\lambda_k \leftarrow \max(\sqrt{\lambda_L \lambda_U}, \lambda_L + 0.01(\lambda_U - \lambda_L))$.

Επιπρόσθετα, έστω $\mathcal{F} = \mathcal{L} \cup \mathcal{G}$. Το σύνολο $\mathcal{N}$ περιέχει τις επαναλήψεις για τις οποίες η κατά Cholesky παραγοντοποίηση μπορεί να μην υπάρχει. Τα σύνολα $\mathcal{L}$ και $\mathcal{G}$ περιέχουν τις άλλες επαναλήψεις, για τις οποίες η $\psi(\lambda)$

είναι μη θετική και θετική, αντίστοιχα. Υποθέτουμε ότι το $\Delta > 0$ είναι δοσμένο και έχουμε ένα διάστημα $[\lambda_L, \lambda_U]$ εμπιστοσύνης του λ , καθώς και μια πρώτη εκτίμηση $\lambda \in [\lambda_L, \lambda_U]$.

Θεώρημα 2.11. *Έστω ότι το κριτήριο τερματισμού στο Βήμα 4 του Αλγορίθμου 2.6 δεν εφαρμόζεται. Τότε οι επαναλήψεις λ_k του αλγορίθμου συγκλίνουν στο λ^* και το όριο του d είναι το d^* . Ο αλγόριθμος συγκλίνει είτε σε πεπερασμένο πλήθος βημάτων ή, εκτός από τη δύσκολη περίπτωση, τελικά με Q -τετραγωνικό ρυθμό.*

Η μέθοδος των Moré & Sorensen [64] για την επίλυση του ΥΠΕ (2.52), θεωρείται μέχρι και σήμερα η καλύτερη μέθοδος για την επίλυση του υποπροβλήματος. Η μέθοδος βασίζεται στην κατά Cholesky παραγοντοποίηση του $B + \lambda I$ κάθε φορά που το υποπρόβλημα λύνεται. Επιπρόσθετα, στη δύσκολη περίπτωση χρειάζεται να παραχθεί μια κατεύθυνση αρνητικής καμπυλότητας, η οποία συνδέεται με το ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική τιμή του B . Κατά συνέπεια, η μέθοδος των Moré & Sorensen είναι πολύ ακριβή και απαγορευτική στην εφαρμογή της για προβλήματα μεγάλης διάστασης. Στις μέρες μας, η έρευνα εστιάζεται στην κατάλληλη τροποποίηση της μεθόδου των Moré & Sorensen, ώστε αυτή να μπορεί να εφαρμοστεί και σε προβλήματα με πολύ μεγάλο αριθμό μεταβλητών. Για το λόγο αυτό, αναπτύχθηκαν επαναληπτικές μέθοδοι οι οποίες βασίζονται στη μέθοδο Lanczos καθώς και σε γινόμενα πίνακα-διανύσματος [45, 48, 77, 78, 83].

Οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης

Σε αυτό το κεφάλαιο μελετάμε τα χαρακτηριστικά μεγέθη των Quasi-Newton ενημερώσεων ελάχιστης μνήμης. Με τον όρο “ελάχιστη μνήμη” εννοούμε ότι για την ενημέρωση του προσεγγιστικού εσσιανού πίνακα κάνουμε χρήση μόνο ενός ($m = 1$) διανυσματικού ζεύγους (s_k, y_k) , ενώ σε κάθε επανάληψη, ο αρχικός μας πίνακας θεωρούμε ότι είναι ένα πολλαπλάσιο του μοναδιαίου πίνακα. Βασιζόμενοι στη θεμελιώδη ιδιότητα της κλάσης αυτής των πινάκων (ελάχιστης μνήμης) είμαστε σε θέση να παράγουμε κλειστό τύπο για το χαρακτηριστικό τους πολυώνυμο. Άμεση συνέπεια αυτού είναι ότι μπορούμε να βρούμε αναλυτικά τις ιδιοτιμές τους, καθώς και τον αντίστροφο του πίνακα $(B + \mu I)$, όπου $\mu \in \mathbb{R}$, που απαιτείται για τον υπολογισμό της κατεύθυνσης σε μεθόδους βελτιστοποίησης. Επίσης, παράγουμε κλειστούς τύπους εύρεσης των ιδιοδιανυσμάτων που αντιστοιχούν σε διακριτές ιδιοτιμές. Αυτό επιτυγχάνεται είτε μέσω απλών αλγεβρικών πράξεων ή μέσω της εφαρμογής της μεθόδου της αντίστροφης δύναμης, στον αντίστροφο του πίνακα $(B + \mu I)$. Η προτεινόμενη μέθοδος αποφεύγει πλήρως την παραγοντοποίηση πινάκων, ενώ βασίζεται σε αθροίσματα διανυσμάτων και εσωτερικών γινομένων. Τα ακόλουθα αποτελούν ερευνητικά αποτελέσματα της παρούσας διατριβής, μέρος των οποίων καταγράφονται στις εργασίες των Apostolopoulou et al. [8, 9, 10].

3.1 Εισαγωγή

Θα ξεκινήσουμε την ανάλυσή μας, θεωρώντας τη γενική PSB ενημέρωση, η οποία αποτελεί και τον γεννήτορα της κλάσης των Quasi-Newton ενημερώσεων, δηλαδή των SR1, BFGS, DFP και PSB. Υπενθυμίζουμε, ότι η

γενική PSB ενημέρωση έχει την ακόλουθη έκφραση :

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) w_k^T + w_k (y_k - B_k s_k)^T}{s_k^T w_k} - \frac{(y_k - B_k s_k)^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{s_k^T w_k}, \quad (3.1)$$

όπου $s_i = x_{i+1} - x_i$, και $y_i = g_{i+1} - g_i$. Όπως είδαμε στην Υποενότητα 2.3.4 (σελ. 42), σε αυτή την κλάση ενημερώσεων, αν θέσουμε $w_k = y_k - B_k s_k$, τότε η (3.1) είναι η SR1 ενημέρωση,

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k)(y_k - B_k s_k)^T}{(y_k - B_k s_k)^T s_k}.$$

Αν $w_k = [1/(v_k + 1)]y_k + [v_k/(v_k + 1)]B_k s_k$, όπου $v_k = (s_k^T y_k / s_k^T B_k s_k)^{1/2}$, τότε η (3.1) δίνει την BFGS ενημέρωση,

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}.$$

Αν $w_k = y_k$, η (3.1) εκφράζει την DFP ενημέρωση,

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) y_k^T + y_k (y_k - B_k s_k)^T}{s_k^T y_k} - \frac{(y_k - B_k s_k)^T s_k}{s_k^T y_k} \frac{y_k y_k^T}{s_k^T y_k}.$$

Τέλος, αν $w_k = s_k$, τότε η (3.1) γίνεται

$$B_{k+1} = B_k + \frac{(y_k - B_k s_k) s_k^T + s_k (y_k - B_k s_k)^T}{s_k^T s_k} - \frac{(y_k - B_k s_k)^T s_k}{s_k^T s_k} \frac{s_k s_k^T}{s_k^T s_k},$$

η οποία είναι η PSB ενημέρωση.

Όπως προαναφέραμε στην Υποενότητα 2.4.1 (σελ. 45), οι Quasi-Newton ενημερώσεις ελάχιστης μνήμης, βασίζονται στις περιορισμένης μνήμης Quasi-Newton ενημερώσεις [67]. Δοθέντος αρχικού πίνακα $B^{(0)} = \theta I$, όπου $\theta \in \mathbb{R}$, οι περιορισμένης μνήμης Quasi-Newton μέθοδοι ενημερώνουν τον B_k χρησιμοποιώντας πληροφορία από τις πιο πρόσφατες επαναλήψεις του αλγορίθμου αναζήτησης στον οποίο χρησιμοποιούνται. Οι πληροφορίες αυτές περιλαμβάνονται σε ένα συγκεκριμένο αριθμό m αποθηκευμένων διανυσματικών ζευγών $\{s_i, y_i\}$. Στις Quasi-Newton ενη-

μερώσεις ελάχιστης μνήμης, ο αριθμός των αποθηκευμένων διανυσματικών ζευγών είναι $m = 1$, ενώ ο προηγούμενος προσεγγιστικός πίνακας αντικαθίσταται από το μοναδιαίο ή κάποιο πολλαπλάσιό του. Επομένως, αυτές οι ενημερώσεις χρησιμοποιούν πληροφορία από την πιο πρόσφατη επανάληψη, δηλαδή μία ενημέρωση εκτελείται στον $B^{(0)}$, χρησιμοποιώντας το διανυσματικό ζεύγος $\{s_k, y_k\}$. Οπότε, η έκφραση της γενικής PSB ενημέρωσης ελάχιστης μνήμης (3.1) είναι ο πίνακας

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)w_k^T + w_k(y_k - \theta_{k+1}s_k)^T}{s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \cdot \frac{w_k w_k^T}{s_k^T w_k}, \quad (3.2)$$

όπου το w_k είναι γραμμικός συνδυασμός των s_k, y_k , της μορφής

$$w_k = c_1 y_k + c_2 s_k.$$

Ανάλογα με το γραμμικό αυτόν συνδυασμό, ο πίνακας εκφράζει:

- την SR1 ενημέρωση ελάχιστης μνήμης, για

$$\begin{aligned} w_k &= y_k - \theta_{k+1}s_k \quad \text{και} \\ c_1 &= 1, \\ c_2 &= \theta_{k+1}, \end{aligned}$$

- την BFGS ενημέρωση ελάχιστης μνήμης, για

$$\begin{aligned} w_k &= \frac{1}{v_k + 1}y_k + \frac{\theta_{k+1}v_k}{v_k + 1}s_k \quad \text{όπου} \quad v_k = \left(\frac{s_k^T y_k}{\theta_{k+1}s_k^T s_k} \right)^{1/2} \quad \text{και} \\ c_1 &= \frac{1}{v_k + 1}, \\ c_2 &= \frac{\theta_{k+1}v_k}{v_k + 1}, \end{aligned}$$

- την DFP ενημέρωση ελάχιστης μνήμης, για

$$\begin{aligned} w_k &= y_k \text{ και} \\ c_1 &= 1, \\ c_2 &= 0, \end{aligned}$$

- και τέλος την PSB ενημέρωση ελάχιστης μνήμης, για

$$\begin{aligned} w_k &= s_k \text{ και} \\ c_1 &= 0, \\ c_2 &= 1. \end{aligned}$$

Αναφορικά με την παράμετρο θ_{k+1} , όπως είδαμε στην Ενότητα 2.2, η πιο ευρέως χρησιμοποιούμενη παράμετρος που συναντάται στη διεθνή βιβλιογραφία [68] ορίζεται ως

$$\theta_{k+1} = \frac{y_k^T y_k}{s_k^T y_k}.$$

Τα τελευταία χρόνια όμως, κυρίως στην BFGS ενημέρωση, έχει αρχίσει να χρησιμοποιείται η φασματική παράμετρος των Barzilai & Borwein [15], γνωστή ως *BB spectral parameter*,

$$\theta_{k+1} = \frac{s_k^T y_k}{s_k^T s_k}.$$

Σκοπός μας είναι να μελετήσουμε τα χαρακτηριστικά μεγέθη της γενικής PSB ενημέρωσης ελάχιστης μνήμης, ώστε να μπορέσουμε να λύσουμε μεγάλης κλίμακας προβλήματα βελτιστοποίησης, με μικρό υπολογιστικό κόστος. Στις Ενότητες 2.6 και 2.8 του Κεφαλαίου 2, οι οποίες αναφέρονται στις μεθόδους καμπυλόγραμμης αναζήτησης και στο υποπρόβλημα περιοχής εμπιστοσύνης, αντίστοιχα, είδαμε ότι απαιτείται η παραγοντοποίηση πινάκων. Αυτό έχει ως αποτέλεσμα οι μέθοδοι αυτές να απαιτούν μεγάλο υπολογιστικό κόστος, με αποτέλεσμα η εφαρμογή τους να καθίσταται απαγορευτική για μεγάλης κλίμακας προβλήματα. Για να μπορέσουμε να αποφύγουμε το μεγάλο υπολογιστικό κόστος, θα πρέπει να αποφύγουμε τόσο την παραγοντοποίηση, όσο και την αποθήκευση πινάκων.

Στις μεθόδους της καμπυλόγραμμης αναζήτησης, απαιτείται ο υπολογισμός της πιο αρνητικής ιδιοτιμής του προσεγγιστικού εσσιανού, καθώς και του αντίστοιχου ιδιοδιανύσματος. Στο υποπρόβλημα περιοχής εμπιστοσύνης απαιτείται επιπρόσθετα ο υπολογισμός του n -διάστατου διανύσματος $(B + \mu I)^{-1} g$, $\mu \in \mathbb{R}$. Μέσω της μελέτης των χαρακτηριστικών μεγεθών των Quasi-Newton ενημερώσεων ελάχιστης μνήμης, προκύπτει ότι όλοι οι απαιτούμενοι υπολογισμοί γίνονται μέσω διανυσματικών αθροισμάτων και εσωτερικών γινομένων. Επομένως, δεν απαιτείται παραγοντοποίηση ή αποθήκευση πινάκων, με αποτέλεσμα η προτεινόμενη μέθοδος να απαιτεί ελάχιστο υπολογιστικό κόστος και μνήμη, άρα να είναι κατάλληλη να εφαρμοστεί σε μεγάλης κλίμακας προβλήματα.

Στη συνέχεια αναλύουμε τον τρόπο εύρεσης των ποσοτήτων που μας ενδιαφέρουν. Συγκεκριμένα, βρίσκουμε το χαρακτηριστικό πολυώνυμο των Quasi-Newton ενημερώσεων ελάχιστης μνήμης. Επιλύοντας τη χαρακτηριστική εξίσωση υπολογίζουμε αναλυτικά τόσο τις ιδιοτιμές, όσο και τον πίνακα $(B + \mu I)^{-1}$, ο οποίος αποτελεί θεωρητικό κατασκεύασμα (αφού δεν αποθηκεύεται) για τον υπολογισμό του διανύσματος $(B + \mu I)^{-1} g$. Άμεση συνέπεια αυτών είναι ότι μπορούμε να εκφράσουμε τα ιδιοδιανύσματα που μας ενδιαφέρουν, μέσω κλειστών τύπων, κάνοντας χρήση της μεθόδου της αντίστροφης δύναμης.

3.2 Το χαρακτηριστικό πολυώνυμο

Για να μπορέσουμε να υπολογίσουμε το χαρακτηριστικό πολυώνυμο των Quasi-Newton ενημερώσεων ελάχιστης μνήμης, χρειαζόμαστε τα ίχνη και τις ορίζουσες των αντίστοιχων πινάκων. Από τη γενική Quasi-Newton ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k) w_k^T + w_k (y_k - \theta_{k+1} s_k)^T}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \cdot \frac{w_k w_k^T}{s_k^T w_k}, \quad (3.3)$$

χρησιμοποιώντας τις ιδιότητες του ίχνους (σελ. 13), έχουμε ότι

$$\begin{aligned} \text{tr}(B_{k+1}) &= n\theta_{k+1} + \frac{2(y_k - \theta_{k+1}s_k)^T w_k}{s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \\ &= n\theta_{k+1} + \frac{(y_k - \theta_{k+1}s_k)^T}{s_k^T w_k} \left(2w_k - \frac{w_k^T w_k}{s_k^T w_k} s_k \right). \end{aligned} \quad (3.4)$$

Αντίστοιχα, εφαρμόζοντας τις ιδιότητες της ορίζουσας (σελ. 13), στον πίνακα (3.3), έχουμε ότι

$$\begin{aligned} \det(B_{k+1}) &= \det \left((\theta_{k+1}I) \left[I + \frac{(y_k - \theta_{k+1}s_k) w_k^T + w_k (y_k - \theta_{k+1}s_k)^T}{\theta_{k+1}s_k^T w_k} \right. \right. \\ &\quad \left. \left. - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{\theta_{k+1}s_k^T w_k} \right] \right) \\ &= \det(\theta_{k+1}I) \det \left[I + \left(\frac{y_k - \theta_{k+1}s_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k \right) w_k^T \right. \\ &\quad \left. + w_k \frac{(y_k - \theta_{k+1}s_k)^T}{\theta_{k+1}s_k^T w_k} \right]. \end{aligned} \quad (3.5)$$

Θέτοντας

$$\begin{aligned} v_1 &= \frac{y_k - \theta_{k+1}s_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k \\ v_2 &= w_k \\ v_3 &= w_k \\ v_4 &= \frac{y_k - \theta_{k+1}s_k}{\theta_{k+1}s_k^T w_k} \end{aligned}$$

και εφαρμόζοντας τη σχέση (σελ. 13)

$$\det(I + v_1 v_2^T + v_3 v_4^T) = (1 + v_1^T v_2)(1 + v_3^T v_4) - (v_1^T v_4)(v_2^T v_3),$$

στην εξίσωση (3.5), βρίσκουμε ότι

$$\det(B_{k+1}) = \theta_{k+1}^n \left[\left(1 + \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k^T w_k \right) \right. \\ \left(1 + \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k} \right) - \left(\frac{\|y_k - \theta_{k+1}s_k\|^2}{\theta_{k+1}^2 (s_k^T w_k)^2} \right. \\ \left. \left. - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k} \right) w_k^T w_k \right].$$

Αν στην παραπάνω σχέση θέσουμε την πρώτη παρένθεση ως

$$A = 1 + \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k^T w_k,$$

τη δεύτερη ως

$$B = 1 + \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k},$$

και την τρίτη ως

$$\Gamma = \frac{\|y_k - \theta_{k+1}s_k\|^2}{\theta_{k+1}^2 (s_k^T w_k)^2} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} \frac{(y_k - \theta_{k+1}s_k)^T w_k}{\theta_{k+1}s_k^T w_k},$$

λαμβάνουμε την παρακάτω απλοποιημένη σχέση για την ορίζουσα

$$\det(B_{k+1}) = \theta_{k+1}^n (AB - \Gamma w_k^T w_k). \quad (3.6)$$

Για τις παρενθέσεις A , B και Γ , έχουμε:

$$\begin{aligned} A &= 1 + \frac{y_k^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{\theta_{k+1}s_k^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k^T w_k \\ &= \frac{y_k^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{\theta_{k+1}(s_k^T w_k)^2} w_k^T w_k, \end{aligned} \quad (3.7)$$

$$B = 1 + \frac{y_k^T w_k}{\theta_{k+1}s_k^T w_k} - \frac{\theta_{k+1}s_k^T w_k}{\theta_{k+1}s_k^T w_k} = \frac{y_k^T w_k}{\theta_{k+1}s_k^T w_k} \quad (3.8)$$

και

$$\begin{aligned}
 \Gamma &= \frac{(y_k - \theta_{k+1} s_k)^T}{\theta_{k+1}^2 (s_k^T w_k)^2} \left(y_k - \theta_{k+1} s_k - \frac{(y_k - \theta_{k+1} s_k)^T w_k}{s_k^T w_k} s_k \right) \\
 &= \frac{(y_k - \theta_{k+1} s_k)^T}{\theta_{k+1}^2 (s_k^T w_k)^2} \left(y_k - \theta_{k+1} s_k - \frac{y_k^T w_k}{s_k^T w_k} s_k + \frac{\theta_{k+1} s_k^T w_k}{s_k^T w_k} s_k \right) \\
 &= \frac{(y_k - \theta_{k+1} s_k)^T}{\theta_{k+1}^2 (s_k^T w_k)^2} \left(y_k - \frac{y_k^T w_k}{s_k^T w_k} s_k \right), \tag{3.9}
 \end{aligned}$$

αντίστοιχα. Αντικαθιστώντας τις σχέσεις (3.7), (3.8) και (3.9) στην (3.6), βρίσκουμε ότι

$$\begin{aligned}
 \det(B_{k+1}) &= \theta_{k+1}^n \left[\left(\frac{y_k^T w_k}{\theta_{k+1} s_k^T w_k} \right)^2 - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{\theta_{k+1} s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \frac{y_k^T w_k}{\theta_{k+1} s_k^T w_k} \right. \\
 &\quad \left. - \frac{(y_k - \theta_{k+1} s_k)^T}{\theta_{k+1}^2 (s_k^T w_k)^2} \left(y_k - \frac{y_k^T w_k}{s_k^T w_k} s_k \right) w_k^T w_k \right] \\
 &= \frac{\theta_{k+1}^{n-2}}{(s_k^T w_k)^2} \left[(y_k^T w_k)^2 \right. \\
 &\quad \left. - w_k^T w_k (y_k - \theta_{k+1} s_k)^T \left(\frac{y_k^T w_k}{s_k^T w_k} s_k + y_k - \frac{y_k^T w_k}{s_k^T w_k} s_k \right) \right] \\
 &= \theta_{k+1}^{n-2} \frac{(y_k^T w_k)^2 - w_k^T w_k (y_k - \theta_{k+1} s_k)^T y_k}{(s_k^T w_k)^2}. \tag{3.10}
 \end{aligned}$$

Το επόμενο θεώρημα δίνει το χαρακτηριστικό πολυώνυμο των Quasi-Newton πινάκων ελάχιστης μνήμης και τον αριθμό των διακριτών ιδιοτιμών τους.

Θεώρημα 3.1. Έστω ότι μία ενημέρωση εκτελείται σε έναν αρχικό πίνακα $B^{(0)} = \theta I$, $\theta \in \mathbb{R}$, χρησιμοποιώντας το διανυσματικό ζεύγος (s_k, y_k) και τη γενική PSB ενημέρωση ελάχιστης μνήμης. Τότε, το χαρακτηριστικό πολυώνυμο του Quasi-Newton πίνακα ελάχιστης μνήμης, όπως ορίζεται στην (3.3), έχει τη γενική μορφή

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} (\lambda^2 - \beta_1 \lambda + \beta_2), \tag{3.11}$$

όπου

$$\beta_1 = 2 \frac{y_k^T w_k}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \quad \text{και} \quad (3.12)$$

$$\beta_2 = \left(\frac{y_k^T w_k}{s_k^T w_k} \right)^2 - \frac{(y_k - \theta_{k+1} s_k)^T y_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k}. \quad (3.13)$$

Επιπρόσθετα, αν τα διανύσματα s_k και y_k είναι γραμμικώς ανεξάρτητα,

- i. αν ο $B^{(0)}$ ενημερώνεται από την πρόσθεση ενός πίνακα πρώτης τάξης, τότε ο B_{k+1} έχει ακριβώς μία διακριτή ιδιοτιμή
- ii. αν ο $B^{(0)}$ ενημερώνεται από την πρόσθεση δύο πινάκων πρώτης τάξης, τότε ο B_{k+1} έχει ακριβώς δύο διακριτές ιδιοτιμές, εκ των οποίων η μία είναι η μικρότερη ιδιοτιμή του B_{k+1} .

Απόδειξη. Έστω λ_i , $i = 1, \dots, n$, οι ιδιοτιμές του πίνακα B_{k+1} , όπως ορίζεται στην (3.3). Θεωρούμε τον πίνακα

$$\bar{B} = \theta_{k+1} I + \left[\frac{y_k - \theta_{k+1} s_k}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k}{s_k^T w_k} \right] w_k^T,$$

ο οποίος παράγεται από την πρόσθεση ενός πίνακα πρώτης τάξης στον $B^{(0)} = \theta_{k+1} I$. Έστω ότι οι ιδιοτιμές του $\bar{B}$ είναι οι $\lambda_i^{\bar{B}}$, $i = 1, \dots, n$. Από το θεώρημα συνδιαπλοκής ιδιοτιμών 1.3 (σελ. 14) στον $\bar{B}$, είναι εύκολο να δούμε ότι ο $\bar{B}$ πέρα από τη μηδενική ιδιοτιμή, έχει ακόμα μία, πολλαπλότητας $(n-1)$, η οποία είναι ίση με θ_{k+1} .

Αν ο B_{k+1} είναι θετικά ορισμένος, τότε προσθέτοντας τον πίνακα πρώτης τάξης

$$\frac{w_k (y_k - \theta_{k+1} s_k)^T}{s_k^T w_k}$$

στον $\bar{B}$, εφαρμόζοντας το θεώρημα συνδιαπλοκής ιδιοτιμών συμπεραίνουμε ότι

$$\lambda_n \geq \theta_{k+1} \geq \lambda_{n-1} \geq \theta_{k+1} \geq \dots \geq \lambda_2 \geq \theta_{k+1} \geq \lambda_1 \geq 0.$$

Από τις παραπάνω ανισότητες συνεπάγεται ότι

$$\lambda_2 = \dots = \lambda_{n-1} = \theta_{k+1} \quad \text{και} \quad \lambda_n \geq \theta_{k+1} \geq \lambda_1. \quad (3.14)$$

Έστω τώρα ότι ο B_{k+1} δεν είναι θετικά ορισμένος. Τότε, αν $\theta_{k+1} > 0$, από το θεώρημα συνδιαπλοκής συμπεραίνουμε ότι

$$\theta_{k+1} \geq \lambda_n \geq \theta_{k+1} \geq \lambda_{n-1} \geq \theta_{k+1} \geq \dots \geq \lambda_3 \geq \theta_{k+1} \geq \lambda_2 \geq 0 \geq \lambda_1.$$

Επομένως, έπεται ότι

$$\lambda_3 = \dots = \lambda_n = \theta_{k+1} \quad \text{και} \quad \theta_{k+1} \geq \lambda_2 \geq \lambda_1. \quad (3.15)$$

Αν $\theta_{k+1} < 0$ έχουμε ότι

$$0 \geq \lambda_n \geq \theta_{k+1} \geq \lambda_{n-1} \geq \theta_{k+1} \geq \dots \geq \lambda_2 \geq \theta_{k+1} \geq \lambda_1.$$

Συνεπώς,

$$\lambda_2 = \dots = \lambda_{n-1} = \theta_{k+1} \quad \text{και} \quad \lambda_n \geq \theta_{k+1} \geq \lambda_1. \quad (3.16)$$

Σε όλες τις παραπάνω περιπτώσεις είναι προφανές ότι ο B_{k+1} έχει το πολύ δύο διακριτές ιδιοτιμές και μία ίση με θ_{k+1} , πολλαπλότητας το λιγότερο $(n-2)$. Δηλώνοντας με λ_x και λ_y τις άγνωστες ιδιοτιμές, το χαρακτηριστικό πολυώνυμο του B_{k+1} μπορεί να γραφεί ως ακολούθως:

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - (\lambda_x + \lambda_y)\lambda + \lambda_x \lambda_y \right].$$

Λαμβάνοντας υπόψη ότι

$$\text{tr}(B_{k+1}) = \sum_{i=1}^n \lambda_i = (n-2)\theta_{k+1} + \lambda_x + \lambda_y \quad (3.17)$$

και

$$\det(B_{k+1}) = \prod_{i=1}^n \lambda_i = \theta_{k+1}^{n-2} \lambda_x \lambda_y, \quad (3.18)$$

έχουμε ότι

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left\{ \lambda^2 - \left[\text{tr}(B_{k+1}) - (n-2)\theta_{k+1} \right] \lambda + \frac{\det(B_{k+1})}{\theta_{k+1}^{n-2}} \right\}.$$

Αν το διάνυσμα w_k ορίζεται ως $w_k = y_k - \theta_{k+1}s_k$, τότε ο $B^{(0)}$ ενημερώνε-

ται από την πρόσθεση ενός πίνακα πρώτης τάξης (ελάχιστης μνήμης SR1 ενημέρωση), οπότε από το θεώρημα συνδιαπλοκής ιδιοτιμών είναι εύκολο να δούμε ότι ο B_{k+1} έχει το πολύ μία ιδιοτιμή διακριτή, έστω την λ_x . Σε αυτή λοιπόν την περίπτωση, το χαρακτηριστικό πολυώνυμο γίνεται

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-1} (\lambda - \lambda_x). \quad (3.19)$$

Από τις σχέσεις (3.17) και (3.18), λύνοντας ως προς λ_x βρίσκουμε ότι $\lambda_x = \text{tr}(B_{k+1}) - (n-1)\theta_{k+1} = \det(B_{k+1})/\theta_{k+1}^{n-1}$. Επομένως, η σχέση (3.19) γίνεται

$$\begin{aligned} p(\lambda) &= (\lambda - \theta_{k+1})^{n-1} \{ \lambda - [\text{tr}(B_{k+1}) - (n-1)\theta_{k+1}] \} \\ &= (\lambda - \theta_{k+1})^{n-2} (\lambda - \theta_{k+1}) \{ \lambda - [\text{tr}(B_{k+1}) - (n-1)\theta_{k+1}] \} \\ &= (\lambda - \theta_{k+1})^{n-2} \left\{ \lambda^2 - [\text{tr}(B_{k+1}) - (n-2)\theta_{k+1}] \lambda + \frac{\det(B_{k+1})}{\theta_{k+1}^{n-2}} \right\}. \end{aligned}$$

Άρα, και στις δύο περιπτώσεις (πρώτης και δεύτερης τάξης ενημέρωση), το χαρακτηριστικό πολυώνυμο έχει την ίδια έκφραση. Θέτοντας

$$\beta_1 = \text{tr}(B_{k+1}) - (n-2)\theta_{k+1} \quad \text{και} \quad \beta_2 = \det(B_{k+1})/\theta_{k+1}^{n-2}, \quad (3.20)$$

και αντικαθιστώντας τη σχέση (3.4) στην πρώτη εξίσωση, έχουμε:

$$\begin{aligned} \beta_1 &= n\theta_{k+1} + \frac{(y_k - \theta_{k+1}s_k)^T}{s_k^T w_k} \left(2w_k - \frac{w_k^T w_k}{s_k^T w_k} s_k \right) - (n-2)\theta_{k+1} \\ &= 2 \frac{(y_k - \theta_{k+1}s_k)^T w_k}{s_k^T w_k} - \frac{w_k^T w_k}{s_k^T w_k} \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} + 2\theta_{k+1} \\ &= 2 \left(\frac{y_k^T w_k}{s_k^T w_k} - \frac{\theta_{k+1}s_k^T w_k}{s_k^T w_k} \right) - \frac{w_k^T w_k}{s_k^T w_k} \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} + 2\theta_{k+1} \\ &= 2 \frac{y_k^T w_k}{s_k^T w_k} - \frac{w_k^T w_k}{s_k^T w_k} \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \end{aligned}$$

Ομοίως, αντικαθιστώντας την (3.10) στη δεύτερη εξίσωση της (3.20), έχου-

με:

$$\begin{aligned}
 \beta_2 &= \frac{\det(B_{k+1})}{\theta_{k+1}^{n-2}} \\
 &= \theta_{k+1}^{n-2} \frac{(y_k^T w_k)^2 - w_k^T w_k (y_k - \theta_{k+1} s_k)^T y_k}{(\theta_{k+1}^{n-2} (s_k^T w_k)^2)} \\
 &= \frac{(y_k^T w_k)^2 - w_k^T w_k (y_k - \theta_{k+1} s_k)^T y_k}{(s_k^T w_k)^2}
 \end{aligned}$$

Άρα, καταλήγουμε στις σχέσεις (3.12) και (3.13), αντίστοιχα. Επομένως, η γενική μορφή του $p(\lambda)$ δίνεται από την (3.11) σε κάθε περίπτωση.

Υποθέτουμε τώρα ότι τα διανύσματα s_k και y_k είναι γραμμικώς ανεξάρτητα. Θεωρούμε πρώτα την περίπτωση όπου η ενημέρωση του $B^{(0)}$ γίνεται από την πρόσθεση ενός πίνακα πρώτης τάξης, δηλαδή την SR1 ελάχιστης μνήμης, όπου $w_k = y_k - \theta_{k+1} s_k$. Στην περίπτωση αυτή ο πίνακας έχει το πολύ μία διακριτή ιδιοτιμή και το χαρακτηριστικό πολυώνυμο δίνεται από τη σχέση (3.19). Έστω ότι καμία ιδιοτιμή δεν είναι διακριτή, δηλαδή έστω ότι ισχύει

$$\lambda_1 = \theta_{k+1} = \dots = \lambda_n.$$

Από το ίχνος (3.4) του χωρίς μνήμη SR1 πίνακα έχουμε ότι

$$n\theta_{k+1} = \text{tr}(B_{k+1}) \Rightarrow \frac{(y_k - \theta_{k+1} s_k)^T (y_k - \theta_{k+1} s_k)}{s_k^T (y_k - \theta_{k+1} s_k)} = 0,$$

και άρα $y_k = \theta_{k+1} s_k$, δηλαδή τα διανύσματα s_k και y_k είναι συγγραμμικά. Αυτό όμως είναι άτοπο, επομένως ακριβώς μία ιδιοτιμή του πίνακα είναι διακριτή, λόγω της σχέσης (3.19).

Θεωρούμε τώρα την περίπτωση όπου ο $B^{(0)}$ ενημερώνεται από την πρόσθεση δύο πινάκων πρώτης τάξης, όπου ισχύει ότι $w_k \neq y_k - \theta_{k+1} s_k$. Έστω ϕ η γωνία μεταξύ των διανυσμάτων s_k , y_k και έστω ότι ο B_{k+1} δεν έχει δύο διακριτές ιδιοτιμές, αλλά καμία ή μία. Οπότε είτε $\lambda_x = \theta_{k+1} = \lambda_y$ ή ότι τουλάχιστον μία από τις ισότητες $\lambda_x = \theta_{k+1}$ ή $\lambda_y = \theta_{k+1}$, θα ισχύει. Και στις δύο περιπτώσεις, χρησιμοποιώντας τις σχέσεις (3.4) και (3.10) και αντικαθιστώντας το διάνυσμα w_k με τις αντίστοιχες εκφράσεις του για τις

BFGS, DFP και PSB ενημερώσεις ελάχιστης μνήμης, έχουμε ότι

$$(s_k^T y_k)^2 = s_k^T s_k y_k^T y_k \quad \text{από όπου έχουμε ότι} \quad \cos \phi = \pm 1.$$

Η παραπάνω σχέση δηλώνει ότι τα διανύσματα s_k και y_k είναι συγγραμμικά, πράγμα άτοπο. Άρα, αν τα s_k και y_k είναι γραμμικώς ανεξάρτητα, ο B_{k+1} έχει ακριβώς δύο διακριτές ιδιοτιμές. Συγκρίνοντας τις ανισότητες (3.14), (3.15) και (3.16) συμπεραίνουμε ότι πάντα η μικρότερη ιδιοτιμή του B_{k+1} είναι διακριτή. $\square$

Προφανώς, για να βρούμε τις δύο διακριτές ιδιοτιμές, αρκεί να λύσουμε την δευτεροβάθμια εξίσωση $\lambda^2 - \beta_1 \lambda + \beta_2 = 0$, ως προς λ . Άρα,

$$\lambda_{1,2} = \left(\beta_1 \pm \sqrt{\beta_1^2 - 4\beta_2} \right) / 2.$$

Ας εξετάσουμε λίγο την περίπτωση που τα s_k και y_k είναι συγγραμμικά. Έστω ότι η σχέση που τα συνδέει είναι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$. Τότε θα ισχύει ότι $w_k = \rho s_k$, $\rho \in \mathbb{R}^*$, όπου

$$\rho = \begin{cases} \kappa - \theta_{k+1}, & \text{για την SR1 ενημέρωση ελάχιστης μνήμης;} \\ \sqrt{\kappa \theta_{k+1}}, & \text{για τη BFGS ενημέρωση ελάχιστης μνήμης;} \\ \kappa, & \text{για την DFP ενημέρωση ελάχιστης μνήμης;} \\ 1, & \text{για την PSB ενημέρωση ελάχιστης μνήμης.} \end{cases}$$

Επομένως, ο Quasi-Newton πίνακας ελάχιστης μνήμης B_{k+1} που δίνεται από τη σχέση (3.3) γίνεται

$$B_{k+1} = \theta_{k+1} I + (\kappa - \theta_{k+1}) \frac{s_k s_k^T}{s_k^T s_k}. \quad (3.21)$$

Όπως μπορούμε να δούμε από την παραπάνω σχέση, οι ιδιοτιμές του B_{k+1} είναι οι κ και θ_{k+1} , με την τελευταία να έχει πολλαπλότητα $n - 1$. Άρα, το χαρακτηριστικό πολυώνυμο του B_{k+1} είναι της μορφής

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-1} (\lambda - \kappa).$$

Προφανώς, σε αυτή την περίπτωση, αν ισχύει ότι $\kappa \neq \theta_{k+1}$, πίνακας έχει

μία διακριτή ιδιοτιμή. Αν μάλιστα ισχύει η ανισότητα

$$\kappa < \theta_{k+1},$$

τότε η πιο μικρή ιδιοτιμή είναι και διακριτή.

3.3 Ο αντίστροφος του πίνακα $B + \mu I$

Όπως είδαμε στην Ενότητα 2.8, όπου και περιγράψαμε το υποπρόβλημα περιοχής εμπιστοσύνης, είναι αναγκαίο να λύσουμε ένα γραμμικό σύστημα της μορφής

$$(B + \mu I)u = v,$$

όπου ο B είναι συμμετρικός $n \times n$ πίνακας, u, v n -διάστατα διανύσματα και $\mu \in \mathbb{R}$. Σκοπός μας είναι να υπολογίσουμε αναλυτικά τη λύση του ανωτέρω γραμμικού συστήματος ώστε να αποφύγουμε την αποθήκευση ή την παραγοντοποίηση πίνακα. Στη συνέχεια θεωρούμε ότι ο B είναι ένας Quasi-Newton πίνακας ελάχιστης μνήμης. Για να μπορέσουμε να εκφράσουμε τον πίνακα $(B + \mu I)^{-1}$ σε κλειστή μορφή, θα βασιστούμε στο χαρακτηριστικό του πολυώνυμο.

Λήμμα 3.1. Έστω ότι μία ενημέρωση εκτελείται στον $B^{(0)} = \theta I$ χρησιμοποιώντας το διανυσματικό ζεύγος (s_k, y_k) και τη γενική PSB ενημέρωση. Τότε, το χαρακτηριστικό πολυώνυμο του πίνακα $B_{k+1} + \mu I$, $\mu \in \mathbb{R}$, είναι το

$$q(x; \mu) = [x - (\theta_{k+1} + \mu)]^{n-2} [x^2 - \gamma'(\mu)x + \gamma(\mu)], \quad (3.22)$$

όπου $\gamma(\mu) = \mu^2 + \beta_1\mu + \beta_2$.

Απόδειξη. Από το Θεώρημα 3.1, γνωρίζουμε ότι ο B_{k+1} έχει ως ιδιοτιμές τη θ_{k+1} , πολλαπλότητας $n - 2$ και άλλες δύο, έστω λ_x και λ_y , οι οποίες συνδέονται με τις σχέσεις

$$\lambda_x + \lambda_y = \beta_1 \quad \text{και} \quad \lambda_x \lambda_y = \beta_2.$$

Προσθέτοντας τον πίνακα λI στον B_{k+1} , οι ιδιοτιμές του προκύπτοντος πίνακα $B_{k+1} + \mu I$ είναι οι $x_i = \lambda_i + \mu$. Επομένως, το χαρακτηριστικό

πολυώνυμο του $B_{k+1} + \mu I$ είναι το

$$q(x; \mu) = [x - (\theta_{k+1} + \mu)]^{n-2} [x^2 - (\lambda_x + \lambda_y + 2\mu)x + (\lambda_x + \mu)(\lambda_y + \mu)].$$

Επειδή $\lambda_x + \lambda_y = \beta_1$ και $\lambda_x \lambda_y = \beta_2$, έχουμε ότι $(\lambda_x + \mu)(\lambda_y + \mu) = \mu^2 + \beta_1 \mu + \beta_2$. Όμως, ισχύει ότι

$$\frac{d}{d\mu}(\lambda_x + \mu)(\lambda_y + \mu) = \lambda_x + \lambda_y + 2\mu,$$

οπότε θέτοντας $\gamma(\mu) = \mu^2 + \beta_1 \mu + \beta_2$, η σχέση (3.22) προκύπτει άμεσα. $\square$

Πρόταση 3.1. Έστω Λ το σύνολο των ιδιοτιμών του *Quasi-Newton* πίνακα ελάχιστης μνήμης B_{k+1} . Τότε, για κάθε $\mu \in \mathbb{R} \setminus \Lambda$, ο αντίστροφος του πίνακα $(B_{k+1} + \mu I)$ είναι ο πίνακας

$$(B_{k+1} + \mu I)^{-1} = \frac{1}{(\mu + \theta_{k+1})\gamma(\mu)} \left\{ B_{k+1}^2 - [\beta'(\mu) - \mu] B_{k+1} + \beta(\mu) I \right\}, \quad (3.23)$$

όπου $\beta(\mu) = \gamma(\mu) + [\gamma'(\mu) - \mu] \theta_{k+1}$.

Απόδειξη. Από το Λήμμα 3.1, προκύπτει ότι το ελάχιστο πολυώνυμο του $B_{k+1} + \mu I$ είναι το

$$q_m(x; \mu) = [x - (\theta_{k+1} + \mu)] [x^2 - \gamma'(\mu)x + \gamma(\mu)].$$

Εφαρμόζοντας το Θεώρημα Caley-Hamilton (σελ. 13) έχουμε ότι

$$q_m(B_{k+1} + \mu I; \lambda) = 0,$$

από όπου έπεται ότι

$$[B_{k+1} + \mu I - (\theta_{k+1} + \mu)I] [(B_{k+1} + \mu I)^2 - \gamma'(\mu)(B_{k+1} + \mu I) + \gamma(\mu)I] = 0.$$

Πολλαπλασιάζοντας αμφότερα τα μέλη της παραπάνω σχέσης με τον πί-

νακα $(B_{k+1} + \mu I)^{-1}$ και λύνοντας ως προς $(B_{k+1} + \mu I)^{-1}$, λαμβάνουμε

$$\begin{aligned} (B_{k+1} + \mu I)^{-1} &= \frac{1}{\gamma(\mu)(\theta_{k+1} + \mu)} \left\{ (B_{k+1} + \mu I)^2 - [\gamma'(\mu) + \theta_{k+1} + \mu] \right. \\ &\quad \left. (B_{k+1} + \mu I) + [\gamma(\mu) + (\theta_{k+1} + \mu)\gamma'(\mu)] I \right\} \\ &= \frac{1}{\gamma(\mu)(\theta_{k+1} + \mu)} \left\{ B_{k+1}^2 - [\gamma'(\mu) + \theta_{k+1} - \mu] B_{k+1} \right. \\ &\quad \left. + [\gamma(\mu) + (\gamma'(\mu) - \mu)\theta_{k+1}] I \right\}. \end{aligned}$$

Θέτοντας ως $\beta(\mu) = \gamma(\mu) + (\gamma'(\mu) - \mu)\theta_{k+1}$ και έχοντας υπόψιν ότι $\beta'(\mu) = \gamma'(\mu) + (\gamma''(\mu) - 1)\theta_{k+1}$, με $\gamma''(\mu) = 2$, καταλήγουμε στη σχέση (3.23). $\square$

Λήμμα 3.2. Έστω ότι τα διανύσματα s_k, y_k είναι συγγραμμικά, με $y_k = \kappa s_k, \kappa \in \mathbb{R}^*$. Τότε ο αντιστροφος του $B_{k+1} + \mu I, \mu \in \mathbb{R} \setminus \Lambda$, είναι ο πίνακας

$$(B_{k+1} + \mu I)^{-1} = -\frac{1}{\gamma(\mu)} \{ B_{k+1} - [\gamma'(\mu) - \mu] I \}. \quad (3.24)$$

Απόδειξη. Στην περίπτωση όπου τα διανύσματα s_k, y_k είναι συγγραμμικά, με $y_k = \kappa s_k, \kappa \in \mathbb{R}^*$, ο πίνακας έχει το πολύ μία διακριτή ιδιοτιμή, την $\kappa + \mu$, και άρα το ελάχιστο πολυώνυμο του $B_{k+1} + \mu I$ είναι το

$$q_m(x; \mu) = [x^2 - \gamma'(\mu)x + \gamma(\mu)] = [x - (\theta_{k+1} + \mu)][x - (\kappa + \mu)].$$

Χρησιμοποιώντας το Θεώρημα Caley-Hamilton (σελ. 13) έχουμε ότι

$$(B_{k+1} + \mu I)^2 - \gamma'(\mu)(B_{k+1} + \mu I) + \gamma(\mu)I = 0. \quad (3.25)$$

Πολλαπλασιάζοντας αμφότερα τα μέλη της παραπάνω εξίσωσης με τον πίνακα $(B_{k+1} + \mu I)^{-1}$ έχουμε ότι

$$(B_{k+1} + \mu I) - \gamma'(\mu)I + \gamma(\mu)(B_{k+1} + \mu I)^{-1} = 0,$$

άρα,

$$(B_{k+1} + \mu I)^{-1} = -\frac{1}{\gamma(\mu)} \{ B_{k+1} - [\gamma'(\mu) - \mu] I \}.$$

$\square$

Παρατήρηση 3.1. Η έκφραση (3.23) της Πρότασης 3.1, δίνει τη γενική μορφή του $(B_{k+1} + \mu I)^{-1}$, ανεξάρτητα από τη σχέση που συνδέει τα διανύσματα s_k και y_k . Πράγματι, από τη σχέση (3.25) βρίσκουμε ότι

$$B_{k+1}^2 = [\gamma'(\mu) - 2\mu] B_{k+1} - [\mu^2 - \gamma'(\mu)\mu + \gamma(\mu)] I.$$

Αντικαθιστώντας την παραπάνω ισότητα στο δεξί μέλος της σχέσης (3.23), μετά από μερικές αλγεβρικές πράξεις, υπολογίζουμε τη σχέση (3.24).

3.4 Τα ιδιοδιανύσματα των διακριτών ιδιοτιμών

Σε αυτή την ενότητα μελετάμε την εύρεση των ιδιοδιανυσμάτων που αντιστοιχούν στις διακριτές ιδιοτιμές των Quasi-Newton πινάκων ελάχιστης μνήμης. Στην Ενότητα 3.2 είδαμε ότι ο αριθμός των διακριτών ιδιοτιμών εξαρτάται από τη γραμμική ανεξαρτησία ή μη των διανυσμάτων s_k και y_k . Γι αυτό το λόγο θα εξετάσουμε αυτές τις δύο περιπτώσεις ξεχωριστά.

3.4.1 Γραμμικώς εξαρτημένα διανύσματα ενημέρωσης

Ας εξετάσουμε πρώτα την περίπτωση που τα s_k και y_k είναι συγγραμμικά. Έστω ότι η σχέση που τα συνδέει είναι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$. Τότε θα ισχύει ότι $w_k = \rho s_k$, $\rho \in \mathbb{R}^*$, όπως είδαμε στη σελ. 91. Σε αυτή την περίπτωση, ο Quasi-Newton πίνακας ελάχιστης μνήμης B_{k+1} ορίζεται σύμφωνα με τη σχέση (3.21) και, όπως εύκολα μπορούμε να αποδείξουμε, το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή κ είναι το s_k . Πράγματι, από τη σχέση (3.21) έπεται ότι

$$B_{k+1}s_k = \theta_{k+1}s_k + (\kappa - \theta_{k+1}) \frac{s_k^T s_k}{s_k^T s_k} s_k = \kappa s_k.$$

Προφανώς, ένα γενικευμένο ιδιοδιάνυσμα που θα αντιστοιχεί στην πολλαπλή ιδιοτιμή θ_{k+1} , θα πρέπει να είναι κάθετο στο s_k , αφού αυτό είναι το μοναδικό διάνυσμα το οποίο ενημερώνει τον αρχικό πίνακα $B^{(0)}$, όπως βλέπουμε από την (3.21). Επιπλέον θα πρέπει να ισχύει ότι $B_{k+1}u = \theta_{k+1}u$. Ένα τέτοιο διάνυσμα u είναι της μορφής

$$u = (-s_k^n/s_k^1, 0, \dots, 0, \dots, 1)^T,$$

όπου s_k^i παριστά την i συνιστώσα του διανύσματος s_k . Πράγματι, έχουμε ότι

$$u^T s_k = -\frac{s_k^n}{s_k^1} s_k^{(1)} + 0 + \dots + 0 + s_k^n = 0$$

και

$$B_{k+1}u = \theta_{k+1}u + (\kappa - \theta_{k+1}) \frac{s_k^T u}{s_k^T s_k} s_k = \theta_{k+1}u,$$

αφού $u^T s_k = 0$. Βέβαια, αν τύχει το s_k^1 να είναι ίσο με το μηδέν, τότε προφανώς θα αναζητήσουμε ένα διάνυσμα της μορφής

$$u = (0, \dots, 0, -\frac{s_k^j}{s_k^i}, 0, \dots, 0, 1, 0, \dots, 0)^T,$$

όπου η μονάδα θα βρίσκεται στην j θέση, έτσι ώστε να ισχύει ότι $u^T s_k = 0$. Τέλος, αν ισχύει ότι $\kappa = \theta_{k+1}$, τότε ο πίνακας είναι της μορφής

$$B_{k+1} = \theta_{k+1}I.$$

Επομένως τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή θ_{k+1} είναι οι στήλες του μοναδιαίου πίνακα,

$$e_i = (0, 0, \dots, 1, 0, \dots, 0)^T, \quad i = 1, 2, \dots, n,$$

όπου το 1 εμφανίζεται στην i συνιστώσα του διανύσματος e_i .

3.4.2 Γραμμικώς ανεξάρτητα διανύσματα ενημέρωσης

Στην περίπτωση όπου τα s_k και y_k είναι γραμμικώς ανεξάρτητα, για να υπολογίσουμε τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές, αρκεί να παρατηρήσουμε ότι ο αρχικός πίνακας $B_{k+1}^{(0)}$ ενημερώνεται από τα διανύσματα s_k και y_k .

Πρόταση 3.2. Έστω ότι εκτελείται μία ενημέρωση στο συμμετρικό πίνακα $B^{(0)} = \theta I$, χρησιμοποιώντας το διανυσματικό ζεύγος (s_k, y_k) , και τη γενική PSB ενημέρωσης ελάχιστης μνήμης. Τότε τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές του ενημερωμένου πίνακα B_{k+1} , είναι της μορφής

$$u(\lambda) = \frac{(\lambda s_k - y_k)^T y_k}{(\lambda s_k - y_k)^T s_k} s_k - y_k, \quad (3.26)$$

όπου λ διακριτή ιδιοτιμή του B_{k+1} .

Απόδειξη. Επειδή ο $B^{(0)}$ ενημερώνεται χρησιμοποιώντας το διανυσματικό ζεύγος (s_k, y_k) , τα ιδιοδιανύσματα των διακριτών ιδιοτιμών θα πρέπει να είναι της μορφής

$$u = c_1 s_k + c_2 y_k, \quad c_1, c_2 \in \mathbb{R}.$$

Επομένως, έχουμε ότι

$$B_{k+1}u = \lambda u \Rightarrow B_{k+1}(c_1 s_k + c_2 y_k) = \lambda(c_1 s_k + c_2 y_k). \quad (3.27)$$

Από την εξίσωση της τέμνουσας (σελ. 35)

$$B_{k+1}s_k = y_k, \quad (3.28)$$

η σχέση (3.27) γίνεται

$$c_1 y_k + c_2 B_{k+1}y_k = \lambda c_1 s_k + \lambda c_2 y_k.$$

Πολλαπλασιάζοντας την παραπάνω εξίσωση με s_k^T και λαμβάνοντας υπόψη τη σχέση (3.28), υπολογίζουμε το c_1 συναρτήσει του c_2 ως εξής:

$$\begin{aligned} c_1 s_k^T y_k + c_2 y_k^T y_k &= \lambda c_1 s_k^T s_k = \lambda c_2 s_k^T y_k \\ c_1 (s_k^T y_k - \lambda s_k^T s_k) &= c_2 (\lambda s_k^T y_k - y_k^T y_k) \\ c_1 &= -c_2 \frac{\lambda s_k^T y_k - y_k^T y_k}{\lambda s_k^T s_k - s_k^T y_k}. \end{aligned}$$

Θέτοντας $c_2 = -1$ και αντικαθιστώντας στην εξίσωση $u = c_1 s_k + c_2 y_k$, βρίσκουμε τη σχέση (3.26). $\square$

Ένας άλλος τρόπος υπολογισμού των ιδιοδιανυσμάτων που αντιστοιχούν σε διακριτές ιδιοτιμές, είναι μέσω της μεθόδου της αντίστροφης δύναμης. Όπως είδαμε στο Κεφάλαιο 1 (σελ. 14), η μέθοδος αυτή χρησιμοποιεί το επαναληπτικό σχήμα

$$u_i = (B - \sigma I)^{-1} \frac{u_{i-1}}{\|u_{i-1}\|},$$

με $\sigma = \lambda \pm \epsilon$, όπου $\epsilon \rightarrow 0$. Ως γνωστόν, αν η ιδιοτιμή λ είναι εκ των

προτέρων γνωστή, κάτι το οποίο ισχύει στην περίπτωση μας, το σχήμα αυτό συγκλίνει σε μία και μόνο επανάληψη. Κατά συνέπεια, το ανωτέρω επαναληπτικό σχήμα μας παρέχει ένα κλειστό τύπο για τον υπολογισμό του ιδιοδιανύσματος.

Χρησιμοποιώντας την Πρόταση 3.1 στη μέθοδο της αντίστροφης δύναμης, μπορούμε να υπολογίσουμε το ιδιοδιάνυσμα που αντιστοιχεί σε μια διακριτή ιδιοτιμή λ του πίνακα B_{k+1} , το οποίο είναι το $u(\sigma) = (B_{k+1} - \sigma I)^{-1} u$, όπου $\sigma = \lambda - \epsilon$. Αν θέσουμε όπου $\hat{\lambda} = -\sigma$, τότε το προκύπτον ιδιοδιάνυσμα είναι το

$$\begin{aligned} u(\hat{\lambda}) &= (B_{k+1} + \hat{\lambda} I)^{-1} u \\ &= \frac{B_{k+1}^2 u - (\beta'(\hat{\lambda}) - \hat{\lambda}) B_{k+1} u + \beta(\hat{\lambda}) u}{(\hat{\lambda} + \theta_{k+1}) \gamma(\hat{\lambda})}, \end{aligned} \quad (3.29)$$

όπου $\hat{\lambda} = -\lambda + \epsilon$, είναι μια διαταραγμένη διακριτή ιδιοτιμή του B_{k+1} με αντίθετο πρόσημο και u είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$.

Τα διανύσματα $B_{k+1} u$ και $B_{k+1}^2 u$ μπορούν να υπολογιστούν από το επαναληπτικό σχήμα

$$\begin{aligned} B_{k+1} v_i &= \theta_{k+1} v_i + \frac{w_k^T v_i}{s_k^T w_k} (y_k - \theta_{k+1} s_k) \\ &+ \left(\frac{(y_k - \theta_{k+1} s_k)^T v_i}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k^T v_i}{s_k^T w_k} \right) w_k, \end{aligned}$$

με $i = 0, 1$ και $v_0 = u$. Μετά από μερικούς αλγεβρικούς υπολογισμούς, καταλήγουμε στην έκφραση

$$u(\hat{\lambda}) = \frac{1}{\gamma(\hat{\lambda})(\hat{\lambda} + \theta_{k+1})} \left(c_1(\hat{\lambda}) u + c_2(\hat{\lambda}) z_k + c_3(\hat{\lambda}) w_k \right), \quad (3.30)$$

όπου $z_k = y_k - \theta_{k+1} s_k$, ενώ οι συντελεστές των διανυσμάτων u , z_k και w_k ,

ορίζονται από τις σχέσεις

$$c_1(\hat{\lambda}) = \gamma(\hat{\lambda}) \quad (3.31)$$

$$c_2(\hat{\lambda}) = - \left(\hat{\lambda} + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + \frac{w_k^T w_k}{s_k^T w_k} \frac{z_k^T u}{s_k^T w_k} \quad (3.32)$$

$$c_3(\hat{\lambda}) = \left[\frac{z_k^T z_k}{s_k^T w_k} + \frac{z_k^T s_k}{s_k^T w_k} \left(\hat{\lambda} + \theta_{k+1} \right) \right] \frac{w_k^T u}{s_k^T w_k} - \left(\hat{\lambda} + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{z_k^T u}{s_k^T w_k}. \quad (3.33)$$

Η σχέση (3.30), είναι ο κλειστός τύπος ο οποίος εκφράζει το ιδιοδιάνυσμα που αντιστοιχεί σε μια διακριτή ιδιοτιμή του B_{k+1} .

Ας δούμε πιο αναλυτικά τον όλο υπολογισμό: αν θέσουμε $z_k = y_k - \theta_{k+1} s_k$ στον πίνακα (3.3), αυτός γράφεται ως εξής:

$$B_{k+1} = \theta_{k+1} I + \frac{z_k w_k^T + w_k z_k^T}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{s_k^T w_k}. \quad (3.34)$$

Το διάνυσμα $B_{k+1} u$ είναι το

$$\begin{aligned} B_{k+1} u &= \theta_{k+1} u + \frac{w_k^T u}{s_k^T w_k} z_k + \left(\frac{z_k^T u}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T u}{s_k^T w_k} \right) w_k \\ &= \theta_{k+1} u + \frac{w_k^T u}{s_k^T w_k} z_k + a w_k, \end{aligned} \quad (3.35)$$

όπου η σταθερά a είναι η ποσότητα

$$a = \frac{z_k^T u}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T u}{s_k^T w_k}. \quad (3.36)$$

Επομένως, επειδή $B_{k+1}^2 u = B_{k+1} B_{k+1} u$, έχουμε ότι

$$\begin{aligned} B_{k+1}^2 u &= \theta_{k+1} B_{k+1} u + \frac{w_k^T B_{k+1} u}{s_k^T w_k} z_k + \left(\frac{z_k^T B_{k+1} u}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T B_{k+1} u}{s_k^T w_k} \right) w_k \\ &= \theta_{k+1} B_{k+1} u + \frac{w_k^T B_{k+1} u}{s_k^T w_k} z_k + b w_k, \end{aligned} \quad (3.37)$$

όπου ως b έχουμε θέσει την έκφραση

$$\begin{aligned}
 b &= \frac{z_k^T B_{k+1} u}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T B_{k+1} u}{s_k^T w_k} \\
 &= \theta_{k+1} \frac{z_k^T u}{s_k^T w_k} + \left(\frac{z_k^T z_k}{s_k^T w_k} - \theta_{k+1} \frac{z_k^T s_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} \\
 &\quad + \left(\frac{z_k^T w_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \right) a. \tag{3.38}
 \end{aligned}$$

Αντικαθιστώντας τις εκφράσεις (3.35) και (3.37) στη σχέση (3.29), έχουμε ότι

$$\begin{aligned}
 u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ B_{k+1}^2 u - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \theta_{k+1} B_{k+1} u + \frac{w_k^T B_{k+1} u}{s_k^T w_k} z_k + b w_k \right. \\
 &\quad \left. - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ [\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda})] B_{k+1} u \right. \\
 &\quad \left. + \frac{w_k^T B_{k+1} u}{s_k^T w_k} z_k + b w_k + \beta(\hat{\lambda}) u \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ [\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda})] \left(\theta_{k+1} u + \frac{w_k^T u}{s_k^T w_k} z_k + a w_k \right) \right. \\
 &\quad \left. + \frac{w_k^T}{s_k^T w_k} \left(\theta_{k+1} u + \frac{w_k^T u}{s_k^T w_k} z_k + a w_k \right) z_k + b w_k + \beta(\hat{\lambda}) u \right\}
 \end{aligned}$$

Αντικαθιστώντας τις σχέσεις (3.36) και (3.38) στην παραπάνω εξίσωση, κάνοντας κάποιες πράξεις και ομαδοποιώντας ως προς τα διανύσματα u ,

z_k και w_k , καταλήγουμε στο εξής:

$$\begin{aligned}
 u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \left[\left(\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda}) \right) \theta_{k+1} + \beta(\hat{\lambda}) \right] u \right. \\
 &\quad + \left[\left(\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda}) \right) \frac{w_k^T u}{s_k^T w_k} \right. \\
 &\quad + \left. \frac{w^T}{s_k^T w_k} \left(\theta_{k+1} u + \frac{w_k^T u}{s_k^T w_k} z_k + a w_k \right) \right] z_k \\
 &\quad + \left. \left[\left(\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda}) \right) a + b \right] w_k \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ c_1(\hat{\lambda}) u + c_2(\hat{\lambda}) z_k + c_3(\hat{\lambda}) w_k \right\}
 \end{aligned}$$

Για να υπολογίσουμε το $c_1(\hat{\lambda})$, λαμβάνουμε υπόψιν τις σχέσεις $\beta(\lambda) = \gamma(\lambda) + [\gamma'(\lambda) - \lambda]\theta_{k+1}$ και $\gamma(\lambda) = \lambda^2 + \beta_1\lambda + \beta_2$, οπότε βρίσκουμε ότι

$$\begin{aligned}
 c_1(\hat{\lambda}) &= \left[\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda}) \right] \theta_{k+1} + \beta(\hat{\lambda}) \\
 &= \left[\theta_{k+1} + \hat{\lambda} - \gamma'(\hat{\lambda}) - \theta_{k+1} \right] \theta_{k+1} + \gamma(\hat{\lambda}) + \left[\gamma'(\hat{\lambda}) - \hat{\lambda} \right] \theta_{k+1} \\
 &= \gamma(\hat{\lambda}),
 \end{aligned}$$

η οποία είναι η έκφραση (3.31). Για να υπολογίσουμε το $c_2(\hat{\lambda})$, λαμβάνουμε επιπλέον υπόψιν τη σχέση (3.35). Έτσι, έχουμε ότι

$$\begin{aligned}
 c_2(\hat{\lambda}) &= \left[\hat{\lambda} - \gamma'(\hat{\lambda}) \right] \frac{w_k^T u}{s_k^T w_k} + \theta_{k+1} \frac{w_k^T u}{s_k^T w_k} + \frac{w_k^T u}{s_k^T w_k} \frac{z_k^T w_k}{s_k^T w_k} + a \frac{w_k^T w_k}{s_k^T w_k} \\
 &= \left[\hat{\lambda} - \gamma'(\hat{\lambda}) + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right] \frac{w_k^T u}{s_k^T w_k} + a \frac{w_k^T w_k}{s_k^T w_k} \\
 &= \left(-\hat{\lambda} - \beta_1 + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + a \frac{w_k^T w_k}{s_k^T w_k}.
 \end{aligned}$$

Όμως, από τη σχέση (3.12) και την $z_k = y_k - \theta_{k+1}s_k$, έχουμε ότι

$$\begin{aligned}
 \beta_1 &= 2 \frac{y_k^T w_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \\
 &= 2 \frac{z_k^T w_k}{s_k^T w_k} + 2\theta_{k+1} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k}
 \end{aligned}$$

οπότε,

$$\begin{aligned} c_2(\hat{\lambda}) &= \left(-\hat{\lambda} - \beta_1 + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + a \frac{w_k^T w_k}{s_k^T w_k} \\ &= \left(-\hat{\lambda} - \theta_{k+1} - \frac{z_k^T w_k}{s_k^T w_k} + \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + a \frac{w_k^T w_k}{s_k^T w_k}. \end{aligned}$$

Αντικαθιστώντας στην τελευταία εξίσωση τη σχέση (3.36), καταλήγουμε στην έκφραση

$$\begin{aligned} c_2(\hat{\lambda}) &= \left(-\hat{\lambda} - \theta_{k+1} - \frac{z_k^T w_k}{s_k^T w_k} + \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} \\ &\quad + \left(\frac{z_k^T u}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T u}{s_k^T w_k} \right) \frac{w_k^T w_k}{s_k^T w_k} \\ &= - \left(\hat{\lambda} + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + \frac{z_k^T u}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k}, \end{aligned}$$

η οποία είναι η σχέση (3.32). Τέλος, για το συντελεστή $c_3(\hat{\lambda})$ του διανύσματος w_k , δουλεύοντας με παρόμοιο τρόπο, καταλήγουμε στη σχέση (3.33). Συγκεκριμένα, έχουμε ότι

$$c_3(\hat{\lambda}) = \left(\theta_{k+1} + \hat{\lambda} - \beta'(\hat{\lambda}) \right) a + b,$$

οπότε, χρησιμοποιώντας τη σχέσεις (3.36) και (3.38), βρίσκουμε ότι

$$\begin{aligned} c_3(\hat{\lambda}) &= \left(\hat{\lambda} - \gamma'(\hat{\lambda}) + \frac{z_k^T w_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \right) a \\ &\quad + \theta_{k+1} \frac{z_k^T u}{s_k^T w_k} + \left(\frac{z_k^T z_k}{s_k^T w_k} - \theta_{k+1} \frac{z_k^T s_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} \end{aligned}$$

Λαμβάνοντας πάλι υπόψιν τις σχέσεις $z_k = y_k - \theta_{k+1} s_k$, $\beta(\lambda) = \gamma(\lambda) +$

$[\gamma'(\lambda) - \lambda]\theta_{k+1}$ και $\gamma(\lambda) = \lambda^2 + \beta_1\lambda + \beta_2$, βρίσκουμε ότι

$$\begin{aligned} c_3(\hat{\lambda}) &= \left(\hat{\lambda} - \gamma'(\hat{\lambda}) - \frac{z_k^T w_k}{s_k^T w_k} + \beta_1 - 2\theta_{k+1} \right) a + \theta_{k+1} \frac{z_k^T u}{s_k^T w_k} \\ &\quad + \left(\frac{z_k^T z_k}{s_k^T w_k} - \theta_{k+1} \frac{z_k^T s_k}{s_k^T w_k} - \frac{z_k^T s_k}{s_k^T w_k} \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} \\ &= - \left(\hat{\lambda} + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{z_k^T u}{s_k^T w_k} \\ &\quad + \left[\frac{z_k^T z_k}{s_k^T w_k} + \frac{z_k^T s_k}{s_k^T w_k} \left(\hat{\lambda} + \theta_{k+1} \right) \right] \frac{w_k^T u}{s_k^T w_k}, \end{aligned}$$

δηλαδή τη σχέση (3.33). Επομένως, για $\sigma = -\hat{\lambda}$, καταλήγουμε στο παρακάτω πόρισμα :

Πόρισμα 3.1. Έστω $\sigma = \lambda - \epsilon$, $\epsilon \rightarrow 0^+$, μια ϵ -διαταραγμένη διακριτή ιδιοτιμή του B_{k+1} και $u_0 = (1, 1, \dots, 1)^T$. Τότε, το ιδιοδιάνυσμα που αντιστοιχεί στη σ είναι το κανονικοποιημένο διάνυσμα

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)z_k + c_3(\sigma)w_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)s_k^T w_k}, \quad (3.39)$$

όπου οι συντελεστές των διανυσμάτων u_0 , z_k και w_k ορίζονται ως ακολούθως:

$$\begin{aligned} c_1(\sigma) &= s_k^T w_k (\sigma^2 - \beta_1\sigma + \beta_2), \\ c_2(\sigma) &= - \left(\theta_{k+1} - \sigma + \frac{z_k^T w_k}{s_k^T w_k} \right) \sum_{i=1}^n w_k^i + \frac{w_k^T w_k}{s_k^T w_k} \sum_{i=1}^n z_k^i, \quad \text{και} \\ c_3(\sigma) &= \left[\frac{z_k^T z_k}{s_k^T w_k} + (\theta_{k+1} - \sigma) \frac{z_k^T s_k}{s_k^T w_k} \right] \sum_{i=1}^n w_k^i - \left(\theta_{k+1} - \sigma + \frac{z_k^T w_k}{s_k^T w_k} \right) \sum_{i=1}^n z_k^i. \end{aligned}$$

Όπως μπορούμε εύκολα να παρατηρήσουμε, για να υπολογίσουμε το παραπάνω ιδιοδιάνυσμα χρειάζεται να προσθέσουμε δύο διανύσματα, τα s_k και y_k , μιας και τα z_k , w_k είναι γραμμικός συνδυασμός αυτών των δύο διανυσμάτων. Επίσης τα εσωτερικά γινόμενα που απαιτούνται είναι εννιά, τα εξής:

Πίνακας 3.1: Απαιτούμενα εσωτερικά γινόμενα

$s_k^T s_k$	$y_k^T w_k$	$s_k^T y_k$	$w_k^T w_k$	$s_k^T w_k$
$z_k^T w_k$	$s_k^T z_k$	$z_k^T z_k$	$y_k^T y_k$	

Προφανώς, επειδή το διάνυσμα w_k είναι γραμμικός συνδυασμός των s_k και y_k , της μορφής $w_k = c_1 s_k + c_2 y_k$, ο υπολογισμός του ιδιοδιανύσματος του ιδιοδιανύσματος $u(\sigma)$ στη γενική PSB ενημέρωση ελάχιστης μνήμης, απαιτεί περισσότερους υπολογισμούς από ότι σε καθεμιά από τις τέσσερις Quasi-Newton ενημερώσεις ελάχιστης μνήμης.

Να σημειώσουμε εδώ ότι η γενική PSB ενημέρωση εκφράζει την κλάση των τεσσάρων Quasi-Newton ενημερώσεων και αποτελεί το θεωρητικό υπόβαθρο για τη μελέτη καθεμιάς ξεχωριστά. Αντικαθιστώντας το w_k με τις αντίστοιχες εκφράσεις που ανάγουν τη γενική PSB ενημέρωση ελάχιστης μνήμης σε καθεμιά από τις τέσσερις Quasi-Newton ενημερώσεις ελάχιστης μνήμης, βρίσκουμε τις αντίστοιχες εκφράσεις των τελευταίων.

Στην πράξη χρησιμοποιείται καθεμιά από τις τέσσερις Quasi-Newton ενημερώσεις ξεχωριστά. Για το λόγο αυτό, θα εξετάσουμε αναλυτικά τις ιδιότητες των SR1, BFGS, PSB και DFP ενημερώσεων ελάχιστης μνήμης. Θα βασιστούμε στις ιδιότητες του ίχνους και της ορίζουσας κάθε πίνακα, καθώς και στα θεωρητικά αποτελέσματα της παραπάνω ανάλυσης.

3.5 Η SR1 ενημέρωση ελάχιστης μνήμης

Η SR1 ενημέρωση ελάχιστης μνήμης υπολογίζεται αν στη γενική PSB ενημέρωση ελάχιστης μνήμης (3.3), θέσουμε $w_k = y_k - \theta_{k+1} s_k$. Ο προκύπτων πίνακας είναι της μορφής

$$B_{k+1} = \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k)(y_k - \theta_{k+1} s_k)^T}{(y_k - \theta_{k+1} s_k)^T s_k}. \quad (3.40)$$

Όσον αφορά το ίχνος και την ορίζουσα του πίνακα αυτού, χρησιμοποιώντας τις σχέσεις (3.4) και (3.10) (σελ. 84 και 86), αντίστοιχα, βρίσκουμε ότι

$$\text{tr}(B_{k+1}) = n\theta_{k+1} + \frac{(y_k - \theta_{k+1} s_k)^T (y_k - \theta_{k+1} s_k)}{(y_k - \theta_{k+1} s_k)^T s_k},$$

και

$$\det(B_{k+1}) = \theta_{k+1}^{n-1} \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k},$$

αντίστοιχα. Από το Θεώρημα 3.1, γνωρίζουμε ότι το χαρακτηριστικό πολυώνυμο της γενικής PSB ενημέρωσης ελάχιστης μνήμης, είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} (\lambda^2 - \beta_1 \lambda + \beta_2),$$

όπου

$$\beta_1 = \lambda_x + \lambda_y \quad \text{και} \quad \beta_2 = \lambda_x \lambda_y,$$

είναι το άθροισμα και το γινόμενο των δύο άγνωστων ιδιοτιμών λ_x και λ_y , αντίστοιχα. Γνωρίζουμε επίσης ότι η SR1 ενημέρωση ελάχιστης μνήμης έχει το πολύ μια διακριτή ιδιοτιμή. Έστω λοιπόν ότι $\lambda_x = \lambda$ και $\lambda_y = \theta_{k+1}$. Άρα πολύ εύκολα μπορούμε να συμπεράνουμε ότι αφού

$$\text{tr}(B_{k+1}) = (n-1)\theta_{k+1} + \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} = (n-1)\theta_{k+1} + \lambda,$$

και

$$\det(B_{k+1}) = \theta_{k+1}^{n-1} \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} = \theta_{k+1}^{n-1} \lambda,$$

τότε

$$\lambda = \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k}.$$

Οπότε,

$$\beta_1 = \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} + \theta_{k+1},$$

$$\beta_2 = \frac{\theta_{k+1} (y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k},$$

με αποτέλεσμα το χαρακτηριστικό πολυώνυμο της SR1 ενημέρωσης ελάχιστης μνήμης (3.40) να εκφράζεται ως εξής:

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-1} \left(\lambda - \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} \right). \quad (3.41)$$

Η έκφραση του αντίστροφου του πίνακα $B_{k+1} + \mu I$, $\mu \in \mathbb{R}^n$, ορίζε-

ται στην Πρόταση 3.1. Επειδή όμως η SR1 είναι μια ενημέρωση πρώτης τάξης, περιμένουμε ότι η εξίσωση (3.23) της Πρότασης 3.1, θα απλουστεύεται. Πράγματι, από το Λήμμα 3.1, γνωρίζουμε ότι το χαρακτηριστικό πολυώνυμο του $B_{k+1} + \mu I$ ορίζεται στην εξίσωση (3.5), δηλαδή την εξίσωση

$$q(x; \mu) = [x - (\theta_{k+1} + \mu)]^{n-2} [x^2 - \gamma'(\mu)x + \gamma(\mu)].$$

Όμως, επειδή

$$\gamma(\mu) = \mu^2 + \beta_1\mu + \beta_2 = (\mu + \theta_{k+1}) \left[\mu + \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} \right],$$

έχουμε ότι

$$x^2 - \gamma'(\mu)x + \gamma(\mu) = [x - (\mu + \theta_{k+1})] \left[x - \left(\mu + \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} \right) \right].$$

Οπότε, το ελάχιστο πολυώνυμο είναι το

$$q_m(x; \mu) = [x^2 - \gamma'(\mu)x + \gamma(\mu)].$$

Χρησιμοποιώντας το θεώρημα Caley-Hamilton βρίσκουμε ότι

$$(B_{k+1} + \mu I)^2 - \gamma'(\mu)(B_{k+1} + \mu I) + \gamma(\mu)I = 0,$$

οπότε πολλαπλασιάζοντας με $(B_{k+1} + \mu I)^{-1}$, έχουμε ότι

$$(B_{k+1} + \mu I) - \gamma'(\mu)I + \gamma(\mu)(B_{k+1} + \mu I)^{-1} = 0.$$

Λύνοντας ως προς $(B_{k+1} + \mu I)^{-1}$, βρίσκουμε ότι

$$(B_{k+1} + \mu I)^{-1} = -\frac{1}{\gamma(\mu)} \{B_{k+1} - [\gamma'(\mu) - \mu]I\}. \quad (3.42)$$

Η γενική μορφή των ιδιοδιανύσμων της SR1 ενημέρωσης ελάχιστης μνήμης (3.40), που αντιστοιχούν σε διακριτές ιδιοτιμές, δίνεται στην Πρόταση 3.2. Επειδή όμως η SR1 είναι ενημέρωση πρώτης τάξης, μπορούμε να υπολογίσουμε τα συγκεκριμένα ιδιοδιανύσματα με πιο απλό τρόπο: πολλαπλασιάζουμε αμφότερα τα μέλη της εξίσωσης (3.40) με το διάνυσμα

$y_k - \theta_{k+1}s_k$ και έχουμε ότι

$$\begin{aligned} B_{k+1}(y_k - \theta_{k+1}s_k) &= \left[\theta_{k+1} + \frac{(y_k - \theta_{k+1}s_k)^T (y_k - \theta_{k+1}s_k)}{(y_k - \theta_{k+1}s_k)^T s_k} \right] (y_k - \theta_{k+1}s_k) \\ &= \lambda (y_k - \theta_{k+1}s_k). \end{aligned}$$

Επομένως, το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή

$$\lambda = \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k},$$

είναι το

$$u = y_k - \theta_{k+1}s_k.$$

Τα γενικευμένα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή θ_{k+1} είναι γραμμικά εξαρτημένα με το $u = y_k - \theta_{k+1}s_k$ και κάθετα σε αυτό. Τα πιο απλά διανύσματα u που ικανοποιούν τις σχέσεις $u_i = \alpha(y_k - \theta_{k+1}s_k)$ και $u_i \perp (y_k - \theta_{k+1}s_k)$, είναι της μορφής

$$u_i = \left(-\frac{y_k^i - \theta_{k+1}s_k^i}{y_k^1 - \theta_{k+1}s_k^1}, 0, \dots, 0, 1, 0, \dots, 0 \right)^T,$$

όπου το $i = 1, \dots, n-1$ δηλώνει την i συνιστώσα των y_k, s_k , ενώ η μονάδα βρίσκεται στην i γραμμή του u_i .

3.6 Η BFGS ενημέρωση ελάχιστης μνήμης

Η BFGS ενημέρωση ελάχιστης μνήμης υπολογίζεται από τη γενική PSB ενημέρωση ελάχιστης μνήμης (3.3), θέτοντας

$$w_k = \frac{1}{v_k + 1} y_k + \frac{\theta_{k+1} v_k}{v_k + 1} s_k, \quad \text{όπου} \quad v_k = \left(\frac{s_k^T y_k}{\theta_{k+1} s_k^T s_k} \right)^{1/2}.$$

Ο πίνακας που προκύπτει είναι ο

$$B_{k+1} = \theta_{k+1} I - \theta_{k+1} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k}. \quad (3.43)$$

Από τις ιδιότητες του ίχνους και της ορίζουσας (σελ. 13), για τον παραπάνω πίνακα έχουμε ότι

$$\text{tr}(B_{k+1}) = (n-1)\theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k}, \quad \text{και} \quad \det(B_{k+1}) = \theta_{k+1}^{n-1} \frac{s_k^T y_k}{s_k^T s_k}. \quad (3.44)$$

Από τις εκφράσεις αυτές μπορούμε να υπολογίσουμε τους συντελεστές του τετραγωνικού όρου του χαρακτηριστικού πολυωνύμου (3.11). Οι σχέσεις (3.20) (σελ. 89) γίνονται

$$\beta_1 = \theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \quad \text{και} \quad \beta_2 = \theta_{k+1} \frac{s_k^T y_k}{s_k^T s_k},$$

οπότε το χαρακτηριστικό πολυώνυμο είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(\theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \right) \lambda + \theta_{k+1} \frac{s_k^T y_k}{s_k^T s_k} \right]. \quad (3.45)$$

Λύνοντας τη δευτεροβάθμια εξίσωση

$$\lambda^2 - \left[\theta_{k+1} + (y_k^T y_k)/(s_k^T y_k) \right] \lambda + \theta_{k+1} (s_k^T y_k)/(s_k^T s_k) = 0$$

ως προς λ , βρίσκουμε τις διακριτές ιδιοτιμές του B_{k+1} , οι οποίες είναι οι

$$\lambda_{1,2} = \left(\theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \pm \sqrt{\left(\theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \right)^2 - 4\theta_{k+1} \frac{s_k^T y_k}{s_k^T s_k}} \right) / 2.$$

Στην περίπτωση που τα s_k και y_k είναι γραμμικά ανεξάρτητα, τα ιδιοδιάνυσμα που αντιστοιχούν στις διακριτές ιδιοτιμές του B_{k+1} βρίσκονται χρησιμοποιώντας την Πρόταση 3.2. Εναλλακτικά μπορούν να υπολογιστούν χρησιμοποιώντας την Πρόταση 3.1, μαζί με τη μέθοδο της αντίστροφης δύναμης. Η μεθοδολογία που ακολουθούμε για την παραγωγή των αναλυτικών εκφράσεων είναι παρόμοια με αυτήν της γενικής PSB ενημέρωσης ελάχιστης μνήμης. Έτσι, καταλήγουμε στο παρακάτω πόρισμα.

Πόρισμα 3.2. Έστω $\sigma = \lambda + \epsilon$, $\epsilon \rightarrow 0^+$, μια ϵ -διαταραγμένη διακριτή ιδιοτιμή του BFGS πίνακα ελάχιστης μνήμης B_{k+1} και το κανονικοποιημένο διάνυσμα $u_0 = (1, 1, \dots, 1)^T$. Τότε, το ιδιοδιάνυσμα που αντιστοιχεί στη σ

είναι το διάνυσμα

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)}, \quad (3.46)$$

όπου οι συντελεστές των διανυσμάτων u_0 , s_k και y_k ορίζονται από τις σχέσεις

$$c_1(\sigma) = \sigma^2 - \beta_1\sigma + \beta_2, \quad (3.47)$$

$$c_2(\sigma) = \left[(\beta_1 - \sigma) \sum_{i=1}^n s_k^i - \sum_{i=1}^n y_k^i \right] \frac{\theta_{k+1}}{s_k^T s_k}, \quad \text{και} \quad (3.48)$$

$$c_3(\sigma) = \frac{\sigma}{s_k^T y_k} \sum_{i=1}^n y_k^i - \frac{\theta_{k+1}}{s_k^T s_k} \sum_{i=1}^n s_k^i. \quad (3.49)$$

Απόδειξη. Από την Πρόταση 3.1, για $\hat{\lambda} = -\sigma$, έχουμε ότι

$$\begin{aligned} u(\sigma) &= (B_{k+1} - \sigma I)^{-1} u \\ u(\hat{\lambda}) &= (B_{k+1} + \hat{\lambda} I)^{-1} u \\ &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ B_{k+1}^2 u - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\}. \end{aligned}$$

Όμως,

$$B_{k+1} u = \theta_{k+1} u - \theta_{k+1} \frac{s_k^T u}{s_k^T s_k} s_k + \frac{y_k^T u}{s_k^T y_k} y_k$$

και

$$B_{k+1}^2 u = \theta_{k+1} B_{k+1} u - \theta_{k+1} \frac{s_k^T B_{k+1} u}{s_k^T s_k} s_k + \frac{y_k^T B_{k+1} u}{s_k^T y_k} y_k,$$

οπότε

$$\begin{aligned}
 u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \theta_{k+1} B_{k+1} u - \theta_{k+1} \frac{s_k^T B_{k+1} u}{s_k^T s_k} s_k + \frac{y_k^T B_{k+1} u}{s_k^T y_k} y_k \right. \\
 &\quad \left. - \left[\beta'(\hat{\lambda}) - \hat{\lambda} \right] B_{k+1} u + \beta(\hat{\lambda}) u \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \left[\theta_{k+1} \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) + \beta(\hat{\lambda}) \right] u \right. \\
 &\quad \left. + (-\theta_{k+1}) \left[\left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{s_k^T u}{s_k^T s_k} + \frac{s_k^T B_{k+1} u}{s_k^T s_k} \right] s_k \right. \\
 &\quad \left. + \left[\left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{y_k^T u}{s_k^T s_k} + \frac{y_k^T B_{k+1} u}{s_k^T y_k} \right] y_k \right\} \\
 &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left[c_1(\hat{\lambda}) u + c_2(\hat{\lambda}) s_k + c_3(\hat{\lambda}) y_k \right].
 \end{aligned}$$

Υπενθυμίζοντας ότι $\beta(\lambda) = \gamma(\lambda) + (\gamma'(\lambda) - \lambda)\theta_{k+1}$, για το συντελεστή του διανύσματος u έχουμε:

$$\begin{aligned}
 c_1(\hat{\lambda}) &= \theta_{k+1} \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) + \beta(\hat{\lambda}) \\
 &= \theta_{k+1} \left(\hat{\lambda} - \gamma'(\hat{\lambda}) \right) + \gamma(\hat{\lambda}) + (\gamma'(\hat{\lambda}) - \hat{\lambda})\theta_{k+1} \\
 &= \gamma(\hat{\lambda}).
 \end{aligned}$$

Για τον $c_2(\hat{\lambda})$, εκμεταλλευόμενοι την εξίσωση της τέμνουσας $B_{k+1} s_k = y_k$, έχουμε:

$$\begin{aligned}
 c_2(\hat{\lambda}) &= -\theta_{k+1} \left[\left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{s_k^T u}{s_k^T s_k} + \frac{s_k^T B_{k+1} u}{s_k^T s_k} \right] \\
 &= -\theta_{k+1} \left[\left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{s_k^T u}{s_k^T s_k} + \frac{y_k^T u}{s_k^T s_k} \right] \\
 &= \left[\left(\gamma'(\hat{\lambda}) - \hat{\lambda} \right) s_k^T u - y_k^T u \right] \frac{\theta_{k+1}}{s_k^T s_k}.
 \end{aligned}$$

Τέλος, για το συντελεστή του y_k , λαμβάνοντας υπόψη ότι $\gamma(\lambda) = \lambda^2 +$

$\beta_1 \lambda + \beta_2$ και $\beta_1 = \theta_{k+1} + y_k^T y_k / s_k^T y_k$, βρίσκουμε ότι

$$\begin{aligned}
 c_3(\hat{\lambda}) &= \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{y_k^T u}{s_k^T s_k} + \frac{y_k^T B_{k+1} u}{s_k^T y_k} \\
 &= \left(\hat{\lambda} - \gamma'(\hat{\lambda}) \right) \frac{y_k^T u}{s_k^T y_k} + \frac{y_k^T}{s_k^T y_k} \left(\theta_{k+1} u - \theta_{k+1} \frac{s_k^T u}{s_k^T s_k} s_k + \frac{y_k^T u}{s_k^T y_k} y_k \right) \\
 &= \left(\hat{\lambda} - \gamma'(\hat{\lambda}) + \theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \right) \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{s_k^T u}{s_k^T s_k} \\
 &= -\hat{\lambda} \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{s_k^T u}{s_k^T s_k}.
 \end{aligned}$$

Δοθέντος ότι το u είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$, και αντικαθιστώντας όπου $\hat{\lambda}$ το $-\sigma$, καταλήγουμε στη σχέση (3.46), με τους συντελεστές c_1 , c_2 και c_3 να ορίζονται όπως στις εκφράσεις (3.47), (3.48) και (3.49), αντίστοιχα. $\square$

Στον παρακάτω πίνακα φαίνονται συγκεντρωτικά οι ποσότητες που απαιτούνται για τον υπολογισμό του $u(\sigma)$. Μπορούμε να δούμε ότι χρειαζόμαστε δύο διανύσματα και τρία εσωτερικά γινόμενα, όσα ακριβώς χρειαζόμαστε για να υπολογίσουμε το ιδιοδιάνυσμα και μέσω της Πρότασης 3.1. Προφανώς όμως, χρειάζονται περισσότεροι πολλαπλασιασμοί και προσθαφαιρέσεις, από αυτές που χρειάζονται στην εξίσωση (3.26).

Πίνακας 3.2: Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στη BFGS ενημέρωση ελάχιστης μνήμης

διανύσματα	εσωτερικά γινόμενα
s_k	$s_k^T s_k$
y_k	$s_k^T y_k$
	$y_k^T y_k$

3.7 Η DFP ενημέρωση ελάχιστης μνήμης

Θέτοντας $w_k = y_k$ στον πίνακα (3.3) της PSB ενημέρωσης ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)w_k^T + w_k(y_k - \theta_{k+1}s_k)^T}{s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{s_k^T w_k},$$

υπολογίζουμε τη DFP ενημέρωση ελάχιστης μνήμης,

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)y_k^T + y_k(y_k - \theta_{k+1}s_k)^T}{s_k^T y_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T y_k} \frac{y_k y_k^T}{s_k^T y_k}. \quad (3.50)$$

Από τις ιδιότητες του ίχνους και της ορίζουσας πινάκων (σελ. 13), βρίσκουμε ότι το ίχνος του παραπάνω πίνακα είναι το

$$\begin{aligned} \text{tr}(B_{k+1}) &= n\theta_{k+1} + 2 \frac{(y_k - \theta_{k+1}s_k)^T y_k}{s_k^T y_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T y_k} \frac{y_k^T y_k}{s_k^T y_k} \\ &= (n-2)\theta_{k+1} + \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k}\right) \frac{y_k^T y_k}{s_k^T y_k}. \end{aligned} \quad (3.51)$$

Αντίστοιχα, για την ορίζουσα έχουμε ότι

$$\begin{aligned} \det(B_{k+1}) &= \det(\theta_{k+1}I) \det\left(I + \frac{(y_k - \theta_{k+1}s_k)y_k^T + y_k(y_k - \theta_{k+1}s_k)^T}{\theta_{k+1}s_k^T y_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T y_k} \frac{y_k y_k^T}{\theta_{k+1}s_k^T y_k}\right) \\ &= \det(\theta_{k+1}I) \det\left[I + \left(\frac{s_k^T s_k}{s_k^T y_k} y_k - s_k\right) \frac{y_k^T}{s_k^T y_k} + \frac{y_k(y_k - \theta_{k+1}s_k)^T}{\theta_{k+1}s_k^T y_k}\right] \\ &= \theta_{k+1}^n \left[1 + \left(\frac{s_k^T s_k}{s_k^T y_k} y_k - s_k\right)^T \frac{y_k}{s_k^T y_k}\right] \left[1 + \frac{y_k^T (y_k - \theta_{k+1}s_k)}{\theta_{k+1}s_k^T y_k}\right] \end{aligned}$$

$$\begin{aligned}
& - \left(\frac{s_k^T s_k}{s_k^T y_k} y_k - s_k \right)^T \frac{(y_k - \theta_{k+1} s_k)}{s_k^T y_k} \frac{y_k^T y_k}{\theta_{k+1} s_k^T y_k} \\
& = \theta_{k+1}^{n-1} \frac{y_k^T y_k}{s_k^T y_k}.
\end{aligned} \tag{3.52}$$

Από το Θεώρημα 3.1 ξέρουμε ότι το χαρακτηριστικό πολυώνυμο του B_{k+1} είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} [\lambda^2 - \beta_1 \lambda + \beta_2].$$

Για να υπολογίσουμε τα β_1 και β_2 αντικαθιστούμε στις σχέσεις (3.20) (σελ. 89), τις σχέσεις (3.51) και (3.52). Εύκολα βρίσκουμε ότι

$$\beta_1 = \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k} \tag{3.53}$$

και

$$\beta_2 = \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}, \tag{3.54}$$

οπότε το χαρακτηριστικό πολυώνυμο γίνεται

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k} \lambda + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k} \right] \tag{3.55}$$

Προφανώς, για να υπολογίσουμε τις διακριτές ιδιοτιμές, αρκεί να λύσουμε τη δευτεροβάθμια εξίσωση

$$\lambda^2 - \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k} \lambda + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k} = 0,$$

από όπου υπολογίζουμε τις ιδιοτιμές

$$\lambda_{1,2} = \frac{y_k^T y_k}{2s_k^T y_k} \left[1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \pm \sqrt{\left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right)^2 - 4\theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k}} \right].$$

Για να βρούμε τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές του B_{k+1} , μπορούμε να εφαρμόσουμε είτε την Πρόταση 3.2, ή εναλλακτικά, την Πρόταση 3.1 σε συνδυασμό με τη μέθοδο της αντίστροφης δύναμης, προκειμένου να οδηγηθούμε σε κλειστό τύπο. Για τον δεύτερο

τρόπο, έχουμε:

$$u(\sigma) = (B_{k+1} - \sigma I)^{-1} u.$$

Για $\hat{\lambda} = -\sigma$, η παραπάνω σχέση γίνεται

$$\begin{aligned} u(\hat{\lambda}) &= (B_{k+1} + \hat{\lambda} I)^{-1} u \\ &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ B_{k+1}^2 u - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\}. \end{aligned}$$

Όμως, επειδή

$$\begin{aligned} B_{k+1} &= \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k) y_k^T + y_k (y_k - \theta_{k+1} s_k)^T}{s_k^T y_k} \\ &\quad - \frac{(y_k - \theta_{k+1} s_k)^T s_k y_k y_k^T}{s_k^T y_k} \frac{s_k^T y_k}{s_k^T y_k} \\ &= \theta_{k+1} I - \theta_{k+1} \frac{y_k s_k^T + s_k y_k^T}{s_k^T y_k} + \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k y_k^T}{s_k^T y_k} \\ &= \theta_{k+1} I - \theta_{k+1} \frac{y_k s_k^T + s_k y_k^T}{s_k^T y_k} + a \frac{y_k y_k^T}{s_k^T y_k}, \end{aligned}$$

όπου έχουμε θέσει $a = 1 + \theta_{k+1} s_k^T s_k / s_k^T y_k$, έχουμε ότι

$$\begin{aligned} B_{k+1} u &= \theta_{k+1} u - \theta_{k+1} \frac{y_k^T u}{s_k^T y_k} s_k + \left(-\theta_{k+1} \frac{s_k^T u}{s_k^T y_k} + a \frac{y_k^T u}{s_k^T y_k} \right) y_k \\ &= \theta_{k+1} u - \theta_{k+1} \frac{s_k^T u}{s_k^T y_k} s_k + a_1 y_k, \end{aligned} \quad (3.56)$$

με

$$a_1 = -\theta_{k+1} \frac{s_k^T u}{s_k^T y_k} + a \frac{y_k^T u}{s_k^T y_k}. \quad (3.57)$$

Παρόμοια, βρίσκουμε ότι

$$\begin{aligned} B_{k+1}^2 u &= \theta_{k+1} B_{k+1} u - \theta_{k+1} \frac{y_k^T B_{k+1} u}{s_k^T y_k} s_k \\ &\quad + \left(-\theta_{k+1} \frac{s_k^T B_{k+1} u}{s_k^T y_k} + a \frac{y_k^T B_{k+1} u}{s_k^T y_k} \right) y_k, \end{aligned}$$

$$= \theta_{k+1} B_{k+1} u - \theta_{k+1} \frac{y_k^T B_{k+1} u}{s_k^T y_k} s_k + a_2 y_k,$$

όπου

$$a_2 = -\theta_{k+1} \frac{s_k^T B_{k+1} u}{s_k^T y_k} + a \frac{y_k^T B_{k+1} u}{s_k^T y_k}. \quad (3.58)$$

Οπότε, έχουμε ότι

$$\begin{aligned} u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ B_{k+1}^2 u - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\} \\ &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \theta_{k+1} B_{k+1} u - \theta_{k+1} \frac{y_k^T B_{k+1} u}{s_k^T y_k} s_k + a_2 y_k \right. \\ &\quad \left. - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\} \end{aligned}$$

Αντικαθιστώντας το διάνυσμα $B_{k+1} u$ στην παραπάνω σχέση, βρίσκουμε ότι

$$\begin{aligned} u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ [\theta_{k+1} (\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) + \beta(\hat{\lambda})] u \right. \\ &\quad \left. + \left[-\theta_{k+1} (\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{y_k^T B_{k+1} u}{s_k^T y_k} \right] s_k \right. \\ &\quad \left. + [(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) a_1 + a_2] y_k \right\} \\ &= \frac{c_1(\hat{\lambda})u + c_2(\hat{\lambda})s_k + c_3(\hat{\lambda})y_k}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})}. \end{aligned} \quad (3.59)$$

Δοθέντος της σχέσης $\beta(\hat{\lambda}) = \gamma(\hat{\lambda}) + (\gamma'(\hat{\lambda}) - \lambda) \theta_{k+1}$, για το συντελεστή του διανύσματος u , υπολογίζουμε το εξής:

$$\begin{aligned} c_1(\hat{\lambda}) &= \theta_{k+1} [\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}] + \beta(\hat{\lambda}) \\ &= \theta_{k+1} [-\gamma'(\hat{\lambda}) + \lambda] + \gamma(\hat{\lambda}) + [\gamma'(\hat{\lambda}) - \lambda] \theta_{k+1} \\ &= \gamma(\hat{\lambda}). \end{aligned} \quad (3.60)$$

Για να υπολογίσουμε το $c_2(\hat{\lambda})$, χρειαζόμαστε το εσωτερικό γινόμενο

$$y_k^T B_{k+1} u,$$

για το οποίο έχουμε ότι

$$\begin{aligned} \frac{y_k^T B_{k+1} u}{s_k^T y_k} &= \frac{y_k^T}{s_k^T y_k} \left(\theta_{k+1} u - \theta_{k+1} \frac{y_k^T u}{s_k^T y_k} s_k + a_1 y_k \right) = a_1 \frac{y_k^T y_k}{s_k^T y_k} \\ &= \left(-\theta_{k+1} \frac{s_k^T u}{s_k^T y_k} + a \frac{y_k^T u}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k}. \end{aligned}$$

Όμως, επειδή $a = 1 + \theta_{k+1} s_k^T s_k / s_k^T y_k$, από τις εκφράσεις (3.53) και (3.54), των β_1 και β_2 , αντίστοιχα, η παραπάνω σχέση γίνεται

$$\frac{y_k^T B_{k+1} u}{s_k^T y_k} = -\beta_2 \frac{s_k^T u}{s_k^T y_k} + \beta_1 \frac{y_k^T u}{s_k^T y_k}. \quad (3.61)$$

Επομένως, για το $c_2(\hat{\lambda})$ υπολογίζουμε:

$$\begin{aligned} c_2(\hat{\lambda}) &= -\theta_{k+1} \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{y_k^T B_{k+1} u}{s_k^T y_k} \\ &= -\theta_{k+1} \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \left(-\beta_2 \frac{s_k^T u}{s_k^T y_k} + \beta_1 \frac{y_k^T u}{s_k^T y_k} \right) \end{aligned}$$

Εκμεταλλευόμενοι τις σχέσεις $\beta(\hat{\lambda}) = \gamma(\hat{\lambda}) + (\gamma'(\hat{\lambda}) - \lambda) \theta_{k+1}$ και $\gamma(\hat{\lambda}) = \hat{\lambda}^2 + \beta_1 \hat{\lambda} + \beta_2$, βρίσκουμε τελικά ότι

$$c_2(\hat{\lambda}) = \left(\hat{\lambda} y_k^T u + \beta_2 s_k^T u \right) \frac{\theta_{k+1}}{s_k^T y_k}. \quad (3.62)$$

Τέλος, για να υπολογίσουμε το $c_3(\hat{\lambda})$, χρειάζεται πρώτα να βρούμε το εσωτερικό γινόμενο $s_k^T B_{k+1} u$, για το οποίο, λόγω της εξίσωσης της τέμνουσας $B_{k+1} s_k = y_k$, έχουμε:

$$\frac{s_k^T B_{k+1} u}{s_k^T y_k} = \frac{y_k^T u}{s_k^T y_k}. \quad (3.63)$$

Επομένως, χρησιμοποιώντας τις σχέσεις (3.57), (3.58), (3.61) και (3.63), για τον $c_3(\hat{\lambda})$ έχουμε:

$$c_3(\hat{\lambda}) = \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) a_1 + a_2$$

$$\begin{aligned}
&= \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \left(a \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{s_k^T u}{s_k^T y_k} \right) \\
&\quad - \theta_{k+1} \frac{s_k^T B_{k+1} u}{s_k^T y_k} + a \frac{y_k^T B_{k+1} u}{s_k^T y_k} \\
&= \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \left(a \frac{y_k^T u}{s_k^T y_k} - \theta_{k+1} \frac{s_k^T u}{s_k^T y_k} \right) - \theta_{k+1} \frac{y_k^T u}{s_k^T y_k} \\
&\quad + a \left(\beta_1 \frac{y_k^T u}{s_k^T y_k} - \beta_2 \frac{s_k^T u}{s_k^T y_k} \right) \\
&= - \left[\theta_{k+1} \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} + a\beta_2 \right) \right] \frac{s_k^T u}{s_k^T y_k} + \\
&\quad + \left[a \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} + \beta_1 \right) - \theta_{k+1} \right] \frac{y_k^T u}{s_k^T y_k}.
\end{aligned}$$

Όμως, $\beta'(\hat{\lambda}) = 2\hat{\lambda} + \beta_1 + \theta_{k+1}$. Επίσης, χρησιμοποιώντας τη σχέση (3.53), βρίσκουμε ότι

$$a = 1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} = \beta_1 \frac{s_k^T y_k}{y_k^T y_k}.$$

Οπότε, το $c_3(\hat{\lambda})$ μπορεί να γραφεί ως εξής:

$$c_3(\hat{\lambda}) = \hat{\lambda} \theta_{k+1} \frac{s_k^T u}{s_k^T y_k} - \left(\hat{\lambda} \beta_1 \frac{s_k^T y_k}{y_k^T y_k} + \theta_{k+1} \right) \frac{y_k^T u}{s_k^T y_k} \quad (3.64)$$

Αντικαθιστώντας όπου $\hat{\lambda}$ το $-\sigma$ στις σχέσεις (3.59), (3.60), (3.62) και (3.64) και λαμβάνοντας υπόψιν ότι το u είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$, καταλήγουμε στο εξής πόρισμα:

Πόρισμα 3.3. Έστω $\sigma = \lambda + \epsilon$, $\epsilon \rightarrow 0^+$, μια ϵ -διαταραγμένη διακριτή ιδιοτιμή του DFP πίνακα ελάχιστης μνήμης B_{k+1} και το κανονικοποιημένο διάνυσμα $u_0 = (1, 1, \dots, 1)^T$. Τότε, το ιδιοδιάνυσμα που αντιστοιχεί στη σ είναι το διάνυσμα

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)}, \quad (3.65)$$

όπου οι συντελεστές των διανυσμάτων u_0 , s_k και y_k ορίζονται από τις σχέσεις

$$\begin{aligned} c_1(\sigma) &= \sigma^2 - \beta_1 \sigma + \beta_2, \\ c_2(\sigma) &= \left(\beta_2 \sum_{i=1}^n s_i - \sigma \sum_{i=1}^n y_i \right) \frac{\theta_{k+1}}{s_k^T y_k} \quad \text{και} \\ c_3(\sigma) &= -\frac{\sigma \theta_{k+1}}{s_k^T y_k} \sum_{i=1}^n s_i - \left(\frac{\theta_{k+1}}{s_k^T y_k} - \frac{\sigma \beta_1}{y_k^T y_k} \right) \sum_{i=1}^n y_i. \end{aligned}$$

Στον παρακάτω πίνακα φαίνεται ότι για τον υπολογισμό του $u(\sigma)$ χρειαζόμαστε δύο διανύσματα και τρία εσωτερικά γινόμενα, όσα και για να υπολογίσουμε το ιδιοδιάνυσμα μέσω της Πρότασης 3.2. Η μόνη διαφορά είναι ότι γίνονται περισσότεροι πολλαπλασιασμοί και προσθαφαιρέσεις.

Πίνακας 3.3: Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στη DFP ενημέρωση ελάχιστης μνήμης

διανύσματα	εσωτερικά γινόμενα
s_k	$s_k^T s_k$
y_k	$s_k^T y_k$
	$y_k^T y_k$

3.8 Η PSB ενημέρωση ελάχιστης μνήμης

Η PSB ενημέρωση ελάχιστης μνήμης ορίζεται ως

$$\begin{aligned} B_{k+1} = \theta_{k+1} I &+ \frac{(y_k - \theta_{k+1} s_k) s_k^T + s_k (y_k - \theta_{k+1} s_k)^T}{s_k^T s_k} \\ &- \frac{(y_k - \theta_{k+1} s_k)^T s_k s_k s_k^T}{s_k^T s_k s_k^T s_k}, \end{aligned}$$

ή αλλιώς,

$$B_{k+1} = \theta_{k+1} I + \frac{y_k s_k^T + s_k y_k^T}{s_k^T s_k} - \left(\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k} \right) \frac{s_k s_k^T}{s_k^T s_k}. \quad (3.66)$$

Ο παραπάνω πίνακας υπολογίζεται θέτοντας $w_k = s_k$ στον πίνακα της γενικής PSB ενημέρωσης ελάχιστης μνήμης. Για το ίχνος του πίνακα (3.66), εύκολα βρίσκουμε ότι

$$\text{tr} B_{k+1} = (n-1)\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k}. \quad (3.67)$$

Όσον αφορά την ορίζουσα, από τη σχέση (3.66), έχουμε ότι

$$\det B_{k+1} = \det(\theta_{k+1} I) \det \left[I + \frac{y_k s_k^T + s_k y_k^T}{\theta_{k+1} s_k^T s_k} - \left(\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k} \right) \frac{s_k s_k^T}{\theta_{k+1} s_k^T s_k} \right]$$

Εφαρμόζοντας την ιδιότητα

$$\det(I + v_1 v_2^T + v_3 v_4^T) = (1 + v_1^T v_2)(1 + v_3^T v_4) - (v_1^T v_4)(v_2^T v_3),$$

βρίσκουμε ότι

$$\det B_{k+1} = \theta_{k+1}^{n-2} \frac{(\theta_{k+1} s_k^T y_k - y_k^T y_k) s_k^T s_k + (s_k^T y_k)^2}{(s_k^T s_k)^2}. \quad (3.68)$$

Επομένως, οι συντελεστές β_1 και β_2 του δευτεροβάθμιου όρου του χαρακτηριστικού πολυωνύμου

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} (\lambda^2 - \beta_1 \lambda + \beta_2),$$

γίνονται

$$\beta_1 = \text{tr}(B_{k+1}) - (n-2)\theta_{k+1} = \theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k}$$

και

$$\beta_2 = \det(B_{k+1}) / \theta_{k+1}^{n-2} = [(\theta_{k+1} s_k^T y_k - y_k^T y_k) s_k^T s_k + (s_k^T y_k)^2] / (s_k^T s_k)^2,$$

αντίστοιχα. Άρα, το χαρακτηριστικό πολυώνυμο της PSB ενημέρωσης ελάχιστης μνήμης, είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k} \right) \lambda + \frac{(\theta_{k+1} s_k^T y_k - y_k^T y_k) s_k^T s_k + (s_k^T y_k)^2}{(s_k^T s_k)^2} \right],$$

από όπου βρίσκουμε ότι οι διακριτές ιδιοτιμές είναι οι

$$\lambda_{1,2} = \theta_{k+1} + \frac{s_k^T y_k \pm \sqrt{(\theta_{k+1} s_k^T s_k - s_k^T y_k)^2 - 4[(s_k^T y_k)^2 - y_k^T y_k s_k^T s_k]}}{2s_k^T s_k}.$$

Τέλος, τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές του B_{k+1} , στην περίπτωση που τα s_k και y_k είναι γραμμικά ανεξάρτητα, υπολογίζονται χρησιμοποιώντας την Πρόταση 3.2 ή μέσω της Πρότασης 3.1 και της μεθόδου της αντίστροφης δύναμης. Για την τελευταία, έχουμε ότι

$$u(\sigma) = (B_{k+1} - \sigma I)^{-1} u,$$

και χρησιμοποιώντας την Πρόταση 3.1, για $\hat{\lambda} = -\sigma$ υπολογίζουμε ότι

$$\begin{aligned} u(\hat{\lambda}) &= (B_{k+1} + \hat{\lambda} I)^{-1} u \\ &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ B_{k+1}^2 u - [\beta'(\hat{\lambda}) - \hat{\lambda}] B_{k+1} u + \beta(\hat{\lambda}) u \right\}. \end{aligned}$$

Όμως, από τη σχέση (3.66) έχουμε ότι

$$B_{k+1} u = \theta_{k+1} u + \frac{s_k^T u}{s_k^T s_k} y_k + \left[\frac{y_k^T u}{s_k^T s_k} - \left(\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k} \right) \frac{s_k^T u}{s_k^T s_k} \right] s_k.$$

Επειδή όμως $\beta_1 = \theta_{k+1} + s_k^T y_k / s_k^T s_k$, η παραπάνω εξίσωση μπορεί να γραφεί ως εξής:

$$B_{k+1} u = \theta_{k+1} u + \frac{s_k^T u}{s_k^T s_k} y_k + \left(\frac{y_k^T u}{s_k^T s_k} - \beta_1 \frac{s_k^T u}{s_k^T s_k} \right) s_k.$$

Παρόμοια, για το διάνυσμα $B_{k+1}^2 u$ έχουμε ότι

$$B_{k+1}^2 u = \theta_{k+1} B_{k+1} u + \frac{s_k^T B_{k+1} u}{s_k^T s_k} y_k + \left(\frac{y_k^T B_{k+1} u}{s_k^T s_k} - \beta_1 \frac{s_k^T B_{k+1} u}{s_k^T s_k} \right) s_k.$$

Αντικαθιστώντας τις παραπάνω ισότητες στην έκφραση του ιδιοδιανύσματος, βρίσκουμε ότι

$$\begin{aligned} u(\hat{\lambda}) &= \frac{1}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})} \left\{ \left[\theta_{k+1} (\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) + \beta(\hat{\lambda}) \right] u \right. \\ &\quad + \left[(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) \left(\frac{y_k^T u}{s_k^T s_k} - \beta_1 \frac{s_k^T u}{s_k^T s_k} \right) \right. \\ &\quad + \left. \left(\frac{y_k^T B_{k+1} u}{s_k^T s_k} - \beta_1 \frac{s_k^T B_{k+1} u}{s_k^T s_k} \right) \right] s_k \\ &\quad + \left. \left[(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda}) \frac{s_k^T u}{s_k^T s_k} + \frac{s_k^T B_{k+1} u}{s_k^T s_k} \right] y_k \right\} \\ &= \frac{c_1(\hat{\lambda})u + c_2(\hat{\lambda})s_k + c_3(\hat{\lambda})y_k}{(\hat{\lambda} + \theta_{k+1})\gamma(\hat{\lambda})}. \end{aligned} \quad (3.69)$$

Επειδή ισχύει ότι $\beta(\hat{\lambda}) = \gamma(\hat{\lambda}) + (\gamma'(\hat{\lambda}) - \lambda) \theta_{k+1}$, για το συντελεστή του διανύσματος u , βρίσκουμε ότι

$$\begin{aligned} c_1(\hat{\lambda}) &= \theta_{k+1} \left[\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right] + \beta(\hat{\lambda}) \\ &= \theta_{k+1} \left[-\gamma'(\hat{\lambda}) + \lambda \right] + \gamma(\hat{\lambda}) + \left[\gamma'(\hat{\lambda}) - \lambda \right] \theta_{k+1} \\ &= \gamma(\hat{\lambda}). \end{aligned} \quad (3.70)$$

Για να υπολογίσουμε το $c_2(\hat{\lambda})$ χρειάζεται πρώτα να υπολογίσουμε τα εσωτερικά γινόμενα $s_k^T B_{k+1} u$ και $y_k^T B_{k+1} u$, για τα οποία έχουμε:

$$\frac{s_k^T B_{k+1} u}{s_k^T s_k} = \frac{y_k^T u}{s_k^T s_k}$$

λόγω της εξίσωσης της τέμνουσας $s_k^T B_{k+1} = y_k$, και

$$\frac{y_k^T B_{k+1} u}{s_k^T s_k} = \beta_1 \frac{y_k^T u}{s_k^T s_k} + \left(\frac{y_k^T y_k}{s_k^T s_k} - \beta_1 \frac{s_k^T y_k}{s_k^T s_k} \right) \frac{s_k^T u}{s_k^T y_k}$$

αντίστοιχα. Επομένως, για το $c_2(\hat{\lambda})$ βρίσκουμε ότι

$$\begin{aligned} c_2(\hat{\lambda}) &= \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{y_k^T u}{s_k^T y_k} \\ &\quad + \left[\frac{y_k^T y_k}{s_k^T y_k} - \beta_1 \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} + \frac{s_k^T y_k}{s_k^T y_k} \right) \right] \frac{s_k^T u}{s_k^T s_k}. \end{aligned}$$

Εκμεταλλευόμενοι τις σχέσεις $\beta(\hat{\lambda}) = \gamma(\hat{\lambda}) + \left(\gamma'(\hat{\lambda}) - \lambda \right) \theta_{k+1}$, $\gamma(\hat{\lambda}) = \hat{\lambda}^2 + \beta_1 \hat{\lambda} + \beta_2$ και $\beta_1 = \theta_{k+1} + s_k^T y_k / s_k^T s_k$, έχουμε ότι

$$\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} = -\hat{\lambda} - \beta_1 \quad \text{και} \quad \frac{s_k^T y_k}{s_k^T s_k} = \beta_1 - \theta_{k+1}.$$

Οπότε η έκφραση του συντελεστή $c_2(\hat{\lambda})$ του s_k γίνεται τελικά

$$c_2(\hat{\lambda}) = \left[\frac{y_k^T y_k}{s_k^T s_k} + \beta_1 \left(\hat{\lambda} + \theta_{k+1} \right) \right] \frac{s_k^T u}{s_k^T s_k} - \left(\hat{\lambda} + \beta_1 \right) \frac{y_k^T u}{s_k^T s_k}. \quad (3.71)$$

Τέλος, για το $c_3(\hat{\lambda})$ έχουμε ότι

$$\begin{aligned} c_3(\hat{\lambda}) &= \left(\theta_{k+1} - \beta'(\hat{\lambda}) + \hat{\lambda} \right) \frac{s_k^T u}{s_k^T y_k} + \frac{s_k^T B_{k+1} u}{s_k^T y_k} \\ &= -(\hat{\lambda} + \beta_1) \frac{s_k^T u}{s_k^T s_k} + \frac{y_k^T u}{s_k^T s_k}. \end{aligned} \quad (3.72)$$

Συνοψίζοντας τα παραπάνω, αντικαθιστώντας όπου $\hat{\lambda}$ το $-\sigma$ στις σχέσεις (3.69), (3.70), (3.71) και (3.72) και λαμβάνοντας υπόψιν ότι το u είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$, καταλήγουμε στο εξής πόρισμα:

Πόρισμα 3.4. Έστω $\sigma = \lambda + \epsilon$, $\epsilon \rightarrow 0^+$, μια ϵ -διαταραγμένη διακριτή ιδιοτιμή του PSB πίνακα ελάχιστης μνήμης B_{k+1} και το κανονικοποιημένο διάνυσμα $u_0 = (1, 1, \dots, 1)^T$. Τότε, το ιδιοδιάνυσμα που αντιστοιχεί στη σ είναι το διάνυσμα

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)}, \quad (3.73)$$

όπου οι συντελεστές των διανυσμάτων u_0 , s_k και y_k ορίζονται από τις σχέσεις

$$\begin{aligned} c_1(\sigma) &= \sigma^2 - \beta_1\sigma + \beta_2, \\ c_2(\sigma) &= \frac{1}{s_k^T s_k} \left[\frac{y_k^T y_k}{s_k^T s_k} - \beta_1 (\sigma - \theta_{k+1}) \right] \sum_{i=1}^n s_i + \frac{\sigma - \beta_1}{s_k^T s_k} \sum_{i=1}^n y_i \quad \text{και} \\ c_3(\sigma) &= \frac{\sigma - \beta_1}{s_k^T s_k} \sum_{i=1}^n s_i + \frac{1}{s_k^T s_k} \sum_{i=1}^n y_i. \end{aligned}$$

Στον παρακάτω πίνακα φαίνονται συγκεντρωτικά οι ποσότητες που απαιτούνται για τον υπολογισμό του $u(\sigma)$.

Πίνακας 3.4: Απαιτούμενα διανύσματα και εσωτερικά γινόμενα για τον υπολογισμό του ιδιοδιανύσματος στην PSB ενημέρωση ελάχιστης μνήμης

διανύσματα	εσωτερικά γινόμενα
s_k	$s_k^T s_k$
y_k	$s_k^T y_k$
	$y_k^T y_k$

3.9 Συμπεράσματα

Στο κεφάλαιο αυτό παρουσιάσαμε τα ερευνητικά μας αποτελέσματα αναφορικά με τα χαρακτηριστικά μεγέθη των Quasi-Newton ενημερώσεων ελάχιστης μνήμης. Παρουσιάσαμε αναλυτικά τον τρόπο εύρεσης του χαρακτηριστικού πολυωνύμου, καθώς και των ιδιοτιμών της γενικής PSB ενημέρωσης ελάχιστης μνήμης. Βασιζόμενοι σε αυτή τη μελέτη, εξετάσαμε τα χαρακτηριστικά μεγέθη και των τεσσάρων Quasi-Newton ενημερώσεων ελάχιστης μνήμης, και συγκεκριμένα της SR1, BFGS, DFP και PSB, ξεχωριστά. Με τη βοήθεια του χαρακτηριστικού πολυωνύμου υπολογίσαμε σε κλειστή μορφή τον αντίστροφο του πίνακα $B + \mu I$, όπου $\mu \in \mathbb{R}$. Τέλος, δώσαμε αναλυτικά τον τύπο υπολογισμού των ιδιοδιανυσμάτων που αντιστοιχούν σε διακριτές ιδιοτιμές με δύο τρόπους, ένας εκ των οποίων κάνει χρήση της μεθόδου της αντίστροφης δύναμης.

Είδαμε ότι η ελάχιστη μνήμη αυτής της κλάσης των ενημερώσεων, μας

έδωσε τη δυνατότητα να παράγουμε κλειστούς τύπους για όλα αυτά τα μεγέθη, χρησιμοποιώντας μόνο αθροίσματα διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό τους. Επομένως, η αποθήκευση ή η παραγοντοποίηση οποιουδήποτε πίνακα αποφεύχθηκε πλήρως. Αυτό έχει ως αποτέλεσμα οι απαιτούμενοι υπολογισμοί να έχουν αμελητέο υπολογιστικό κόστος. Ως συνέπεια αυτού, η προτεινόμενη μέθοδος μπορεί να εφαρμοστεί μέσα σε ένα πλαίσιο επίλυσης μεγάλης κλίμακας προβλημάτων βελτιστοποίησης.

Η BFGS ενημέρωση μικρής μνήμης

Σε αυτό το κεφάλαιο μελετάμε τα χαρακτηριστικά μεγέθη της BFGS ενημέρωσης μικρής μνήμης, λόγω της σπουδαιότητας που έχει η BFGS ενημέρωση στις μεθόδους βελτιστοποίησης. Στη συγκεκριμένη ενημέρωση μικρής μνήμης, ο αρχικός πίνακας είναι ο μοναδιαίος πίνακας κλιμακούμενος με τη BB φασματική παράμετρο. Τα αποθηκευμένα διανυσματικά ζεύγη που χρησιμοποιούνται για την ενημέρωση του αρχικού πίνακα είναι είτε ένα, όπου λαμβάνουμε την BFGS ενημέρωσης ελάχιστης μνήμης, ή δύο, από όπου προκύπτει η BFGS ενημέρωσης μικρής μνήμης. Μελετάμε τις ιδιοτιμές και τα ιδιοδιανύσματα που αντιστοιχούν σε διακριτές ιδιοτιμές. Αποδεικνύουμε ότι ο ακριβής αριθμός των διακριτών ιδιοτιμών εξαρτάται άμεσα από τη σχέση (γραμμική ανεξαρτησία ή μη) που συνδέει τα διανυσματικά ζεύγη ενημέρωσης. Τα ακόλουθα αποτελούν ερευνητικά αποτελέσματα της παρούσας διατριβής, τα οποία καταγράφονται στις εργασίες των Apostolopoulou et al. [6] και [12].

4.1 Οι 1BFGS και 2BFGS ενημερώσεις

Είναι γνωστό ότι η επικρατέστερη Quasi-Newton ενημέρωση είναι η BFGS. Για το λόγο αυτό θα μελετήσουμε αναλυτικά τα χαρακτηριστικά μεγέθη της BFGS ενημέρωσης, όπου ο αριθμός των αποθηκευμένων διανυσματικών ζευγών ενημέρωσης $\{(s_i, y_i)\}_{i=k-m+1}^k$ είναι $m = 1$ ή $m = 2$, ενώ ο αρχικός πίνακας $B^{(0)}$ κλιμακώνεται με μία συγκεκριμένη παράμετρο, τη BB φασματική παράμετρο.

Στο Κεφάλαιο 3 αναλύσαμε τα χαρακτηριστικά μεγέθη των Quasi-Newton ενημερώσεων ελάχιστης μνήμης. Σε όλες τις ενημερώσεις υποθέσαμε ότι ο αρχικός πίνακας $B^{(0)}$ είναι είτε ο μοναδιαίος ή κάποιο πολλαπλάσιό του, δηλαδή $B^{(0)} = \theta I$, όπου $\theta \in \mathbb{R}^*$. Ειδικά για τη BFGS

ενημέρωση, η πιο ευρέως χρησιμοποιούμενη παράμετρος θ που συναντάται στη διεθνή βιβλιογραφία [68] ορίζεται ως

$$\theta_{k+1} = \frac{y_k^T y_k}{s_k^T y_k}.$$

Τα τελευταία χρόνια όμως έχει αρχίσει να χρησιμοποιείται η φασματική παράμετρος των Barzilai & Borwein [15], γνωστή ως *BB spectral parameter*. Η παράμετρος αυτή είναι η αντίστροφη της ποσότητας Rayleigh $s^T G s / s^T s$, που παίρνει τιμές ανάμεσα στη μικρότερη και στη μεγαλύτερη ιδιοτιμή του μέσου εσσιανού πίνακα G και ορίζεται ως

$$\theta_{k+1} = \frac{s_k^T s_k}{s_k^T y_k}. \quad (4.1)$$

Η παράμετρος αυτή χρησιμοποιείται στον αντίστροφο του BFGS πίνακα,

$$B_{k+1}^{-1} \equiv H_{k+1} = H_k + \left(1 + \frac{y_k^T H_k y_k}{s_k^T y_k}\right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T H_k + H_k y_k s_k^T}{s_k^T y_k},$$

με αρχικό πίνακα τον $H^{(0)} = \theta I$. Επομένως, ο BFGS πίνακας

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

θα έχει ως αρχικό πίνακα τον $B^{(0)} = (1/\theta) I$.

Όταν $m = 1$, θα αναφερόμαστε στη συγκεκριμένη αυτή ενημέρωση ως 1BFGS και θα συμβολίζουμε τον αντίστοιχο πίνακα με $B^{(1)}$. Ο πίνακας αυτός έχει την έκφραση

$$B_{k+1}^{(1)} = \frac{1}{\theta_{k+1}} I - \frac{s_k s_k^T}{\theta_{k+1} s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \quad \text{όπου} \quad B_{k+1}^{(0)} = (1/\theta_{k+1}) I.$$

Το ζεύγος των διανυσμάτων που χρησιμοποιείται για την ενημέρωση του $B_{k+1}^{(0)}$ είναι το (s_k, y_k) .

Αντίστοιχα, όταν $m = 2$, θα ονομάζουμε την ενημέρωση ως 2BFGS και θα συμβολίζουμε τον αντίστοιχο πίνακα με $B^{(2)}$, του οποίου η έκφραση

είναι η εξής:

$$B_{k+1}^{(2)} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

όπου η έκφραση του $B_k^{(1)}$ είναι η ακόλουθη

$$B_k^{(1)} = \frac{1}{\theta_k} I - \frac{s_{k-1} s_{k-1}^T}{\theta_k s_{k-1}^T s_{k-1}} + \frac{y_{k-1} y_{k-1}^T}{s_{k-1}^T y_{k-1}},$$

στην οποία έχουμε αντικαταστήσει τον $B^{(0)}$ με $B_k^{(0)} = (1/\theta_k) I$. Είναι προφανές ότι η συγκεκριμένη ενημέρωση 2BFGS απαιτεί δύο ζεύγη διανυσμάτων. Τα ζεύγη αυτά είναι τα (s_{k-1}, y_{k-1}) και (s_k, y_k) , αντίστοιχα. Στο πρώτο ζεύγος είναι αποθηκευμένη η πληροφορία διόρθωσης από την προηγούμενη επανάληψη, ενώ στο δεύτερο είναι αποθηκευμένη η πληροφορία διόρθωσης της τρέχουσας επανάληψης.

Σκοπός μας είναι να μελετήσουμε τις συγκεκριμένες αυτές BFGS ενημερώσεις, λόγω της σπουδαιότητας τους στην επίλυση προβλημάτων βελτιστοποίησης. Απώτερος στόχος είναι να λύσουμε μεγάλης κλίμακας προβλήματα και ταυτόχρονα να ελαχιστοποιήσουμε το υπολογιστικό κόστος που απαιτείται. Πιο συγκεκριμένα, ενδιαφερόμαστε να υπολογίσουμε χωρίς αποθήκευση ή παραγοντοποίηση πινάκων τις ιδιοτιμές και τα ιδιοδιανύσματα των διακριτών ιδιοτιμών. Ακόμα θέλουμε να βρούμε σε κλειστή μορφή τους αντιστρόφους των πινάκων $B_{k+1}^{(1)} + \mu I$ και $B_{k+1}^{(2)} + \mu I$, $\mu \in \mathbb{R}$, ώστε να είμαστε σε θέση να λύσουμε γραμμικά συστήματα της μορφής $(B + \mu I)u = v$, που u, v n -διάστατα διανύσματα.

Στη συνέχεια αναλύουμε και τεκμηριώνουμε τον τρόπο εύρεσης των ποσοτήτων που μας ενδιαφέρουν. Βασίζομαστε στο ίχνος και την ορίζουσα των πινάκων που μελετάμε, καθώς και στον τρόπο που συνδέονται τα διανύσματα στα ζεύγη διανυσμάτων ενημέρωσης.

4.2 Η 1BFGS ενημέρωση

Η 1BFGS ενημέρωση ορίζεται ως

$$B_{k+1}^{(1)} = \frac{1}{\theta_{k+1}} I - \frac{s_k s_k^T}{\theta_{k+1} s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \quad (4.2)$$

όπου $B_{k+1}^{(0)} = (1/\theta_{k+1}) I$, με

$$\theta_{k+1} = \frac{s_k^T s_k}{s_k^T y_k},$$

την BB φασματική παράμετρο. Ο αντίστροφος αυτού του πίνακα είναι ο

$$(B_{k+1})^{-1} \equiv H_{k+1} = H_k - \frac{H_k y_k s_k^T + s_k y_k^T H_k}{s_k^T y_k} + \frac{s_k s_k^T}{s_k^T y_k}, \quad (4.3)$$

όπου ο αρχικός πίνακας $H^{(0)}$ ορίζεται ως $H^{(0)} = \theta I$.

Για να βρούμε το χαρακτηριστικό πολυώνυμο του πίνακα $B_{k+1}^{(1)}$ και κατά συνέπεια τις ιδιοτιμές του, χρειάζεται πρώτα να υπολογίσουμε το ίχνος και την ορίζουσά του. Χρησιμοποιώντας τις ιδιότητες του ίχνους (σελ. 13), εύκολα βρίσκουμε ότι

$$\text{tr} B_{k+1}^{(1)} = \frac{n-1}{\theta_{k+1}} + \frac{y_k^T y_k}{s_k^T y_k}. \quad (4.4)$$

Για να υπολογίσουμε την ορίζουσα, γράφουμε τον πίνακα (4.2) ως εξής:

$$B_{k+1}^{(1)} = \frac{1}{\theta_{k+1}} I \left(I - \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{\theta_{k+1} s_k^T y_k} \right).$$

Οπότε, από τις ιδιότητες της ορίζουσας (σελ. 13), έχουμε ότι

$$\det(B_{k+1}^{(1)}) = \det\left(\frac{1}{\theta_{k+1}} I\right) \det\left(I - \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{\theta_{k+1} s_k^T y_k}\right).$$

Θέτοντας

$$\begin{aligned} v_1 &= -\frac{s_k}{s_k^T s_k} \\ v_2 &= s_k \\ v_3 &= \frac{y_k}{\theta_{k+1} s_k^T y_k} \\ v_4 &= y_k \end{aligned}$$

και εφαρμόζοντας τη σχέση

$$\det(I + v_1 v_2^T + v_3 v_4^T) = (1 + v_1^T v_2)(1 + v_3^T v_4) - (v_1^T v_4)(v_2^T v_3),$$

έχουμε ότι

$$\det(B_{k+1}^{(1)}) = \frac{1}{\theta_{k+1}^n} \frac{s_k^T y_k}{\theta_{k+1} s_k^T s_k}.$$

Όμως, λόγω του ορισμού της φασματικής παραμέτρου $\theta_{k+1} = s_k^T s_k / s_k^T y_k$, που έχουμε υιοθετήσει, λαμβάνουμε ότι

$$\det(B_{k+1}^{(1)}) = 1/\theta_{k+1}^n, \quad (4.5)$$

Με τη βοήθεια των σχέσεων (4.4) και (4.5), θα υπολογίσουμε στη συνέχεια το χαρακτηριστικό πολυώνυμο του πίνακα $B_{k+1}^{(1)}$.

4.2.1 Το χαρακτηριστικό πολυώνυμο της 1BFGS ενημέρωσης

Η μεθοδολογία της εύρεσης του χαρακτηριστικού πολυωνύμου και των ιδιοτιμών της 1BFGS ενημέρωσης, είναι παρόμοια με αυτήν που αναπτύξαμε στο Κεφάλαιο 3, για τις Quasi-Newton ενημερώσεις ελάχιστης μνήμης. Όμως, όπως θα αποδείξουμε στη συνέχεια, η χρήση της BB φασματικής παραμέτρου, έχει ως αποτέλεσμα να απλοποιούνται οι εκφράσεις που αποτελούν το χαρακτηριστικό πολυώνυμο. Επιπλέον, όλες οι ιδιοτιμές του πίνακα $B_{k+1}^{(1)}$ έχουν το ίδιο πρόσημο. Κατά συνέπεια, η γνώση του προσήμου της παραμέτρου θ_{k+1} , μας εξασφαλίζει αν ο πίνακας $B_{k+1}^{(1)}$ είναι θετικά ορισμένος ή όχι. Επιπλέον, ένα σημαντικό συμπέρασμα που απορρέει από την ακόλουθη ανάλυση, είναι ότι οι ακραίες ιδιοτιμές του εν λόγω πίνακα, είναι πάντα διακριτές.

Λήμμα 4.1. Έστω ότι μία ενημέρωση εφαρμόζεται στο συμμετρικό πίνακα $B^{(0)} = (1/\theta)I$, χρησιμοποιώντας το διανυσματικό ζεύγος (s_k, y_k) και τον BFGS τύπο. Το χαρακτηριστικό πολυώνυμο του ενημερωμένου πίνακα $B_{k+1}^{(1)}$, είναι το

$$p_1(\lambda) = \left(\lambda - \frac{1}{\theta_{k+1}} \right)^{n-2} \left(\lambda^2 - \frac{a_k}{\theta_{k+1}} \lambda + \frac{1}{\theta_{k+1}^2} \right), \quad (4.6)$$

όπου $a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}$. Επιπλέον, αν $a_k > 2$, τότε

$$\lambda_1 < 1/\theta_{k+1} < \lambda_n,$$

όπου και οι λ_1 και λ_n είναι η μικρότερη και μεγαλύτερη ιδιοτιμή του $B_{k+1}^{(1)}$, αντίστοιχα.

Απόδειξη. Πρώτα θα δείξουμε ότι ο $B_{k+1}^{(1)}$ έχει το πολύ δύο διακριτές ιδιοτιμές. Γι αυτό το λόγο θεωρούμε τον πίνακα

$$\bar{B} = (1/\theta_{k+1})I - s_k s_k^T / \theta_{k+1} s_k^T s_k,$$

του οποίου η τάξη είναι $(n-1)$. Εφαρμόζοντας το θεώρημα συνδιαπλοκής ιδιοτιμών (σελ. 14) στον $\bar{B}$, είναι εύκολο να δούμε ότι ο $\bar{B}$ εκτός από τη μηδενική ιδιοτιμή, έχει ακόμα μία ιδιοτιμή, την $1/\theta_{k+1}$, πολλαπλότητας $(n-1)$. Αν ο $B_{k+1}^{(0)}$ είναι θετικά ορισμένος, τότε η πρόσθεση του πίνακα

$$\frac{y_k y_k^T}{s_k^T y_k}$$

στον πίνακα $\bar{B}$ έχει ως αποτέλεσμα τις παρακάτω ανισότητες ιδιοτιμών:

$$\lambda_n^{B^{(1)}} \geq 1/\theta_{k+1} \geq \lambda_{n-1}^{B^{(1)}} \geq 1/\theta_{k+1} \geq \dots \geq \lambda_2^{B^{(1)}} \geq 1/\theta_{k+1} \geq \lambda_1^{B^{(1)}} \geq 0,$$

όπου τα $\lambda_i^{B^{(1)}}$ δηλώνουν τις ιδιοτιμές του $B_{k+1}^{(1)}$. Αν ο $B_{k+1}^{(0)}$ δεν είναι θετικά ορισμένος, έχουμε ότι

$$0 \geq \lambda_n^{B^{(1)}} \geq 1/\theta_{k+1} \geq \lambda_{n-1}^{B^{(1)}} \geq 1/\theta_{k+1} \geq \dots \geq \lambda_2^{B^{(1)}} \geq 1/\theta_{k+1} \geq \lambda_1^{B^{(1)}}.$$

Και στις δύο περιπτώσεις, είναι προφανές ότι ισχύει

$$\lambda_2^{B^{(1)}} = \dots = \lambda_{n-1}^{B^{(1)}} = 1/\theta_{k+1}$$

και

$$\lambda_1^{B^{(1)}} \leq 1/\theta_{k+1} \leq \lambda_n^{B^{(1)}}. \quad (4.7)$$

Η σχέση (4.7) υποδηλώνει ότι ο $B_{k+1}^{(1)}$ έχει το πολύ δύο διακριτές ιδιοτιμές και μία πολλαπλή που ισούται με $1/\theta_{k+1}$, πολλαπλότητας το λιγότερο $(n-2)$. Μάλιστα, οι πιθανές διακριτές ιδιοτιμές, θα είναι οι ακραίες.

Δηλώνοντας με λ_x και λ_y τις δύο πιθανώς διακριτές ιδιοτιμές, έχουμε ότι το χαρακτηριστικό πολυώνυμο του $B_{k+1}^{(1)}$ είναι της μορφής

$$p_1(\lambda) = (\lambda - 1/\theta_{k+1})^{n-2} [\lambda^2 - (\lambda_x + \lambda_y)\lambda + \lambda_x \lambda_y]. \quad (4.8)$$

Επειδή όμως

$$\text{tr}(B_{k+1}^{(1)}) = \lambda_x + \lambda_y + \frac{n-2}{\theta_{k+1}},$$

έχουμε ότι

$$\lambda_x + \lambda_y = \text{tr}(B_{k+1}^{(1)}) - \frac{n-2}{\theta_{k+1}}.$$

Από τη σχέση (4.4) υπολογίζουμε ότι

$$\lambda_x + \lambda_y = \frac{1}{\theta_{k+1}} + \frac{y_k^T y_k}{s_k^T y_k} = \frac{1}{\theta_{k+1}} \left(1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k} \right).$$

Θέτοντας

$$a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k},$$

λαμβάνουμε ότι

$$\lambda_x + \lambda_y = a_k / \theta_{k+1}.$$

Για να υπολογίσουμε το γινόμενο των δύο άγνωστων ιδιοτιμών, λαμβάνουμε υπόψη μας ότι ισχύει η ισότητα

$$\det(B_{k+1}^{(1)}) = \lambda_x \lambda_y / \theta_{k+1}^{n-2},$$

οπότε έχουμε ότι

$$\lambda_x \lambda_y = \det(B_{k+1}^{(1)}) \theta_{k+1}^{n-2}.$$

Από τη σχέση (4.5) προκύπτει ότι

$$\lambda_x \lambda_y = 1 / \theta_{k+1}^2.$$

Συνεπώς, λαμβάνοντας υπόψη την ισότητα (4.8), η σχέση (4.6) έπεται αμέσως.

Όσον αφορά το δεύτερο μέρος του λήμματος, αναφορικά με την παράμετρο $a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}$, εύκολα συμπεραίνουμε ότι είναι κάτω φραγμένη από το 2. Πράγματι, έχουμε ότι

$$a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k} = 1 + \frac{\|s_k\|^2 \|y_k\|^2}{(s_k^T y_k)^2} = 1 + \frac{1}{\cos^2 \phi} \geq 2, \quad (4.9)$$

όπου ϕ είναι η γωνία των διανυσμάτων s_k και y_k . Αν $a_k = 2$, τότε τα

διανύσματα είναι συγγραμμικά και το χαρακτηριστικό πολυώνυμο παίρνει τη μορφή

$$p_1(\lambda) = (\lambda - 1/\theta_{k+1})^n,$$

ενώ ο πίνακας $B_{k+1}^{(1)}$ εκφυλίζεται στον κλιμακούμενο μοναδιαίο πίνακα

$$B_{k+1}^{(1)} = (1/\theta_{k+1})I = B_{k+1}^{(0)}.$$

Αν τώρα $a_k > 2$, τότε προφανώς τα διανύσματα είναι γραμμικώς ανεξάρτητα, ενώ το χαρακτηριστικό πολυώνυμο σε παραγοντοποιημένη μορφή γράφεται ως

$$p_1(\lambda) = (\lambda - 1/\theta_{k+1})^{n-2}(\lambda - \lambda_x)(\lambda - \lambda_y),$$

όπου οι ιδιοτιμές λ_x και λ_y , όπου

$$\lambda_{x,y} = \frac{a_k \pm \sqrt{a_k^2 - 4}}{2\theta_{k+1}}, \quad (4.10)$$

είναι προφανώς διακριτές, από την υπόθεση ότι $a_k > 2$. Από τις ανισότητες (4.7) έπεται ότι

$$\min(\lambda_x, \lambda_y) < \frac{1}{\theta_{k+1}} < \max(\lambda_x, \lambda_y),$$

από όπου θέτοντας $\lambda_1 = \min(\lambda_x, \lambda_y)$ και $\lambda_n = \max(\lambda_x, \lambda_y)$, προκύπτει ότι οι ακραίες ιδιοτιμές λ_1 και λ_n είναι διακριτές. $\square$

Είναι σαφές ότι το πρόσημο των διακριτών ιδιοτιμών εξαρτάται από την BB φασματική παράμετρο. Κατά συνέπεια, λαμβάνοντας υπόψη τη σχέση (4.10), μπορούμε να συμπεράνουμε ότι όλες οι ιδιοτιμές έχουν το ίδιο πρόσημο με τη φασματική παράμετρο θ_{k+1} .

Η τιμή της παραμέτρου a_k υποδηλώνει αν τα διανύσματα s_k, y_k είναι γραμμικώς ανεξάρτητα ή όχι. Συγκεκριμένα, αν $a_k > 2$, τότε είναι γραμμικώς ανεξάρτητα, ενώ αν $a_k = 2$, τα s_k, y_k είναι γραμμικώς εξαρτημένα διανύσματα. Στην πρώτη περίπτωση, ($a_k > 2$), οι δύο διακριτές ιδιοτιμές δίνονται από τη σχέση (4.10), ενώ στη δεύτερη περίπτωση ($a_k = 2$), η μόνη ιδιοτιμή του πίνακα είναι η $1/\theta_{k+1}$.

4.2.2 Ο πίνακας $(B_{k+1}^{(1)} + \mu I)^{-1}$

Αντικείμενο μελέτης της παρούσας υποενότητας είναι η εύρεση του αντιστρόφου του πίνακα $B_{k+1}^{(1)} + \mu I$, όπου $\mu \in \mathbb{R}$. Το παρακάτω λήμμα τεκμηριώνει ότι ο αντίστροφος του συγκεκριμένου πίνακα μπορεί να γραφεί σε κλειστή μορφή.

Λήμμα 4.2. Έστω Λ το σύνολο των ιδιοτιμών του $B_{k+1}^{(1)}$ με αντίθετα πρόσημα. Τότε, για κάθε $\mu \in \mathbb{R} \setminus \Lambda$, ο πίνακας $(B_{k+1}^{(1)} + \mu I)$ είναι αντιστρέψιμος και ο αντίστροφός του είναι ο πίνακας

$$\left(B_{k+1}^{(1)} + \mu I\right)^{-1} = \frac{1}{\gamma} \sum_{i=0}^2 (-1)^i \gamma_i(\mu) \left(B_{k+1}^{(1)}\right)^i, \quad (4.11)$$

όπου οι ποσότητες

$$\begin{aligned} \gamma(\mu) &= (1/\theta_{k+1} + \mu)(\mu^2 + a_k\mu, / \theta_{k+1} + 1/\theta_{k+1}^2), \\ \gamma_2(\mu) &= 1, \\ \gamma_1(\mu) &= \mu + (a_k + 1)/\theta_{k+1} \quad \text{και} \\ \gamma_0(\mu) &= \mu^2 + (a_k + 1)\mu/\theta_{k+1} + (a_k + 1)/\theta_{k+1}^2, \end{aligned}$$

είναι συναρτήσεις του πραγματικού αριθμού μ .

Απόδειξη. Έστω $\lambda_i, i = 1, \dots, n$ οι ιδιοτιμές του $B_{k+1}^{(1)}$ και έστω $\mu \in \mathbb{R}$. Αν $\mu \in \Lambda$, τότε προφανώς ο πίνακας

$$B_{k+1}^{(1)} + \mu I = B_{k+1}^{(1)} - \lambda_i I$$

είναι ιδιάζον. Οπότε, για να είναι ο $B_{k+1}^{(1)} + \mu I$ αντιστρέψιμος, θα πρέπει να ισχύει η σχέση $\mu \in \mathbb{R} \setminus \Lambda$. Χωρίς βλάβη της γενικότητας, υποθέτουμε ότι ο $B_{k+1}^{(1)}$ έχει δύο διακριτές ιδιοτιμές, έστω λ_x και λ_y . Συνεπώς, ο $(B_{k+1}^{(1)} + \mu I)$ έχει επίσης δύο διακριτές ιδιοτιμές, τις $\lambda_x + \mu$ και $\lambda_y + \mu$. Βασιζόμενοι στο Λήμμα 4.1, εύκολα βρίσκουμε ότι το χαρακτηριστικό πολυώνυμο είναι της μορφής

$$q(x) = \left[x - \left(\mu + \frac{1}{\theta_{k+1}} \right) \right]^{n-2} \left[x^2 - (\lambda_x + \lambda_y + 2\mu)x + \mu^2 + (\lambda_x + \lambda_y)\mu + \lambda_x \lambda_y \right].$$

Από το Λήμμα 4.1 γνωρίζουμε ότι

$$\lambda_x + \lambda_y = a_k/\theta_{k+1} \quad \text{και} \quad \lambda_x \lambda_y = 1/\theta_{k+1}^2.$$

Επομένως, το χαρακτηριστικό πολυώνυμο του $(B_{k+1}^{(1)} + \mu I)$ εκφράζεται από την παρακάτω σχέση:

$$q(x) = (x - c)^{n-2} (x^2 - c_1 x + c_0),$$

όπου

$$\begin{aligned} c &= 1/\theta_{k+1} + \mu \\ c_1 &= a_k/\theta_{k+1} + 2\mu \quad \text{και} \\ c_0 &= \mu^2 + a_k \mu / \theta_{k+1} + 1/\theta_{k+1}^2 \end{aligned}$$

Οπότε, το ελάχιστο πολυώνυμο είναι το

$$q_m(x) = (x - c) (x^2 - c_1 x + c_0).$$

Από το Θεώρημα Caley-Hamilton έχουμε ότι

$$q_m(B_{k+1}^{(1)} + \mu I) = 0.$$

Πολλαπλασιάζοντας και τα δύο μέλη της παραπάνω ισότητας με $(B_{k+1}^{(1)} + \mu I)^{-1}$, μετά από κάποιες αλγεβρικές πράξεις, βρίσκουμε ότι

$$\begin{aligned} cc_0 (B_{k+1}^{(1)} + \mu I)^{-1} &= (B_{k+1}^{(1)} + \mu I)^2 - (c + c_1)(B_{k+1}^{(1)} + \mu I) + (c_0 + cc_1)I \\ &= (B_{k+1}^{(1)})^2 - (c + c_1 - 2\mu)B_{k+1}^{(1)} \\ &\quad + [\mu^2 - (c + c_1)\mu + cc_0 c_1] I. \end{aligned}$$

Λύνοντας ως προς $(B_{k+1}^{(1)} + \mu I)^{-1}$ και αντικαθιστώντας τις εκφράσεις των c , c_1 και c_0 στην τελευταία εξίσωση, καταλήγουμε στη σχέση (4.11). $\square$

Όταν $a_k = 2$, δηλαδή όταν τα διανύσματα s_k και y_k είναι συγγραμμικά, είδαμε ότι η μόνη ιδιοτιμή του πίνακα είναι η $1/\theta_{k+1}$. Προφανώς, επειδή ο πίνακας $B_{k+1}^{(1)}$ είναι της μορφής $B_{k+1}^{(1)} = (1/\theta_{k+1})I$, ο αντίστροφος του $B_{k+1}^{(1)} + \mu I$ θα είναι ο πίνακας

$$\left(B_{k+1}^{(1)} + \mu I\right)^{-1} = \frac{\theta_{k+1}}{1 + \mu\theta_{k+1}}I.$$

4.2.3 Τα ιδιοδιανύσματα των ακραίων ιδιοτιμών

Στη συνέχεια θα μελετήσουμε τα ιδιοδιανύσματα των ακραίων ιδιοτιμών της 1BFGS ενημέρωσης. Επειδή το πλήθος των διακριτών ιδιοτιμών εξαρτάται από την τιμή της παραμέτρου a_k , διακρίνουμε δύο περιπτώσεις.

Πρώτα θεωρούμε την περίπτωση όπου $a_k = 2$, δηλαδή την περίπτωση όπου τα s_k , y_k είναι συγγραμμικά. Χρησιμοποιώντας την ιδιοανάλυση του $B^{(1)}$ έχουμε ότι

$$B^{(1)} = U\Lambda U^T,$$

όπου

$$U = I \quad \text{και} \quad \Lambda = \text{diag}(\lambda_1, \lambda_1, \dots, \lambda_1).$$

Είναι εύκολο να δούμε ότι το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 είναι το

$$u_1 = e_1 = (1, 0, \dots, 0)^T.$$

Όσον αφορά τη δεύτερη περίπτωση, όπου $a_k > 2$, από το Λήμμα 4.1 γνωρίζουμε ότι οι ακραίες ιδιοτιμές είναι διακριτές. Οπότε, για να βρούμε τα ιδιοδιανύσματα των ακραίων ιδιοτιμών, μπορούμε είτε να χρησιμοποιήσουμε την Πρόταση 3.2 της Ενότητας 3.3 (σελίδα 96) ή να κάνουμε χρήση της μεθόδου της αντίστροφης δύναμης. Στην πρώτη περίπτωση, η Πρόταση 3.2 τεκμηριώνει ότι τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές είναι της μορφής

$$u(\lambda) = \frac{(\lambda s_k - y_k)^T y_k}{(\lambda s_k - y_k)^T s_k} s_k - y_k,$$

όπου λ διακριτή ιδιοτιμή. Στη δεύτερη περίπτωση, η χρήση της μεθόδου της αντίστροφης δύναμης συνδυάζεται με το Λήμμα 4.2. Χρησιμοποιώντας αυτά τα δύο, συμπεραίνουμε ότι τα ιδιοδιανύσματα που αντιστοιχούν στις διακριτές ιδιοτιμές του $B_{k+1}^{(1)}$ είναι της μορφής

$$\begin{aligned}\hat{u}(\hat{\lambda}) &= \sum_{i=0}^2 (-1)^i \gamma_i(\hat{\lambda}) \left(B_{k+1}^{(1)}\right)^i \frac{u}{\gamma(\hat{\lambda})} \\ &= -\gamma_u(\hat{\lambda}) u + \gamma_{us}(\hat{\lambda}) s_k - \gamma_{uy}(\hat{\lambda}) y_k,\end{aligned}\quad (4.12)$$

όπου $\hat{\lambda} = -\lambda + \epsilon$, λ διακριτή ιδιοτιμή, u είναι το κανονικοποιημένο διάνυσμα $(1, 1, \dots, 1)^T$ και οι συντελεστές γ_u , γ_{us} και γ_{uy} ορίζονται ως

$$\begin{aligned}\gamma_u(\hat{\lambda}) &= \left[1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + \gamma_0(\hat{\lambda}) \theta_{k+1}^2\right] / \left[\gamma(\hat{\lambda}) \theta_{k+1}^2\right], \\ \gamma_{us}(\hat{\lambda}) &= \left\{ \left[1 - \gamma_1(\hat{\lambda}) \theta_{k+1}\right] s_k^T u + \theta_{k+1} y_k^T u \right\} / \left[\gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T s_k\right] \text{ και,} \\ \gamma_{uy}(\hat{\lambda}) &= \left\{ \left[1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + a_k\right] \theta_{k+1} y_k^T u - s_k^T u \right\} / \left[\gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T y_k\right].\end{aligned}$$

Οι συντελεστές αυτοί υπολογίζονται αν αντικαταστήσουμε στη σχέση

$$\hat{u}(\hat{\lambda}) = \sum_{i=0}^2 (-1)^i \gamma_i(\hat{\lambda}) \left(B_{k+1}^{(1)}\right)^i \frac{u}{\gamma(\hat{\lambda})},$$

τις παρακάτω εξισώσεις:

$$B_{k+1}^{(1)} u = \frac{u}{\theta_{k+1}} - \frac{s_k^T u}{\theta_{k+1} s_k^T s_k} s_k + \frac{y_k^T u}{s_k^T y_k} y_k,$$

και

$$(B_{k+1}^{(1)})^2 u = \frac{B_{k+1}^{(1)} u}{\theta_{k+1}} - \frac{s_k^T B_{k+1}^{(1)} u}{\theta_{k+1} s_k^T s_k} s_k + \frac{y_k^T B_{k+1}^{(1)} u}{s_k^T y_k} y_k,$$

με $s_k^T B_{k+1}^{(1)} u = y_k^T u$ και

$$y_k^T B_{k+1}^{(1)} u = \frac{y_k^T u}{\theta_{k+1}} - \frac{s_k^T u}{\theta_{k+1} s_k^T s_k} s_k^T y_k + \frac{y_k^T u}{s_k^T y_k} y_k^T y_k.$$

Επίσης, θεωρώντας ότι το $u = \frac{1}{\sqrt{n}}(1, 1, \dots, 1)^T$ είναι το κανονικοποιημένο αρχικό διάνυσμα με $\|u\| = 1$, καταλήγουμε στο παρακάτω πόρισμα:

Πόρισμα 4.1. Έστω $\hat{\lambda} = -\lambda + \epsilon$, όπου $\epsilon \rightarrow 0$, μια ϵ -διαταραγμένη διακριτή

ιδιοτιμή του 1BFGS πίνακα με αντίθετο πρόσημο, και το κανονικοποιημένο διάνυσμα $u = \frac{1}{\sqrt{n}}(1, 1, \dots, 1)^T$. Τότε, το ιδιοδιάνυσμα που αντιστοιχεί στη $\hat{\lambda}$ είναι το διάνυσμα

$$\hat{u}(\hat{\lambda}) = -\gamma_u(\hat{\lambda}) u + \gamma_{us}(\hat{\lambda}) s_k - \gamma_{uy}(\hat{\lambda}) y_k,$$

όπου οι συντελεστές των διανυσμάτων u , s_k και y_k ορίζονται από τις σχέσεις

$$\begin{aligned} \gamma_u(\hat{\lambda}) &= \frac{1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + \gamma_0(\hat{\lambda}) \theta_{k+1}^2}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2}, \\ \gamma_{us}(\hat{\lambda}) &= \frac{1 - \gamma_1(\hat{\lambda}) \theta_{k+1}}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T s_k} \sum_{i=1}^n s_i + \frac{1}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T s_k} \sum_{i=1}^n y_i, \quad \text{και} \\ \gamma_{uy}(\hat{\lambda}) &= \frac{[1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + a_k] \theta_{k+1}}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T y_k} \sum_{i=1}^n y_i - \frac{1}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T y_k} \sum_{i=1}^n s_i, \end{aligned}$$

αντίστοιχα, ενώ οι ποσότητες $\gamma_0(\hat{\lambda})$, $\gamma_1(\hat{\lambda})$ και $\gamma(\hat{\lambda})$ ορίζονται όπως στο Λήμμα 4.2.

4.3 Η 2BFGS ενημέρωση

Στην ενότητα αυτή θα μελετήσουμε τη BFGS ενημέρωση μικρής μνήμης, δηλαδή την περίπτωση όπου ο αριθμός των αποθηκευμένων διανυσματικών ζευγών είναι $m = 2$. Την ενημέρωση αυτή θα την ονομάζουμε 2BFGS, της οποίας η έκφραση είναι

$$B_{k+1}^{(2)} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k} + \frac{y_k y_k^T}{s_k^T y_k}. \quad (4.13)$$

Η πληροφορία παρέχεται από τις δύο προηγούμενες επαναλήψεις και είναι αποθηκευμένη στα διανυσματικά ζεύγη $\{s_{k-1}, y_{k-1}\}$ και $\{s_k, y_k\}$. Ο πίνακας $B_k^{(1)}$ είναι προφανώς η 1BFGS ενημέρωση, που ορίζεται στη σχέση (4.2). Ο αντίστροφος της 2BFGS είναι ο πίνακας

$$H_{k+1}^{(2)} = H_k^{(1)} + \left(1 + \frac{y_k^T H_k^{(1)} y_k}{s_k^T y_k} \right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T H_k^{(1)} + H_k^{(1)} y_k s_k^T}{s_k^T y_k}, \quad (4.14)$$

που προκύπτει άμεσα μέσω της σχέσης (2.23) (σελ. 42).

Στη συνέχεια θα ακολουθήσουμε τα ίδια βήματα ανάλυσης που προηγήθηκαν και για την 1BFGS ενημέρωση, προκειμένου να παράγουμε σε κλειστούς τύπους τα χαρακτηριστικά μεγέθη που μας ενδιαφέρουν. Είναι γνωστό (βλέπε [68]) ότι το ίχνος και η ορίζουσα της BFGS ενημέρωσης, δίνονται από τους τύπους

$$\text{tr}(B_{k+1}) = \text{tr}(B_k) - \frac{\|B_k s_k\|^2}{s_k^T B_k s_k} + \frac{y_k^T y_k}{s_k^T y_k} \quad (4.15)$$

και

$$\det(B_{k+1}) = \det(B_k) \frac{s_k^T y_k}{s_k^T B_k s_k}, \quad (4.16)$$

αντίστοιχα. Προφανώς, οι ίδιοι τύποι ισχύουν και για την 2BFGS ενημέρωση.

4.3.1 Οι ιδιοτιμές της 2BFGS ενημέρωσης

Στη συνέχεια θα υπολογίσουμε το χαρακτηριστικό πολυώνυμο του πίνακα (4.13) και θα αποδείξουμε ότι ο αριθμός των διακριτών ιδιοτιμών του εξαρτάται από τη γραμμική ανεξαρτησία των διανυσματικών ζευγών $\{s_{k-1}, y_{k-1}\}$ και $\{s_k, y_k\}$. Επιπλέον, θα τεκμηριώσουμε ότι αν ο πίνακας έχει τουλάχιστον δύο διακριτές ιδιοτιμές, τότε αυτές είναι οι ακραίες ιδιοτιμές του $B_{k+1}^{(2)}$.

Προφανώς, αφού ο $B_{k+1}^{(2)}$ δημιουργείται χρησιμοποιώντας τέσσερα διανύσματα για την ενημέρωση του αρχικού πίνακα $B^{(0)}$, περιμένουμε ότι η 2BFGS ενημέρωση θα έχει το πολύ τέσσερις (4) διακριτές ιδιοτιμές και μία πολλαπλή, την $1/\theta_{k+1}$, πολλαπλότητας το λιγότερο $n - 4$. Αυτό το συμπέρασμα αποδεικνύεται στο παρακάτω λήμμα.

Λήμμα 4.3. Έστω ότι δύο ενημερώσεις εφαρμόζονται στο συμμετρικό πίνακα $B^{(0)}$ χρησιμοποιώντας τα διανυσματικά ζεύγη $\{s_i, y_i\}_{i=k-1}^k$ και τον τύπο της BFGS. Το χαρακτηριστικό πολυώνυμο του ενημερωμένου πίνακα $B_{k+1}^{(2)}$ έχει τη γενική μορφή

$$p_2(\lambda) = (\lambda - 1/\theta_k)^{n-4} (\lambda^4 - \beta_3 \lambda^3 + \beta_2 \beta_0 \lambda^2 - \beta_1 \beta_0 \lambda + \beta_0), \quad (4.17)$$

όπου οι συντελεστές β_i , $i = 0, \dots, 3$ ορίζονται ως

$$\beta_0 = \frac{1}{\theta_k^4} \frac{s_k^T y_k}{s_k^T B_k^{(1)} s_k}, \quad \beta_1 = (a_{k-1} + 2)\theta_k - 2 \frac{s_k^T w_k}{s_k^T y_k} + b_k \theta_{k+1}, \quad (4.18)$$

$$\begin{aligned} \beta_2 = & \theta_k (\beta_1 - a_{k-1} \theta_k) (a_{k-1} + 2) - 2\theta_k^2 + \left(\frac{s_k^T w_k}{s_k^T y_k} \right)^2 - \theta_{k+1} \frac{w_k^T w_k}{s_k^T y_k} \\ & + 2 \frac{s_k^T H_k w_k}{s_k^T y_k} - b_k \frac{s_k^T H_k s_k}{s_k^T y_k} \quad \text{και} \end{aligned} \quad (4.19)$$

$$\beta_3 = \frac{a_{k-1} + 2}{\theta_k} - \frac{\|B_k^{(1)} s_k\|^2}{s_k^T B_k^{(1)} s_k} + \frac{\|y_k\|^2}{s_k^T y_k}, \quad (4.20)$$

$$\mu_k H_k = \left(B_k^{(1)} \right)^{-1}, \quad w_k = H_k y_k \quad \text{και} \quad b_k = 1 + \frac{y_k^T w_k}{s_k^T y_k}.$$

Απόδειξη. Πρώτα θα δείξουμε ότι ο $B_{k+1}^{(2)}$ έχει το πολύ τέσσερις διακριτές ιδιοτιμές. Χωρίς βλάβη της γενικότητας, θεωρούμε ότι οι $B_k^{(1)}$ και $B_{k+1}^{(2)}$ είναι θετικά ορισμένοι πίνακες (σε διαφορετική περίπτωση, η απόδειξη γίνεται με παρόμοιο τρόπο). Έστω $\lambda_i^{B^{(1)}}$ και $\lambda_i^{\bar{B}}$ οι ιδιοτιμές των πινάκων $B_k^{(1)}$ και

$$\bar{B} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k},$$

αντίστοιχα. Επειδή

$$\lambda_2^{B^{(1)}} = \dots = \lambda_{n-1}^{B^{(1)}} = 1/\theta_k,$$

από το θεώρημα συνδιαπλοκής ιδιοτιμών έχουμε ότι

$$\lambda_n^{B^{(1)}} \geq \lambda_n^{\bar{B}} \geq 1/\theta_k \geq \lambda_{n-1}^{\bar{B}} \geq 1/\theta_k \geq \dots \geq 1/\theta_k \geq \lambda_2^{\bar{B}} \geq \lambda_1^{B^{(1)}} \geq \lambda_1^{\bar{B}}.$$

Προσθέτοντας τον πίνακα πρώτης τάξης $(y_k y_k^T)/(s_k^T y_k)$ στον πίνακα $\bar{B}$, παίρνουμε

$$\lambda_n^{B^{(2)}} \geq \lambda_n^{\bar{B}} \geq \lambda_{n-1}^{B^{(2)}} \geq 1/\theta_k \geq \dots \geq 1/\theta_k \geq \lambda_2^{B^{(2)}} \geq \lambda_2^{\bar{B}} \geq \lambda_1^{B^{(2)}} \geq \lambda_1^{\bar{B}},$$

όπου $\lambda_i^{B^{(2)}}$ είναι οι ιδιοτιμές του $B_{k+1}^{(2)}$. Από τις πιο πάνω ανισώσεις, προ-

κύπτει ότι

$$\lambda_n^{B^{(2)}} \geq \lambda_{n-1}^{B^{(2)}} \geq 1/\theta_k \geq \lambda_2^{B^{(2)}} \geq \lambda_1^{B^{(2)}}, \quad (4.21)$$

και άρα ο πίνακας $B_{k+1}^{(2)}$ έχει το πολύ τέσσερις διακριτές ιδιοτιμές. Αν συμβολίσουμε με $\lambda_1, \lambda_2, \lambda_3$ και λ_4 αυτές τις ιδιοτιμές, τότε το χαρακτηριστικό πολυώνυμο του $B_{k+1}^{(2)}$ μπορεί να γραφεί ως εξής:

$$p_2(\lambda) = (\lambda - 1/\theta_k)^{n-4} (\lambda^4 - c_3\lambda^3 + c_2\lambda^2 - c_1\lambda + c_0),$$

όπου

$$c_3 = \sum_{i=1}^4 \lambda_i, \quad c_2 = \sum_{\substack{i,j \\ i < j}} \lambda_i \lambda_j, \quad c_1 = \sum_{\substack{i,j,\ell \\ i < j < \ell}} \lambda_i \lambda_j \lambda_\ell, \quad \text{και} \quad c_0 = \prod_{i=1}^4 \lambda_i.$$

Επιπλέον, λόγω ότι $\sum_{i=1}^4 \lambda_i = \text{tr} \left(B_{k+1}^{(2)} \right) - \frac{n-4}{\theta_k}$, προκύπτει άμεσα ότι

$$c_3 = \text{tr} \left(B_{k+1}^{(2)} \right) - \frac{n-4}{\theta_k}. \quad (4.22)$$

Ενώ, επειδή $\prod_{i=1}^4 \lambda_i = \det \left(B_{k+1}^{(2)} \right) \theta_k^{n-4}$, συνεπάγεται ότι,

$$c_0 = \det \left(B_{k+1}^{(2)} \right) \theta_k^{n-4}. \quad (4.23)$$

Με απλή εφαρμογή της ταυτότητας

$$(a + b + c + d)^2 = a^2 + b^2 + c^2 + d^2 + 2(ab + ac + ad + bc + bd + cd),$$

για $c_3 = \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4$, έχουμε ότι

$$2c_2 = c_3^2 - \text{tr} \left(B_{k+1}^{(2)} \right)^2 - \frac{n-4}{\theta_k^2}. \quad (4.24)$$

Τέλος, για το c_1 έχουμε ότι

$$\begin{aligned} c_1 &= \det \left(B_{k+1}^{(2)} \right) \theta_k^{n-4} \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \frac{1}{\lambda_4} \right) \\ &= \det \left(B_{k+1}^{(2)} \right) \theta_k^{n-4} \left[\text{tr} \left(H_{k+1}^{(2)} \right) - (n-4)\theta_k \right]. \end{aligned} \quad (4.25)$$

Χρησιμοποιώντας τις εξισώσεις (4.15) και (4.16), καθώς και τις (4.22)-(4.25), προκύπτει η έκφραση του χαρακτηριστικού πολυωνύμου (4.17) με τις παραμέτρους $\beta_0, \beta_1, \beta_2$ και β_3 όπως ορίζονται στη διατύπωση του Λήμματος. $\square$

Στην περίπτωση όπου ο $B_{k+1}^{(2)}$ έχει ακριβώς τέσσερις διακριτές ιδιοτιμές, αυτές μπορούν να υπολογιστούν αναλυτικά, λύνοντας την τετάρτου βαθμού εξίσωση

$$\lambda^4 - \beta_3 \lambda^3 + \beta_2 \beta_0 \lambda^2 - \beta_1 \beta_0 \lambda + \beta_0 = 0,$$

χρησιμοποιώντας γνωστές μεθόδους [19, 32]. Εντούτοις, ο ακριβής αριθμός των διακριτών ιδιοτιμών εξαρτάται αποκλειστικά από τη γραμμική ανεξαρτησία ή μη, των διανυσματικών ζευγών $\{s_i, y_i\}_{i=k-1}^k$, για την οποία μπορούμε εύκολα να αποφανθούμε μέσω της τιμής της παραμέτρου a_i ,

$$a_i = 1 + \theta_{i+1} \frac{y_i^T y_i}{s_i^T y_i} = 1 + \frac{1}{\cos^2 \phi},$$

όπου ϕ η γωνία των διανυσμάτων s_i και y_i . Προφανώς, όταν $a_i = 2$, τότε τα διανύσματα s_i, y_i είναι γραμμικά εξαρτημένα. Αλλιώς, αν $a_i > 2$, τότε τα s_i, y_i είναι γραμμικά ανεξάρτητα. Στην πρόταση που ακολουθεί διερευνούμε σχολαστικά όλες τις δυνατές περιπτώσεις που μπορούν να προκύψουν όσον αφορά τις τιμές των παραμέτρων a_{k-1} και a_k . Σε κάθε περίπτωση εξάγουμε τον ακριβή αριθμό των διακριτών ιδιοτιμών του $B_{k+1}^{(2)}$.

Πρόταση 4.1. Έστω ο συμμετρικός πίνακας $B_{k+1}^{(2)} \in \mathbb{R}^{n \times n}$. Τότε,

1. Αν $a_{k-1} = 2$, ο $B_{k+1}^{(2)}$ έχει το πολύ δύο διακριτές ιδιοτιμές. Συγκεκριμένα,
 - (i) αν $a_k > 2$, τότε υπάρχουν ακριβώς δύο διακριτές ιδιοτιμές,
 - (ii) αν $a_k = 2$ και $\theta_k \neq \theta_{k+1}$, τότε υπάρχει ακριβώς μία διακριτή ιδιοτιμή, και
 - (iii) αν $a_k = 2$ και $\theta_k = \theta_{k+1}$, τότε όλες οι ιδιοτιμές είναι ίσες με $1/\theta_k$ (δεν υπάρχουν διακριτές ιδιοτιμές).
2. Αν $a_{k-1} > 2$, τότε ο $B_{k+1}^{(2)}$ έχει τουλάχιστον τρεις διακριτές ιδιοτιμές. Συγκεκριμένα,

- (i) αν $a_k = 2$, τότε υπάρχουν ακριβώς τρεις διακριτές ιδιοτιμές, ενώ
(ii) αν $a_k > 2$, τότε ο $B_{k+1}^{(2)}$ έχει ακριβώς τέσσερις διακριτές ιδιοτιμές.

Απόδειξη. Έστω ότι ισχύει $a_{k-1} = 2$. Από το Λήμμα 4.1 έχουμε ότι σε αυτή την περίπτωση ο πίνακας $B_k^{(1)}$ είναι της μορφής

$$B_k^{(1)} = \left(\frac{1}{\theta_k} \right) I,$$

και συνεπώς,

$$B_{k+1}^{(2)} = \frac{1}{\theta_k} I - \frac{1}{\theta_k} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k}.$$

Σε αυτή την περίπτωση, το χαρακτηριστικό πολυώνυμο του $B_{k+1}^{(2)}$ λαμβάνει τη μορφή

$$p_2(\lambda) = \left(\lambda - \frac{1}{\theta_k} \right)^{n-2} \left[\lambda^2 - \left(\frac{a_k - 1}{\theta_{k+1}} + \frac{1}{\theta_k} \right) \lambda + \frac{1}{\theta_k \theta_{k+1}} \right], \quad (4.26)$$

από όπου έπεται ότι ο $B_{k+1}^{(2)}$ έχει το πολύ δύο διακριτές ιδιοτιμές. Θα διερευνήσουμε στη συνέχεια τον ακριβή αριθμό των διακριτών ιδιοτιμών εξετάζοντας την παράμετρο a_k . Για το σκοπό αυτό θεωρούμε τη δευτεροβάθμια εξίσωση

$$\lambda^2 - \left(\frac{a_k - 1}{\theta_{k+1}} + \frac{1}{\theta_k} \right) \lambda + \frac{1}{\theta_k \theta_{k+1}} = 0,$$

η οποία έχει τις ακόλουθες λύσεις:

$$\lambda_{1,2} = \frac{\theta_{k+1} + \theta_k (a_k - 1) \pm \sqrt{(\theta_k - \theta_{k+1})^2 + \theta_k (a_k - 2)(a_k \theta_k + 2\theta_{k+1})}}{2\theta_k \theta_{k+1}}.$$

Όπως μπορεί να παρατηρήσει κανείς, η έκφραση των λύσεων είναι συνάρτηση των τριών παραμέτρων a_k , θ_k και θ_{k+1} . Έστω ότι $a_k > 2$ και υποθέτουμε ότι καμία από τις λ_1 και λ_2 δεν ισούται με $1/\theta_k$. Λύνοντας τις εξισώσεις

$$\lambda_i - 1/\theta_k = 0$$

ως προς a_k , βρίσκουμε ότι $a_k = 2$, το οποίο είναι άτοπο. Οπότε, όταν $a_k > 2$, ο πίνακας έχει ακριβώς δύο διακριτές ιδιοτιμές. Αν $a_k = 2$, τότε

το χαρακτηριστικό πολυώνυμο γίνεται

$$p_2(\lambda) = (\lambda - 1/\theta_k)^{n-1}(\lambda - 1/\theta_{k+1}).$$

Άρα, στην περίπτωση όπου $\theta_k \neq \theta_{k+1}$, η μόνη διακριτή ιδιοτιμή είναι η

$$\frac{1}{\theta_{k+1}}.$$

Σε αντίθετη περίπτωση, ο $B_{k+1}^{(2)}$ δεν έχει διακριτές ιδιοτιμές.

Θεωρούμε τώρα την περίπτωση όπου $a_{k-1} > 2$. Στο Λήμμα 4.1 έχουμε αποδείξει ότι τότε ο $B_k^{(1)}$ έχει δύο διακριτές ιδιοτιμές. Άρα, ο πίνακας

$$\bar{B} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k}$$

εκτός από τη μηδενική ιδιοτιμή, έχει άλλες δύο διακριτές. Συνεπώς, ο

$$B_{k+1}^{(2)} = \bar{B} + \frac{y_k y_k^T}{s_k^T y_k},$$

έχει το πολύ τρεις διακριτές ιδιοτιμές. Αν $a_k = 2$, τότε $s_k = \theta_{k+1} y_k$ και ο πίνακας $B_{k+1}^{(2)}$ γίνεται

$$B_{k+1}^{(2)} = B_k^{(1)} - \frac{B_k^{(1)} y_k y_k^T B_k^{(1)}}{y_k^T B_k^{(1)} y_k} + \frac{y_k y_k^T}{\theta_{k+1} y_k^T y_k}.$$

Οπότε,

$$\bar{B} = B_k^{(1)} - \frac{B_k^{(1)} y_k y_k^T B_k^{(1)}}{y_k^T B_k^{(1)} y_k}.$$

Επομένως, η πρόσθεση του πίνακα $(y_k y_k^T)/(\theta_{k+1} y_k^T y_k)$ στον $\bar{B}$ έχει ως αποτέλεσμα τη μεταβολή της μηδενικής ιδιοτιμής του $\bar{B}$, σε $1/\theta_{k+1}$, ενώ οι υπόλοιπες ιδιοτιμές μένουν αμετάβλητες. Άρα, ο $B_{k+1}^{(2)}$ έχει τόσες διακριτές ιδιοτιμές όσες και ο $\bar{B}$, δηλαδή τρεις. Επειδή μία από αυτές είναι η $1/\theta_{k+1}$, χρησιμοποιώντας το Λήμμα 4.3, βρίσκουμε ότι το χαρακτηριστικό πολυώνυμο του $B_{k+1}^{(2)}$ είναι το

$$p_2(\lambda) = (\lambda - 1/\theta_k)^{n-3} (\lambda - 1/\theta_{k+1}) (\lambda^2 - c_1 \lambda + c_2),$$

όπου

$$\begin{aligned} c_1 &= a_{k-1}/\theta_k - \|B_k^{(1)} y_k\|^2 / y_k^T B_k^{(1)} y_k \quad \text{και} \\ c_2 &= y_k^T y_k / (\theta_k^2 \theta_{k+1} y_k^T B_k^{(1)} y_k). \end{aligned}$$

Λύνοντας τη δευτεροβάθμια εξίσωση

$$\lambda^2 - c_1 \lambda + c_2 = 0,$$

βρίσκουμε τις άλλες δύο διακριτές ιδιοτιμές. Τέλος, αν $a_k > 2$, τότε το y_k δεν είναι ιδιοδιάνυσμα του

$$\bar{B} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k}.$$

Άρα, η πρόσθεση του πίνακα $(y_k y_k^T) / (s_k^T y_k)$ έχει ως αποτέλεσμα να προκύψει ακόμα μια διακριτή ιδιοτιμή του $B_{k+1}^{(2)}$. $\square$

Το παρακάτω λήμμα τεκμηριώνει ότι αν ο πίνακας $B_{k+1}^{(2)}$ έχει τουλάχιστον δύο διακριτές ιδιοτιμές, αυτές θα είναι οι ακραίες ιδιοτιμές του.

Λήμμα 4.4. *Αν ο πίνακας $B_{k+1}^{(2)}$ έχει το λιγότερο δύο διακριτές ιδιοτιμές, τότε ισχύει ότι*

$$\lambda_1 < 1/\theta_k < \lambda_n, \quad (4.27)$$

όπου οι λ_1 και λ_n είναι η μικρότερη και μεγαλύτερη ιδιοτιμή του $B_{k+1}^{(2)}$, αντίστοιχα.

Απόδειξη. Όταν ο $B_{k+1}^{(2)}$ έχει το λιγότερο τρεις διακριτές ιδιοτιμές, από τη σχέση (4.21) είναι αυταπόδεικτο ότι η σχέση (4.27) ισχύει. Αν ο $B_{k+1}^{(2)}$ έχει δύο διακριτές ιδιοτιμές, τότε θα πρέπει ένα ζεύγος από το διανυσματικό σύνολο $\{(s_{k-1}, y_{k-1}), (s_k, y_k)\}$ να αποτελείται από γραμμικά εξαρτημένα διανύσματα. Λόγω της Πρότασης 4.1, αυτά τα δύο διανύσματα θα πρέπει να είναι τα s_{k-1} και y_{k-1} , πράγμα που σημαίνει ότι ο $B_k^{(1)}$ δεν έχει διακριτές ιδιοτιμές. Υπό αυτές τις συνθήκες, εφαρμόζοντας το θεώρημα συνδιαπλοκής ιδιοτιμών, λαμβάνουμε τη σχέση (4.27). $\square$

4.3.2 Ο πίνακας $(B_{k+1}^{(2)} + \mu I)^{-1}$

Στη συνέχεια θα υπολογίσουμε τον αντίστροφο του πίνακα $B_{k+1}^{(2)} + \mu I$, $\mu \in \mathbb{R}$. Όπως ήδη έχουμε δει, ο πίνακας αυτός χρειάζεται τόσο στην επίλυση γραμμικών συστημάτων της μορφής $(B + \mu I)u = v$, $u, v \in \mathbb{R}^n$, όσο και κατά τον υπολογισμό ιδιοδιανυσμάτων διακριτών ιδιοτιμών.

Λήμμα 4.5. Έστω Λ το σύνολο των ιδιοτιμών του $B_{k+1}^{(2)}$ με αντίθετα πρόσημα. Τότε, για κάθε $\mu \in \mathbb{R} \setminus \Lambda$, ο πίνακας $(B_{k+1}^{(2)} + \mu I)$ είναι αντιστρέψιμος. Ο αντίστροφός του μπορεί να εκφραστεί από τον τύπο

$$(B_{k+1}^{(2)} + \mu I)^{-1} = \frac{1}{\nu(\mu)} \sum_{i=0}^4 (-1)^i \nu_i(\mu) (B_{k+1}^{(2)})^i, \quad (4.28)$$

όπου οι ποσότητες

$$\begin{aligned} \nu_4(\mu) &= 1, \\ \nu_3(\mu) &= \mu + \beta_3 + 1/\theta_k, \\ \nu_2(\mu) &= \mu\nu_3(\mu) + \beta_2\beta_0 + \beta_3/\theta_k, \\ \nu_1(\mu) &= \mu\nu_2(\mu) + \beta_1\beta_0 + \beta_2\beta_0/\theta_k, \\ \nu_0(\mu) &= \mu\nu_1(\mu) + \beta_0 + \beta_0\beta_1/\theta_k \quad \text{και} \\ \nu(\mu) &= \mu\nu_0(\mu) + \beta_0/\theta_k, \end{aligned}$$

είναι συναρτήσεις του μ , ενώ οι σταθερές β_0 , β_1 , β_2 και β_3 ορίζονται στο Λήμμα 4.3.

Απόδειξη. Προφανώς, όπως και στην απόδειξη του Λήμματος 4.2 (σελ. 133), για να είναι ο $B_{k+1}^{(2)} + \mu I$ αντιστρέψιμος, η σχέση $\mu \in \mathbb{R} \setminus \Lambda$ πρέπει να ισχύει. Από την Πρόταση 4.1, γνωρίζουμε ότι ο $B_{k+1}^{(2)}$ έχει το πολύ τέσσερις διακριτές ιδιοτιμές, έστω τις λ_i , $i = 1, \dots, 4$. Συνεπώς, ο πίνακας $(B_{k+1}^{(2)} + \mu I)$ έχει επίσης το πολύ τέσσερις διακριτές ιδιοτιμές, τις $\lambda_i + \mu$. Επομένως, αν x είναι ιδιοτιμή του $(B_{k+1}^{(2)} + \mu I)$, το χαρακτηριστικό του πολυώνυμο θα είναι της μορφής

$$q(x) = (x - c)^{n-4} (x^4 - c_3x^3 + c_2x^2 - c_1x + c_0).$$

Χρησιμοποιώντας το Λήμμα 4.3 (σελίδα 138), βρίσκουμε ότι

$$\begin{aligned} c &= \theta_k^{-1} + \mu, \\ c_0 &= \mu^4 + \beta_3\mu^3 + \beta_2\beta_0\mu^2 + \beta_1\beta_0\mu + \beta_0, \\ c_1 &= 4\mu^3 + 3\beta_3\mu^2 + 2\beta_2\beta_0\mu + \beta_1\beta_0 = c'_0, \\ c_2 &= 6\mu^2 + 3\beta_3\mu + \beta_2\beta_0 = c''_0/2 \quad \text{και} \\ c_3 &= 4\mu + \beta_3 = c'''_0/6, \end{aligned}$$

όπου οι σταθερές β_i ορίζονται στο Λήμμα 4.3. Στη γενική περίπτωση, το ελάχιστο πολυώνυμο είναι το

$$q_m(x) = (x - c)(x^4 - c_3x^3 + c_2x^2 - c_1x + c_0).$$

Από το Θεώρημα Caley-Hamilton, ξέρουμε ότι ισχύει η ισότητα

$$q_m(B_{k+1}^{(2)} + \mu I) = 0.$$

Πολλαπλασιάζοντας και τα δύο μέλη της παραπάνω εξίσωσης με τον αντίστροφο του πίνακα $B_{k+1}^{(2)} + \mu I$, μετά από πράξεις βρίσκουμε ότι

$$\begin{aligned} cc_0(B_{k+1}^{(2)} + \mu I)^{-1} &= (B_{k+1}^{(2)} + \mu I)^4 - (c + c_3)(B_{k+1}^{(2)} + \mu I)^3 \\ &\quad + (cc_3 + c_2)(B_{k+1}^{(2)} + \mu I)^2 - (cc_2 + c_1)(B_{k+1}^{(2)} + \mu I) \\ &\quad + (cc_1 + c_0)I. \end{aligned}$$

Από την παραπάνω ισότητα έχουμε ότι

$$\begin{aligned} cc_0(B_{k+1}^{(2)} + \mu I)^{-1} &= (B_{k+1}^{(2)})^4 - (c + c_3 - 4\mu)(B_{k+1}^{(2)})^3 \\ &\quad + [cc_3 + c_2 - 3(c + c_3)\mu + 6\mu^2](B_{k+1}^{(2)})^2 \\ &\quad - [cc_2 + c_1 - 2(cc_3 + c_2)\mu + 3(c + c_3)\mu^2 - 4\mu^3]B_{k+1}^{(2)} \\ &\quad + [cc_1 + c_0 - (cc_2 + c_1)\mu + (cc_3 + c_2)\mu^2 \\ &\quad - (c + c_3)\mu^3 + \mu^4]I. \end{aligned}$$

Γράφοντας τα c_i συναρτήσει των β_i , $i = 0, 1, 2, 3$, καταλήγουμε στη σχέση

$$\begin{aligned}
\left(B_{k+1}^{(2)} + \mu I\right)^{-1} &= \frac{1}{\mu\nu_0 + \beta_0/\theta_k} \left[\left(B_{k+1}^{(2)}\right)^4 - (\mu + \beta_3 + 1/\theta_k) \left(B_{k+1}^{(2)}\right)^3 \right. \\
&\quad + (\mu\nu_3 + \beta_2\beta_0 + \beta_3/\theta_k) \left(B_{k+1}^{(2)}\right)^2 \\
&\quad - (\mu\nu_2 + \beta_1\beta_0 + \beta_2\beta_0/\theta_k) B_{k+1}^{(2)} \\
&\quad \left. + (\mu\nu_1 + \beta_0 + \beta_0\beta_1/\theta_k) I \right],
\end{aligned}$$

η οποία μπορεί να γραφεί όπως η σχέση (4.28). $\square$

Ας εξετάσουμε την περίπτωση όπου $a_{k-1} = a_k = 2$. Τότε, ο $B_{k+1}^{(2)}$ έχει τη μορφή

$$B_{k+1}^{(2)} = \frac{1}{\theta_k} I + \left(\frac{1}{\theta_{k+1}} - \frac{1}{\theta_k} \right) \frac{y_k y_k^T}{y_k^T y_k},$$

και δύο ιδιοτιμές, τις $1/\theta_{k+1}$ και $1/\theta_k$. Συνεπώς, το χαρακτηριστικό πολυώνυμο του $B_{k+1}^{(2)} + \mu I$, θα είναι το

$$q(x; \mu) = \left[x - \left(\frac{1}{\theta_k} + \mu \right) \right]^{n-1} \left[x - \left(\frac{1}{\theta_{k+1}} + \mu \right) \right].$$

Άρα, το ελάχιστο πολυώνυμο θα είναι το

$$q_m(x)(x; \mu) = \left[x - \left(\frac{1}{\theta_k} + \mu \right) \right] \left[x - \left(\frac{1}{\theta_{k+1}} + \mu \right) \right].$$

Εφαρμόζοντας το Θεώρημα Caley-Hamilton, έχουμε ότι

$$\begin{aligned}
\left(B_{k+1}^{(2)} + \mu I\right)^{-1} &= -\frac{1}{\left(\frac{1}{\theta_k} + \mu\right) \left(\frac{1}{\theta_{k+1}} + \mu\right)} \left[B_{k+1}^{(2)} \right. \\
&\quad \left. - \left(\mu + \frac{1}{\theta_k} + \frac{1}{\theta_{k+1}} \right) I \right]. \tag{4.29}
\end{aligned}$$

Πόρισμα 4.2. Έστω Λ το σύνολο των ιδιοτιμών του $B_{k+1}^{(2)}$ με αντίθετα πρόσημα και $a_{k-1} = a_k = 2$. Τότε, για κάθε $\mu \in \mathbb{R} \setminus \Lambda$, ο αντίστροφος του πίνακα $\left(B_{k+1}^{(2)} + \mu I\right)$, εκφράζεται από τον τύπο (4.29).

4.3.3 Ιδιοδιανύσματα διακριτών ιδιοτιμών

Η Πρόταση 4.1 μαζί με το Λήμμα 4.4, δείχνουν ότι αν η πιο μικρή ιδιοτιμή λ_1 του $B^{(2)}$ είναι διακριτή, τότε ο $B^{(2)}$ είτε έχει τουλάχιστον δύο διακριτές ιδιοτιμές ή ακριβώς μία, η οποία ισούται με

$$\lambda_1 = \frac{1}{\theta_{k+1}}.$$

Στην πρώτη περίπτωση, όπου έχουμε τουλάχιστον δύο διακριτές ιδιοτιμές, μπορούμε να βρούμε το ιδιοδιάνυσμα που αντιστοιχεί σε διακριτή ιδιοτιμή, εφαρμόζοντας τη μέθοδο της αντίστροφης δύναμης. Το προκύπτον ιδιοδιάνυσμα είναι το κανονικοποιημένο διάνυσμα

$$\hat{u} = \left[\sum_{i=0}^4 (-1)^i \nu_i(\hat{\lambda}) \left(B_{k+1}^{(2)} \right)^i \right] \frac{u}{\nu(\hat{\lambda})}, \quad (4.30)$$

όπου

$$\hat{\lambda} = -\lambda + \epsilon,$$

είναι μια διαταραγμένη διακριτή ιδιοτιμή με αντίθετο πρόσημο, ενώ ως αρχικό διάνυσμα u θεωρούμε το κανονικοποιημένο διάνυσμα

$$u = \frac{1}{\sqrt{n}}(1, \dots, 1)^T.$$

Ο υπολογισμός του διανύσματος $\hat{u}$ μπορεί να επιτευχθεί χρησιμοποιώντας την παρακάτω διαδικασία:

Διαδικασία 4.1. (Υπολογισμός ιδιοδιανύσματος $\hat{u}$)

$\hat{u}^{(4)} = u;$

for $j = 4, \dots, 1$ **do**

$v \leftarrow B^{(2)} \hat{u}^{(j)};$

$\hat{u}^{(j-1)} \leftarrow v + (-1)^{j-1} \nu_{j-1}(\hat{\lambda}) u;$

end for

return $\hat{u}^{(0)} / \nu(\hat{\lambda});$

Όσον αφορά το βοηθητικό διάνυσμα $v \in \mathbb{R}^n$ της παραπάνω διαδικα-

σίας, μπορούμε να το υπολογίσουμε από την έκφραση

$$v \equiv B_{k+1}^{(2)} \hat{u}^{(j)} = B_k^{(1)} \hat{u}^{(j)} - \frac{s_k^T B_k^{(1)} \hat{u}^{(j)}}{s_k^T B_k^{(1)} s_k} B_k^{(1)} s_k + \frac{y_k^T \hat{u}^{(j)}}{s_k^T y_k} y_k.$$

Τα διανύσματα

$$B_k^{(1)} \hat{u}^{(j)} \quad \text{και} \quad B_k^{(1)} s_k,$$

υπολογίζονται εύκολα, ως εξής:

$$B_k^{(1)} \hat{u}^{(j)} = \frac{1}{\theta_k} \hat{u}^{(j)} - \frac{s_{k-1}^T \hat{u}^{(j)}}{\theta_k s_{k-1}^T s_{k-1}} s_{k-1} + \frac{y_{k-1}^T \hat{u}^{(j)}}{s_{k-1}^T y_{k-1}} y_{k-1}$$

και

$$B_k^{(1)} s_k = \frac{1}{\theta_k} s_k - \frac{s_{k-1}^T s_k}{\theta_k s_{k-1}^T s_{k-1}} s_{k-1} + \frac{y_{k-1}^T s_k}{s_{k-1}^T y_{k-1}} y_{k-1}$$

αντίστοιχα.

Στη δεύτερη περίπτωση, δηλαδή στην περίπτωση που ο πίνακας $B_{k+1}^{(2)}$ έχει μόνο μία διακριτή ιδιοτιμή, είναι εύκολο να δούμε ότι το ιδιοδιάνυσμα που αντιστοιχεί σε αυτή είναι το κανονικοποιημένο διάνυσμα y_k . Πράγματι, σύμφωνα με την Πρόταση 4.1, στη συγκεκριμένη περίπτωση ισχύει ότι $\theta_k \neq \theta_{k+1}$ και

$$a_k = 2 \Rightarrow s_k = \theta_{k+1} y_k \quad \text{και} \quad a_{k-1} = 2 \Rightarrow s_{k-1} = \theta_k y_{k-1},$$

οπότε ο $B_{k+1}^{(2)}$ είναι της μορφής

$$B_{k+1}^{(2)} = \left(\frac{1}{\theta_k} \right) I - \frac{y_k y_k^T}{\theta_k y_k^T y_k} + \frac{y_k y_k^T}{\theta_{k+1} y_k^T y_k}.$$

Επομένως, πολλαπλασιάζοντας και τα δύο μέλη της παραπάνω εξίσωσης με y_k , βρίσκουμε ότι

$$B_{k+1}^{(2)} y_k = (1/\theta_{k+1}) y_k.$$

Επίσης, ένα γενικευμένο ιδιοδιάνυσμα που αντιστοιχεί στην πολλαπλή ιδιοτιμή είναι της μορφής

$$\hat{u}_1 = \left(-\frac{y_k^{(n)}}{y_k^{(1)}}, 0, \dots, 0, 1 \right)^T,$$

όπου το $y_k^{(i)}$ δηλώνει την i συνιστώσα του y_k . Αν βέβαια τύχει και το y_k^1 είναι ίσο με το μηδέν, τότε προφανώς θα αναζητήσουμε ένα διάνυσμα της μορφής

$$u = (0, \dots, 0, -\frac{y_k^j}{y_k^i}, 0, \dots, 0, 1, 0, \dots, 0)^T,$$

όπου η μονάδα βρίσκεται στην j θέση, έτσι ώστε να ισχύει ότι $\hat{u}_1^T y_k = 0$.

Αν ο πίνακας δεν έχει καμία διακριτή ιδιοτιμή, είμαστε στην περίπτωση όπου $a_{k-1} = a_k = 2$ και $\theta_k = \theta_{k+1}$. Τότε, τα ιδιοδιανύσματα του πίνακα είναι τα μοναδιαία διανύσματα

$$e_i = (0, 0, \dots, 1, 0, \dots, 0)^T,$$

όπου η μονάδα βρίσκεται στην i θέση.

4.4 Συμπεράσματα

Σε αυτό το κεφάλαιο παρουσιάσαμε τα ερευνητικά μας αποτελέσματα αναφορικά με την BFGS ενημέρωση ελάχιστης και μικρής μνήμης, στην περίπτωση όπου ο αρχικός πίνακας είναι ο μοναδιαίος, κλιμακούμενος με την BB φασματική παράμετρο. Στην αρχή μελετήσαμε την 1BFGS ενημέρωση, η οποία είναι η BFGS ενημέρωση ελάχιστης μνήμης, με την ποσότητα θ να ορίζεται ως η BB φασματική παράμετρος. Χρησιμοποιώντας αυτή την παράμετρο, αποδείξαμε κάποιες επιπλέον ιδιότητες του BFGS πίνακα ελάχιστης μνήμης. Στην ενημέρωση αυτή αποδείξαμε ότι η τιμή της συγκεκριμένης αυτής παραμέτρου μας επιτρέπει να ξέρουμε εκ των προτέρων αν ο πίνακας είναι θετικά ορισμένος ή όχι. Η εύρεση της αναλυτικής έκφρασης του χαρακτηριστικού πολυωνύμου μας πληροφορεί ότι οι ιδιοτιμές έχουν όλες το ίδιο πρόσημο, ενώ οι ακραίες ιδιοτιμές είναι διακριτές. Επίσης, η εισαγωγή της ποσότητας a_k , μας έδωσε τη δυνατότητα να εξετάσουμε με εύκολο υπολογιστικά τρόπο αν τα διανύσματα s_k και y_k είναι γραμμικά ανεξάρτητα ή όχι.

Στη συνέχεια μελετήσαμε την 2BFGS ενημέρωση, στην οποία χρησιμοποιούνται δύο διανυσματικά ζεύγη για την ενημέρωση του πίνακα, τα $\{s_{k-1}, y_{k-1}\}$ και $\{s_k, y_k\}$. Οι ιδιοτιμές και σε αυτή την περίπτωση βρίσκονται αναλυτικά, λύνοντας μια εξίσωση τετάρτου βαθμού, η οποία προκύπτει από την παραγόμενη έκφραση του χαρακτηριστικού πολυωνύμου.

Επιπλέον, βασιζόμενοι στην πληροφορία που παρέχεται από τις ποσότητες a_{k-1} και a_k , τεκμηριώσαμε ικανές συνθήκες για τον ακριβή αριθμό των διακριτών ιδιοτιμών του πίνακα $B_{k+1}^{(2)}$. Αυτό μας έδωσε τη δυνατότητα να ξέρουμε πότε μια ιδιοτιμή είναι διακριτή, ώστε να μπορούμε να εφαρμόσουμε την μέθοδο της αντίστροφης δύναμης για τον υπολογισμό του αντίστοιχου ιδιοδιανύσματος. Ο τρόπος υπολογισμού του τελευταίου βασίστηκε στην αναλυτική έκφραση του αντιστρόφου του πίνακα $B_{k+1}^{(2)} + \mu I$, με αποτέλεσμα να παραχθεί ένας κλειστός τύπος υπολογισμού ιδιοδιανυσμάτων. Επιπρόσθετα, η αναλυτική αυτή έκφραση μπορεί να χρησιμοποιηθεί και για την επίλυση γραμμικών συστημάτων. Τέλος, παρόλο που η 2BFGS ενημέρωση απαιτεί περισσότερες θέσεις μνήμης από ότι απαιτεί η 1BFGS, ο τρόπος που προτείναμε για τον υπολογισμό των ιδιοτιμών και των ιδιοδιανυσμάτων δεν απαιτεί υπολογιστικό κόστος. Η μέθοδος αποφεύγει τόσο την αποθήκευση όσο και την παραγοντοποίηση πινάκων, ενώ στηρίζεται στον υπολογισμό εσωτερικών γινομένων και διανυσματικών αθροισμάτων.

Εν κατακλείδι, η προτεινόμενη μέθοδος είναι υπολογιστικά ανέξοδη, ενώ υπολογίζει αναλυτικά όλα τα χαρακτηριστικά μεγέθη που μας ενδιαφέρουν. Συνεπώς, μπορεί να εφαρμοστεί σε μεγάλης κλίμακας προβλήματα βελτιστοποίησης, η επίλυση των οποίων απαιτεί τον υπολογισμό ιδιοτιμών και ιδιοδιανυσμάτων, ή ακόμα και την επίλυση γραμμικών συστημάτων.

Επιλύοντας το Υποπρόβλημα Περιοχής Εμπιστοσύνης

Σε αυτό το κεφάλαιο αναπτύσσουμε τη δική μας προσέγγιση για την επίλυση μεγάλης διάστασης υποπροβλημάτων περιοχής εμπιστοσύνης (ΥΠΕ), γνωστών στη διεθνή βιβλιογραφία ως TRS (Trust Region Subproblems). Η προτεινόμενη μέθοδος, η οποία αποτελεί παραλλαγή της μεθόδου των Moré & Sorensen, που περιγράψαμε στο Κεφάλαιο 2, υπολογίζει μια σχεδόν ακριβή λύση του ΥΠΕ. Βασίζεται τόσο στις ιδιότητες των Quasi-Newton ενημερώσεων ελάχιστης μνήμης, όσο και στις ιδιότητες των 1BFGS και 2BFGS ενημερώσεων, ενώ αποφεύγει πλήρως την παραγοντοποίηση Cholesky. Η προτεινόμενη μέθοδος μπορεί να εφαρμοστεί σε προβλήματα μεγάλης κλίμακας, διότι στηρίζεται σε εσωτερικά γινόμενα και διανυσματικά αθροίσματα, με αποτέλεσμα να μην απαιτεί μεγάλο υπολογιστικό κόστος. Τα ακόλουθα αποτελούν ερευνητικά αποτελέσματα της παρούσας διατριβής, μέρος των οποίων καταγράφονται στις εργασίες των Apostoloroulou et al. [6, 8, 10, 12].

5.1 Η προτεινόμενη μέθοδος

Στο Κεφάλαιο 2 αναφερθήκαμε στο υποπρόβλημα περιοχής εμπιστοσύνης (ΥΠΕ)

$$\min_{d \in \mathbb{R}^n} \phi(d) = g^T d + \frac{1}{2} d^T B d \quad \text{υπό τον περιορισμό} \quad \|d\| \leq \Delta, \quad (5.1)$$

όπου B είναι ένας συμμετρικός $n \times n$ πίνακας (θετικά ορισμένος ή όχι), $g \in \mathbb{R}^n$, $d \in \mathbb{R}^n$ είναι το άγνωστο διάνυσμα και $\Delta \in \mathbb{R}$, η ακτίνα περιοχής εμπιστοσύνης. Το συγκεκριμένο πρόβλημα αποτελεί από μόνο του ένα πρόβλημα βελτιστοποίησης υπό περιορισμό. Είναι αρκετά σημαντικό πρό-

βλημα, γιατί συναντάται σε πολλά προβλήματα βελτιστοποίησης, όπως σε επίλυση προβλημάτων με μεθόδους περιοχής εμπιστοσύνης, προβλήματα βελτιστοποίησης με περιορισμούς ισότητας και ανισότητας, προβλήματα αποκατάστασης εικόνας, διαχωρισμού γραφημάτων κ.α.

Στην Ενότητα 2.8 του Κεφαλαίου 2, αναπτύξαμε τη μέθοδο των Moré & Sorensen [64] για την επίλυση του υποπροβλήματος περιοχής εμπιστοσύνης (5.1), η οποία θεωρείται μέχρι και σήμερα η καλύτερη μέθοδος για την επίλυση του υποπροβλήματος. Οι Moré & Sorensen βρίσκουν μια σχεδόν ακριβή λύση του ΥΠΕ (5.1), βασιζόμενοι στις πρώτης τάξης ΚΚΤ αναγκαίες συνθήκες βελτιστότητας. Κατά τη διάρκεια επίλυσης του ΥΠΕ (5.1), αν ο βέλτιστος πολλαπλασιαστής Lagrange λ^* , ανήκει στο διάστημα $(-\lambda_1, +\infty)$, όπου λ_1 η μικρότερη ιδιοτιμή του B , έχουμε την τυπική περίπτωση, ενώ αν $\lambda^* = -\lambda_1$, έχουμε τη δύσκολη περίπτωση. Οι περιπτώσεις αυτές είναι γνωστές στη διεθνή βιβλιογραφία ως standard και hard case, αντίστοιχα.

Η συγκεκριμένη μέθοδος χρησιμοποιεί την κατά Cholesky παραγοντοποίηση πινάκων, κάτι που την καθιστά πολύ ακριβή και ίσως απαγορευτική στην εφαρμογή της για μεγάλης διάστασης προβλήματα. Όπως ήδη έχουμε αναφέρει στην Ενότητα 2.8, η έρευνα γύρω από το ΥΠΕ εστιάζεται σήμερα στο πως η μέθοδος των Moré & Sorensen μπορεί να τροποποιηθεί με τέτοιο τρόπο, ώστε να εφαρμόζεται και σε μεγάλης διάστασης προβλήματα. Τα πιο πρόσφατα ερευνητικά αποτελέσματα [45, 48, 77, 78, 83] βασίζονται στη μέθοδο Lanczos (σελ. 15) καθώς και σε γινόμενα πίνακα-διανύσματος.

Η δική μας προσέγγιση για την επίλυση μεγάλης διάστασης ΥΠΕ, παραλλάσσει τη μέθοδο των Moré & Sorensen, με τέτοιο τρόπο, ώστε να αποφεύγει τόσο την κατά Cholesky παραγοντοποίηση πινάκων, όσο και την αποθήκευση πίνακα. Ακόμα, βρίσκει με μεγάλη ακρίβεια τις ποσότητες που μας ενδιαφέρουν, χωρίς να χρησιμοποιεί εσωτερικές επαναλήψεις ή γινόμενα πίνακα-διανύσματος. Ο λόγος είναι ότι χρησιμοποιεί τις ιδιότητες των Quasi-Newton ενημερώσεων ελάχιστης μνήμης, καθώς και τις ιδιότητες των 1BFGS και 2BFGS ενημερώσεων. Να υπενθυμίσουμε εδώ ότι η τελευταία αποτελεί μια Quasi-Newton ενημέρωση μικρής μνήμης.

Χρησιμοποιώντας τις ιδιότητες των Quasi-Newton ενημερώσεων ελάχιστης και μικρής μνήμης, όπως αποδείξαμε στα Κεφάλαια 3 και 4, μπορούμε να υπολογίσουμε αναλυτικά το χαρακτηριστικό πολυώνυμο του πίνακα

B . Άμεση συνέπεια αυτού είναι ότι μπορούμε να βρούμε με μεγάλη ακρίβεια (αναλυτικά) τις ιδιοτιμές του καθώς και τον αντίστροφο του $(B + \lambda I)$, με χρήση του θεωρήματος των Caley-Hamilton. Επομένως, όσον αφορά τη διαδικασία επίλυσης του ΥΠΕ, η εύρεση του δοκιμαστικού βήματος d , γίνεται με τέτοιο τρόπο ώστε να αποφευχθεί πλήρως η κατά Cholesky παραγοντοποίηση πινάκων. Επιπρόσθετα, ο υπολογισμός του ιδιοδιανύσματος που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του πίνακα βασίζεται είτε σε απλούς υπολογισμούς ή στην εφαρμογή της μεθόδου της αντίστροφης δυνάμεως [14, 52]. Συνεπώς, δεν απαιτείται η μέθοδος Lanczos για την εύρεση ιδιοδιανυσμάτων, οπότε αποφεύγεται τόσο μια εσωτερική επαναληπτική διαδικασία, όσο και η αποθήκευση του πίνακα με στήλες τα παραγόμενα διανύσματα Lanczos.

5.2 Η τυπική περίπτωση

Όπως είδαμε στο Κεφάλαιο 2, Ενότητα 2.8.1, στην τυπική περίπτωση (standard case), ο πολλαπλασιαστής Lagrange λ ανήκει στο διάστημα $(-\lambda_1, \infty)$, όπου λ_1 είναι η πιο μικρή ιδιοτιμή του πίνακα B . Το δοκιμαστικό βήμα

$$d(\lambda) = -(B + \lambda I)^{-1} g$$

θα πρέπει να ικανοποιεί τον περιορισμό $\|d(\lambda)\| \leq \Delta$. Για να λύσουμε το υποπρόβλημα σε αυτή την περίπτωση, αν ο βέλτιστος πολλαπλασιαστής Lagrange δεν είναι το μηδέν, χρησιμοποιούμε τη μέθοδο Newton για την επίλυση της εξίσωσης

$$\psi(\lambda) = 0,$$

ως προς λ , γνωστής ως secular equation, όπου

$$\psi(\lambda) := \frac{1}{\|d(\lambda)\|} - \frac{1}{\Delta},$$

έτσι ώστε να ικανοποιείται ο περιορισμός της περιοχής εμπιστοσύνης. Το αναδρομικό σχήμα της Newton είναι το εξής:

$$\begin{aligned}\lambda_{\ell+1} &= \lambda_{\ell} - \frac{\psi(\lambda_{\ell})}{\psi'(\lambda_{\ell})} \\ &= \lambda_{\ell} + \frac{\|d(\lambda_{\ell})\|^2}{d(\lambda_{\ell})^T (B_{k+1} + \lambda_{\ell} I)^{-1} d(\lambda_{\ell})} \left(\frac{\|d(\lambda_{\ell})\| - \Delta}{\Delta} \right), \quad (5.2)\end{aligned}$$

όπου $\ell = 0, 1, 2, \dots$

Τα προβλήματα που συναντάμε στην τυπική περίπτωση, είναι τα ακόλουθα: πρώτον, είναι απαραίτητο να λύσουμε ένα γραμμικό σύστημα της μορφής

$$(B + \lambda I) d = -g,$$

ως προς το δοκιμαστικό βήμα d , άρα χρειάζεται να υπολογίσουμε τον πίνακα $(B + \lambda I)^{-1}$, για τις διάφορες τιμές του πολλαπλασιαστή Lagrange, λ . Δεύτερον, πρέπει να υπολογίσουμε τον παρανομαστή της επανάληψης Newton,

$$d(\lambda_{\ell})^T (B + \lambda_{\ell} I)^{-1} d(\lambda_{\ell}).$$

Τέλος, το λ θα πρέπει να ανήκει στο διάστημα $(-\lambda_1, \infty)$, κάτι που σημαίνει ότι σε κάθε επανάληψη της Newton χρειάζεται να ελέγχουμε αν το λ που παράγεται είναι μη αρνητικός αριθμός και διάφορος του $-\lambda_1$. Θα πρέπει δηλαδή να ισχύει ότι

$$\lambda \neq -\lambda_1,$$

έτσι ώστε ο πίνακας $B + \lambda I$ να μην γίνει ιδιάζον, αλλά να παραμένει θετικά ορισμένος.

5.2.1 Εφαρμογή της γενικής PSB ενημέρωσης ελάχιστης μνήμης

Θεωρούμε ότι ο συμμετρικός πίνακας B στο ΥΠΕ (5.1) είναι ο πίνακας της γενικής PSB ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k) w_k^T + w_k (y_k - \theta_{k+1} s_k)^T}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{s_k^T w_k}.$$

Για να λύσουμε το γραμμικό σύστημα $(B + \lambda I) d(\lambda) = -g$, βασιζόμαστε στην Πρόταση 3.1 (σελ. 93). Στη συγκεκριμένη πρόταση, τεκμηριώσαμε ότι ο πίνακας $(B + \lambda I)^{-1}$ δίνεται από την έκφραση

$$(B_{k+1} + \mu I)^{-1} = \frac{1}{(\mu + \theta_{k+1})\gamma(\mu)} \left\{ B_{k+1}^2 - [\beta'(\mu) - \mu] B_{k+1} + \beta(\mu) I \right\}, \quad (5.3)$$

όπου $\mu \in \mathbb{R} \setminus \Lambda$, $\Lambda = \{-\lambda_1, -\lambda_2, \dots, -\lambda_n\}$, με

$$\begin{aligned} \beta(\mu) &= \gamma(\mu) + [\gamma'(\mu) - \mu] \theta_{k+1}, \\ \gamma(\mu) &= \mu^2 + \beta_1 \mu + \beta_2, \\ \beta_1 &= 2 \frac{y_k^T w_k}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \quad \text{και} \\ \beta_2 &= \left(\frac{y_k^T w_k}{s_k^T w_k} \right)^2 - \frac{(y_k - \theta_{k+1} s_k)^T y_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k}. \end{aligned}$$

Το δοκιμαστικό βήμα $d_{k+1}(\lambda) = -(B_{k+1} + \lambda I)^{-1} g_{k+1}$ στην τυπική περίπτωση, μπορεί να υπολογιστεί άμεσα, από την εξίσωση (5.3). Συγκεκριμένα, για $\mu = \lambda$, έχουμε ότι

$$d_{k+1}(\lambda) = -\frac{1}{(\lambda + \theta_{k+1})\gamma(\lambda)} \left[B_{k+1}^2 g_{k+1} - [\beta'(\lambda) - \lambda] B_{k+1} g_{k+1} + \beta(\lambda) g_{k+1} \right]. \quad (5.4)$$

Για να υπολογίσουμε τα διανύσματα $B_{k+1} g_{k+1}$ και $(B_{k+1})^2 g_{k+1}$, χρησι-

μποποιούμε το επαναληπτικό σχήμα

$$\begin{aligned} B_{k+1}v_i &= \theta_{k+1}v_i + \frac{w_k^T v_i}{s_k^T w_k} (y_k - \theta_{k+1}s_k) \\ &+ \left(\frac{(y_k - \theta_{k+1}s_k)^T v_i}{s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \frac{w_k^T v_i}{s_k^T w_k} \right) w_k, \end{aligned}$$

$i = 0, 1$, με $v_0 = g_{k+1}$. Μετά από πράξεις (οι απαιτούμενοι υπολογισμοί έχουν αναλυθεί στην Ενότητα 3.4), η σχέση (5.4) γράφεται ως εξής:

$$d_{k+1}(\lambda) = -\frac{1}{\gamma(\lambda)(\lambda + \theta_{k+1})} [c_g(\lambda)g_{k+1} + c_z(\lambda)z_k + c_w(\lambda)w_k], \quad (5.5)$$

όπου $z_k = y_k - \theta_{k+1}s_k$ και

$$\begin{aligned} c_g(\lambda) &= \gamma(\lambda), \\ c_z(\lambda) &= -\left(\lambda + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{w_k^T u}{s_k^T w_k} + \frac{w_k^T w_k}{s_k^T w_k} \frac{z_k^T u}{s_k^T w_k}, \\ c_w(\lambda) &= \left[\frac{z_k^T z_k}{s_k^T w_k} + \frac{z_k^T s_k}{s_k^T w_k} (\lambda + \theta_{k+1}) \right] \frac{w_k^T u}{s_k^T w_k} - \left(\lambda + \theta_{k+1} + \frac{z_k^T w_k}{s_k^T w_k} \right) \frac{z_k^T u}{s_k^T w_k}. \end{aligned}$$

Στην περίπτωση όπου τα διανύσματα s_k και y_k είναι συγγραμμικά, δηλαδή ισχύει $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, ξέρουμε ότι ο πίνακας B_{k+1} έχει την έκφραση

$$B_{k+1} = \theta_{k+1}I + (\kappa - \theta_{k+1}) \frac{s_k s_k^T}{s_k^T s_k}.$$

Επιπρόσθετα, το Λήμμα 3.2 (σελ. 94) τεκμηριώνει ότι ο αντίστροφος του $B_{k+1} + \lambda I$ σε αυτή την περίπτωση είναι ο πίνακας

$$(B_{k+1} + \lambda I)^{-1} = -\frac{1}{\gamma(\lambda)} \{ B_{k+1} - [\gamma'(\lambda) - \lambda] I \}. \quad (5.6)$$

Άμεση συνέπεια αυτού είναι ότι, όταν $y_k = \kappa s_k$, το δοκιμαστικό βήμα ισούται με

$$d_{k+1}(\lambda) = \frac{1}{\gamma(\lambda)} \left[-(\lambda + \kappa) g_{k+1} + (\kappa - \theta_{k+1}) \frac{s_k^T g_{k+1}}{s_k^T s_k} s_k \right]. \quad (5.7)$$

Για τον υπολογισμό του πολλαπλασιαστή Lagrange λ , ακολουθούμε τη μέθοδο των Moré & Sorensen [64]. Εφαρμόζουμε δηλαδή τη μέθοδο Newton, της οποίας ο επαναληπτικός τύπος δίνεται στη σχέση (5.2). Ο παρανομαστής της (5.2),

$$\Pi = d_{k+1}(\lambda)^T (B_{k+1} + \lambda I)^{-1} d_{k+1}(\lambda)$$

μπορεί να υπολογιστεί εύκολα, αν λάβουμε υπόψη μας τη σχέση (5.3), καθώς και το γεγονός ότι $(B_{k+1} + \lambda I) d_{k+1}(\lambda) = -g_{k+1}$, από το οποίο έπεται ότι

$$B_{k+1} d_{k+1}(\lambda) = -[\lambda d_{k+1}(\lambda) + g_{k+1}].$$

Οπότε, χρησιμοποιώντας την παραπάνω ισότητα και τη σχέση (5.3), βρίσκουμε ότι

$$\begin{aligned} \Pi &= \frac{\|B_{k+1} d_{k+1}(\lambda)\|^2 - [\beta'(\lambda) - \lambda] d_{k+1}(\lambda)^T B_{k+1} d_{k+1}(\lambda)}{\gamma(\lambda)(\lambda + \theta_{k+1})} \\ &\quad + \frac{\beta(\lambda) \|d_{k+1}(\lambda)\|^2}{\gamma(\lambda)(\lambda + \theta_{k+1})} \\ &= \frac{\|\lambda d_{k+1}(\lambda) + g_{k+1}\|^2 + [\beta'(\lambda) - \lambda] [\lambda \|d_{k+1}(\lambda)\|^2 + g_{k+1}^T d_{k+1}(\lambda)]}{\gamma(\lambda)(\lambda + \theta_{k+1})} \\ &\quad + \frac{\beta(\lambda) \|d_{k+1}(\lambda)\|^2}{\gamma(\lambda)(\lambda + \theta_{k+1})}. \end{aligned}$$

Επομένως, ο παρανομαστής $d_{k+1}(\lambda)^T (B_{k+1} + \lambda I)^{-1} d_{k+1}(\lambda)$ της επανάληψης Newton (5.2), υπολογίζεται από την έκφραση

$$\begin{aligned} \Pi &= \frac{[\beta(\lambda) + \lambda \beta'(\lambda)] \|d_{k+1}(\lambda)\|^2 + [\beta'(\lambda) + \lambda] d_{k+1}(\lambda)^T g_{k+1} + \|g_{k+1}\|^2}{\gamma(\lambda)(\lambda + \theta_{k+1})} \\ &\quad + \frac{\|g_{k+1}\|^2}{\gamma(\lambda)(\lambda + \theta_{k+1})}. \end{aligned} \tag{5.8}$$

Αν $y_k = \kappa s_k$, τότε ο πίνακας $(B_{k+1} + \lambda I)^{-1}$ ορίζεται στην (5.6). Άρα,

έχουμε ότι

$$\begin{aligned}\Pi &= -\frac{1}{\gamma(\lambda)} \{d_{k+1}(\lambda)^T B_{k+1} d_{k+1}(\lambda)^T - [\gamma'(\lambda) - \lambda] \|d_{k+1}(\lambda)\|^2\} \\ &= \frac{1}{\gamma(\lambda)} \{d_{k+1}(\lambda)^T [\lambda d_{k+1}(\lambda) + g_{k+1}] + [\gamma'(\lambda) - \lambda] \|d_{k+1}(\lambda)\|^2\}.\end{aligned}$$

Συνεπώς, καταλήγουμε στο ότι ο παρανομαστής της επανάληψης Newton (5.2), στη συγκεκριμένη περίπτωση, ισούται με

$$\Pi = \frac{\gamma'(\lambda) \|d_{k+1}(\lambda)\|^2 + d_{k+1}(\lambda)^T g_{k+1}}{\gamma(\lambda)}. \quad (5.9)$$

Τέλος, όσον αφορά τον έλεγχο του πολλαπλασιαστική Lagrange αναφορικά με τη $-\lambda_1$, όπου λ_1 η μικρότερη ιδιοτιμή του B , δοθέντος ότι γνωρίζουμε το χαρακτηριστικό πολυώνυμο του πίνακα, είναι τετριμμένο να ελέγχουμε αν το λ που υπολογίζεται από την τρέχουσα επανάληψη Newton, συμπίπτει ή όχι με τη $-\lambda_1$.

Στη συνέχεια, θα εξάγουμε τις εκφράσεις που μας δίνουν το δοκιμαστικό βήμα $d_{k+1}(\lambda)$ και τον παρανομαστή της επανάληψης Newton, Π , για καθεμιά από τις τέσσερις Quasi-Newton ενημερώσεις ελάχιστης μνήμης.

5.2.1.1 Η SR1 ενημέρωση ελάχιστης μνήμης. Στην SR1 ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k)(y_k - \theta_{k+1} s_k)^T}{(y_k - \theta_{k+1} s_k)^T s_k},$$

για να λύσουμε το γραμμικό σύστημα $(B + \lambda I) d = -g$, χρησιμοποιούμε την έκφραση του αντιστρόφου του $(B + \lambda I)$, όπως την ορίσαμε στη σχέση (3.42) (σελ. 106) της Ενότητας 3.5. Ο πίνακας $(B + \lambda I)^{-1}$ ορίζεται ως

$$(B_{k+1} + \lambda I)^{-1} = -\frac{1}{\gamma(\lambda)} \{B_{k+1} - [\gamma'(\lambda) - \lambda] I\}. \quad (5.10)$$

Επομένως, το δοκιμαστικό βήμα είναι της μορφής

$$d_{k+1}(\lambda) = \frac{1}{\gamma(\lambda)} \{B_{k+1} g_{k+1} - [\gamma'(\lambda) - \lambda] g_{k+1}\},$$

όπου

$$B_{k+1}g_{k+1} = \theta_{k+1}g_{k+1} + \frac{(y_k - \theta_{k+1}s_k)^T g_{k+1}}{(y_k - \theta_{k+1}s_k)^T s_k} (y_k - \theta_{k+1}s_k).$$

Άρα, έχουμε ότι

$$\begin{aligned} d_{k+1}(\lambda) = & \frac{\lambda - \gamma'(\lambda) + \theta_{k+1}}{\gamma(\lambda)} g_{k+1} \\ & + \frac{(y_k - \theta_{k+1}s_k)^T g_{k+1}}{\gamma(\lambda)(y_k - \theta_{k+1}s_k)^T s_k} (y_k - \theta_{k+1}s_k). \end{aligned} \quad (5.11)$$

Για να υπολογίσουμε τον παρανομαστή Π των Newton επαναλήψεων (5.4), λαμβάνουμε υπόψη μας την εξίσωση

$$B_{k+1} d_{k+1}(\lambda) = -[\lambda d_{k+1}(\lambda) + g_{k+1}]$$

και την εξίσωση (5.10). Έχουμε ότι

$$\begin{aligned} \Pi &= d_{k+1}(\lambda)^T (B_{k+1} + \lambda I)^{-1} d_{k+1}(\lambda) \\ &= -\frac{1}{\gamma(\lambda)} \{d_{k+1}(\lambda)^T B_{k+1} d_{k+1}(\lambda) - [\gamma'(\lambda) - \lambda] d_{k+1}(\lambda)^T d_{k+1}(\lambda)\} \\ &= \frac{1}{\gamma(\lambda)} \{d_{k+1}(\lambda)^T [\lambda d_{k+1}(\lambda) + g_{k+1}] + [\gamma'(\lambda) - \lambda] d_{k+1}(\lambda)^T d_{k+1}(\lambda)\}. \end{aligned}$$

Συνεπώς, ο παρανομαστής Π των Newton επαναλήψεων (5.4) βρίσκεται από την παρακάτω σχέση:

$$\Pi = \frac{1}{\gamma(\lambda)} [\gamma'(\lambda) \|d_{k+1}(\lambda)\|^2 + d_{k+1}(\lambda)^T g_{k+1}].$$

Αν τα διανύσματα s_k και y_k είναι συγγραμμικά, τότε το $d_{k+1}(\lambda)$ και το Π υπολογίζονται μέσω των σχέσεων (5.7) και (5.9), αντίστοιχα.

5.2.1.2 Η BFGS ενημέρωση ελάχιστης μνήμης.

Στην BFGS ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I - \theta_{k+1} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

το δοκιμαστικό βήμα βρίσκεται σε συνάρτηση με την έκφραση (5.3). Κά-

νοντας παρόμοιους υπολογισμούς με αυτούς της απόδειξης του Πορίσματος 3.2 (σελ. 108), καταλήγουμε στον τύπο

$$d_{k+1}(\lambda) = -\frac{1}{\gamma(\lambda)(\lambda + \theta_{k+1})} [c_g(\lambda)g_{k+1} + c_s(\lambda)s_k + c_y(\lambda)y_k], \quad (5.12)$$

όπου

$$\begin{aligned} c_g(\lambda) &= \gamma(\lambda) \\ c_s(\lambda) &= [(\gamma'(\lambda) - \lambda) s_k^T g_{k+1} - y_k^T g_{k+1}] \frac{\theta_{k+1}}{s_k^T s_k} \quad \text{και} \\ c_y(\lambda) &= -\frac{y_k^T g_{k+1}}{s_k^T y_k} \lambda - \frac{s_k^T g_{k+1}}{s_k^T s_k} \theta_{k+1}. \end{aligned}$$

Ο παρανομαστής Π της Newton επανάληψης (5.4) υπολογίζεται από τον τύπο (5.8). Τέλος, αν τα διανύσματα s_k , y_k είναι γραμμικά εξαρτημένα, τότε τα $d(\lambda)$ και Π , βρίσκονται από τους τύπους (5.7) και (5.9), αντίστοιχα.

5.2.1.3 Η DFP ενημέρωση ελάχιστης μνήμης. Στην DFP ενημέρωση ελάχιστης μνήμης

$$\begin{aligned} B_{k+1} = \theta_{k+1}I &+ \frac{(y_k - \theta_{k+1}s_k) y_k^T + y_k (y_k - \theta_{k+1}s_k)^T}{s_k^T y_k} \\ &- \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T y_k} \frac{y_k y_k^T}{s_k^T y_k}, \end{aligned}$$

το δοκιμαστικό βήμα $d(\lambda) = -(B + \lambda I)^{-1}g$ υπολογίζεται μέσω της έκφρασης (5.3). Μετά από υπολογισμούς ανάλογους των υπολογισμών του Πορίσματος 3.3 (σελ. 117), βρίσκουμε ότι το δοκιμαστικό βήμα εκφράζεται από τον τύπο (5.12), όπου

$$\begin{aligned} c_g(\lambda) &= \gamma(\lambda) \\ c_s(\lambda) &= (\lambda y_k^T g_{k+1} + \beta_2 s_k^T g_{k+1}) \frac{\theta_{k+1}}{s_k^T y_k} \quad \text{και} \\ c_y(\lambda) &= \lambda \theta_{k+1} \frac{s_k^T g_{k+1}}{s_k^T y_k} - \left(\lambda \beta_1 \frac{s_k^T y_k}{y_k^T y_k} + \theta_{k+1} \right) \frac{y_k^T g_{k+1}}{s_k^T y_k}, \end{aligned}$$

με

$$\beta_1 = \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k}\right) \frac{y_k^T y_k}{s_k^T y_k} \quad \text{και} \quad \beta_2 = \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}.$$

Ο παρανομαστής Π της Newton επανάληψης (5.4), υπολογίζεται μέσω της εξίσωσης (5.8). Τέλος, αν τα διανύσματα s_k , y_k είναι γραμμικά εξαρτημένα, τότε τα $d(\lambda)$ και Π βρίσκονται από τους τύπους (5.7) και (5.9), αντίστοιχα.

5.2.1.4 Η PSB ενημέρωση ελάχιστης μνήμης. Στην PSB ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1} I + \frac{(y_k - \theta_{k+1} s_k) s_k^T + s_k (y_k - \theta_{k+1} s_k)^T}{s_k^T s_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k s_k^T s_k^T}{s_k^T s_k s_k^T s_k},$$

το $d(\lambda) = -(B + \lambda I)^{-1} g$ υπολογίζεται μέσω της (5.12), όπου

$$\begin{aligned} c_g(\lambda) &= \gamma(\lambda) \\ c_s(\lambda) &= \left[\frac{y_k^T y_k}{s_k^T s_k} + \beta_1 (\lambda + \theta_{k+1}) \right] \frac{s_k^T g_{k+1}}{s_k^T s_k} - (\lambda + \beta_1) \frac{y_k^T g_{k+1}}{s_k^T s_k} \quad \text{και} \\ c_y(\lambda) &= -(\lambda + \beta_1) \frac{s_k^T g_{k+1}}{s_k^T s_k} + \frac{y_k^T g_{k+1}}{s_k^T s_k}, \end{aligned}$$

με $\beta_1 = \theta_{k+1} + (s_k^T y_k)/(s_k^T s_k)$.

Στην έκφραση αυτή καταλήγουμε αν χρησιμοποιήσουμε τη σχέση (5.3) και κάνουμε παρόμοιους υπολογισμούς με αυτούς του Πορίσματος 3.4 (σελίδα 122). Τέλος, ο παρανομαστής Π της επανάληψης Newton (5.4), υπολογίζεται μέσω της εξίσωσης (5.8), ενώ αν τα διανύσματα s_k , y_k είναι γραμμικά εξαρτημένα, τα $d(\lambda)$ και Π βρίσκονται από τους τύπους (5.7) και (5.9), αντίστοιχα.

5.2.2 Η 1BFGS ενημέρωση στην τυπική περίπτωση

Στην Ενότητα 4.2 του Κεφαλαίου 4, αποδείξαμε ότι στην 1BFGS ενημέρωση, η παράμετρος

$$a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k},$$

μας επιτρέπει με πολύ εύκολο τρόπο, να αποφανθούμε αν τα διανύσματα s_k και y_k είναι γραμμικά ανεξάρτητα ή όχι. Όταν $a_k > 2$, δηλαδή τα διανύσματα s_k και y_k είναι γραμμικώς ανεξάρτητα, τότε χρησιμοποιούμε το Λήμμα 4.2, για να υπολογίσουμε το δοκιμαστικό βήμα $d(\lambda)$. Συγκεκριμένα, ο αντίστροφος του πίνακα $B + \lambda I$, είναι ο πίνακας

$$\left(B_{k+1}^{(1)} + \mu I\right)^{-1} = \frac{1}{\gamma} \sum_{i=0}^2 (-1)^i \gamma_i(\mu) \left(B_{k+1}^{(1)}\right)^i, \quad (5.13)$$

όπου

$$\begin{aligned} \gamma(\lambda) &= (1/\theta_{k+1} + \lambda)(\lambda^2 + a_k \lambda / \theta_{k+1} + 1/\theta_{k+1}^2) \\ \gamma_2(\lambda) &= 1 \\ \gamma_1(\lambda) &= \lambda + (a_k + 1)/\theta_{k+1} \quad \text{και} \\ \gamma_0(\lambda) &= \lambda^2 + (a_k + 1)\lambda / \theta_{k+1} + (a_k + 1)/\theta_{k+1}^2. \end{aligned}$$

Συνεπώς, το δοκιμαστικό βήμα είναι το

$$d_{k+1}(\lambda) = -\frac{1}{\gamma(\lambda)} \sum_{i=0}^2 (-1)^i \gamma_i(\lambda) \left(B_{k+1}^{(1)}\right)^i g_{k+1}.$$

Τα διανύσματα $B_{k+1}^{(1)} g_{k+1}$ και $\left(B_{k+1}^{(1)}\right)^2 g_{k+1}$ υπολογίζονται από το επαναληπτικό σχήμα

$$v_{i+1} \equiv B_{k+1}^{(1)} v_i = \frac{1}{\theta_{k+1}} v_i - \frac{s_k^T v_i}{\theta_{k+1} s_k^T s_k} s_k + \frac{y_k^T v_i}{s_k^T y_k} y_k, \quad i = 0, 1,$$

με $v_0 = g_{k+1}$. Μετά από μερικές πράξεις (ανάλογες με αυτές της Υποενότητας 4.2.3, σελίδα 135), καταλήγουμε στην ακόλουθη γενική έκφραση

του δοκιμαστικού βήματος:

$$d_{k+1}(\lambda) = -\gamma_g(\lambda) g_{k+1} + \gamma_s(\lambda) s_k - \gamma_y(\lambda) y_k,$$

όπου

$$\begin{aligned}\gamma_g(\lambda) &= [1 - \gamma_1(\lambda) \theta_{k+1} + \gamma_0(\lambda) \theta_{k+1}^2] / [\gamma(\lambda) \theta_{k+1}^2], \\ \gamma_s(\lambda) &= \{ [1 - \gamma_1(\lambda) \theta_{k+1}] s_k^T g_{k+1} + \theta_{k+1} y_k^T g_{k+1} \} / [\gamma(\lambda) \theta_{k+1}^2 s_k^T s_k], \text{ και} \\ \gamma_y(\lambda) &= \{ [1 - \gamma_1(\lambda) \theta_{k+1} + a_k] \theta_{k+1} y_k^T g_{k+1} - s_k^T g_{k+1} \} / [\gamma(\lambda) \theta_{k+1}^2 s_k^T y_k].\end{aligned}$$

Επιπρόσθετα, για $\lambda = 0$, υπολογίζοντας τους συντελεστές, βρίσκουμε ότι

$$\begin{aligned}\gamma_g(0) &= \theta_{k+1} \\ \gamma_s(0) &= (\theta_{k+1} y_k^T g_{k+1} - a_k s_k^T g_{k+1}) / s_k^T y_k \text{ και} \\ \gamma_y(0) &= -\theta_{k+1} s_k^T g_{k+1} / s_k^T y_k.\end{aligned}$$

Επομένως,

$$d_{k+1}(0) \equiv - \left(B_{k+1}^{(1)} \right)^{-1} g_{k+1} = -\gamma_g(0) g_{k+1} + \gamma_s(0) s_k - \gamma_y(0) y_k.$$

Όταν $a_k = 2$, δηλαδή τα διανύσματα s_k και y_k είναι γραμμικά εξαρτημένα, από την Υποενότητα 4.2.2 γνωρίζουμε ότι επειδή ο πίνακας $B_{k+1}^{(1)}$ είναι ένα συγκεκριμένο πολλαπλάσιο του μοναδιαίου πίνακα I της μορφής $B_{k+1}^{(1)} = (1/\theta_{k+1})I$, ο αντίστροφος του $B_{k+1}^{(1)} + \lambda I$ είναι ο πίνακας

$$\left(B_{k+1}^{(1)} + \lambda I \right)^{-1} = \frac{\theta_{k+1}}{1 + \lambda \theta_{k+1}} I. \quad (5.14)$$

Αυτό έχει ως συνέπεια, το δοκιμαστικό βήμα να απλοποιείται κατά πολύ και να ισούται με

$$d_{k+1}(\lambda) = -\frac{\theta_{k+1}}{1 + \lambda \theta_{k+1}} g_{k+1}. \quad (5.15)$$

Στη μέθοδο Newton (5.4), για τον υπολογισμό του πολλαπλασιαστή Lagrange, χρησιμοποιώντας την εξίσωση (5.13), βρίσκουμε ότι ο παρανο-

μαστής του επαναληπτικού σχήματος (5.4), ισούται με

$$\begin{aligned} \Pi &= d_{k+1}(\lambda)^T (B_{k+1} + \lambda I)^{-1} d_{k+1}(\lambda) \\ &= -\frac{\gamma_1(\lambda)}{\gamma(\lambda)} d_{k+1}(\lambda)^T B_{k+1}^{(1)} d_{k+1}(\lambda) + \frac{\gamma_0(\lambda)}{\gamma(\lambda)} \|d_{k+1}(\lambda)\|^2 + \frac{\|B_{k+1}^{(1)} d_{k+1}(\lambda)\|^2}{\gamma(\lambda)} \end{aligned}$$

Λόγω της σχέσης

$$B_{k+1}^{(1)} d_{k+1}(\lambda) = -[\lambda d_{k+1}(\lambda) + g_{k+1}],$$

ο παρανομαστής Π της επανάληψης Newton (5.4), μπορεί να γραφεί ως

$$\Pi = \sum_{i=0}^2 \frac{\gamma_i(\lambda) \lambda^i}{\gamma(\lambda)} \|d_{k+1}(\lambda)\|^2 + \sum_{i=1}^2 \frac{i \gamma_i(\lambda) \lambda^{i-1}}{\gamma(\lambda)} d_{k+1}(\lambda)^T g_{k+1} + \frac{\|g_{k+1}\|^2}{\gamma(\lambda)}.$$

Τώρα, στην ειδική περίπτωση όπου $a_k = 2$, δηλαδή όταν τα s_k και y_k είναι γραμμικά εξαρτημένα, από τη σχέση (5.14), έχουμε ότι

$$\Pi = \frac{\theta_{k+1}}{1 + \lambda \theta_{k+1}} \|d_{k+1}(\lambda)\|^2.$$

Αντικαθιστώντας στην παραπάνω έκφραση τη σχέση (5.15), βρίσκουμε τελικά ότι ο παρανομαστής της επανάληψης Newton (5.4), ισούται με

$$\Pi = \left(\frac{\theta_{k+1}}{1 + \lambda \theta_{k+1}} \right)^3 \|g_{k+1}\|^2.$$

5.2.3 Η συνάρτηση βήματος στην 2BFGS ενημέρωση

Ας εξετάσουμε την περίπτωση όπου χρησιμοποιούμε την 2BFGS ενημέρωση μικρής μνήμης για την επίλυση του ΥΠΕ στην τυπική περίπτωση. Όταν $\lambda \in (-\lambda_1, \infty)$ χρησιμοποιούμε το Λήμμα 4.5 της Υποενότητας 4.3.2 (σελ. 145). Σύμφωνα με αυτό, ο αντίστροφος του πίνακα $(B_{k+1}^{(2)} + \lambda I)$ δίνεται από τον τύπο

$$(B_{k+1}^{(2)} + \lambda I)^{-1} = \frac{1}{\nu(\lambda)} \sum_{i=0}^4 (-1)^i \nu_i(\lambda) (B_{k+1}^{(2)})^i \quad (5.16)$$

όπου $\nu_4(\lambda) = 1$, $\nu_3(\lambda) = \lambda + \beta_3 + 1/\theta_k$, $\nu_2(\lambda) = \lambda\nu_3(\lambda) + \beta_2\beta_0 + \beta_3/\theta_k$,
 $\nu_1(\lambda) = \lambda\nu_2(\lambda) + \beta_1\beta_0 + \beta_2\beta_0/\theta_k$, $\nu_0(\lambda) = \lambda\nu_1(\lambda) + \beta_0 + \beta_0\beta_1/\theta_k$ και
 $\nu(\lambda) = \lambda\nu_0(\lambda) + \beta_0/\theta_k$, ενώ

$$\begin{aligned}\beta_0 &= \frac{1}{\theta_k^4} \frac{s_k^T y_k}{s_k^T B_k^{(1)} s_k}, \\ \beta_1 &= (a_{k-1} + 2)\theta_k - 2 \frac{s_k^T w_k}{s_k^T y_k} + b_k \theta_{k+1}, \\ \beta_2 &= \theta_k(\beta_1 - a_{k-1}\theta_k)(a_{k-1} + 2) - 2\theta_k^2 + \left(\frac{s_k^T w_k}{s_k^T y_k} \right)^2 \\ &\quad - \theta_{k+1} \frac{w_k^T w_k}{s_k^T y_k} + 2 \frac{s_k^T H_k w_k}{s_k^T y_k} - b_k \frac{s_k^T H_k s_k}{s_k^T y_k}, \\ \beta_3 &= \frac{a_{k-1} + 2}{\theta_k} - \frac{\|B_k^{(1)} s_k\|^2}{s_k^T B_k^{(1)} s_k} + \frac{\|y_k\|^2}{s_k^T y_k},\end{aligned}$$

όπου $H_k = (B_k^{(1)})^{-1}$, $w_k = H_k y_k$ και $b_k = 1 + \frac{y_k^T w_k}{s_k^T y_k}$. Επομένως, το δοκιμαστικό βήμα

$$d_{k+1}(\lambda) = - \left(B_{k+1}^{(2)} + \lambda I \right)^{-1} g_{k+1} = \frac{1}{\nu(\lambda)} \sum_{i=0}^4 (-1)^{i+1} \nu_i(\lambda) \left(B_{k+1}^{(2)} \right)^i g_{k+1}$$

μπορεί να υπολογιστεί χρησιμοποιώντας την Διαδικασία 4.1 (σελ. 148). Στη συγκεκριμένη περίπτωση, θα έχουμε:

Διαδικασία 5.1. (Υπολογισμός δοκιμαστικού βήματος)

```

 $d^{(4)} = g;$ 
for  $j = 4, \dots, 1$  do
   $v \leftarrow B^{(2)} d^{(j)};$ 
   $d^{(j-1)} \leftarrow v + (-1)^{j-1} \nu_{j-1}(\lambda) g;$ 
end for
return  $-d^{(0)}/\nu(\lambda);$ 

```

Το βοηθητικό διάνυσμα $v \in \mathbb{R}^n$ μπορεί να υπολογιστεί μέσω του τύπου

$$v \equiv B_{k+1}^{(2)} d^{(j)} = B_k^{(1)} d^{(j)} - \frac{s_k^T B_k^{(1)} d^{(j)}}{s_k^T B_k^{(1)} s_k} B_k^{(1)} s_k + \frac{y_k^T d^{(j)}}{s_k^T y_k} y_k,$$

ενώ για τα διανύσματα $B_k^{(1)} d^{(j)}$ και $B_k^{(1)} s_k$, έχουμε ότι

$$B_k^{(1)} d^{(j)} = \frac{1}{\theta_k} d^{(j)} - \frac{s_{k-1}^T d^{(j)}}{\theta_k s_{k-1}^T s_{k-1}} s_{k-1} + \frac{y_{k-1}^T d^{(j)}}{s_{k-1}^T y_{k-1}} y_{k-1}$$

και

$$B_k^{(1)} s_k = \frac{1}{\theta_k} s_k - \frac{s_{k-1}^T s_k}{\theta_k s_{k-1}^T s_{k-1}} s_{k-1} + \frac{y_{k-1}^T s_k}{s_{k-1}^T y_{k-1}} y_{k-1}$$

αντίστοιχα.

Στην περίπτωση όπου ο πολλαπλασιαστή Lagrange είναι μηδέν, δηλαδή $\lambda = 0$, η διαδικασία υπολογισμού του δοκιμαστικού βήματος απλοποιείται περαιτέρω, ως εξής:

$$d_{k+1}(0) \equiv -H_{k+1} g_{k+1} = -H_k g_{k+1} + \frac{w_k^T g_{k+1} - b_k s_k^T g_{k+1}}{s_k^T y_k} s_k + \frac{s_k^T g_{k+1}}{s_k^T y_k} w_k,$$

όπου $H_{k+1} = (B_{k+1}^{(2)})^{-1}$ και $H_k = (B_k^{(1)})^{-1}$.

Αν $a_{k-1} = a_k = 2$, τότε από το Πόρισμα 4.2 (σελ. 147), γνωρίζουμε ότι ο αντίστροφος του $B_{k+1}^{(2)} + \lambda I$ έχει τη μορφή

$$\begin{aligned} (B_{k+1}^{(2)} + \lambda I)^{-1} &= -\frac{1}{\left(\frac{1}{\theta_k} + \lambda\right) \left(\frac{1}{\theta_{k+1}} + \lambda\right)} \left[B_{k+1}^{(2)} \right. \\ &\quad \left. - \left(\lambda + \frac{1}{\theta_k} + \frac{1}{\theta_{k+1}} \right) I \right]. \end{aligned} \quad (5.17)$$

Επομένως, το δοκιμαστικό βήμα υπολογίζεται από την ακόλουθη σχέση:

$$\begin{aligned} d_{k+1}(\lambda) &= -\left(B_{k+1}^{(2)} + \lambda I \right)^{-1} g_{k+1} \\ &= \frac{1}{\left(\frac{1}{\theta_k} + \lambda\right) \left(\frac{1}{\theta_{k+1}} + \lambda\right)} \left[B_{k+1}^{(2)} g_{k+1} - \left(\lambda + \frac{1}{\theta_k} + \frac{1}{\theta_{k+1}} \right) g_{k+1} \right]. \end{aligned}$$

Όσον αφορά τη μέθοδο Newton για τον προσδιορισμό του βέλτιστου λ^* , η έκφραση του παρανομαστή της επαναληπτικής διαδικασίας (5.4),

$$\Pi = d(\lambda)^T (B + \lambda I)^{-1} d(\lambda),$$

γίνεται μέσω της σχέσης (5.16). Βρίσκουμε ότι

$$\begin{aligned} \Pi &= \frac{\nu_0(\lambda)}{\nu(\lambda)} \|d_{k+1}(\lambda)\|^2 - \frac{\nu_1(\lambda)}{\nu(\lambda)} d_{k+1}(\lambda)^T B_{k+1}^{(2)} d_{k+1}(\lambda) \\ &\quad + \frac{\nu_2(\lambda)}{\nu(\lambda)} \|B_{k+1}^{(2)} d_{k+1}(\lambda)\|^2 - \frac{\nu_3(\lambda)}{\nu(\lambda)} d_{k+1}(\lambda)^T \left(B_{k+1}^{(2)}\right)^3 d_{k+1}(\lambda) \\ &\quad + \frac{\nu_4(\lambda)}{\nu(\lambda)} \left\| \left(B_{k+1}^{(2)}\right)^2 d_{k+1}(\lambda) \right\|^2. \end{aligned} \quad (5.18)$$

Όμως, επειδή ισχύει η σχέση $B_{k+1}^{(2)} d_{k+1}(\lambda) = -[\lambda d_{k+1}(\lambda) + g_{k+1}]$, έχουμε ότι

$$\begin{aligned} d_{k+1}(\lambda)^T B_{k+1}^{(2)} d_{k+1}(\lambda) &= -\lambda \|d_{k+1}(\lambda)\|^2 - d_{k+1}(\lambda)^T g_{k+1}, \\ \|B_{k+1}^{(2)} d_{k+1}(\lambda)\|^2 &= \lambda^2 \|d_{k+1}(\lambda)\|^2 + 2\lambda d_{k+1}(\lambda)^T g_{k+1} + \|g_{k+1}\|^2, \\ d_{k+1}(\lambda)^T \left(B_{k+1}^{(2)}\right)^3 d_{k+1}(\lambda) &= -\lambda^3 \|d_{k+1}(\lambda)\|^2 - 3\lambda^2 d_{k+1}(\lambda)^T g_{k+1} \\ &\quad - 2\lambda \|g_{k+1}\|^2 + g_{k+1}^T B_{k+1}^{(2)} g_{k+1} \quad \text{και}, \\ \left\| \left(B_{k+1}^{(2)}\right)^2 d_{k+1}(\lambda) \right\|^2 &= \lambda^4 \|d_{k+1}(\lambda)\|^2 + 4\lambda^3 d_{k+1}(\lambda)^T g_{k+1} \\ &\quad + 3\lambda^2 \|g_{k+1}\|^2 - 2\lambda g_{k+1}^T B_{k+1}^{(2)} g_{k+1} \\ &\quad + \|B_{k+1}^{(2)} g_{k+1}\|^2. \end{aligned}$$

Αντικαθιστώντας τις παραπάνω ιδιότητες στη σχέση (5.18), βρίσκουμε ότι

$$\begin{aligned} \Pi &= \frac{1}{\nu(\lambda)} \left\{ \sum_{i=0}^4 [\lambda^i \nu_i(\lambda)] \|d_{k+1}(\lambda)\|^2 + \sum_{i=1}^4 [i \lambda^{i-1} \nu_i(\lambda)] d_{k+1}(\lambda)^T g_{k+1} \right. \\ &\quad + \sum_{i=2}^4 [(i-1) \lambda^{i-2} \nu_i(\lambda)] \|g_{k+1}\|^2 \\ &\quad - \sum_{i=3}^4 [(i-2) \lambda^{i-3} \nu_i(\lambda)] g_{k+1}^T B_{k+1}^{(2)} g_{k+1} \\ &\quad \left. + \|B_{k+1}^{(2)} g_{k+1}\|^2 \right\}. \end{aligned}$$

Αν $a_{k-1} = a_k = 2$, τότε από τη σχέση (5.17) έχουμε ότι

$$\Pi = -\frac{d_{k+1}(\lambda)^T}{(1/\theta_k + \lambda)(1/\theta_{k+1} + \lambda)} \left[B_{k+1}^{(2)} - (\lambda + 1/\theta_k + 1/\theta_{k+1}) I \right] d_{k+1}(\lambda).$$

Από την εξίσωση $B_{k+1}^{(2)} d_{k+1}(\lambda) = -[\lambda d_{k+1}(\lambda) + g_{k+1}]$, βρίσκουμε ότι η παραπάνω σχέση μπορεί να γραφεί ως εξής:

$$\Pi = \frac{(2\lambda + 1/\theta_k + 1/\theta_{k+1}) \|d_{k+1}(\lambda)\|^2 + d_{k+1}(\lambda)^T g_{k+1}}{(\lambda + 1/\theta_k)(\lambda + 1/\theta_{k+1})}.$$

Από την ανωτέρω ανάλυση γίνεται σαφές ότι όταν χρησιμοποιούμε την 2BFGS ενημέρωση, μπορούμε πολύ εύκολα σε κάθε περίπτωση να εφαρμόσουμε το επαναληπτικό σχήμα της Newton για τον υπολογισμό του πολλαπλασιαστή Lagrange και κατά συνέπεια τη συνάρτηση του δοκιμαστικού βήματος $d(\lambda)$.

5.3 Η δύσκολη περίπτωση

Η δύσκολη περίπτωση (hard case) εμφανίζεται όταν ο πίνακας B είναι μη θετικά ορισμένος και το διάνυσμα της κλίσης g είναι κάθετο σε όλα τα ιδιοδιανύσματα που αντιστοιχούν στην πιο αρνητική ιδιοτιμή του πίνακα. Αναλυτικά, μπορεί κανείς να ανατρέξει στο Κεφάλαιο 2 και συγκεκριμένα στην Υποενότητα 2.8.2.

Στην περίπτωση αυτή, ο βέλτιστος πολλαπλασιαστής Lagrange είναι ίσος με $\lambda^* = -\lambda_1$, όπου λ_1 η πιο αρνητική ιδιοτιμή του πίνακα. Συνεπώς, ο πίνακας $(B - \lambda_1 I)$ είναι ιδιάζον και χρειάζεται να υπολογιστεί μια κατεύθυνση αρνητικής καμπυλότητας, και συγκεκριμένα το ιδιοδιάνυσμα u_1 , το οποίο αντιστοιχεί στην λ_1 . Σε αυτή την περίπτωση, σύμφωνα με τους Moré & Sorensen, το βήμα d ισούται με

$$d = -(B - \lambda_1 I)^\dagger g + \tau u,$$

που $(B - \lambda_1 I)^\dagger$ είναι ο Moore-Penrose ψευδοαντίστροφος (σελ. 5), ενώ η παράμετρος $\tau \in \mathbb{R}$ ισούται με

$$\tau = \frac{\Delta^2 - \|d\|^2}{d^T u + \operatorname{sgn}(d^T u)^2 \sqrt{(d^T u)^2 + (\Delta^2 - \|d\|^2)}},$$

έτσι ώστε να ικανοποιείται ο περιορισμός της περιοχής εμπιστοσύνης

$$\|d\| = \Delta.$$

Οι δυσκολίες που παρουσιάζει η δύσκολη περίπτωση (hard case), είναι αφενός μεν η εύρεση της πιο μικρής ιδιοτιμής του πίνακα B , αφετέρου δε, ο υπολογισμός του αντίστοιχου ιδιοδιανύσματος. Όταν ο B είναι ένας Quasi-Newton πίνακας ελάχιστης ή μικρής μνήμης, είναι εύκολο να υπολογίσουμε αυτές τις ποσότητες. Στα Κεφάλαια 3 και 4 αναλύσαμε διεξοδικά τον τρόπο υπολογισμού του χαρακτηριστικού πολυωνύμου και των ιδιοδιανυσμάτων που αντιστοιχούν σε διακριτές ιδιοτιμές, για κάθε Quasi-Newton πίνακα ελάχιστης ή μικρής μνήμης. Επομένως, μπορούμε να βρούμε με μεγάλη ακρίβεια τόσο την πιο αρνητική ιδιοτιμή, όσο και το αντίστοιχο ιδιοδιάνυσμα. Επίσης, μέσω του κλειστού τύπου που εκφράστηκε ο αντίστροφος του πίνακα $B + \mu I$, μπορούμε να υπολογίσουμε επακριβώς τον πίνακα $(B + \hat{\lambda}I)^{-1}$, όπου $\hat{\lambda} = -\lambda_1 + \epsilon$, είναι η κατά ϵ -διαταραγμένη μικρότερη ιδιοτιμή του B , με αντίθετο πρόσημο.

Στη συνέχεια, μέσω της ανάλυσης των Κεφαλαίων 3 και 4, θα συνοψίσουμε για κάθε ενημέρωση ξεχωριστά, τον τρόπο υπολογισμού των ποσοτήτων που μας ενδιαφέρουν για την υλοποίηση της προτεινόμενης μεθόδου. Συγκεκριμένα θα δώσουμε τους τύπους των χαρακτηριστικών πολυωνύμων και των ιδιοδιανυσμάτων που αντιστοιχούν στην πιο αρνητική ιδιοτιμή του B .

5.3.1 Η γενική PSB ενημέρωση ελάχιστης μνήμης

Σύμφωνα με την ανάλυση της Ενότητας 3.2, το χαρακτηριστικό πολυώνυμο της γενικής PSB ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)w_k^T + w_k(y_k - \theta_{k+1}s_k)^T}{s_k^T w_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T w_k} \frac{w_k w_k^T}{s_k^T w_k},$$

ορίζεται στο Θεώρημα 3.1 (σελίδα 86). Σύμφωνα με αυτό, το χαρακτηριστικό πολυώνυμο είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} (\lambda^2 - \beta_1 \lambda + \beta_2),$$

όπου

$$\begin{aligned} \beta_1 &= 2 \frac{y_k^T w_k}{s_k^T w_k} - \frac{(y_k - \theta_{k+1} s_k)^T s_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k} \quad \text{και} \\ \beta_2 &= \left(\frac{y_k^T w_k}{s_k^T w_k} \right)^2 - \frac{(y_k - \theta_{k+1} s_k)^T y_k}{s_k^T w_k} \frac{w_k^T w_k}{s_k^T w_k}. \end{aligned}$$

Όσον αφορά το ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του B , από την Πρόταση 3.2 (σελίδα 96), γνωρίζουμε ότι αν η λ_1 είναι διακριτή, τότε το αντίστοιχο ιδιοδιάνυσμα είναι το

$$u(\lambda_1) = \frac{(\lambda_1 s_k - y_k)^T y_k}{(\lambda_1 s_k - y_k)^T s_k} s_k - y_k. \quad (5.19)$$

Εναλλακτικά, μπορούμε να υπολογίσουμε το ιδιοδιάνυσμα όπως αυτό δόθηκε στο Πόρισμα 4.1 (σελίδα 136). Σύμφωνα με αυτό, έχουμε ότι

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)z_k + c_3(\sigma)w_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)s_k^T w_k},$$

όπου $\sigma = \lambda_1 - \epsilon$ και

$$\begin{aligned} c_1(\sigma) &= s_k^T w_k (\sigma^2 - \beta_1\sigma + \beta_2), \\ c_2(\sigma) &= - \left(\theta_{k+1} - \sigma + \frac{z_k^T w_k}{s_k^T w_k} \right) \sum_{i=1}^n w_k^i + \frac{w_k^T w_k}{s_k^T w_k} \sum_{i=1}^n z_k^i, \quad \text{και} \\ c_3(\sigma) &= \left[\frac{z_k^T z_k}{s_k^T w_k} + (\theta_{k+1} - \sigma) \frac{z_k^T s_k}{s_k^T w_k} \right] \sum_{i=1}^n w_k^i - \left(\theta_{k+1} - \sigma + \frac{z_k^T w_k}{s_k^T w_k} \right) \sum_{i=1}^n z_k^i. \end{aligned}$$

Αν επιπρόσθετα τα διανύσματα s_k και y_k είναι συγγραμμικά και έστω $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, τότε και $w_k = \rho s_k$, $\rho \in \mathbb{R}^*$. Σε αυτή την περίπτωση έχουμε ότι το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 είναι το

$$u = s_k.$$

Αν η λ_1 είναι πολλαπλή ιδιοτιμή και $\kappa \neq \theta_{k+1}$, τότε ένα γενικευμένο ιδιοδιάνυσμα που αντιστοιχεί στην πολλαπλή ιδιοτιμή θ_{k+1} είναι το

$$u = (-s_k^n/s_k^1, 0, \dots, 0, \dots, 1)^T.$$

Αν ισχύει ότι $\kappa = \theta_{k+1}$, τότε

$$u \equiv e_i = (0, 0, \dots, 1, 0, \dots, 0)^T, \quad i = 1, 2, \dots, n,$$

όπου το 1 εμφανίζεται στην i συνιστώσα του διανύσματος e_i .

5.3.1.1 Η SR1 ενημέρωση ελάχιστης μνήμης. Στην SR1 ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)(y_k - \theta_{k+1}s_k)^T}{(y_k - \theta_{k+1}s_k)^T s_k},$$

σύμφωνα με την ανάλυση της Ενότητας 3.5, το χαρακτηριστικό πολυώνυμο είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-1} \left(\lambda - \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k} \right).$$

Αν η λ_1 είναι διακριτή, δηλαδή αν

$$\lambda_1 = \frac{(y_k - \theta_{k+1}s_k)^T y_k}{(y_k - \theta_{k+1}s_k)^T s_k},$$

τότε το αντίστοιχο ιδιοδιάνυσμα είναι το

$$u = y_k - \theta_{k+1}s_k.$$

Αν η λ_1 είναι πολλαπλή ιδιοτιμή, δηλαδή αν $\lambda_1 = \theta_{k+1}$, τότε ένα γενικευμένο ιδιοδιάνυσμα που αντιστοιχεί σε αυτή είναι το

$$u_i = \left(-\frac{y_k^i - \theta_{k+1}s_k^i}{y_k^1 - \theta_{k+1}s_k^1}, 0, \dots, 0, 1, 0, \dots, 0 \right)^T,$$

όπου το $i = 1, \dots, n-1$ δηλώνει την i συνιστώσα των y_k, s_k , ενώ η μονάδα

βρίσκεται στην i γραμμή του u_i .

5.3.1.2 Η BFGS ενημέρωση ελάχιστης μνήμης. Στην BFGS ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I - \theta_{k+1} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

το χαρακτηριστικό πολυώνυμο είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(\theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k} \right) \lambda + \theta_{k+1} \frac{s_k^T y_k}{s_k^T s_k} \right].$$

Το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 , αν αυτή είναι διακριτή, μπορεί να υπολογιστεί είτε μέσω του τύπου (5.19) ή μέσω του Πορίσματος 3.2 (σελίδα 108). Σύμφωνα με το τελευταίο, έχουμε ότι

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)},$$

όπου $\sigma = \lambda_1 + \epsilon$

$$\begin{aligned} c_1(\sigma) &= \sigma^2 - \beta_1\sigma + \beta_2, \\ c_2(\sigma) &= \left[(\beta_1 - \sigma) \sum_{i=1}^n s_k^i - \sum_{i=1}^n y_k^i \right] \frac{\theta_{k+1}}{s_k^T s_k}, \quad \text{και} \\ c_3(\sigma) &= \frac{\sigma}{s_k^T y_k} \sum_{i=1}^n y_k^i - \frac{\theta_{k+1}}{s_k^T s_k} \sum_{i=1}^n s_k^i, \end{aligned}$$

και

$$\beta_1 = \theta_{k+1} + \frac{y_k^T y_k}{s_k^T y_k}, \quad \beta_2 = \theta_{k+1} \frac{s_k^T y_k}{s_k^T s_k}.$$

Αν ισχύει ότι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, τότε το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 είναι το

$$u = s_k.$$

Αν η λ_1 είναι πολλαπλή ιδιοτιμή, τότε ένα αντίστοιχο γενικευμένο ι-

διοδιάνυσμα είναι το

$$u = \begin{cases} (-s_k^n/s_k^1, 0, \dots, 0, \dots, 1)^T, & \text{αν } \kappa \neq \theta_{k+1}; \\ (1, 0, \dots, 0, 0, \dots, 0)^T, & \text{αν } \kappa = \theta_{k+1}. \end{cases} \quad (5.20)$$

5.3.1.3 Η DFP ενημέρωση ελάχιστης μνήμης.

Στην DFP ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)y_k^T + y_k(y_k - \theta_{k+1}s_k)^T}{s_k^T y_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T y_k} \frac{y_k y_k^T}{s_k^T y_k},$$

το χαρακτηριστικό πολυώνυμο, όπως δίνεται στην Ενότητα 3.7 (σελίδα 112), είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k} \lambda + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k} \right].$$

Αν η λ_1 είναι διακριτή, το αντίστοιχο ιδιοδιάνυσμα μπορούμε να το υπολογίσουμε από τον τύπο (5.19). Ακόμα, μπορούμε να το υπολογίσουμε χρησιμοποιώντας το Πόρισμα 3.3 (σελίδα 117). Σύμφωνα με αυτό, έχουμε ότι

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)},$$

όπου $\sigma = \lambda_1 + \epsilon$

$$c_1(\sigma) = \sigma^2 - \beta_1\sigma + \beta_2,$$

$$c_2(\sigma) = \left(\beta_2 \sum_{i=1}^n s_i - \sigma \sum_{i=1}^n y_i \right) \frac{\theta_{k+1}}{s_k^T y_k} \quad \text{και}$$

$$c_3(\sigma) = -\frac{\sigma\theta_{k+1}}{s_k^T y_k} \sum_{i=1}^n s_i - \left(\frac{\theta_{k+1}}{s_k^T y_k} - \frac{\sigma\beta_1}{y_k^T y_k} \right) \sum_{i=1}^n y_i$$

και

$$\beta_1 = \left(1 + \theta_{k+1} \frac{s_k^T s_k}{s_k^T y_k} \right) \frac{y_k^T y_k}{s_k^T y_k}, \quad \beta_2 = \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}.$$

Αν ισχύει ότι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, τότε το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 είναι το

$$u = s_k.$$

Αν η λ_1 είναι πολλαπλή ιδιοτιμή, τότε ένα αντίστοιχο γενικευμένο ι-διοδιάνυσμα υπολογίζεται από τις σχέσεις (5.20).

5.3.1.4 Η PSB ενημέρωση ελάχιστης μνήμης. Στην PSB ενημέρωση ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I + \frac{(y_k - \theta_{k+1}s_k)s_k^T + s_k(y_k - \theta_{k+1}s_k)^T}{s_k^T s_k} - \frac{(y_k - \theta_{k+1}s_k)^T s_k}{s_k^T s_k} \frac{s_k s_k^T}{s_k^T s_k},$$

το χαρακτηριστικό πολυώνυμο είναι το

$$p(\lambda) = (\lambda - \theta_{k+1})^{n-2} \left[\lambda^2 - \left(\theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k} \right) \lambda + \frac{(\theta_{k+1}s_k^T y_k - y_k^T y_k)s_k^T s_k + (s_k^T y_k)^2}{(s_k^T s_k)^2} \right]$$

(βλέπε Ενότητα 3.8, σελ. 118). Αν η λ_1 είναι διακριτή, το αντίστοιχο ιδιοδιάνυσμα μπορούμε να το υπολογίσουμε είτε από τον τύπο (5.19), ή μέσω του Πορίσματος 3.4 (σελίδα 122), από όπου έχουμε ότι

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n}(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)},$$

όπου $\sigma = \lambda_1 + \epsilon$,

$$c_1(\sigma) = \sigma^2 - \beta_1\sigma + \beta_2,$$

$$c_2(\sigma) = \frac{1}{s_k^T s_k} \left[\frac{y_k^T y_k}{s_k^T s_k} - \beta_1(\sigma - \theta_{k+1}) \right] \sum_{i=1}^n s_i + \frac{\sigma - \beta_1}{s_k^T s_k} \sum_{i=1}^n y_i \quad \text{και}$$

$$c_3(\sigma) = \frac{\sigma - \beta_1}{s_k^T s_k} \sum_{i=1}^n s_i + \frac{1}{s_k^T s_k} \sum_{i=1}^n y_i,$$

και

$$\beta_1 = \theta_{k+1} + \frac{s_k^T y_k}{s_k^T s_k}, \quad \beta_2 = [(\theta_{k+1} s_k^T y_k - y_k^T y_k) s_k^T s_k + (s_k^T y_k)^2] / (s_k^T s_k)^2.$$

Αν ισχύει ότι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, τότε το ιδιοδιάνυσμα που αντιστοιχεί στη λ_1 είναι το

$$u = s_k.$$

Αν τέλος η λ_1 είναι πολλαπλή ιδιοτιμή, τότε ένα αντίστοιχο γενικευμένο ιδιοδιάνυσμα υπολογίζεται από τις σχέσεις (5.20).

5.3.2 Η 1BFGS ενημέρωση στη δύσκολη περίπτωση

Στην 1BFGS ενημέρωση ελάχιστης μνήμης

$$B_{k+1}^{(1)} = \frac{1}{\theta_{k+1}} I - \frac{s_k s_k^T}{\theta_{k+1} s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \quad \text{όπου} \quad B_{k+1}^{(0)} = (1/\theta_{k+1}) I,$$

όπου $\theta_{k+1} = (s_k^T s_k) / (s_k^T y_k)$ είναι η BB φασματική παράμετρος, γνωρίζουμε από την ανάλυση της Ενότητας 4.2, ότι το χαρακτηριστικό πολυώνυμο βρίσκεται μέσω του Λήμματος 4.1 (σελίδα 129), και είναι το

$$p_1(\lambda) = \left(\lambda - \frac{1}{\theta_{k+1}}\right)^{n-2} \left(\lambda^2 - \frac{a_k}{\theta_{k+1}} \lambda + \frac{1}{\theta_{k+1}^2}\right),$$

$$\text{όπου } a_k = 1 + \theta_{k+1} \frac{y_k^T y_k}{s_k^T y_k}.$$

Αν η ιδιοτιμή λ_1 είναι διακριτή, τότε το αντίστοιχο ιδιοδιάνυσμα μπορεί να υπολογιστεί μέσω του τύπου (5.19), ή μέσω του Πορίσματος 3.2 (σελ. 108). Από το πόρισμα αυτό έχουμε ότι

$$\hat{u}(\hat{\lambda}) = -\gamma_u(\hat{\lambda}) u + \gamma_{us}(\hat{\lambda}) s_k - \gamma_{uy}(\hat{\lambda}) y_k,$$

όπου $\hat{\lambda} = -\lambda_1 + \epsilon$ και

$$\begin{aligned}\gamma_u(\hat{\lambda}) &= \frac{1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + \gamma_0(\hat{\lambda}) \theta_{k+1}^2}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2}, \\ \gamma_{us}(\hat{\lambda}) &= \frac{1 - \gamma_1(\hat{\lambda}) \theta_{k+1}}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T s_k} \sum_{i=1}^n s_i + \frac{1}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T s_k} \sum_{i=1}^n y_i, \quad \text{και} \\ \gamma_{uy}(\hat{\lambda}) &= \frac{\left[1 - \gamma_1(\hat{\lambda}) \theta_{k+1} + a_k\right] \theta_{k+1}}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T y_k} \sum_{i=1}^n y_i - \frac{1}{\sqrt{n} \gamma(\hat{\lambda}) \theta_{k+1}^2 s_k^T y_k} \sum_{i=1}^n s_i.\end{aligned}$$

Αν η λ_1 είναι πολλαπλή, τότε ένα γενικευμένο ιδιοδιάνυσμα είναι το

$$u_1 = e_1 = (1, 0, \dots, 0)^T.$$

5.3.3 Η 2BFGS ενημέρωση στη δύσκολη περίπτωση

Στην 2BFGS ενημέρωση μικρής μνήμης

$$B_{k+1}^{(2)} = B_k^{(1)} - \frac{B_k^{(1)} s_k s_k^T B_k^{(1)}}{s_k^T B_k^{(1)} s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

το χαρακτηριστικό πολυώνυμο είναι το

$$p_2(\lambda) = (\lambda - 1/\theta_k)^{n-4} (\lambda^4 - \beta_3 \lambda^3 + \beta_2 \beta_0 \lambda^2 - \beta_1 \beta_0 \lambda + \beta_0),$$

όπου

$$\begin{aligned}\beta_0 &= \frac{1}{\theta_k^4} \frac{s_k^T y_k}{s_k^T B_k^{(1)} s_k}, \quad \beta_1 = (a_{k-1} + 2) \theta_k - 2 \frac{s_k^T w_k}{s_k^T y_k} + b_k \theta_{k+1}, \\ \beta_2 &= \theta_k (\beta_1 - a_{k-1} \theta_k) (a_{k-1} + 2) - 2 \theta_k^2 + \left(\frac{s_k^T w_k}{s_k^T y_k} \right)^2 - \theta_{k+1} \frac{w_k^T w_k}{s_k^T y_k} \\ &\quad + 2 \frac{s_k^T H_k w_k}{s_k^T y_k} - b_k \frac{s_k^T H_k s_k}{s_k^T y_k}, \\ \beta_3 &= \frac{a_{k-1} + 2}{\theta_k} - \frac{\|B_k^{(1)} s_k\|^2}{s_k^T B_k^{(1)} s_k} + \frac{\|y_k\|^2}{s_k^T y_k},\end{aligned}$$

με $H_k = \left(B_k^{(1)}\right)^{-1}$, $w_k = H_k y_k$ και $b_k = 1 + \frac{y_k^T w_k}{s_k^T y_k}$ (βλέπε Λήμμα 4.3, σελ. 138).

Στην περίπτωση όπου η λ_1 είναι διακριτή ιδιοτιμή, έχουμε ότι το αντίστοιχο ιδιοδιάνυσμα είναι το

$$\hat{u} = \left[\sum_{i=0}^4 (-1)^i \nu_i(\hat{\lambda}) \left(B_{k+1}^{(2)}\right)^i \right] \frac{u}{\nu(\hat{\lambda})},$$

όπου $\hat{\lambda} = -\lambda_1 + \epsilon$ και $\nu_4(\hat{\lambda}) = 1$, $\nu_3(\hat{\lambda}) = \hat{\lambda} + \beta_3 + 1/\theta_k$, $\nu_2(\hat{\lambda}) = \hat{\lambda}\nu_3(\hat{\lambda}) + \beta_2\beta_0 + \beta_3/\theta_k$, $\nu_1(\hat{\lambda}) = \hat{\lambda}\nu_2(\hat{\lambda}) + \beta_1\beta_0 + \beta_2\beta_0/\theta_k$, $\nu_0(\hat{\lambda}) = \hat{\lambda}\nu_1(\hat{\lambda}) + \beta_0 + \beta_0\beta_1/\theta_k$ και $\nu(\hat{\lambda}) = \hat{\lambda}\nu_0(\hat{\lambda}) + \beta_0/\theta_k$. Το συγκεκριμένο ιδιοδιάνυσμα υπολογίζεται μέσω της Διαδικασίας 4.1 (σελίδα 148). Αν η λ_1 είναι η μόνη διακριτή ιδιοτιμή, τότε το ιδιοδιάνυσμα που αντιστοιχεί σε αυτή είναι το κανονικοποιημένο διάνυσμα

$$u = y_k.$$

Αν η λ_1 είναι πολλαπλή ιδιοτιμή, τότε ένα γενικευμένο ιδιοδιανύσματα που αντιστοιχεί σε αυτή είναι της μορφής

$$\hat{u}_1 = \left(-\frac{y_k^{(n)}}{y_k^{(1)}}, 0, \dots, 0, 1\right)^T, \quad (5.21)$$

όπου το $y_k^{(i)}$ δηλώνει την i συνιστώσα του y_k . Τέλος, αν η λ_1 είναι η μόνη ιδιοτιμή του πίνακα, τότε

$$u = e_i = (0, 0, \dots, 1, 0, \dots, 0)^T,$$

όπου η μονάδα βρίσκεται στην i θέση.

5.4 Ο προτεινόμενος αλγόριθμος επίλυσης του ΥΠΕ

Στις προηγούμενες ενότητες αναλύσαμε τον τρόπο επίλυσης του ΥΠΕ (5.1), στην περίπτωση όπου ο προσεγγιστικός εσοιανός πίνακας είναι ένας ελάχιστης ή μικρής μνήμης Quasi-Newton πίνακας. Η προτεινόμενη μέθοδος μπορεί να συνοψιστεί σε έναν αλγόριθμο για την επίλυση του ΥΠΕ.

Αλγόριθμος 5.1 Υπολογισμός του δοκιμαστικού βήματος

Βήμα 1: Υπολόγισε τις ιδιοτιμές λ_i του B ; δοθέντος $\epsilon \rightarrow 0^+$, θέσε για άκρα του διαστήματος αβεβαιότητας τα

$$\lambda_L := \max(0, -\lambda_1 + \epsilon) \quad \text{και} \quad \lambda_U := \max |\lambda_i| + (1 + \epsilon)\|g\|/\Delta.$$

Βήμα 2: Αν $\lambda_1 > 0$, τότε αρχικοποίησε τον πολλαπλασιαστή Lagrange λ , θέτοντας $\lambda := 0$ και υπολόγισε το βήμα $d(\lambda)$; αν $\|d\| \leq \Delta$ σταμάτα; αλλιώς πήγαινε στο Βήμα 4.

Βήμα 3: Αρχικοποίησε το λ θέτοντας $\lambda := -\lambda_1 + \epsilon$, τέτοιο ώστε ο πίνακας $B + \lambda I$ να είναι θετικά ορισμένος και υπολόγισε το βήμα $d(\lambda)$;

- a. αν $\|d\| > \Delta$ πήγαινε στο Βήμα 4;
- b. αν $\|d\| = \Delta$ σταμάτα;
- c. αν $\|d\| < \Delta$ υπολόγισε τον πραγματικό αριθμό τ και το ιδιοδιάνυσμα u , έτσι ώστε να ισχύει

$$\| - (B + \hat{\lambda}I)g + \tau u \| = \Delta;$$

θέσε $d := d + \tau u$ και σταμάτα;

Βήμα 4: Χρησιμοποίησε τη μέθοδο Newton για να βρεις ένα πολλαπλασιαστή Lagrange τέτοιο ώστε να ισχύει

$$\lambda \in [\lambda_L, \lambda_U]$$

και υπολόγισε το βήμα $d(\lambda)$;

Βήμα 5: Αν $\|d\| \leq \Delta$ σταμάτα; αλλιώς ενημέρωσε τα άκρα του διαστήματος αβεβαιότητας λ_L και λ_U έτσι ώστε

$$\lambda_L \leq \lambda \leq \lambda_U$$

και πήγαινε στο Βήμα 4.

Ο προτεινόμενος αλγόριθμος είναι ένας ενοποιημένος αλγόριθμος, ο οποίος ενσωματώνει τόσο την τυπική, όσο και τη δύσκολη περίπτωση. Η διαδικασία ελέγχου του πολλαπλασιαστή Lagrange που παράγεται από τις

επαναλήψεις Newton

$$\lambda_{\ell+1} = \lambda_{\ell} - \frac{\psi(\lambda_{\ell})}{\psi'(\lambda_{\ell})}, \quad \ell = 0, 1, 2, \dots$$

χρησιμοποιεί δύο παραμέτρους, τις λ_L και λ_U . Η λ_L αποτελεί το κάτω φράγμα και η λ_U το άνω φράγμα ενός διαστήματος αβεβαιότητας, έτσι ώστε το $[\lambda_L, \lambda_U]$ να περιέχει το βέλτιστο πολλαπλασιαστή λ^* . Παρακινούμενοι από τους Nocedal & Yuan [69], έχουμε θέσει

$$\lambda_L = \max(0, -\lambda_1 + \epsilon)$$

και

$$\lambda_U = \max_{1 \leq i \leq n} |\lambda_i| + (1 + \epsilon)\|g\|/\Delta.$$

Προφανώς, το κάτω φράγμα λ_L είναι μεγαλύτερο από το $-\lambda_1$, κάτι το οποίο βεβαιώνει ότι ο πίνακας $B + \lambda I$ είναι πάντα θετικά ορισμένος.

Η πιο σημαντική εφαρμογή του υποπροβλήματος περιοχής εμπιστοσύνης, είναι στις μεθόδους περιοχής εμπιστοσύνης (βλέπε Κεφάλαιο 2, Ενότητα 2.7). Στο πλαίσιο αυτών των μεθόδων θεμελιώθηκε το ακόλουθο λήμμα από τον Powell [74].

Λήμμα 5.1. *Αν d^* είναι λύση του υποπροβλήματος (5.1), τότε*

$$\phi(0) - \phi(d^*) \geq \frac{1}{2}\|g\| \min\{\Delta, \|g\|/\|B\|\}.$$

Το λήμμα αυτό δίνει ένα κάτω φράγμα για τη μέγιστη μείωση $\phi(0) - \phi(d^*)$ του τετραγωνικού μοντέλου μέσα στην περιοχή εμπιστοσύνης $\|d\| \leq \Delta$. Για την ολική σύγκλιση των αλγορίθμων περιοχής εμπιστοσύνης απαιτείται το δοκιμαστικό βήμα d_k να ικανοποιεί την ανισότητα

$$\phi_k(0) - \phi(d_k) \geq c\|g_k\| \min\{\Delta_k, \|g_k\|/\|B_k\|\} \quad (5.22)$$

για όλα τα k , όπου c είναι μια θετική σταθερά [25, 68]. Το ακόλουθο θεώρημα δείχνει ότι η σχέση (5.22) ικανοποιείται, αν το δοκιμαστικό βήμα παραχθεί μέσω του προτεινόμενου Αλγορίθμου 5.4.

Θεώρημα 5.1. *Υποθέτουμε ότι για τον Quasi-Newton πίνακα B ισχύει $\|B_k\| \leq M < \infty$, για όλα τα k , όπου M είναι μια θετική σταθερά. Αν το*

δοκιμαστικό βήμα d_k υπολογίζεται από τον Αλγόριθμο 5.4, τότε ισχύει ότι

$$\phi_k(0) - \phi_k(d_k(\lambda)) \geq \frac{1}{8} \|g_k\| \min\{\Delta_k, \|g_k\|/\|k\|\}. \quad (5.23)$$

Απόδειξη. Υπενθυμίζουμε ότι σε κάθε επανάληψη k , το τετραγωνικό μοντέλο είναι της μορφής

$$\phi_k(d_k) = d_k^T g_k + \frac{1}{2} d_k^T B_k d_k. \quad (5.24)$$

Θεωρούμε τις ακόλουθες περιπτώσεις:

(i) Αν $B_k > 0$, τότε, από το Βήμα 2 έχουμε ότι αν $\lambda = 0$ και $\|d_k\| \leq \Delta_k$, ο αλγόριθμος σταματά. Σε αυτή την περίπτωση, ισχύει ότι

$$g_k^T d_k = -g_k (B_k)^{-1} g_k \leq \frac{-\|g_k\|^2}{\|B_k\|}. \quad (5.25)$$

Τότε, από τις σχέσεις (5.24) και (5.25) έχουμε ότι

$$\begin{aligned} \phi_k(0) - \phi_k(d_k(\lambda)) &= -d_k(\lambda)^T g_k - \frac{1}{2} d_k(\lambda)^T B_k d_k(\lambda) \\ &= -d_k(\lambda)^T g_k - \frac{1}{2} d_k(\lambda)^T (B_k + \lambda I) d_k(\lambda) + \frac{1}{2} \lambda \|d_k(\lambda)\|^2 \\ &= -\frac{1}{2} d_k(\lambda)^T g_k + \frac{1}{2} \lambda \|d_k(\lambda)\|^2 \\ &\geq \frac{1}{2} \frac{\|g_k\|^2}{\|B_k\|}. \end{aligned} \quad (5.26)$$

Αν $\lambda = 0$ και $\|d_k\| > \Delta_k$, τα Βήματα 4 και 5 του Αλγορίθμου 5.4 υπολογίζουν ένα λ και ένα d , τέτοια ώστε να ισχύουν οι σχέσεις

$$0 < \lambda \leq \lambda_U \quad \text{και} \quad \|d_k(\lambda)\| \leq \Delta_k,$$

αντίστοιχα. Αφού ο B_k είναι θετικά ορισμένος, έχουμε ότι

$$\|B_k + \lambda I\| = \lambda_n + \lambda = \|B_k\| + \lambda,$$

όπου η λ_n είναι η μεγαλύτερη ιδιοτιμή του πίνακα B_k . Επομένως, λαμ-

βάνοντας υπόψη ότι ισχύουν οι ανισότητες

$$\lambda \leq \|B_k\| + (1 + \epsilon)\|g_k\|/\Delta_k$$

και

$$\|B_k + \lambda I\| = \|B_k\| + \lambda \leq 2\|B_k\| + \frac{(1 + \epsilon)\|g_k\|}{\Delta_k} \leq 2 \left(\|B_k\| + \frac{\|g_k\|}{\Delta_k} \right),$$

έχουμε ότι

$$\begin{aligned} g_k^T d_k(\lambda) &\leq -\frac{\|g_k\|^2}{\|(B_k + \lambda I)^{-1}\|} \\ &\leq -\frac{1}{2} \frac{\|g_k\|^2}{\|B_k\| + \|g_k\|/\Delta_k} \\ &\leq -\frac{1}{4} \|g_k\| \min(\Delta_k, \|g_k\|/\|B_k\|). \end{aligned} \quad (5.27)$$

Επομένως, από την μείωση του μοντέλου βρίσκουμε ότι

$$\begin{aligned} \phi_k(0) - \phi_k(d_k(\lambda)) &= -\frac{1}{2} d_k(\lambda)^T g_k + \frac{1}{2} \lambda \|d_k(\lambda)\|^2 \\ &\geq \frac{1}{8} \|g_k\| \min(\Delta_k, \|g_k\|/\|B_k\|). \end{aligned} \quad (5.28)$$

(ii) Αν $B_k \leq 0$, τότε $d_k(\lambda)^T B_k d_k(\lambda) \leq 0$. Από το Βήμα 3(a) μαζί με τα Βήματα 4 και 5, έχουμε ότι

$$\hat{\lambda} \leq \lambda \leq \lambda_U \quad \text{και} \quad \|d_k(\lambda)\| \leq \Delta_k,$$

όπου $\hat{\lambda} = -\lambda_1 + \epsilon$. Άρα, χρησιμοποιώντας τη σχέση (5.27), για τη μείωση του μοντέλου υπολογίζουμε

$$\begin{aligned} \phi_k(0) - \phi_k(d_k(\lambda)) &= -d_k(\lambda)^T g_k - \frac{1}{2} d_k(\lambda)^T B_k d_k(\lambda) \\ &\geq -d_k(\lambda)^T g_k \geq \frac{1}{4} \|g_k\| \min(\Delta_k, \|g_k\|/\|B_k\|) \end{aligned} \quad (5.29)$$

Τέλος, στο Βήμα 3(c), λαμβάνοντας υπόψη ότι $\lambda = \hat{\lambda}$, $u_1^T g_k = 0$ και

$$\|B_k + \hat{\lambda} I\| \leq \lambda_n + \hat{\lambda} \leq 2\|B_k\|,$$

όπου η λ_n είναι η μεγαλύτερη ιδιοτιμή του B_k , βρίσκουμε ότι

$$g_k^T d_k(\hat{\lambda}) = -g_k \left(B_k + \hat{\lambda} I \right)^{-1} g_k \leq -\frac{\|g_k\|^2}{\|B_k + \hat{\lambda}\|} \leq -\frac{\|g_k\|^2}{2\|B_k\|}. \quad (5.30)$$

Συνεπώς, χρησιμοποιώντας τη σχέση (5.30) έχουμε ότι

$$\begin{aligned} \phi_k(0) - \phi_k(d_k(\hat{\lambda})) &= -d_k(\hat{\lambda})^T g_k - \frac{1}{2} d_k(\hat{\lambda})^T B_k d_k(\hat{\lambda}) \\ &\geq -d_k(\hat{\lambda})^T g_k \geq \frac{1}{2} \frac{\|g_k\|^2}{\|B_k\|}. \end{aligned} \quad (5.31)$$

Συνδυάζοντας τις σχέσεις (5.26), (5.28), (5.29) και (5.31), η σχέση (5.23) έπεται άμεσα. $\square$

5.5 Συμπεράσματα

Σε αυτό το κεφάλαιο αναπτύξαμε τη δική μας προσέγγιση για την επίλυση του υποπροβλήματος περιοχής εμπιστοσύνης. Η προτεινόμενη μέθοδος βασίστηκε στις ιδιότητες των Quasi-Newton ενημερώσεων ελάχιστης και μικρής μνήμης. Οι ιδιότητες αυτές μας βοήθησαν στο να μην αποθηκεύσουμε ούτε να παραγοντοποιήσουμε πίνακα. Μπορέσαμε να λύσουμε το γραμμικό σύστημα $(B + \lambda I)d = -g$ και κατά συνέπεια να υπολογίσουμε το βήμα, μέσω της εκτέλεσης μιας σειράς εσωτερικών γινομένων και διανυσματικών αθροισμάτων. Επίσης, ο ακριβής υπολογισμός της μικρότερης ιδιοτιμής, αποτέλεσε τη βάση τόσο της εύρεσης του αντίστοιχου ιδιοδιανύσματος με μεγάλη ακρίβεια και χωρίς υπολογιστικό κόστος, όσο και της απλοποίησης των υπολογισμών στις επαναλήψεις Newton.

Ο αλγόριθμος που προτάθηκε για την επίλυση του ΥΠΕ, ενσωματώνει τόσο την τυπική, όσο και τη δύσκολη περίπτωση. Ειδικά στην τελευταία περίπτωση, η οποία θεωρείται και η πιο απαιτητική υπολογιστικά, η προτεινόμενη μέθοδος στηρίζεται μόνο σε αθροίσματα διανυσμάτων. Μάλιστα, η μεθοδολογία που εφαρμόζεται στην συγκεκριμένη περίπτωση είναι τέτοια, ώστε η δύσκολη περίπτωση να γίνεται τελικά "εύκολη" για το συγκεκριμένο αλγόριθμο. Τα αριθμητικά αποτελέσματα που θα παρουσιάσουμε στο επόμενο κεφάλαιο, επιβεβαιώνουν τα θεωρητικά αποτελέσματά μας. Τέλος, αποδείξαμε ότι ο προτεινόμενος αλγόριθμος διατηρεί τις καλές ιδιότητες σύγκλισης, όταν το ΥΠΕ χρησιμοποιείται στα πλαίσια ενός

αλγορίθμου περιοχής εμπιστοσύνης.

Αριθμητικά αποτελέσματα του ΥΠΕ

Στο κεφάλαιο αυτό παρουσιάζουμε μία εμπειρισιασμένη αριθμητική μελέτη όπου αποτιμούμε υπολογιστικά την κλάση των μεθόδων που βασίζονται στις ελάχιστης και μικρής μνήμης Quasi-Newton ενημερώσεις, για την επίλυση του ΥΠΕ. Ο προτεινόμενος αλγόριθμος που παρουσιάστηκε στο προηγούμενο κεφάλαιο αποτιμήθηκε αφενός μεν μέσα σε έναν τυπικό αλγόριθμο περιοχής εμπιστοσύνης για την βελτιστοποίηση μη γραμμικών συναρτήσεων, αφετέρου δε σε τυχαία στιγμιότυπα του ΥΠΕ, έτσι ώστε να καλυφθούν όλες οι δυνατές περιπτώσεις που προκύπτουν στην πράξη. Τα ακόλουθα αποτελούν ερευνητικά αποτελέσματα της παρούσας διατριβής, μέρος των οποίων καταγράφονται στις εργασίες των Apostolopoulou et al., [6, 8, 10, 12].

6.1 Επίλυση τυχαίων στιγμιότυπων του ΥΠΕ

Θα ξεκινήσουμε την αποτίμηση της προτεινόμενης μεθόδου σε τυχαία παραγόμενα στιγμιότυπα του ΥΠΕ μεγάλης διάστασης. Όπως έχουμε αναφέρει στην Ενότητα 2.8 του Κεφαλαίου 2, το υποπρόβλημα περιοχής εμπιστοσύνης,

$$\min_{d \in \mathbb{R}^n} \phi(d) = g^T d + \frac{1}{2} d^T B d \quad \text{υπό τον περιορισμό} \quad \|d\| \leq \Delta,$$

αποτελεί από μόνο του ένα πρόβλημα βελτιστοποίησης με περιορισμούς, το οποίο συναντάται ως επί το πλείστον στις μεθόδους περιοχής εμπιστοσύνης. Για να δείξουμε την αποτελεσματικότητα και την ακρίβεια της μεθόδου μας, δημιουργήσαμε τυχαία στιγμιότυπα του ΥΠΕ τόσο στην τυπική όσο και στη δύσκολη περίπτωση. Τα στιγμιότυπα αυτά δημιουργήθηκαν χρησιμοποιώντας τυχαία n -διάστατα διανύσματα s και y για την ενημέρωση του προσεγγιστικού εσσιανού πίνακα, για καθεμιά από τις τέσσερις

Quasi-Newton ενημερώσεις. Επίσης, δημιουργήσαμε τυχαία στιγμιότυπα και στην περίπτωση όπου χρησιμοποιούνται οι 1BFGS και 2BFGS ενημερώσεις ελάχιστης και μικρής μνήμης, αντίστοιχα, για να προσεγγίσουν τον πίνακα B . Συγκρίναμε την προτεινόμενη μέθοδο, με άλλες τρεις μεθόδους. Οι δύο από αυτές υπολογίζουν μια σχεδόν ακριβή λύση του ΥΠΕ, ενώ η τρίτη υπολογίζει μια προσεγγιστική λύση, που σαφέστατα είναι υποδεέστερη από μία σχεδόν ακριβή. Οι αλγόριθμοι με τους οποίους συγκρινόμαστε θεωρούνται μέχρι και σήμερα, οι καλύτεροι στη διεθνή βιβλιογραφία για την επίλυση του ΥΠΕ. Στη συνέχεια, θα περιγράψουμε λεπτομερώς τα αποτελέσματα αυτών των συγκρίσεων.

6.1.1 Χρησιμοποιώντας τις Quasi-Newton ενημερώσεις ελάχιστης μνήμης

Σε αυτή την υποενότητα αποτιμούμε την προτεινόμενη μέθοδο στις τέσσερις Quasi-Newton ενημερώσεις ελάχιστης μνήμης, SR1, BFGS, DFP και PSB. Όσον αφορά την παράμετρο κλιμάκωσης του αρχικού πίνακα, θ , θεωρήσαμε ότι είναι σταθερή και ίση με

$$\theta = 1,$$

μιας και τα στιγμιότυπα είναι τυχαία (τεχνητά). Σε διαφορετική περίπτωση θα είχαμε μία επιπλέον τυχαία μεταβλητή, η οποία δεν επηρεάζει την προτεινόμενη μέθοδο.

Τα τυχαία στιγμιότυπα του ΥΠΕ δημιουργήθηκαν στην πλατφόρμα MATLAB (έκδοση R2008b) και έχουν διαστάσεις $n = 100, 200, \dots, 1000$. Οι διαστάσεις που επιλέξαμε ήταν σχετικά μικρές προκειμένου να είναι εφικτή η σύγκριση με τους υπόλοιπους καθιερωμένους αλγορίθμους της διεθνούς βιβλιογραφίας. Αξίζει να τονίσουμε ότι η προτεινόμενη μέθοδος ουσιαστικά αποσκοπεί στην επίλυση προβλημάτων της τάξης των εκατομμυρίων μεταβλητών και ο μόνος περιορισμός είναι η μνήμη RAM του εκάστοτε υπολογιστικού συστήματος που χρησιμοποιείται, μιας και η χωρική πολυπλοκότητα του αλγορίθμου είναι $O(n)$, με μικρή πολλαπλασιαστική σταθερά.

Για κάθε διάσταση n , δημιουργήσαμε 1000 τυχαία στιγμιότυπα του ΥΠΕ, με τις συντεταγμένες των διανυσμάτων g , s και y , επιλεγμένες ανεξάρτητα από ομοιόμορφη κατανομή στο διάστημα $(-10^5, +10^5)$. Σε κάθε

ενημέρωση συνολικά έχουμε 10 000 τυχαία στιγμιότυπα (1 000 σε κάθε μία από τις 10 διαστάσεις), οπότε το σύνολο των τυχαίων στιγμιότυπων είναι 40 000.

Υλοποιήσαμε τον Αλγόριθμο 5.4, (MQN), για κάθε μία από τις τέσσερις ενημερώσεις, σε FORTRAN 90. Συγκρίναμε τη μέθοδό μας με τους GQTPAR [3] και LSTRS [79] αλγορίθμους.

- GQTPAR¹, των Moré & Sorensen [64]. Ο αλγόριθμος GQTPAR είναι υλοποιημένος σε FORTRAN 77 και είναι διαθέσιμος ως η υπορουτίνα `dsgqt.f` στο πακέτο MINPACK-2. Βασίζεται στη μέθοδο των Moré & Sorensen [64], που αναπτύχθηκε στο Κεφάλαιο 2, Ενότητα 2.8, για τον υπολογισμό μιας σχεδόν ακριβούς λύσης του ΥΠΕ. Ο αλγόριθμος χρησιμοποιεί την παραγοντοποίηση Cholesky κατά τη διάρκεια επίλυσης του ΥΠΕ.
- LSTRS², Rojas et al. [79]. Ο αλγόριθμος LSTRS έχει προταθεί από τους Rojas et al. [78], και αντιμετωπίζει το ΥΠΕ ως ένα παραμετρικοποιημένο πρόβλημα ιδιοτιμών. Βρίσκει και αυτός μια σχεδόν ακριβή λύση του ΥΠΕ και είναι υλοποιημένος σε MATLAB. Χρησιμοποιεί τη μέθοδο Lanczos για να υπολογίσει τη μικρότερη ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα και βασίζεται σε γινόμενα πίνακα-διανύσματος.

Για όλους τους παραπάνω αλγορίθμους η πλατφόρμα που χρησιμοποιήθηκε ήταν MATLAB-R2008b και όλα τα αριθμητικά πειράματα έγιναν σε προσωπικό υπολογιστή Pentium 1.86 GHz, με 2GB RAM, χρησιμοποιώντας το λειτουργικό σύστημα Linux.

Κάθε αλγόριθμος αποτιμήθηκε στο ίδιο σύνολο τυχαίων στιγμιότυπων. Επιπλέον, θεωρήσαμε ότι ένα στιγμιότυπο του ΥΠΕ λύθηκε επιτυχώς, στην περίπτωση όπου ικανοποιήθηκε το ακόλουθο κριτήριο τερματισμού

$$\|(B + \lambda I) d - g\| \leq 10^{-3}, \quad (6.1)$$

¹The Army High Performance Computing Research Center at the University of Minnesota and the Mathematics and Computer Science Division at Argonne National Laboratory.

²Algorithm 873: LSTRS: MATLAB software for large scale trust region subproblems and regularization, ACM Transactions on Mathematical Software, volume 34, No. 2, 1-28, 2008.

τόσο στην τυπική, όσο και στη δύσκολη περίπτωση. Σε όλους τους αλγορίθμους, για την επίλυση της secular equation, θεωρήσαμε ότι ο επιτρεπόμενος μέγιστος αριθμός επαναλήψεων Newton ήταν 200 (δηλαδή $\text{maxiter} = 200$). Επίσης, για όλα τα πειράματα, η ακτίνα Δ της εκάστοτε περιοχής εμπιστοσύνης θεωρήθηκε σταθερή και ίση με

$$\Delta = 10.$$

Για την αποτίμηση και τη συγκριτική μελέτη των μεθόδων που χρησιμοποιήθηκαν, βασιστήκαμε σε τρεις δείκτες απόδοσης, από τους οποίους εξάγαμε στατιστικά αποτελέσματα. Συγκεκριμένα, οι δείκτες αυτοί είναι:

- ο αριθμός των επαναλήψεων Newton (iterations),
- η ακρίβεια της λύσης του ΥΠΕ, $\|(B + \lambda I) d - g\|$ (accuracy),
- η σχετική ακρίβεια λύσης, $|\|d\| - \Delta| / \Delta$ (relative accuracy), δηλαδή η ακρίβεια με την οποία ικανοποιείται ο περιορισμός εμπιστοσύνης (secular equation).

Ο χρόνος εκτέλεσης (CPU time) δεν αποτελεί στην περίπτωσή μας έναν αξιόπιστο δείκτη απόδοσης, λόγω του ότι οι αλγόριθμοι είναι υλοποιημένοι σε διαφορετικές γλώσσες προγραμματισμού, ενώ η πλατφόρμα διασύνδεσής τους ήταν το MATLAB, όπου και αναπτύχθηκε ο γεννήτορας των τυχαίων στιγμιotypών του ΥΠΕ.

Μετά το πέρας της εκτέλεσης όλων των πειραμάτων, τα αποτελέσματα συγκεντρώθηκαν σε μία ενιαία βάση δεδομένων του SPSS (.sav). Για την στατιστική ανάλυση των δεδομένων, λάβαμε υπόψη μόνο τα δεδομένα που θεωρήθηκαν για κάθε αλγόριθμο, επιτυχώς λυμένα (success), βασιζόμενοι στο κριτήριο τερματισμού (6.1). Για τη σύγκριση των μεθόδων αναφορικά με κάθε δείκτη απόδοσης, τα στατιστικά μέτρα που χρησιμοποιήθηκαν είναι

- η μέση τιμή (Mean),
- η τυπική απόκλιση (Std. Div.),
- η ελάχιστη τιμή (Min),
- η μέγιστη τιμή (Max) και,

- το 95% διάστημα εμπιστοσύνης (C.I.) της μέσης τιμής.

6.1.1.1 Η τυπική περίπτωση.

Τα αποτελέσματα των πειραμάτων στην τυπική περίπτωση, καταγράφονται στους Πίνακες 6.1-6.9.

Πίνακας 6.1: Τυπική περίπτωση: επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επιτυχία			
update	MQN	LSTRS	GQTPAR
SR1	100.00%	57.27%	99.95%
BFGS	100.00%	55.36%	99.44%
DFP	100.00%	70.10%	95.07%
PSB	100.00%	62.95%	99.98%
total succ.(%)	100.00%	61.42%	98.61%
# inst.	40 000	24 568	39 444

Στον Πίνακα 6.1, καταγράφεται το ποσοστό των επιτυχώς λυμένων στιγμιοτύπων ανά ενημέρωση, για κάθε μέθοδο ξεχωριστά. Όπως μπορούμε να παρατηρήσουμε, ο αλγόριθμος MQN είχε 100% επιτυχία σε όλες τις ενημερώσεις. Αντιθέτως, ο αλγόριθμος LSTRS παρουσίασε το μεγαλύτερο ποσοστό αποτυχιών (38.58%). Έλυσε επιτυχώς μόνο 24 568, σε σύνολο 40 000 προβλημάτων. Τέλος, ο αλγόριθμος GQTPAR παρουσίασε πολύ μεγάλο ποσοστό επιτυχίας, λύνοντας το 98.61% των προβλημάτων. Να υπενθυμίσουμε ότι στο σύνολο των 40 000 προβλημάτων, η διάσταση των στιγμιοτύπων κυμαίνεται σε σχετικά μικρές διαστάσεις, από $n = 100$ έως και $n = 1\,000$. Επιλέξαμε αυτές τις διαστάσεις, προκειμένου να είναι εφικτή η επίλυση των στιγμιοτύπων από τον αλγόριθμο GQTPAR, ο οποίος έχει τετραγωνική χωρική πολυπλοκότητα, λόγω των πινάκων που συντηρεί.

Δεδομένου ότι ο προτεινόμενος αλγόριθμος MQN είχε 100% επιτυχία, επικεντρώνουμε την περαιτέρω ανάλυσή μας στους άλλους δύο αλγορίθμους, LSTRS και GQTPAR. Για να δούμε αναλυτικότερα τα ποσοστά των επιτυχιών αυτών των δύο αλγορίθμων, θεωρήσαμε ανά διάσταση, τον αριθμό των επιτυχώς λυμένων στιγμιοτύπων, για κάθε ενημέρωση.

Ο Πίνακας 6.2 αναφέρεται στα ποσοστά επιτυχίας ανά διάσταση, για καθεμιά από τις τέσσερις ενημερώσεις, που εφοδιάσαμε τον αλγόριθμο

LSTRS. Οι ενημερώσεις που παρουσίασε ο συγκεκριμένος αλγόριθμος τις περισσότερες αποτυχίες, είναι οι SR1 και BFGS ενημερώσεις, με ποσοστά αποτυχίας 42.7% και 44.6%, αντίστοιχα. Αξίζει να παρατηρήσει κανείς ότι η LSTRS παρουσίασε το μικρότερο ποσοστό επιτυχίας (55.7%) όταν η διάσταση των υποπροβλημάτων ήταν $n = 1\,000$. Επίσης, τη χειρότερη επίδοση, ως προς τον αριθμό των επιτυχώς λυμένων στιγμιοτύπων, την παρουσίασε στις 400 διαστάσεις, στην SR1 ενημέρωση ελάχιστης μνήμης.

Πίνακας 6.2: Τυπική περίπτωση: λυμένα στιγμιότυπα στην LSTRS

Δείκτης απόδοσης: Επιτυχία					
n	succ				Overall succ
	SR1	BFGS	DFP	PSB	
100	673	595	730	806	2804
	67.3%	59.5%	73.0%	80.6%	70.1%
200	724	641	839	473	2677
	72.4%	64.1%	83.9%	47.3%	66.9%
300	414	595	723	569	2301
	41.4%	59.5%	72.3%	56.9%	57.5%
400	313	551	699	667	2230
	31.3%	55.1%	69.9%	66.7%	55.8%
500	626	564	706	749	2645
	62.6%	56.4%	70.6%	74.9%	66.1%
600	607	537	731	610	2485
	60.7%	53.7%	73.1%	61.0%	62.1%
700	617	562	717	621	2517
	61.7%	56.2%	71.7%	62.1%	62.9%
800	641	522	667	582	2412
	64.1%	52.2%	66.7%	58.2%	60.3%
900	539	488	624	617	2268
	53.9%	48.8%	62.4%	61.7%	56.7%
1 000	573	481	574	601	2229
	57.3%	48.1%	57.4%	60.1%	55.7%
σύνολο	5727	5536	7010	6295	24568
	57.3%	55.4%	70.1%	63.0%	61.4%

Ο Πίνακας 6.3 αναφέρεται στα ποσοστά των επιτυχώς λυμένων στιγμιοτύπων από τη μέθοδο GQTPAR, η οποία χρησιμοποιεί την κατά Cholesky

παραγοντοποίηση πινάκων. Παρατηρούμε ότι το μικρότερο ποσοστό επιτυχίας (91.7%) παρουσιάστηκε στις 100 διαστάσεις. Επιπλέον, εστιάζοντας την προσοχή μας στη συγκεκριμένη γραμμή του πίνακα και αναφορικά με την εκάστοτε ενημέρωση, μπορούμε να δούμε ότι η ενημέρωση με τις λιγότερες επιτυχίες (70.8%) φαίνεται να είναι η DFP. Κατά συνέπεια, από την τελευταία γραμμή του πίνακα, στην οποία αναφέρονται συγκεντρωτικά οι επιτυχίες σε κάθε ενημέρωση, συνάγεται ότι η DFP ενημέρωση ελάχιστης μνήμης παρουσίασε το μικρότερο ποσοστό επιτυχίας (95.1%) στη συγκεκριμένη μέθοδο.

Πίνακας 6.3: Τυπική περίπτωση: λυμένα στιγμιότυπα στην GQTPAR

Δείκτης απόδοσης: Επιτυχία					
<i>n</i>	succ				Overall succ
	SR1	BFGS	DFP	PSB	
100	1000	960	708	1000	3668
	100.0%	96.0%	70.8%	100.0%	91.7%
200	1000	984	825	1000	3809
	100.0%	98.4%	82.5%	100.0%	95.2%
300	1000	1000	974	1000	3974
	100.0%	100.0%	97.4%	100.0%	99.4%
400	1000	1000	1000	999	3999
	100.0%	100.0%	100.0%	99.9%	100.0%
500	1000	1000	1000	999	3999
	100.0%	100.0%	100.0%	99.9%	100.0%
600	1000	1000	1000	1000	4000
	100.0%	100.0%	100.0%	100.0%	100.0%
700	999	1000	1000	1000	3999
	99.9%	100.0%	100.0%	100.0%	100.0%
800	998	1000	1000	1000	3998
	99.8%	100.0%	100.0%	100.0%	100.0%
900	1000	1000	1000	1000	4000
	100.0%	100.0%	100.0%	100.0%	100.0%
1 000	998	1000	1000	1000	3998
	99.8%	100.0%	100.0%	100.0%	100.0%
σύνολο	9995	9944	9507	9998	39444
	100.0%	99.4%	95.1%	100.0%	98.6%

Από αμφότερους τους Πίνακες 6.2 και 6.3, βλέπουμε ότι η διάσταση n και το ποσοστό επιτυχίας, είναι ανεξάρτητες μεταξύ τους. Επίσης, κάθε μέθοδος συμπεριφέρθηκε διαφορετικά σε κάθε ενημέρωση. Μάλιστα, μπορούμε να δούμε ότι η DFP ενημέρωση ελάχιστης μνήμης είχε το μεγαλύτερο ποσοστό επιτυχίας στην LSTRS, ενώ αντίθετα παρουσίασε το μικρότερο ποσοστό στην GQTPAR.

Πίνακας 6.4: Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επαναλήψεις					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	4.00	1.44	0	9	(3.98, 4.01)
LSTRS	8.22	5.31	3	149	(8.16, 8.29)
GQTPAR	9.06	3.13	3	26	(9.03, 9.10)

Ο Πίνακας 6.4 συνοψίζει τα συγκεντρωτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων Newton που έκανε κάθε μέθοδος, στα επιτυχώς λυμένα στιγμιότυπα. Μπορούμε να παρατηρήσουμε ότι, η μέθοδος που παρουσίασε την καλύτερη απόδοση, είναι η προτεινόμενη μέθοδος, MQN. Η συγκεκριμένη μέθοδος υπερτερεί σε όλα τα στατιστικά μέτρα που αναφέρονται, έναντι των άλλων δύο μεθόδων. Στη συνέχεια θα αναλύσουμε περαιτέρω τον παραπάνω πίνακα αναφορικά με τις τέσσερις ενημερώσεις (update).

Τα αποτελέσματα ως προς τον αριθμό των επαναλήψεων Newton ανά ενημέρωση, σε κάθε μία από τις τρεις μεθόδους, συνοψίζονται στον Πίνακα 6.5. Οι αλγόριθμοι MQN και LSTRS είχαν την καλύτερη απόδοση στην BFGS ενημέρωση ελάχιστης μνήμης, ενώ ο GQTPAR στην DFP ενημέρωση ελάχιστης μνήμης. Μπορούμε να δούμε ότι ο MQN είχε την καλύτερη απόδοση συγκριτικά με τους υπόλοιπους αλγορίθμους, σε όλες τις ενημερώσεις.

Πίνακας 6.5: Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επαναλήψεις						
Μέθοδος	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	4.70	0.89	3	7	(4.68, 4.71)
	BFGS	2.93	1.22	0	9	(2.90, 2.95)
	DFP	3.49	1.46	0	8	(3.46, 3.51)
	PSB	4.89	1.08	2	9	(4.86, 4.90)
LSTRS	SR1	10.76	3.79	3	128	(10.67, 10.86)
	BFGS	5.93	3.52	3	72	(5.84, 6.02)
	DFP	6.12	5.32	3	119	(6.00, 6.25)
	PSB	10.27	5.90	3	149	(10.13, 10.42)
GQTPAR	SR1	10.82	1.75	6	26	(10.78, 10.85)
	BFGS	9.52	3.74	3	24	(9.45, 9.60)
	DFP	5.70	1.53	3	16	(5.67, 5.73)
	PSB	10.05	2.12	5	22	(10.01, 10.10)

Επίσης, η BFGS ενημέρωση σημείωσε την καλύτερη απόδοση από ότι οι υπόλοιπες ενημερώσεις, σε δύο από τις τρεις μεθόδους. Τέλος, παρατηρώντας το 95% διάστημα εμπιστοσύνης, μπορούμε πολύ εύκολα να κατατάξουμε τις ενημερώσεις, για κάθε μέθοδο, ως ακολούθως:

MQN: 1η: BFGS, 2η: DFP, 3η: SR1, 4η: PSB

LSTRS: 1η: BFGS, 2η: DFP, 3η: PSB, 4η: SR1

GQTPAR: 1η: DFP, 2η: BFGS, 3η: PSB, 4η: SR1

Τα αποτελέσματα του Πίνακα 6.5, αναφορικά με τη μέση τιμή των επαναλήψεων Newton ανά ενημέρωση, αποτυπώνονται στο Σχήμα 6.1, όπου φαίνεται ξεκάθαρα και η σειρά κατάταξης που δώσαμε πιο πάνω. Στο συγκεκριμένο σχήμα φαίνεται ευκρινώς ότι ο MQN χρειάστηκε κατά μέσο όρο, λιγότερο από τις μισές επαναλήψεις Newton για να λύσει όλα τα υποπροβλήματα και στις τέσσερις ενημερώσεις που χρησιμοποιήθηκαν.

Σχήμα 6.1: Τυπική περίπτωση: μέσος αριθμός επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Ο Πίνακας 6.6 συνοψίζει τα συγκριτικά αποτελέσματα αναφορικά με την ακρίβεια της λύσης στα επιτυχώς λυμένα στιγμιότυπα. Και σε αυτόν το δείκτη απόδοσης, μπορούμε να παρατηρήσουμε ότι ο αλγόριθμος MQN παρουσίασε την καλύτερη απόδοση, σε όλα τα στατιστικά μέτρα.

Πίνακας 6.6: Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Ακρίβεια λύσης					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	2.4E-09	4.2E-07	5.8E-16	8.2E-05	(-1.6E-09, 6.6E-09)
LSTRS	2.7E-05	2.8E-05	2.6E-15	1.0E-04	(2.6E-05, 2.7E-05)
GQTPAR	1.1E-08	9.0E-07	3.0E-15	9.2E-05	(2.2E-09, 2.0E-08)

Πίνακας 6.7: Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Ακρίβεια λύσης						
	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	2.4E-13	1.1E-11	1.5E-15	1.1E-09	(1.9E-14, 4.7E-13)
	BFGS	2.4E-13	7.0E-12	5.8E-16	6.6E-10	(1.0E-13, 3.8E-13)
	DFP	9.9E-09	8.4E-07	8.5E-16	8.2E-05	(-6.5E-09, 2.6E-08)
	PSB	4.7E-15	6.6E-14	1.5E-15	6.3E-12	(3.4E-15, 6.0E-15)
LSTRS	SR1	3.4E-05	2.8E-05	2.3E-07	1.0E-04	(3.4E-05, 3.5E-05)
	BFGS	2.3E-05	2.7E-05	3.0E-15	9.9E-05	(2.2E-05, 2.4E-05)
	DFP	1.8E-05	2.5E-05	2.6E-15	1.0E-04	(1.8E-05, 1.9E-05)
	PSB	3.2E-05	2.7E-05	3.9E-15	1.0E-04	(3.2E-05, 3.3E-05)
GQTPAR	SR1	1.9E-09	1.9E-07	4.3E-15	1.9E-05	(-1.8E-09, 5.6E-09)
	BFGS	8.7E-13	1.8E-11	3.0E-15	1.6E-09	(5.1E-13, 1.2E-12)
	DFP	1.5E-08	1.0E-06	6.2E-15	9.2E-05	(-5.6E-09, 3.6E-08)
	PSB	2.7E-08	1.4E-06	3.4E-15	8.7E-05	(-1.1E-09, 5.5E-08)

Τα αναλυτικά αποτελέσματα ανά ενημέρωση και ανά μέθοδο, φαίνονται στον Πίνακα 6.7. Αξιοσημείωτο είναι το γεγονός ότι η καλύτερη μέση ακρίβεια λύσης στον MQN αλγόριθμο είναι της τάξης του 10^{-15} , στην PSB ενημέρωση ελάχιστης μνήμης, ενώ ακολουθούν οι ενημερώσεις SR1 και BFGS, οι οποίες πέτυχαν την ίδια τάξη ακρίβειας, με την BFGS να υπερτερεί σχετικά στα στατιστικά μέτρα. Αντίθετα, ο αλγόριθμος LSTRS παρουσίασε μέση ακρίβεια λύσης της τάξης του 10^{-05} σε όλες τις ενημερώσεις. Τέλος, ο GQTPAR έλυσε με τη μεγαλύτερη μέση ακρίβεια (τάξης του 10^{-13}), τα στιγμιότυπα που χρησιμοποίησαν την BFGS ενημέρωση για να προσεγγίσουν τον εσσιανό πίνακα.

Τα αποτελέσματα που καταγράφονται σε αυτό τον πίνακα, αναφορικά με τη μέση τιμή της ακρίβειας λύσης σε κάθε ενημέρωση, φαίνονται στο Σχήμα 6.2.

Σχήμα 6.2: Τυπική περίπτωση: μέση ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Πίνακας 6.8: Τυπική περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Σχετική ακρίβεια λύσης					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	7.5E-10	1.8E-09	0.0E+00	1.0E-08	(7.3E-10, 7.7E-10)
LSTRS	2.7E-10	1.1E-09	0.0E+00	1.0E-08	(2.5E-10, 2.8E-10)
GQTPAR	6.9E-10	1.7E-09	0.0E+00	1.0E-08	(6.7E-10, 7.1E-10)

Τα συγκριτικά αποτελέσματα ως προς την σχετική ακρίβεια λύσης, συνοψίζονται στον Πίνακα 6.8. Μπορούμε να δούμε ότι όλες οι μέθοδοι συμπεριφέρθηκαν με παρόμοιο τρόπο. Η μέση σχετική ακρίβεια λύσης είναι της τάξης του 10^{-10} σε όλες τις μεθόδους, με την LSTRS να υπερτερεί ελάχιστα μεν, αλλά, θα πρέπει να υπενθυμίσουμε ότι έλυσε μόλις το

61.42% των στιγμιοτύπων, από όπου εξάγουμε τα στατιστικά αποτελέσματα.

Πίνακας 6.9: Τυπική περίπτωση: αναλυτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Σχετική ακρίβεια λύσης						
	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	8.8E-10	1.9E-09	0.0E+00	1.0E-08	(8.5E-10, 9.2E-10)
	BFGS	7.4E-10	1.7E-09	0.0E+00	9.9E-09	(7.0E-10, 7.7E-10)
	DFP	6.0E-10	1.6E-09	0.0E+00	1.0E-08	(5.7E-10, 6.3E-10)
	PSB	7.8E-10	1.9E-09	0.0E+00	9.9E-09	(7.5E-10, 8.2E-10)
LSTRS	SR1	1.9E-16	1.7E-16	0.0E+00	1.2E-15	(1.9E-16, 2.0E-16)
	BFGS	4.0E-10	1.4E-09	0.0E+00	9.9E-09	(3.7E-10, 4.4E-10)
	DFP	5.9E-10	1.6E-09	0.0E+00	1.0E-08	(5.5E-10, 6.3E-10)
	PSB	3.7E-11	4.7E-10	0.0E+00	9.5E-09	(2.5E-11, 4.8E-11)
GQTPAR	SR1	7.9E-10	1.8E-09	0.0E+00	9.9E-09	(7.5E-10, 8.2E-10)
	BFGS	5.8E-10	1.5E-09	0.0E+00	1.0E-08	(5.5E-10, 6.1E-10)
	DFP	6.1E-10	1.6E-09	0.0E+00	1.0E-08	(5.7E-10, 6.4E-10)
	PSB	7.8E-10	1.8E-09	0.0E+00	9.9E-09	(7.5E-10, 8.2E-10)

Το ίδιο μπορούμε να παρατηρήσουμε και στη σχετική ακρίβεια λύσης ανά ενημέρωση. Στον Πίνακα 6.9 βλέπουμε ότι οι αλγόριθμοι MQN και GQTPAR πέτυχαν στο συγκεκριμένο δείκτη απόδοσης μέση τιμή γύρω στο 10^{-10} σε όλες τις ενημερώσεις. Η LSTRS πέτυχε την καλύτερη απόδοση, στην SR1 ενημέρωση, δοθέντος ότι κατάφερε να λύσει μόλις το 57.3% των στιγμιοτύπων. Αντίθετα, η συγκεκριμένη ενημέρωση παρουσίασε τη χειρότερη σχετική ακρίβεια λύσης, στους δύο άλλους αλγορίθμους. Επίσης, παρατηρούμε ότι οι ενημερώσεις BFGS και DFP είχαν την καλύτερη απόδοση τόσο στην MQN, όσο και στη GQTPAR μέθοδο. Όμως στην LSTRS, οι δύο αυτές ενημερώσεις, παρουσίασαν τις χειρότερες επιδόσεις.

Στο Σχήμα 6.3 φαίνονται τα αποτελέσματα που καταγράφονται στον Πίνακα 6.9, ως προς τη μέση τιμή της σχετικής ακρίβειας λύσης, ανά ενημέρωση και ανά μέθοδο.

Σχήμα 6.3: Τυπική περίπτωση: μέση σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Συνοψίζοντας τα παραπάνω, ο αλγόριθμος που παρουσίασε την καλύτερη απόδοση σε όλους τους δείκτες απόδοσης που χρησιμοποιήθηκαν, είναι ο MQN. Ο προτεινόμενος αλγόριθμος φάνηκε να λύνει το 100% των στιγμιότυπων με την καλύτερη ακρίβεια, κάνοντας τις λιγότερες επαναλήψεις Newton. Ακόμα, θα μπορούσαμε να πούμε ότι, κατά μέσο όρο, σε όλους τους δείκτες απόδοσης και σε όλους τους αλγορίθμους, η καλύτερη απόδοση εμφανίστηκε στην BFGS ενημέρωση ελάχιστης μνήμης.

6.1.1.2 Η δύσκολη περίπτωση. Για να δημιουργήσουμε στιγμιότυπα στη δύσκολη περίπτωση, χρησιμοποιήσαμε τη συνάρτηση `eigs` του Matlab. Μέσω της συγκεκριμένης ρουτίνας, υπολογίσαμε τη μικρότερη ιδιοτιμή λ_1 και το αντίστοιχο ιδιοδιάνυσμα u του πίνακα B , για κάθε ενημέρωση, σε καθεμιά από τις 10 διαφορετικές διαστάσεις. Αρχικοποιήσαμε την ακτίνα περιοχής με $\Delta = 10.0\Delta_{hc}$, όπου

$$\Delta_{hc} = \|(B - \lambda_1 I)^\dagger g\|,$$

όπου με το σύμβολο $\dagger$ παριστάνουμε τον ψευδοαντίστροφο πίνακα. Για να είναι το g κάθετο στο u , θέσαμε το g ως

$$g = \left(-\frac{u(n)}{u(1)}, 0, \dots, 0, 1 \right)^T, \quad \text{όπου } u_1 \neq 0.$$

Να σημειώσουμε εδώ ότι, για τον αλγόριθμο LSTRS, πολλά από τα στιγμιότυπα που παράχθηκαν, δεν ανταποκρίνονται στη δύσκολη περίπτωση, αλλά στην τυπική. Αυτό συνέβη γιατί η συγκεκριμένη μέθοδος χρησιμοποιεί τη μέθοδο Lanczos για τον υπολογισμό των ιδιοτιμών και των αντίστοιχων ιδιοδιανυσμάτων, με αποτέλεσμα στον παραγόμενο τριδιαγώνιο πίνακα, το g να μην είναι πάντα κάθετο σε όλα τα ιδιοδιανύσματα που αντιστοιχούν στην πιο αρνητική ιδιοτιμή. Η ενημέρωση στην οποία παρουσιάστηκε δύσκολη περίπτωση, στο μεγαλύτερο ποσοστό των στιγμιοτύπων, ήταν η SR1 ενημέρωση ελάχιστης μνήμης. Για το λόγο αυτό, μας ενδιαφέρει να συγκρίνουμε την προτεινόμενη μέθοδο κυρίως με την GQTPAR, επειδή και οι δύο μέθοδοι αντιμετώπισαν τη δύσκολη περίπτωση στο 100% των στιγμιοτύπων.

Πίνακας 6.10: Δύσκολη περίπτωση: επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επιτυχία			
update	MQN	LSTRS	GQTPAR
SR1	100.00%	0.01%	100.00%
BFGS	100.00%	100.00%	100.00%
DFP	100.00%	99.59%	100.00%
PSB	100.00%	99.39%	100.00%
total succ.(%)	100.00%	74.75%	100.00%
# inst.	40 000	29 899	40 000

Στον Πίνακα 6.10 καταγράφεται συνολικά το ποσοστό των επιτυχώς λυμένων στιγμιοτύπων από όλες τις μεθόδους, ανά διάσταση, σε όλες τις ενημερώσεις. Όπως μπορούμε να δούμε, οι MQN και GQTPAR παρουσίασαν 100% επιτυχία, ενώ αντίθετα, ο αλγόριθμος LSTRS κατάφερε να λύσει μόνο τα 29 899 από τα 40 000 στιγμιότυπα.

Για να δούμε πιο αναλυτικά στη μέθοδο LSTRS πού παρουσιάστηκε μεγαλύτερη αποτυχία, εξετάσαμε την κάθε ενημέρωση και την κάθε διά-

σταση ξεχωριστά. Τα αποτελέσματα συγκεντρώνονται στον Πίνακα 6.11.

Πίνακας 6.11: Δύσκολη περίπτωση: επιτυχώς λυμένα στιγμιότυπα στην LSTRS

<i>n</i>	Δείκτης απόδοσης: Επιτυχία				Overall succ
	SR1	BFGS	DFP	PSB	
100	1	1000	999	997	2997
	0.1%	100.0%	99.9%	99.7%	74.90%
200	0	1000	998	996	2994
	0.0%	100.0%	99.8%	99.6%	74.90%
300	0	1000	997	994	2991
	0.0%	100.0%	99.7%	99.4%	74.80%
400	0	1000	997	996	2993
	0.0%	100.0%	99.7%	99.6%	74.80%
500	0	1000	997	996	2993
	0.0%	100.0%	99.7%	99.6%	74.80%
600	0	1000	993	994	2987
	0.0%	100.0%	99.3%	99.4%	74.70%
700	0	1000	996	992	2988
	0.0%	100.0%	99.6%	99.2%	74.70%
800	0	1000	994	992	2986
	0.0%	100.0%	99.4%	99.2%	74.70%
900	0	1000	996	991	2987
	0.0%	100.0%	99.6%	99.1%	74.70%
1 000	0	1000	992	991	2983
	0.0%	100.0%	99.2%	99.1%	74.60%
σύνολο	1	10000	9959	9939	29899
	0.0%	100.0%	99.6%	99.4%	74.70%

Όπως παρατηρούμε, στην ενημέρωση SR1, η συγκεκριμένη μέθοδος κατάφερε να λύσει μόνο ένα από τα συνολικά 10 000 στιγμιότυπα που αντιστοιχούν σε αυτή την ενημέρωση, και μάλιστα μόνο στη μικρότερη διάσταση. Να υπενθυμίσουμε εδώ, ότι στην SR1 είχαμε το μεγαλύτερο ποσοστό εμφάνισης της δύσκολης περίπτωσης, από ότι στις υπόλοιπες ενημερώσεις, για τη συγκεκριμένη μέθοδο.

Τα συγκριτικά αποτελέσματα ως προς τον πρώτο δείκτη απόδοσης, δηλαδή τις επαναλήψεις Newton, συνοψίζονται στον Πίνακα 6.12. Παρα-

τηρούμε ότι ο αλγόριθμος MQN δεν χρειάστηκε καμία επανάληψη Newton, σε κανένα στιγμιότυπο από τα 40 000, που παράχθηκαν συνολικά.

Πίνακας 6.12: Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επαναλήψεις					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	0.00	0.0	0	0	0.00
LSTRS	5.15	0.59	3	7	(5.14, 5.16)
GQTPAR	28.50	2.10	13	30	(28.48, 28.52)

Πίνακας 6.13: Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς τον αριθμό των επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Επαναλήψεις						
Μέθοδος	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	0.00	0.00	0	0	0.00
	BFGS	0.00	0.00	0	0	0.00
	DFP	0.00	0.00	0	0	0.00
	PSB	0.00	0.00	0	0	0.00
LSTRS	SR1	4.00	—	4	4	—
	BFGS	5.72	0.54	3	7	(5.71, 5.73)
	DFP	4.88	0.34	4	6	(4.87, 4.89)
	PSB	4.85	0.36	4	6	(4.84, 4.86)
GQTPAR	SR1	28.50	2.10	16	30	(28.46, 28.54)
	BFGS	28.64	1.99	13	30	(28.60, 28.68)
	DFP	28.43	2.16	13	30	(28.39, 28.47)
	PSB	28.44	2.15	16	30	(28.40, 28.48)

Τα αναλυτικότερα αποτελέσματα ως προς τις επαναλήψεις Newton σε κάθε ενημέρωση, για όλους τους αλγορίθμους, καταγράφονται στον Πίνακα 6.13. Λόγω του γεγονότος ότι στην LSTRS μέθοδο δεν προέκυψε δύσκολη περίπτωση σε όλα στα επιτυχώς λυμένα στιγμιότυπα, μπορούμε ουσιαστικά να συγκρίνουμε την προτεινόμενη μέθοδο, μόνο με την

GQTPAR. Στην τελευταία, σε όλα τα στιγμιότυπα εμφανίστηκε η δύσκολη περίπτωση. Προφανώς, ο αλγόριθμος MQN υπερτερεί σημαντικά του GQTPAR, ως προς όλα τα στατιστικά μέτρα, σε κάθε ενημέρωση. Ο μέσος όρος των επαναλήψεων Newton που καταγράφεται στον Πίνακα 6.13, για κάθε ενημέρωση, αναπαρίσταται στο Σχήμα 6.4, στις τρεις συγκρίσιμες μεθόδους.

Σχήμα 6.4: Δύσκολη περίπτωση: μέσος αριθμός επαναλήψεων στα επιτυχώς λυμένα στιγμιότυπα

Πίνακας 6.14: Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Ακρίβεια λύσης					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	6.4E-07	1.9E-07	2.2E-07	9.9E-07	(6.4E-07, 6.5E-07)
LSTRS	1.7E-09	7.7E-09	9.3E-15	9.1E-08	(1.6E-09, 1.8E-09)
GQTPAR	9.5E-07	3.5E-07	2.7E-07	2.1E-06	(9.5E-07, 9.6E-07)

Τα συγκριτικά αποτελέσματα ως προς την ακρίβεια που επιτεύχθηκε από τους τρεις συγκρινόμενους αλγόριθμους, στη λύση των στιγμιοτύπων, καταγράφονται στον Πίνακα 6.14. Οι αλγόριθμοι MQN και GQTPAR, οι οποίοι αντιμετώπισαν τη δύσκολη περίπτωση στο σύνολο των προβλημάτων που είχαν να επιλύσουν, εμφανίζουν την ίδια περίπου συμπεριφορά, με μια μικρή υπερέρηση του MQN.

Πίνακας 6.15: Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς την ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Ακρίβεια λύσης						
	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	6.5E-07	1.9E-07	2.2E-07	9.9E-07	(6.4E-07,6.5E-07)
	BFGS	6.2E-07	1.8E-07	2.3E-07	9.3E-07	(6.2E-07,6.3E-07)
	DFP	6.6E-07	1.9E-07	2.3E-07	9.9E-07	(6.5E-07,6.6E-07)
	PSB	6.5E-07	1.9E-07	2.4E-07	9.7E-07	(6.5E-07,6.5E-07)
LSTRS	SR1	1.7E-14	—	1.7E-14	1.7E-14	—
	BFGS	1.3E-09	5.6E-09	9.3E-13	9.0E-08	(1.2E-09,1.4E-09)
	DFP	1.7E-09	7.8E-09	1.3E-14	8.8E-08	(1.6E-09,1.9E-09)
	PSB	2.2E-09	9.3E-09	1.1E-14	9.1E-08	(2.0E-09,2.3E-09)
GQTPAR	SR1	9.5E-07	3.5E-07	2.8E-07	2.1E-06	(9.5E-07,9.6E-07)
	BFGS	9.9E-07	3.8E-07	2.9E-07	2.1E-06	(9.9E-07,1.0E-06)
	DFP	9.3E-07	3.3E-07	2.8E-07	2.0E-06	(9.2E-07,9.4E-07)
	PSB	9.4E-07	3.4E-07	2.7E-07	2.0E-06	(9.3E-07,9.5E-07)

Στον Πίνακα 6.15, μπορούμε να δούμε ότι, στον MQN αλγόριθμο, η BFGS ενημέρωση ελάχιστης μνήμης είχε ελαφρώς καλύτερη απόδοση από ότι οι άλλες ενημερώσεις. Να σημειώσουμε εδώ ότι η ακρίβεια λύσης που πέτυχε η LSTRS στην SR1 ενημέρωση, αναφέρεται στο μοναδικό στιγμιότυπο που κατάφερε να λύσει η συγκεκριμένη μέθοδος. Για το λόγο αυτό, δεν ορίζεται τόσο η τυπική απόκλιση (Std. Div), όσο και το 95% του διαστήματος εμπιστοσύνης (95% C.I.).

Η μέση τιμή της ακρίβειας λύσης ανά μέθοδο και ανά ενημέρωση, φαίνεται στο Σχήμα 6.5. Αναφορικά με τις μεθόδους MQN και GQTPAR, μπορούμε να δούμε ότι η ακρίβεια λύσης είναι περίπου της ίδιας τάξης, σε κάθε ενημέρωση. Δηλαδή, το ΥΠΕ λύνεται με μεγάλη ακρίβεια, ανε-

ξάρτητα από την ενημέρωση που έχουμε διαλέξει για να προσεγγίσουμε τον ερσιανό πίνακα.

Σχήμα 6.5: Δύσκολη περίπτωση: μέση ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Ο τελευταίος δείκτης απόδοσης που χρησιμοποιήσαμε για να ελέγξουμε την αξιοπιστία και αποδοτικότητα της προτεινόμενης μεθόδου, είναι η σχετική ακρίβεια λύσης, δηλαδή η ακρίβεια που λύνεται η εξίσωση

$$\frac{1}{\Delta} - \frac{1}{d(\lambda)} = 0,$$

η οποία εκφράζει τον περιορισμό της περιοχής εμπιστοσύνης.

Πίνακας 6.16: Δύσκολη περίπτωση: συγκεντρωτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Σχετική ακρίβεια λύσης					
Μέθοδος	Mean	Std.Div.	Min	Max	95% C.I.
MQN	1.6E-16	1.4E-16	0.0E+00	1.2E-15	(1.5E-16,1.6E-16)
LSTRS	2.0E-16	1.7E-16	0.0E+00	1.7E-15	(2.0E-16,2.0E-16)
GQTPAR	3.8E-16	2.9E-16	0.0E+00	2.2E-15	(3.4E-16,3.1E-16)

Τα συγκριτικά αποτελέσματα για το συγκεκριμένο δείκτη, φαίνονται στον Πίνακα 6.16, από όπου βλέπουμε ότι και οι τρεις μέθοδοι επέδειξαν παρόμοια συμπεριφορά, με την MQN να υπερτερεί ελαφρώς.

Πίνακας 6.17: Δύσκολη περίπτωση: αναλυτικά αποτελέσματα ως προς τη σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Δείκτης απόδοσης: Σχετική ακρίβεια λύσης						
	update	Mean	Std.Div.	Min	Max	95% C.I.
MQN	SR1	1.5E-16	1.4E-16	0.0E+00	8.7E-16	(1.5E-16,1.6E-16)
	BFGS	1.6E-16	1.4E-16	0.0E+00	1.0E-15	(1.5E-16,1.6E-16)
	DFP	1.6E-16	1.4E-16	0.0E+00	1.2E-15	(1.6E-16,1.6E-16)
	PSB	1.6E-16	1.4E-16	0.0E+00	8.8E-16	(1.5E-16,1.6E-16)
LSTRS	SR1	3.7E-16	—	3.7E-16	3.7E-16	—
	BFGS	2.0E-16	1.6E-16	0.0E+00	1.1E-15	(1.9E-16,2.0E-16)
	DFP	2.0E-16	1.7E-16	0.0E+00	1.3E-15	(2.0E-16,2.1E-16)
	PSB	2.0E-16	1.7E-16	0.0E+00	1.7E-15	(2.0E-16,2.1E-16)
GQTPAR	SR1	3.4E-16	2.9E-16	0.0E+00	1.9E-15	(3.4E-16,3.5E-16)
	BFGS	3.5E-16	3.0E-16	0.0E+00	1.9E-15	(3.4E-16,3.5E-16)
	DFP	3.4E-16	2.9E-16	0.0E+00	2.1E-15	(3.4E-16,3.5E-16)
	PSB	3.4E-16	3.0E-16	0.0E+00	2.2E-15	(3.3E-16,3.5E-16)

Τα αναλυτικά αποτελέσματα ανά ενημέρωση, σε κάθε μέθοδο, συνοψίζονται στον Πίνακα 6.17. Όπως φαίνεται από τα στατιστικά μέτρα, η BFGS ενημέρωση επέδειξε ελαφρώς καλύτερη συμπεριφορά από τις άλλες ενημερώσεις, στην προτεινόμενη μέθοδο. Αναφορικά με τη μέθοδο LSTRS, στην SR1 ενημέρωση δεν ορίζονται η τυπική απόκλιση (Std. Div)

και το 95% του διαστήματος εμπιστοσύνης (95% C.I.), γιατί, όπως προαναφέραμε, η μέθοδος έλυσε μόνο ένα στιγμιότυπο από τα 40 000. Τέλος, η μέση τιμή της σχετικής ακρίβειας λύσης, όπως αυτή καταγράφεται στον Πίνακα 6.17, αποτυπώνεται στο Σχήμα 6.6.

Συνοψίζοντας τα παραπάνω, συμπεραίνουμε ότι η δύσκολη περίπτωση, δεν είναι δύσκολη για την προτεινόμενη μέθοδο. Όπως είδαμε, δεν χρειάζεται καμία επανάληψη Newton για να λυθεί το ΥΠΕ, ενώ οι ακρίβειες τόσο της λύσης του ΥΠΕ, όσο και της σχετικής λύσης, είναι πολύ μεγάλες.

Σχήμα 6.6: Δύσκολη περίπτωση: μέση σχετική ακρίβεια λύσης στα επιτυχώς λυμένα στιγμιότυπα

Μάλιστα, η σχετική ακρίβεια λύσης στη δύσκολη περίπτωση είναι κατά πολύ μεγαλύτερη από αυτήν που πετυχαίνει ο MQN στην τυπική περίπτωση.

6.1.2 Συγκριτικά αποτελέσματα για την BFGS ενημέρωση ελάχιστης μνήμης

Όπως γνωρίζουμε, η BFGS ενημέρωση θεωρείται η πιο αποδοτική από όλες τις Quasi-Newton ενημερώσεις, στη διεθνή βιβλιογραφία. Πράγματι, η BFGS ενημέρωση ελάχιστης μνήμης φάνηκε να επιδεικνύει την καλύτερη συμπεριφορά στην επίλυση του ΥΠΕ, σε σύγκριση με τις άλλες Quasi-Newton ενημερώσεις ελάχιστης μνήμης, αναφορικά με την προτεινόμενη μέθοδο. Σκοπός μας είναι να διερευνήσουμε περισσότερο την περίπτωση όπου χρησιμοποιείται η συγκεκριμένη ενημέρωση. Επιπρόσθετα, θέλουμε να εξετάσουμε αν η επιλογή της παραμέτρου θ έχει επίπτωση στην επίλυση του ΥΠΕ, όπως επίσης να δούμε αν η σχέση των διανυσμάτων s και y , επηρεάζει την απόδοση του αλγορίθμου. Τέλος, θέλουμε να αξιολογήσουμε την απόδοση της προτεινόμενης μεθόδου ως προς τον χρόνο CPU. Τα αριθμητικά αποτελέσματα αυτής της ενότητας έχουν καταγραφεί στις εργασίες των Apostolopoulou et al., *A practical method for solving large-scale TRS*, του περιοδικού *Optimization Letters* [10] και *A nearly exact method for solving large-scale TRS*. In: *Proceedings of the 7th IEEE International Conference on Industrial Informatics*, pp. 337-342 [8].

Για να διερευνήσουμε καλύτερα αυτές τις περιπτώσεις, δημιουργήσαμε τυχαία στιγμιότυπα του ΥΠΕ, διάστασης από 100, έως και 15 000 000 μεταβλητές. Τα πειράματα δημιουργήθηκαν σε FORTRAN 90 και αποτελούνται από μεσαίας διάστασης στιγμιότυπα ($n = 100, 500, 1000$), από μεγαλύτερης διάστασης ($n = 10^4, 10^5, 10^6$), και από πολύ μεγάλης διάστασης στιγμιότυπα ($n = 2 \cdot 10^6, 3 \cdot 10^6, \dots, 15 \cdot 10^6$).

Για κάθε διάσταση n , δημιουργήθηκαν 1000 τυχαία στιγμιότυπα του ΥΠΕ, με τις συνεταγμένες των διανυσμάτων g , s και y , να έχουν επιλεχθεί ανεξάρτητα από την ομοιόμορφη κατανομή στο διάστημα $(-10^2, +10^2)$. Υλοποιήσαμε τον Αλγόριθμο 5.4 χρησιμοποιώντας την BFGS ενημέρωση ελάχιστης μνήμης (MBFGS) για την ενημέρωση του προσεγγιστικού εσσιανού πίνακα, σε FORTRAN 90. Όλα τα αριθμητικά πειράματα έγιναν σε προσωπικό υπολογιστή Pentium 1.86 GHz, με 2GB RAM, χρησιμοποιώντας το λειτουργικό σύστημα Linux.

Συγκρίναμε τον αλγόριθμό μας με τους GQTPAR και GLTR αλγορίθμους. Ο αλγόριθμος GLTR έχει προταθεί από τους Gould et al. [45], και είναι διαθέσιμος ως ένα module σε FORTRAN 90 στην Harwell Su-

broutine Library [51]. Η συγκεκριμένη μέθοδος βασίζεται στην κατά Lanczos τριδιαγωνιοποίηση του πίνακα B και στην επίλυση μιας ακολουθίας προβλημάτων περιορισμένων σε υπόχωρους Krylov στον $\mathbb{R}^n$. Είναι ένας εναλλακτικός αλγόριθμος της μεθόδου των Steihaug-Toint [83, 85], και υπολογίζει μια προσεγγιστική λύση του ΥΠΕ. Ο αλγόριθμος απαιτεί γινόμενα πίνακα-διανύσματος, ενώ δεν είναι σχεδιασμένος να αντιμετωπίζει τη δύσκολη περίπτωση. Επιλέξαμε τον αλγόριθμο GLTR, τόσο για να δούμε αν η προτεινόμενη σχεδόν ακριβής λύση είναι υπολογιστικά δυσκολότερη από μια προσεγγιστική, όσο και για να έχουμε αλγορίθμους ίδιας γλώσσας προγραμματισμού, ώστε να μπορούμε να συγκρίνουμε χρόνο CPU.

Το ίδιο σύνολο τυχαίων στιγμιότυπων έχει χρησιμοποιηθεί σε κάθε αλγόριθμο. Θεωρήσαμε ότι ένα στιγμιότυπο ΥΠΕ λύθηκε επιτυχώς, αν η λύση ικανοποίησε τη συνθήκη (6.1), δηλαδή την ανισότητα

$$\|(B + \lambda I)d - g\| \leq 10^{-3},$$

τόσο στην τυπική, όσο και στη δύσκολη περίπτωση. Σε κάθε αλγόριθμο, ο μέγιστος επιτρεπόμενος αριθμός επαναλήψεων Newton ορίστηκε ως

$$\text{maxiter} = 200,$$

ενώ η ακτίνα περιοχής εμπιστοσύνης τέθηκε

$$\Delta = 10.$$

Να σημειώσουμε εδώ ότι για διαστάσεις μεγαλύτερες του 1000, αναφέρουμε αποτελέσματα μόνο για τους αλγορίθμους MBFGS και GLTR, λόγω των μεγάλων απαιτήσεων μνήμης που χρειάζεται ο GQTPAR, για να αποθηκεύσει και να παραγοντοποιήσει τον πίνακα B . Επιπλέον, ο αλγόριθμος GLTR δεν αναφέρεται στα συγκριτικά αποτελέσματα στη δύσκολη περίπτωση, επειδή υπολογίζει μια προσεγγιστική και όχι μια σχεδόν ακριβή λύση του ΥΠΕ.

Για κάθε διάσταση υπάρχει ένας συγκεντρωτικός πίνακας, ο οποίος παρουσιάζει την απόδοση του κάθε αλγορίθμου, για τα προβλήματα τα οποία λύθηκαν επιτυχώς. Οι δείκτες απόδοσης που χρησιμοποιήσαμε είναι οι επαναλήψεις Newton, (*it*), η ακρίβεια της λύσης, (*acc*), και ο χρόνος CPU. Για τους δύο πρώτους δείκτες απόδοσης, τα στατιστικά μέτρα

που παρουσιάζονται για κάθε 1000 στιγμιότυπα σε κάθε διάσταση, είναι: η μέση τιμή (Mean), η τυπική απόκλιση (Std. Div.), η ελάχιστη (Min) και η μέγιστη τιμή (Max).

Για να απεικονίσουμε τόσο την αποτελεσματικότητα, όσο και την ακρίβεια της προτεινόμενης μεθόδου, χρησιμοποιήσαμε τα προφίλ απόδοσης των Dolan & Moré [30]. Τα προφίλ απόδοσης κάνουν τη γραφική παράσταση του ποσοστού των προβλημάτων τα οποία έλυσε μία δοθείσα μέθοδος, μέσα σε ένα κλάσμα του καλύτερου χρόνου που επιτεύχθηκε από τις μεθόδους που χρησιμοποιήθηκαν. Ο αριστερός άξονας δείχνει την αποδοτικότητα της εκάστοτε μεθόδου, ενώ το δεξί μέρος του γραφήματος δείχνει τη δυναμικότητα του αλγορίθμου. Για τη δημιουργία αυτών των προφίλ, χρησιμοποιήσαμε τη βιβλιοθήκη Libopt [40], η οποία συνίσταται από cell scripts του Linux και υλοποιεί τη μέθοδο των Dolan & Moré, ενώ παράλληλα τα γραφήματα παράγονται μέσω της πλατφόρμας Matlab.

6.1.2.1 Τυχαία στιγμιότυπα στην τυπική περίπτωση. Πρώτα θα παρουσιάσουμε τη συμπεριφορά της προτεινόμενης μεθόδου, στην τυπική περίπτωση. Για καθεμιά από τις προαναφερόμενες διαστάσεις μεταξύ των $n = 100$ και $n = 10^6$, θεωρήσαμε τις ακόλουθες περιπτώσεις:

- (α) s_k, y_k γραμμικά ανεξάρτητα και $\theta_{k+1} = 1$;
- (β) s_k, y_k γραμμικά ανεξάρτητα και $\theta_{k+1} = (y_k^T y_k) / (s_k^T y_k)$;
- (γ) s_k, y_k συγγραμμικά και $\theta_{k+1} = 1$;
- (δ) s_k, y_k συγγραμμικά και $\theta_{k+1} = (y_k^T y_k) / (s_k^T y_k)$.

Αν η επιλογή της παραμέτρου θ είναι η $\theta_{k+1} = 1$, τότε ο αρχικός πίνακας $B^{(0)}$, είναι ο μοναδιαίος. Αλλιώς, ως θ τέθηκε η παράμετρος που προτείνεται στη διεθνή βιβλιογραφία [68] για την παραμετρικοποίηση του αρχικού πίνακα στην BFGS μέθοδο. Θεωρήσαμε τις τέσσερις παραπάνω περιπτώσεις για να αναλύσουμε την συμπεριφορά της προτεινόμενης μεθόδου σε όλες τις περιπτώσεις που αφορούν τόσο το διανυσματικό ζεύγος $\{s, y\}$, όσο και την παράμετρο θ . Συνολικά, έχουμε 12 000 πειράματα μεσαίας κλίμακας ($n = 100, 500$ και 1000) και 12 000 πειράματα μεγάλης κλίμακας ($n = 10^4, 10^5$ και 10^6).

Οι Πίνακες 6.18-6.23 συνοψίζουν τα αποτελέσματα των πειραμάτων με διαστάσεις από $n = 100$ έως και $n = 10^6$. Σε όλους του πίνακες, η πρώτη

στήλη δηλώνει τη μέθοδο, ενώ η δεύτερη στήλη δείχνει το ποσοστό των επιτυχημένα λυθέντων στιγμιότυπων (*succ*), μέσα στην απαιτούμενη ακρίβεια (10^{-3}). Στους Πίνακες 6.21-6.23, ο GQTPAR αλγόριθμος έχει παραληφθεί από τη σύγκριση, λόγω των πολύ μεγάλων απαιτήσεων μνήμης που έχει.

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.8	1.1E-13	1.3	3.4E-12	0	1.8E-14	8	2.0E-10
GQTPAR	99.6%	3.7	8.6E-07	2.8	5.4E-05	1	3.1E-14	21	3.4E-03
GLTR	91.4%	3.6	1.2E-12	1.5	7.7E-11	1	2.7E-14	7	4.2E-09

Πίνακας 6.18: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 100$ σε 4000 στιγμιότυπα

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.55	4.10E-13	1.10	9.91E-12	0	5.51E-14	7	5.64E-10
GQTPAR	99.9%	3.36	7.57E-12	2.60	5.21E-11	1	8.13E-14	24	1.96E-09
GLTR	85.9%	3.85	6.33E-12	1.45	1.23E-10	1	5.94E-14	7	4.55E-09

Πίνακας 6.19: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 500$ σε 4000 στιγμιότυπα

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.45	2.55E-13	1.03	5.19E-12	0	7.97E-14	7	3.13E-10
GQTPAR	99.7%	3.20	4.17E-11	2.45	5.62E-10	1	1.11E-13	23	2.91E-08
GLTR	86.5%	3.83	4.41E-10	1.46	2.54E-08	1	8.14E-14	7	1.49E-06

Πίνακας 6.20: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 1\,000$ σε 4000 στιγμιότυπα

Τα αποτελέσματα που συνοψίζονται σε όλους τους πίνακες δείχνουν ότι ο αλγόριθμος MBFGS έλυσε όλα τα προβλήματα επιτυχώς, σε όλες τις διαστάσεις. Αντίθετα, οι δύο άλλοι αλγόριθμοι απέτυχαν να λύσουν το 100% των προβλημάτων με την απαιτούμενη ακρίβεια, με χειρότερο όλων τον

αλγόριθμο GLTR. Επιπλέον, μπορούμε να παρατηρήσουμε ότι ο MBFGS υπερτερεί έναντι των αλγορίθμων GQTPAR και GLTR, ως προς τον αριθμό των επαναλήψεων Newton. Αυτό φαίνεται από όλα τα στατιστικά μέτρα, στα οποία ο MBFGS παρουσίασε τη μικρότερη τιμή. Μπορούμε να δούμε ότι ο MBFGS έκανε κατά μέσο όρο, το $1/3$ των επαναλήψεων Newton, από ότι οι άλλοι δύο αλγόριθμοι. Παρόμοια συμπεράσματα μπορούμε να εξαγάγουμε αναφορικά με την ακρίβεια της λύσης του ΥΠΕ. Η προτεινόμενη μέθοδος παρουσίασε τη μεγαλύτερη ακρίβεια, συγκριτικά με τις δύο άλλες μεθόδους, σε όλες τις διαστάσεις. Πιο συγκεκριμένα, στα μεσαίας κλίμακας προβλήματα, ο MBFGS έλυσε το ΥΠΕ με ακρίβεια της τάξης του 10^{-13} . Στα μεγαλύτερης κλίμακας προβλήματα, η λύση επιτεύχθηκε με ακρίβεια της τάξης του 10^{-10} . Συνοψίζοντας τα παραπάνω, μπορούμε να πούμε ότι είναι αξιοσημείωτο το γεγονός ότι, ακόμα και στις μεγάλες διαστάσεις, ο MBFGS έδωσε λύση με τη μεγαλύτερη ακρίβεια και το μικρότερο αριθμό επαναλήψεων Newton.

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.31	5.77E-13	1.06	7.53E-12	0	2.78E-13	8	7.41E-10
GLTR	86.3%	3.84	1.37E-08	1.47	6.93E-07	1	2.79E-13	7	4.04E-05

Πίνακας 6.21: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 10\,000$ σε 4000 στιγμιότυπα

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.14	4.59E-11	1.04	9.48E-10	0	8.19E-13	9	3.88E-08
GLTR	87.5%	3.87	7.06E-07	1.50	1.07E-05	1	8.27E-13	7	2.51E-04

Πίνακας 6.22: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 100\,000$ σε 4000 στιγμιότυπα

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	100.0%	1.00	7.07E-10	1.07	1.22E-08	0	2.74E-12	9	6.38E-07
GLTR	86.8%	3.74	3.74E-05	1.49	1.41E-04	1	2.75E-12	9	8.59E-04

Πίνακας 6.23: Τυπική περίπτωση: Συγκριτικά αποτελέσματα για $n = 1\,000\,000$ σε 4000 στιγμιότυπα

Αναφορικά με τις τέσσερις περιπτώσεις που διακρίναμε, ως προς την επιλογή της παραμέτρου θ και της σχέσης των s και y , παρατηρήσαμε τα ακόλουθα. Η επιλογή της θ δεν έδειξε να επηρεάζει τη συμπεριφορά κανενός αλγορίθμου. Αντίθετα, η γραμμική ανεξαρτησία ή μη, επηρέασε τους αλγορίθμους. Παρατηρήσαμε ότι στις περιπτώσεις (γ) και (δ), στις οποίες τα διανύσματα είναι συγγραμμικά, ο MBFGS παρουσίασε την υψηλότερη ακρίβεια και τον μικρότερο αριθμό επαναλήψεων Newton. Αντίθετα, ο GLTR παρουσίασε τις περισσότερες αποτυχίες, την χαμηλότερη ακρίβεια λύσης και τον μεγαλύτερο αριθμό επαναλήψεων Newton. Βασιζόμενοι σε αυτές τις παρατηρήσεις, στις πολύ μεγάλες διαστάσεις, μεταξύ $2 \cdot 10^6$ και $15 \cdot 10^6$, θεωρήσαμε μόνο την περίπτωση όπου τα διανύσματα είναι γραμμικά ανεξάρτητα, επειδή αυτή η περίπτωση είναι η ευκολότερη για τον αλγόριθμο GLTR.

Τα αριθμητικά αποτελέσματα των Πινάκων 6.18-6.23 απεικονίζονται στα Σχήματα 6.7 και 6.8. Στο Σχήμα 6.9, παρουσιάζεται η απόδοση των MBFGS και GLTR για $n = 10^6, 2 \cdot 10^6, 3 \cdot 10^6, \dots, 15 \cdot 10^6$. Κάθε σχήμα περιλαμβάνει τρία επιμέρους σχήματα, σε καθένα από τα οποία παρουσιάζονται τα προφίλ απόδοσης ως προς τις επαναλήψεις Newton, την ακρίβεια λύσης, καθώς και το χρόνο CPU (σε δευτερόλεπτα), αντίστοιχα.

(α') Απόδοση ως προς τον αριθμό επαναλήψεων

(β') Απόδοση ως προς την ακρίβεια

Σχήμα 6.7: Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS, GQTPAR και GLTR, σε σύνολο 12000 στιγμιοτύπων μεσαίας διάστασης ($n = 100, 500$ και 1000).

(β') Απόδοση ως προς την ακρίβεια

(γ') Απόδοση ως προς το CPU χρόνο

Σχήμα 6.8: Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS και GLTR, σε σύνολο 12000 στιγμιοτύπων μεγάλης κλίμακας ($n = 10^4, 10^5$ και 10^6).

(α') Απόδοση ως προς τον αριθμό επαναλήψεων

(β') Απόδοση ως προς την ακρίβεια

Σχήμα 6.9: Τυπική περίπτωση: Προφίλ απόδοσης για τους MBFGS και GLTR, σε σύνολο 15000 στιγμιοτύπων πολύ μεγάλης κλίμακας ($n = 1 \cdot 10^6, 2 \cdot 10^6, 3 \cdot 10^6, \dots, 15 \cdot 10^6$).

Το Σχήμα 6.7 παρουσιάζει την απόδοση των αλγορίθμων σε προβλήματα μεσαίας διάστασης. Προφανώς, ο πιο δυναμικός και αποτελεσματικός αλγόριθμος εμφανίζεται να είναι ο MBFGS, αφού παρουσιάζει την καλύτερη απόδοση σε όλους ανεξαιρέτως τους δείκτες απόδοσης. Ο αλγόριθμος GQTPAR (Σχ. 6.7(γ)), φαίνεται να είναι ο πιο χρονοβόρος, κάτι το οποίο περιμέναμε, αφού αποθηκεύει πίνακα και χρησιμοποιεί την παραγοντοποίηση Cholesky.

Το Σχήμα 6.8 δείχνει την απόδοση των αλγορίθμων στα μεγάλης διάστασης προβλήματα. Σύμφωνα με τα Σχήματα 6.8(α') και 6.8(β'), η πιθανότητα να είναι ο MBFGS ο καλύτερος αλγόριθμος, ως προς τις επαναλήψεις Newton και την ακρίβεια λύσης, είναι 100%. Στην περίπτωση που θεωρούμε το χρόνο CPU ως δείκτη απόδοσης, το Σχήμα 6.8(γ') δείχνει ότι τα εσωτερικά γινόμενα και τα διανυσματικά αθροίσματα που χρησιμοποιεί ο MBFGS για να υπολογίσει μια σχεδόν ακριβή λύση του υποπροβλήματος, είναι λιγότερο ανέξοδα από ότι η κατά Lanczos τριδιαγωνοποίηση του B , την οποία χρησιμοποιεί ο GLTR, για να υπολογίσει μια προσεγγιστική

λύση του ΥΠΕ. Συμπερασματικά, βλέπουμε ότι η μέθοδος MBFGS είναι προτιμότερη από την GLTR για την επίλυση μεγάλης κλίμακας υποπροβλημάτων, αφού εμφανίζεται να είναι η πιο αποδοτική μέθοδος.

Οι δείκτες απόδοσης για πολύ μεγάλης κλίμακας προβλήματα, παρουσιάζονται στο Σχήμα 6.9. Μπορούμε να παρατηρήσουμε ότι ο MBFGS υπερτερεί έναντι του GLTR, σε όλους τους δείκτες απόδοσης. Τα Σχήματα 6.9(α') και 6.9(γ'), δείχνουν ότι ο MBFGS είναι ο καλύτερος αλγόριθμος επίλυσης του ΥΠΕ, με πιθανότητα 100%. Επιπλέον, το Σχήμα 6.9(γ') δείχνει ότι ο MBFGS είναι ο γρηγορότερος αλγόριθμος από τους δύο που συγκρίνονται. Αυτό οφείλεται στο ότι η MBFGS μέθοδος εκμεταλλεύεται τις χαρακτηριστικές ιδιότητες του B . Αυτή η παρατήρηση υποδηλώνει ότι η προτεινόμενη μέθοδος είναι πολύ πρακτική και αποτελεσματική, αναφορικά με την επίλυση μεγάλης κλίμακας υποπροβλημάτων, της τάξης των εκατομμυρίων μεταβλητών.

6.1.2.2 Τυχαία στιγμιότυπα στη δύσκολη περίπτωση. Για να δημιουργήσουμε στιγμιότυπα στη δύσκολη περίπτωση, χρησιμοποιήσαμε το Matlab. Μέσω της συνάρτησης `eigs`, υπολογίσαμε τη μικρότερη ιδιοτιμή λ_1 και το αντίστοιχο ιδιοδιάνυσμα u , του εκάστοτε πίνακα B . Αρχικοποιήσαμε την ακτίνα περιοχής με $\Delta = 10.0\Delta_{hc}$, όπου

$$\Delta_{hc} = \| (B - \lambda_1 I)^\dagger g \|.$$

Για να είναι το g κάθετο στο u , θέσαμε το g ως

$$g = \left(-\frac{u(n)}{u(1)}, 0, \dots, 0, 1 \right)^T, \quad \text{όπου } u(1) \neq 0.$$

Να σημειώσουμε εδώ ότι, επειδή ο αλγόριθμος GLTR υπολογίζει μια προσεγγιστική και όχι μια σχεδόν ακριβή λύση του ΥΠΕ και δεν είναι σχεδιασμένος να αντιμετωπίζει τη δύσκολη περίπτωση, συγκρίναμε τον MBFGS μόνο με τον GQTPAR. Επίσης, λόγω των μεγάλων απαιτήσεων μνήμης του GQTPAR, συγκρίναμε τους δύο αυτούς αλγόριθμους μόνο σε προβλήματα με διαστάσεις $n = 100, 500$, και 1000 .

Για κάθε διάσταση θεωρήσαμε μόνο τις περιπτώσεις (α), (β) και (γ). Την περίπτωση (δ) την εξαιρέσαμε για τον εξής λόγο: η συγκεκριμένη περίπτωση δείχνει ότι ο B έχει μία ιδιοτιμή ίση με θ , πολλαπλότητας

n . Επομένως, ο B είναι ένα πολλαπλάσιο του μοναδιαίου πίνακα, με ιδιοδιανύσματα τα

$$e_i = (0, 0, \dots, 1, 0, \dots, 0)^T,$$

όπου το 1 είναι στην i συνιστώσα. Επειδή η δύσκολη περίπτωση εμφανίζεται όταν όλα τα ιδιοδιανύσματα που αντιστοιχούν στην πιο αρνητική ιδιοτιμή είναι κάθετα με το g , στην περίπτωση (δ), θα έπρεπε το g να είναι μηδέν. Γι αυτό το λόγο δεν εξετάσαμε τη συγκεκριμένη αυτή περίπτωση. Συνεπώς, έχουμε συνολικά 9 000 στιγμιότυπα μεσαίας κλίμακας.

Μέθοδος	$succ$	Mean		Std. Div.		Min		Max	
		it	acc	it	acc	it	acc	it	acc
MBFGS	100.0%	0	9.13E-06	0.0	1.84E-04	0	8.24E-11	0	7.15E-03
GQTPAR	96.8%	27.52	1.41E-06	3.3	1.56E-05	18	6.45E-14	43	6.88E-04

Πίνακας 6.24: Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n = 100$ σε 3000 στιγμιότυπα

Οι Πίνακες 6.24, 6.25, και 6.26, συνοψίζουν τα αποτελέσματα για τις διαστάσεις 100, 500, 1000, αντίστοιχα. Όλοι οι πίνακες δείχνουν ότι η συμπεριφορά και των δύο αλγορίθμων, του MBFGS και του GQTPAR, είναι παρόμοια, αναφορικά με την ακρίβεια της λύσης. Από την άλλη μεριά, ο MBFGS υπερτερεί σημαντικά του GQTPAR, ως προς τον αριθμό των επαναλήψεων Newton. Όπως μπορούμε να δούμε, η προτεινόμενη μέθοδος, λύνει το υποπρόβλημα στη δύσκολη περίπτωση αναλυτικά και κατά συνέπεια δεν χρειάζεται να κάνει καμία επανάληψη Newton.

Μέθοδος	$succ$	Mean		Std. Div.		Min		Max	
		it	acc	it	acc	it	acc	it	acc
MBFGS	100.0%	0	9.13E-06	0.0	2.11E-04	0	9.31E-11	0	7.94E-03
GQTPAR	96.0%	29.10	1.25E-06	3.7	1.58E-05	17	1.22E-08	52	7.97E-04

Πίνακας 6.25: Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n = 500$ σε 3000 στιγμιότυπα

Μέθοδος	<i>succ</i>	Mean		Std. Div.		Min		Max	
		<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>	<i>it</i>	<i>acc</i>
MBFGS	99.9%	0	1.23E-05	0.0	2.01E-04	0	1.86E-11	0	6.69E-03
GQTPAR	98.5%	27.44	2.50E-06	3.4	4.06E-05	19	2.58E-08	55	1.67E-03

Πίνακας 6.26: Δύσκολη Περίπτωση: Συγκριτικά αποτελέσματα για $n = 1\,000$ σε 3000 στιγμιότυπα

Βασιζόμενοι σε αυτή την παρατήρηση μπορούμε να πούμε ότι για τον αλγόριθμο MBFGS, η δύσκολη περίπτωση ουσιαστικά αποτελεί την “εύκολη” περίπτωση.

Το Σχήμα 6.10 παρουσιάζει τα αποτελέσματα από τους Πίνακες 6.24-6.26. Οι δείκτες απόδοσης είναι η ακρίβεια λύσης (*acc*) και ο χρόνος CPU (σε δευτερόλεπτα). Δεν αναφέραμε την απόδοση ως προς τις επαναλήψεις Newton (*it*), λόγω του ότι ο MBFGS δεν χρειάζεται να κάνει καμία τέτοια επανάληψη για να λύσει το υποπρόβλημα.

(α') Απόδοση ως προς την ακρίβεια

(β) Απόδοση ως προς τον CPU χρόνο

Σχήμα 6.10: Δύσκολη περίπτωση: Προφίλ απόδοσης για τους MBFGS και GQTPAR, σε σύνολο 9000 στιγμιοτύπων μεσαίας κλίμακας ($n = 100, 500$ και 1000).

Το Σχήμα 6.10(α') δείχνει ότι ο αλγόριθμος MBFGS υπερτερεί του αλγορίθμου GQTPAR, ως προς την ακρίβεια της λύσης. Η ίδια παρατήρηση, όπως ήταν αναμενόμενο, μπορεί να γίνει και στο Σχήμα 6.10(β'), λόγω των απαιτήσεων μνήμης του αλγορίθμου GQTPAR.

6.1.3 Συγκριτικά αποτελέσματα για τις 1BFGS και 2BFGS ενημερώσεις

Σε αυτή την υποενότητα παρουσιάζουμε τα αριθμητικά αποτελέσματα που προέκυψαν από τυχαία στιγμιότυπα υποπροβλημάτων, στην περίπτωση όπου ο ερσιανός πίνακας προσεγγίστηκε από τις 1BFGS και 2BFGS ενημερώσεις ελάχιστης και μικρής μνήμης, αντίστοιχα. Τα αριθμητικά αποτελέσματα αυτής της ενότητας έχουν καταγραφεί στην εργασία των Apostolopoulou et al., *Solving the quadratic trust-region subproblem in a low-memory BFGS framework*, *Optimization Methods and Software*, Volume 23, Issue 5, pp. 651-674 [12].

Τα προβλήματα έχουν διάσταση από 100 έως και 100 000 μεταβλητές και δημιουργήθηκαν σε MATLAB. Για να μπορέσουμε να συγκρίνουμε όλους τους αλγορίθμους ως προς το CPU χρόνο, υλοποιήσαμε τον Αλγόριθμο 5.4, για την περίπτωση $m = 1$ (1BFGS) και για $m = 2$ (2BFGS), σε MATLAB 7.3. Όλα τα αριθμητικά πειράματα έγιναν σε προσωπικό υπολογιστή Pentium 1.86 GHz, με 2GB RAM, χρησιμοποιώντας το λειτουργικό σύστημα Linux. Χρησιμοποιήθηκε η αριθμητική κινητής υποδιαστολής IEEE διπλής ακρίβειας περίπου 2.2×10^{-16} .

Συγκρίναμε τη μέθοδό μας με τους GQTPAR [3] και LSTRS [79] αλγορίθμους. Για να έχουμε σύγκριση ως προς το CPU χρόνο, υλοποιήσαμε τον GQTPAR σε MATLAB. Δηλαδή, ο συγκρινόμενος αλγόριθμος GQTPAR είναι μια εκδοχή σε MATLAB της υπορουτίνας `dgqr.f` του πακέτου MINPACK-2. Ας σημειώσουμε εδώ ότι για μεγάλες διαστάσεις ($n \geq 1000$) αναφέρουμε αριθμητικά αποτελέσματα μόνο για του αλγορίθμους 1BFGS, 2BFGS και LSTRS, επειδή ο αλγόριθμος GQTPAR απαιτεί πολλή μνήμη για την παραγοντοποίηση του προσεγγιστικού εσσιανού πίνακα.

Για κάθε διάσταση n , παράχθηκαν 100 τυχαία στιγμιότυπα με τον ακόλουθο τρόπο. Οι συντεταγμένες των διανυσμάτων g , s_m , και y_m ($m = 1, 2$), όπου το m δηλώνει τον αριθμό των αποθηκευμένων διανυσματικών ζευγών που χρειάζονται για την ενημέρωση του πίνακα B , επιλέχθηκαν ανεξάρτητα από ομοιόμορφες κατανομές στο διάστημα $(-10^5, +10^5)$.

Συνολικά, παρήχθησαν 3200 διαφορετικά στιγμιότυπα, διαιρούμενα σε δύο ομάδες, μεσαίας ($n = 100, 200, 300, 400, 500$) και μεγάλης διάστασης ($n = 10^3, 10^4, 10^5$). Τα μισά από αυτά τα στιγμιότυπα αναφέρονται στην τυπική περίπτωση, ενώ τα υπόλοιπα στη δύσκολη περίπτωση. Για κάθε αλγόριθμο χρησιμοποιήσαμε το ίδιο σύνολο τυχαίων στιγμιότυπων. Θεωρήσαμε ότι ένα στιγμιότυπο ΥΠΕ επιλύθηκε επιτυχώς, αν υπολογίστηκε μια λύση που ικανοποιεί την ανίσωση

$$|(\|d\| - \Delta) / \Delta| \leq tol$$

για αμφότερες την τυπική και τη δύσκολη περίπτωση. Σε όλους τους αλγορίθμους, η απαιτούμενη ακρίβεια tol τέθηκε ίση με 10^{-8} . Ο μέγιστος αριθμός επαναλήψεων Newton που επιτράπηκε ήταν

$$maxiter = 200,$$

ενώ η περιοχή εμπιστοσύνης ορίστηκε ως

$$\Delta = 10.$$

Για να παράγουμε στιγμιότυπα στη δύσκολη περίπτωση, για $n \leq 500$, χρησιμοποιήσαμε τη συνάρτηση του Matlab, `eigs`, για να υπολογίσουμε τη μικρότερη ιδιοτιμή λ_1 και το αντίστοιχο ιδιοδιάνυσμα u του πίνακα B , σχηματιζόμενου από τα διανυσματικά ζεύγη. Για τους αλγορίθμους 1BFGS, 2BFGS και GQTPAR, αρχικοποιήσαμε την ακτίνα περιοχών εμπιστοσύνης με $\Delta = 10.0\Delta_{hc}$, όπου

$$\Delta_{hc} = \|(B - \lambda_1 I)^\dagger g\|.$$

Το διάνυσμα της κλίση g υπολογίστηκε ως

$$g = \left(-\frac{u(n)}{u(1)}, 0, \dots, 0, 1 \right)^T,$$

ώστε να είναι κάθετο στο ιδιοδιάνυσμα της μικρότερης ιδιοτιμής. Για τον αλγόριθμο LSTRS, θέσαμε

$$\Delta = 2.0\Delta_{hc} \quad \text{και} \quad g = g - u(g^T u).$$

Για τα προβλήματα με μεγαλύτερη διάσταση, δηλαδή για $n \geq 1000$, θέσαμε

$$\Delta_{hc} = \|(B + \hat{\lambda}I)^{-1}g\|,$$

όπου τα $\hat{\lambda} = -\lambda_1 + tol$ και $(B + \hat{\lambda}I)^{-1}g$, υπολογίστηκαν από τους τύπους του Κεφαλαίου 4. Να σημειώσουμε εδώ ότι επιλέξαμε τα στιγμιότυπα που εμφανίζεται δύσκολη περίπτωση για όλους τους αλγορίθμους. Όπως αναφέραμε στην Υποενότητα 6.1.1, στην LSTRS μέθοδο δεν αντιστοιχούν όλα τα στιγμιότυπα στη δύσκολη περίπτωση, λόγω της μεθόδου Lanczos που χρησιμοποιεί η συγκεκριμένη μέθοδος. Για το λόγο αυτό, απαιτήσαμε από τον γεννήτορα των στιγμιότυπων να ελέγχει αν κάθε παραγόμενο στιγμιότυπο αποτελεί δύσκολη περίπτωση για όλες ανεξαιρέτως τις μεθόδους. Σε διαφορετική περίπτωση, ο γεννήτορας συνέχιζε στην παραγωγή νέου στιγμιότυπου.

Τα αριθμητικά αποτελέσματα συνοψίζονται στους Πίνακες 6.27-6.32.

Κάθε γραμμή αυτών των πινάκων δείχνει τη διάσταση του προβλήματος (n), το μέσο όρο των επαναλήψεων Newton, (it), τη μέση σχετική ακρίβεια (acc) της λύσης περιοχών εμπιστοσύνης

$$|(\|d\| - \Delta) / \Delta|,$$

καθώς και το μέσο CPU χρόνο (cpu) σε δευτερόλεπτα. Στην περίπτωση όπου $\lambda = 0$ και $\|d\| < \Delta$, χρησιμοποιήσαμε την απόλυτη τιμή του γινομένου $\lambda(\|d\| - \Delta)$ αντί της σχετικής ακρίβειας $|(\|d\| - \Delta) / \Delta|$.

Πίνακας 6.27: Συγκριτικά αποτελέσματα για την τυπική περίπτωση ($m = 1$), 100 στιγμιότυπα ανά διάσταση

n	1BFGS			GQTPAR			LSTRS		
	it	acc	cpu	it	acc	cpu	it	acc	cpu
100	1.0	4.1E-15	6.8E-04	0.07	5.8E-07	1.1E-02	2.1	4.2E-10	1.4E-02
200	1.0	5.9E-12	7.2E-04	0.05	3.7E-08	5.1E-02	2.0	4.6E-10	1.4E-02
300	1.0	1.7E-16	6.7E-04	0.06	6.2E-13	1.5E-01	2.0	1.2E-10	1.4E-02
400	1.0	2.1E-16	7.8E-04	0.07	1.9E-12	3.4E-01	2.0	7.9E-11	1.4E-02
500	1.0	2.0E-16	6.9E-04	0.03	2.7E-13	6.2E-01	2.0	3.0E-12	1.4E-02

Πίνακας 6.28: Συγκριτικά αποτελέσματα για τη δύσκολη περίπτωση ($m = 1$), 100 στιγμιότυπα ανά διάσταση

n	1BFGS			GQTPAR			LSTRS		
	it	acc	cpu	it	acc	cpu	it	acc	cpu
100	0.0	1.1E-16	2.4E-04	10.6	8.4E-03	7.9E-02	15.1	4.9E-11	6.5E-02
200	0.0	1.3E-16	6.2E-04	10.0	2.8E-02	4.2E-01	16.5	2.0E-02	6.9E-02
300	0.0	1.4E-16	4.8E-04	7.7	5.2E-02	9.9E-01	17.1	5.7E-11	8.0E-02
400	0.0	1.4E-16	6.0E-04	5.8	1.7E-03	2.0E+00	17.4	1.0E-10	7.4E-02
500	0.0	1.6E-16	5.9E-04	6.7	4.3E-01	4.0E+00	17.9	1.1E-02	7.8E-02

Οι Πίνακες 6.27-6.30 συνοψίζουν τα αποτελέσματα των προβλημάτων μεσαίας διάστασης, ενώ οι Πίνακες 6.31 και 6.32 συνοψίζουν τα αποτελέσματα των προβλημάτων μεγαλύτερης διάστασης. Υπενθυμίζουμε ότι

στους δύο τελευταίους πίνακες δεν αναφέρουμε αποτελέσματα για τον αλγόριθμο GQTPAR, επειδή είναι ανέφικτη η εφαρμογή αυτού του αλγορίθμου σε αυτές τις διαστάσεις.

Τα αποτελέσματα στους Πίνακες 6.27 και 6.29 (τυπική περίπτωση) δείχνουν ότι ο αλγόριθμος GQTPAR υπερτερεί έναντι των αλγορίθμων 1BFGS, 2BFGS και LSTRS μόνο ως προς το μέσο αριθμό επαναλήψεων.

Πίνακας 6.29: Συγκριτικά αποτελέσματα για την τυπική περίπτωση ($m = 2$), 100 στιγμιότυπα ανά διάσταση

n	2BFGS			GQTPAR			LSTRS		
	it	acc	cpu	it	acc	cpu	it	acc	cpu
100	1.0	1.8E-16	1.6E-03	0.14	1.2E-11	1.0E-02	2.0	6.5E-10	1.4E-02
200	1.0	2.9E-16	1.7E-03	0.09	3.3E-12	5.1E-02	2.0	3.6E-10	1.5E-02
300	1.0	3.5E-16	2.0E-03	0.08	1.0E-12	1.5E-01	2.0	2.6E-10	1.5E-02
400	1.0	5.6E-16	2.0E-03	0.09	5.5E-11	3.8E-01	2.0	2.8E-10	1.5E-02
500	1.0	3.1E-12	2.1E-03	0.06	6.1E-09	6.4E-01	2.0	1.3E-10	1.5E-02

Εντούτοις, αμφότεροι οι αλγόριθμοι 1BFGS και 2BFGS παρουσιάζουν τις καλύτερες επιδόσεις έναντι των αλγορίθμων GQTPAR και LSTRS, ως προς τη σχετική ακρίβεια και τον CPU χρόνο.

Πίνακας 6.30: Συγκριτικά αποτελέσματα για τη δύσκολη περίπτωση ($m = 2$), 100 στιγμιότυπα ανά διάσταση

n	2BFGS			GQTPAR			LSTRS		
	it	acc	cpu	it	acc	cpu	it	acc	cpu
100	0.0	1.1E-16	1.3E-03	9.0	1.7E-02	6.7E-02	15.9	1.0E-02	7.3E-02
200	0.0	1.2E-16	1.4E-03	6.1	2.3E-02	2.5E-01	15.9	7.6E-02	7.3E-02
300	0.0	1.3E-16	1.6E-03	7.3	2.4E-01	1.0E+00	17.3	4.5E-12	8.3E-02
400	0.0	1.4E-16	1.7E-03	7.4	7.0E-02	1.7E+00	16.7	2.2E-16	8.3E-02
500	0.0	1.6E-16	1.8E-03	6.6	5.3E-01	3.7E+00	18.5	3.0E-02	1.0E-01

Από την άλλη μεριά, αναφορικά με τη δύσκολη περίπτωση (Πίνακες 6.28 και 6.30), οι αλγόριθμοι 1BFGS και 2BFGS δείχνουν να υπερτερούν σημαντικά έναντι των άλλων δύο αλγορίθμων, ειδικά σε αριθμό

επαναλήψεων και ποιότητας λύσης (τάξεως 10^{-16}). Παρόμοια συμπεράσματα μπορούν να εξαχθούν και από τους Πίνακες 6.31 και 6.32, όπου η διάσταση των στιγμιοτύπων ΥΠΕ είναι πολύ μεγάλη.

Πίνακας 6.31: Συγκριτικά αποτελέσματα για μεγάλης διάστασης ΥΠΕ ($m = 1$), 100 στιγμιότυπα ανά διάσταση

	n	1BFGS			LSTRS		
		it	acc	cpu	it	acc	cpu
τυπική περίπτωση	10^3	1.0	1.2E-13	1.8E-03	2.0	1.6E-10	2.4E-02
	10^4	1.0	6.1E-16	2.0E-03	2.0	6.3E-12	3.2E-02
	10^5	1.0	3.1E-10	1.8E-02	2.0	1.1E-12	4.2E-01
δύσκολη περίπτωση	10^3	0.0	2.0E-16	8.9E-04	18.6	7.2E+00	1.0E-01
	10^4	0.0	5.7E-16	2.4E-03	20.5	3.9E-01	2.8E-01
	10^5	0.0	2.0E-15	1.6E-02	22.8	2.1E+00	5.2E+00

Πίνακας 6.32: Συγκριτικά αποτελέσματα για μεγάλης διάστασης ΥΠΕ ($m = 2$), 100 στιγμιότυπα ανά διάσταση

	n	2BFGS			LSTRS		
		it	acc	cpu	it	acc	cpu
τυπική περίπτωση	10^3	1.0	2.3E-16	5.4E-03	2.0	8.6E-11	1.7E-02
	10^4	0.9	1.6E-11	1.4E-02	2.0	7.0E-12	5.4E-02
	10^5	0.9	2.3E-10	2.0E-01	2.0	1.0E-11	9.6E-01
δύσκολη περίπτωση	10^3	0.0	2.4E-16	2.9E-03	19.9	3.2E+00	1.3E-01
	10^4	0.0	6.2E-16	1.1E-02	22.6	6.9E-01	6.0E-01
	10^5	0.0	1.9E-15	1.5E-01	24.7	4.8E-01	1.2E+01

Τα αποτελέσματα όλων των πινάκων δείχνουν ότι η επίδοση της προτεινόμενης μεθόδου είναι πολλά υποσχόμενη. Η μέθοδος μπορεί να χειριστεί εξίσου καλά και την εύκολη και τη δύσκολη περίπτωση αρκετά γρήγορα, δίνοντας λύσεις μεγάλης ακρίβειας σε μικρό αριθμό επαναλήψεων. Όλα τα στιγμιότυπα ΥΠΕ επιλύθηκαν επιτυχώς από τους αλγόριθμους 1BFGS και 2BFGS. Αντιθέτως, οι άλλοι δύο αλγόριθμοι σε μερικά στιγμιότυπα,

δεν έφτασαν τη απαιτούμενη ακρίβεια, ειδικά στη δύσκολη περίπτωση. Γενικά, κάποιος θα μπορούσε να παρατηρήσει ότι στην τυπική περίπτωση όλοι οι αλγόριθμοι συμπεριφέρθηκαν παρόμοια, αλλά στη δύσκολη περίπτωση, η απόδοση της προτεινόμενης μεθόδου είναι πολύ ενθαρρυντική, δείχνοντας ότι είναι πολλά υποσχόμενη για πρακτικές εφαρμογές.

6.2 Ο προτεινόμενος αλγόριθμος βελτιστοποίησης

Στην ενότητα αυτή χρησιμοποιούμε τον Αλγόριθμο 5.4 επίλυσης του ΥΠΕ (σελίδα 180, Κεφάλαιο 5), μέσα στο πλαίσιο ενός αλγορίθμου περιοχής εμπιστοσύνης (βλέπε Αλγόριθμο 2.4, σελ. 59), για την ελαχιστοποίηση μη γραμμικών αντικειμενικών συναρτήσεων. Ο προσεγγιστικός εσσιανός που χρησιμοποιείται προκύπτει από τις 1BFGS και 2BFGS ενημερώσεις ελάχιστης και μικρής μνήμης αντίστοιχα. Τα αριθμητικά αποτελέσματα αυτής της ενότητας έχουν καταγραφεί στην εργασία των Apostolopoulou et al., *Solving the quadratic trust-region subproblem in a low-memory BFGS framework, Optimization Methods and Software, Volume 23, Issue 5, pp. 651-674* [12].

Πιο συγκεκριμένα, οι αλγόριθμοι 1BFGS, 2BFGS, LSTRS και GQTPAR, χρησιμοποιήθηκαν στη μέθοδο περιοχών εμπιστοσύνης (βλέπε Κεφάλαιο 2, Ενότητα 2.7), για την επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης χωρίς περιορισμούς. Υπενθυμίζουμε ότι το βασικό πρόβλημα σε κάθε βήμα ενός τέτοιου αλγορίθμου, είναι η επίλυση του ΥΠΕ, τουλάχιστον μία φορά. Χρησιμοποιήσαμε το ίδιο πλαίσιο αλγορίθμου περιοχών εμπιστοσύνης, όπου ο L-BFGS τύπος για $m = 1$ και 2 χρησιμοποιήθηκε για την ενημέρωση του πίνακα B_k . Ο κώδικας γράφτηκε σε MATLAB 7.3 και όλα τα αριθμητικά πειράματα έγιναν σε προσωπικό υπολογιστή Pentium 1.86 GHz, με 2GB RAM, χρησιμοποιώντας το λειτουργικό σύστημα Linux.

Επειδή ο B_k μπορεί να είναι μη θετικά ορισμένος, η συνθήκη $s_k^T y_k > 0$ δεν ισχύει πάντα. Έτσι, για να είναι ο B_k καλά ορισμένος, παραλείψαμε την αποθήκευση του διανυσματικού ζεύγους $\{s_k, y_k\}$, όταν ίσχυε η σχέση

$$|s_k^T y_k| \leq 10^{-12} \|s_k\| \|y_k\|.$$

Ως κριτήριο τερματισμού, θέσαμε το $\|g_k\| \leq tol$, όπου $tol = 10^{-5}$. Για την

επίλυση του ΥΠΕ, χρησιμοποιήθηκαν όλοι οι προαναφερθέντες αλγόριθμοι, για λόγους σύγκρισης. Για τους αλγορίθμους 1BFGS και 2BFGS το κριτήριο τερματισμού των επαναλήψεων Newton στο ΥΠΕ ήταν

$$|(\|d\| - \Delta) / \Delta| \leq 10^{-5} \quad \text{ή} \quad |\psi(\lambda)| \leq 10^{-5}.$$

Για τον αλγόριθμο GQTPAR η ακρίβεια ορίστηκε ως $rtol = 10^{-5}$ και $atol = 10^{-20}$, ενώ για τον αλγόριθμο LSTRS θέσαμε

$$epsilon.Delta = epsilon.HC = 10^{-5}.$$

Διαλέξαμε τριάντα δύο (32) προβλήματα βελτιστοποίησης χωρίς περιορισμούς [47, 62], τα οποία μπορούμε να τα γενικεύσουμε σε οποιαδήποτε διάσταση. Αρκεστήκαμε στη γενίκευσή τους μέχρι και 100 000 μεταβλητές:

Πίνακας 6.33: Προβλήματα βελτιστοποίησης

1. Extended Trigonometric	17. Diagonal 6
2. Extended Rosenbrock	18. Diagonal 7
3. Extended White & Holst	19. Diagonal 8
4. Extended Beale	20. Full Hessian
5. Extended Penalty	21. Sincos
6. Extended Tridiagonal 1	22. Broyden tridiagonal
7. Extended Three Expo Terms	23. Nondia
8. Extended Himmelblau	24. Dqdrtic
9. Extended Block-Diagonal BD1	25. Dixmaana
10. Extended Block-Diagonal BD2	26. Dixmaanb
11. Generalized Tridiagonal 1	27. Dixmaanc
12. Generalized Tridiagonal 2	28. Edensch
13. Generalized Quartic GQ1	29. Liarwhd
14. Generalized Quartic GQ2	30. Cosine
15. Raydan 2	31. Extended Denschnb
16. Diagonal 4	32. Extended Denschnf

Για κάθε συνάρτηση διεξήχθησαν πέντε (5) αριθμητικά πειράματα διαφορετικής διάστασης το καθένα, με αριθμό μεταβλητών $n = 100, 500$ (μεσαίας κλίμακας προβλήματα), $10^3, 10^4$ και 10^5 (μεγαλύτερης κλίμακας

Πίνακας 6.34: Συγκριτικά αποτελέσματα για μεσαίας διάστασης προβλήματα

No.	$n = 100$						$n = 500$					
	$m = 1$			$m = 2$			$m = 1$			$m = 2$		
	1BFGS	GQTPAR	LSTRS	2BFGS	GQTPAR	LSTRS	1BFGS	GQTPAR	LSTRS	2BFGS	GQTPAR	LSTRS
1.	74	220	71	73	—	61	75	282	78	63	—	49
2.	81	—	1868	56	—	826	78	—	—	45	—	831
3.	72	—	—	63	—	—	58	—	—	62	—	—
4.	50	—	42	19	—	41	52	—	114	23	1981	42
5.	38	40	42	36	—	28	57	52	57	46	—	46
6.	28	420	23	18	—	15	23	—	10	34	—	—
7.	19	31	19	14	62	12	22	24	23	14	69	14
8.	21	26	22	12	48	16	23	25	23	15	180	14
9.	44	47	38	17	56	16	39	36	35	28	430	19
10.	47	258	53	21	1193	25	26	66	24	19	1921	15
11.	23	41	23	24	39	22	30	36	30	24	40	24
12.	64	62	64	42	77	40	43	436	42	38	160	57
13.	12	12	12	10	35	12	16	24	16	14	190	13
14.	47	87	47	39	262	40	50	110	55	39	214	49
15.	2	2	2	2	2	2	7	7	7	7	17	7
16.	42	1457	89	12	1195	7	42	1435	52	10	1350	14
17.	2	2	2	2	2	2	7	7	7	7	17	7
18.	11	11	10	12	95	11	10	10	10	10	84	10
19.	8	8	8	8	11	8	7	7	7	7	8	7
20.	4	4	4	4	4	4	5	5	5	5	5	5
21.	16	20	16	14	544	12	25	26	25	14	744	14
22.	45	61	44	33	126	35	98	279	103	93	221	77
23.	48	1659	295	29	—	79	52	—	—	30	—	435
24.	280	—	301	98	27	114	199	—	263	100	30	96
25.	8	9	8	9	234	9	11	11	11	10	342	10
26.	9	9	9	10	401	9	11	11	11	12	1964	11
27.	11	11	11	11	813	11	12	12	12	13	1999	12
28.	28	38	28	24	39	27	43	59	42	27	193	28
29.	58	264	49	31	1834	51	66	1243	93	30	—	54
30.	21	14	20	27	120	16	13	17	13	20	66	13
31.	10	12	10	10	15	10	14	16	14	12	29	10
32.	19	25	19	24	270	18	23	26	23	24	951	19

(α) $m = 1$ (β) $m = 2$ **Σχήμα 6.11:** Μέσος όρος επαναλήψεων για μεσαίας διάστασης προβλήματα.

προβλήματα), αντίστοιχα. Η αρχική ακτίνα περιοχής εμπιστοσύνης Δ_0

ορίστηκε, για $n = 100$ και 500 , ως $\Delta_0 = 10$. Για τα μεγαλύτερης διάστασης προβλήματα ($10^3, 10^4$ και 10^5), ορίστηκε ως $\Delta_0 = 100$. Για τον αριθμό μεταβλητών $n = 10^3, 10^4$ και 10^5 , συγκρίναμε τη μέθοδό μας μόνο με τον αλγόριθμο LSTRS, λόγω του υπολογιστικού κόστους του αλγορίθμου περιοχών εμπιστοσύνης που χρησιμοποιεί τον GQTPAR για την επίλυση του ΥΠΕ.

Τα αριθμητικά αποτελέσματα συνοψίζονται στους Πίνακες 6.34-6.37. Η πρώτη στήλη δηλώνει τον αύξοντα αριθμό του προβλήματος, όπως αυτός δόθηκε στον Πίνακα 6.33. Οι δείκτες απόδοσης που αναφέρονται είναι οι εξής: ο αριθμός των επαναλήψεων (*it*) του αλγορίθμου περιοχών εμπιστοσύνης, ο οποίος ισούται με τον αριθμό των συναρτησιακών υπολογισμών και των υπολογισμό κλίσεων ανά επανάληψη, καθώς και ο CPU χρόνος (*cpu*) σε δευτερόλεπτα. Σε όλους τους πίνακες, το σύμβολο ‘—’ χρησιμοποιείται για να δείξει ότι το αντίστοιχο πρόβλημα απέτυχε να συγκλίνει με την προβλεπόμενη ακρίβεια, δηλαδή $\|g_k\| \leq 10^{-5}$, μέσα στο μέγιστο αριθμό επιτρεπόμενων επαναλήψεων (*maxiter* = 2000).

Ο Πίνακας 6.34 συνοψίζει τα αποτελέσματα για $n = 100$ και $n = 500$. Σε αυτό τον πίνακα αναφέρουμε, για κάθε μέθοδο, μόνο τον αριθμό των επαναλήψεων. Οι Πίνακες 6.35-6.37 συνοψίζουν τα αποτελέσματα αναφορικά με τις υπόλοιπες διαστάσεις.

Το Σχήμα 6.11 απεικονίζει τα αποτελέσματα που καταγράφηκαν στον Πίνακα 6.34 και δείχνει το μέσο όρο των επαναλήψεων που έκανε κάθε μέθοδος. Τα αριθμητικά αποτελέσματα που καταγράφονται στους Πίνακες 6.35-6.37 απεικονίζονται στα Σχήματα 6.12 και 6.13. Τα σχήματα αυτά δείχνουν το μέσο όρο των επαναλήψεων, καθώς και το άθροισμα του απαιτούμενου CPU χρόνου για να λυθεί ένα πρόβλημα, για τους αλγορίθμους 1BFGS, 2BFGS και LSTRS αντίστοιχα. Για να υπολογίσουμε αμφοτέρους αυτούς τους δείκτες απόδοσης, λάβαμε υπόψη μας, σε όλους τους αλγορίθμους, μόνο τα προβλήματα που λύθηκαν επιτυχώς με την απαιτούμενη ακρίβεια, μέσα στο όριο που τέθηκε για το μέγιστο αριθμό επαναλήψεων.

Παρατηρώντας τον Πίνακα 6.34 καθώς και το Σχήμα 6.11, βλέπουμε ότι ο προτεινόμενος αλγόριθμος έλυσε το 100% των προβλημάτων μεσαίας διάστασης. Επιπλέον, παρουσίασε την καλύτερη επίδοση ανάμεσα στις άλλες δύο μεθόδους, αφού είχε τον μικρότερο μέσο όρο επαναλήψεων για αμφοτέρους τα $m = 1$ και 2 .

Πίνακας 6.35: Συγκριτικά αποτελέσματα για $n = 1000$

No.	$m = 1$				$m = 2$			
	1BFGS		LSTRS		2BFGS		LSTRS	
	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>
1.	72	8.7E-02	23	3.6E-01	82	2.9E-01	—	—
2.	80	4.4E-02	—	—	64	1.2E-01	780	1.7E+01
3.	93	9.7E-02	—	—	112	9.5E-01	—	—
4.	48	3.4E-02	136	2.2E+00	27	4.6E-02	57	1.5E+00
5.	58	2.4E-02	64	1.2E+00	48	5.6E-02	47	1.1E+00
6.	37	4.2E-02	—	—	23	3.4E-02	—	—
7.	15	1.2E-02	16	2.4E-01	12	1.7E-02	15	2.9E-01
8.	18	6.5E-03	30	4.4E-01	16	1.4E-01	14	3.1E-01
9.	31	1.4E-02	41	7.3E-01	24	1.6E-01	18	3.7E-01
10.	43	3.2E-02	29	4.3E-01	23	3.5E-02	22	4.0E-01
11.	31	2.5E-02	28	4.1E-01	25	3.7E-02	27	5.4E-01
12.	62	8.4E-02	54	1.1E+00	41	7.6E-02	44	9.2E-01
13.	21	8.4E-03	13	1.7E-01	13	1.3E-02	19	3.9E-01
14.	39	1.5E-02	42	7.3E-01	36	4.1E-02	45	9.5E-01
15.	8	4.3E-03	8	1.1E-01	8	1.4E-02	8	1.7E-01
16.	51	2.9E-02	50	7.3E-01	7	5.7E-03	7	1.1E-01
17.	8	7.9E-03	8	1.1E-01	8	1.4E-02	8	1.6E-01
18.	12	7.3E-03	11	1.3E-01	14	1.3E-01	12	2.1E-01
19.	10	6.9E-03	5	5.9E-02	10	1.4E-02	10	1.7E-01
20.	5	3.4E-03	6	5.7E-02	5	5.0E-03	5	5.9E-02
21.	18	1.0E-02	21	3.1E-01	12	1.7E-02	12	2.0E-01
22.	71	4.7E-02	115	2.1E+00	41	4.5E-02	43	8.9E-01
23.	58	9.4E-02	—	—	130	2.5E+00	823	4.3E+01
24.	265	1.5E-02	253	4.6E+00	144	3.3E-01	136	3.1E+00
25.	9	2.0E-02	11	1.6E-01	9	2.7E-02	9	1.7E-01
26.	9	2.1E-02	11	1.7E-01	9	2.4E-02	9	1.6E-01
27.	10	2.5E-02	13	2.0E-01	10	3.0E-02	10	1.7E-01
28.	28	3.5E-02	37	6.1E-01	29	5.2E-02	27	5.4E-01
29.	77	4.7E-02	67	1.4E+00	43	7.5E-02	54	1.2E+00
30.	53	3.5E-02	16	2.3E-01	53	8.0E-02	29	7.3E-01
31.	15	5.0E-03	15	2.2E-01	12	1.2E-02	14	2.3E-01
32.	29	1.4E-02	27	4.2E-01	25	1.8E-01	23	4.8E-01

Πίνακας 6.36: Συγκριτικά αποτελέσματα για $n = 10\,000$

No.	$m = 1$				$m = 2$			
	1BFGS		LSTRS		2BFGS		LSTRS	
	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>
1.	87	5.6E-01	83	3.9E+00	61	5.4E-01	79	7.3E+00
2.	86	1.8E-01	—	—	53	6.4E-01	845	9.0E+01
3.	86	5.4E-01	—	—	82	5.7E+00	—	—
4.	62	2.4E-01	142	8.0E+00	19	1.7E-01	72	7.6E+00
5.	65	1.2E-01	78	3.7E+00	69	3.6E-01	—	—
6.	28	2.0E-01	—	—	18	1.6E-01	—	—
7.	21	9.8E-02	18	8.5E-01	17	1.2E-01	14	1.2E+00
8.	23	3.9E-02	27	1.1E+00	13	5.3E-02	18	1.5E+00
9.	48	1.0E-01	35	1.6E+00	20	9.5E-02	20	1.5E+00
10.	49	2.5E-01	33	1.4E+00	20	2.1E-01	47	3.7E+00
11.	41	1.7E-01	33	1.6E+00	27	1.8E-01	22	2.1E+00
12.	92	8.8E-01	104	6.1E+00	85	1.3E+00	44	4.7E+00
13.	12	2.7E-02	18	7.5E-01	11	4.9E-02	12	7.6E-01
14.	45	6.4E-02	43	1.9E+00	34	1.4E-01	36	3.6E+00
15.	2	4.6E-03	9	3.4E-01	2	5.0E-03	2	4.4E-02
16.	54	7.7E-02	67	3.2E+00	12	4.3E-02	7	4.9E-01
17.	2	4.6E-03	9	3.4E-01	2	5.0E-03	2	4.6E-02
18.	11	3.9E-02	11	3.3E-01	9	9.0E-01	11	1.1E+00
19.	8	3.6E-02	7	2.2E-01	8	5.4E-02	8	5.0E-01
20.	4	2.7E-02	7	2.0E-01	4	2.0E-02	4	1.5E-01
21.	18	5.9E-02	19	8.3E-01	14	7.4E-02	13	1.1E+00
22.	146	3.9E-01	55	2.9E+00	84	9.1E-01	39	4.1E+00
23.	46	8.6E-02	—	—	30	2.7E+00	—	—
24.	325	5.9E-01	247	1.2E+01	76	8.0E-01	74	7.9E+00
25.	9	1.4E-01	13	7.5E-01	9	1.6E-01	9	7.3E-01
26.	10	1.6E-01	18	1.1E+00	10	1.8E-01	10	8.4E-01
27.	12	1.9E-01	21	1.4E+00	12	2.1E-01	11	9.4E-01
28.	41	3.4E-01	30	1.9E+00	28	3.0E-01	24	2.1E+00
29.	82	2.0E-01	128	9.8E+00	55	4.7E+00	111	2.1E+01
30.	17	6.5E-02	54	2.4E+00	339	1.2E+01	17	1.3E+00
31.	11	2.1E-02	17	7.4E-01	11	4.6E-02	10	8.0E-01
32.	20	4.9E-02	22	9.7E-01	26	1.2E-01	20	1.6E+00

Πίνακας 6.37: Συγκριτικά αποτελέσματα για $n = 100\,000$

No.	$m = 1$				$m = 2$			
	1BFGS		LSTRS		2BFGS		LSTRS	
	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>	<i>it</i>	<i>cpu</i>
1.	39	2.8E+00	78	5.3E+01	77	7.6E+00	39	5.9E+01
2.	102	3.9E+00	—	—	59	3.6E+01	734	1.1E+03
3.	96	7.7E+00	—	—	68	1.0E+02	—	—
4.	58	2.9E+00	65	5.9E+01	21	1.7E+00	58	8.2E+01
5.	94	1.9E+00	—	—	77	4.7E+00	79	1.3E+02
6.	28	2.2E+00	—	—	40	6.7E+01	—	—
7.	25	1.2E+00	46	5.2E+01	73	3.4E+02	16	2.1E+01
8.	36	6.9E-01	32	2.2E+01	20	1.4E+00	17	2.0E+01
9.	40	1.1E+00	42	2.9E+01	20	1.7E+01	20	2.6E+01
10.	39	2.2E+00	46	3.5E+01	25	2.2E+00	22	2.9E+01
11.	32	1.6E+00	32	2.5E+01	26	2.1E+00	23	2.8E+01
12.	101	1.2E+01	154	1.8E+02	75	1.6E+01	96	1.6E+02
13.	14	3.0E-01	20	1.2E+01	15	7.4E-01	13	1.4E+01
14.	49	9.6E-01	51	4.0E+01	38	2.0E+00	52	6.1E+01
15.	8	2.1E-01	70	2.2E+02	8	3.9E-01	8	7.2E+00
16.	61	1.6E+00	56	4.5E+01	10	5.5E-01	8	7.5E+00
17.	8	2.1E-01	11	5.6E+00	8	3.9E-01	8	7.2E+00
18.	12	4.2E-01	13	6.6E+00	10	1.6E+01	12	9.1E+00
19.	6	2.5E-01	9	4.0E+00	6	3.5E-01	6	4.4E+00
20.	5	2.2E-01	8	3.4E+00	5	2.7E-01	5	3.1E+00
21.	27	9.1E-01	—	—	81	4.2E+02	16	1.9E+01
22.	97	3.5E+00	44	2.9E+01	83	1.5E+01	83	1.2E+02
23.	264	8.5E+01	—	—	33	8.4E+01	—	—
24.	247	8.9E+00	294	2.5E+02	127	3.5E+01	81	1.3E+02
25.	12	1.9E+00	14	1.2E+01	11	2.1E+00	12	1.5E+01
26.	12	2.0E+00	42	3.2E+01	12	2.3E+00	12	1.4E+01
27.	13	2.1E+00	46	3.2E+01	14	2.6E+00	14	1.7E+01
28.	38	3.5E+00	36	2.8E+01	30	3.7E+00	48	4.6E+01
29.	85	3.3E+00	160	2.6E+02	90	1.8E+02	114	4.3E+02
30.	25	1.1E+00	33	2.9E+01	—	—	—	—
31.	16	3.0E-01	22	1.6E+01	12	5.4E-01	12	1.2E+01
32.	33	8.5E-01	22	1.4E+01	19	1.0E+00	22	2.4E+01

(α) $m = 1$ (β) $m = 2$ **Σχήμα 6.12:** Μέσος όρος επαναλήψεων για μεγάλης διάστασης προβλήματα.

(α) $m = 1$ (β) $m = 2$ **Σχήμα 6.13:** Ολικός CPU χρόνος για μεγάλης διάστασης προβλήματα.

Τα αποτελέσματα των Πινάκων 6.35-6.37 μαζί με το Σχήμα 6.12 δείχνουν ότι για $m = 1$, ο αλγόριθμος περιοχών εμπιστοσύνης που χρησιμο-

ποίησε τους 1BFGS και LSTRS αλγόριθμους για την επίλυση του ΥΠΕ, είχε περίπου την ίδια συμπεριφορά. Εντούτοις, ο 1BFGS έλυσε όλα τα προβλήματα και στις τρεις διαστάσεις, ενώ ο LSTRS παρουσίασε κάποιες αποτυχίες. Για $m = 2$, ο αλγόριθμος 2BFGS πέτυχε το μικρότερο μέσο όρο επαναλήψεων, παρόλο που παρουσίασε μία αποτυχία.

Τέλος, σύμφωνα με το Σχήμα 6.13, μπορούμε να δούμε ότι η μέθοδος μας χρειάστηκε το μικρότερο CPU χρόνο, για να λύσει τα προβλήματα μεγαλύτερης διάστασης. Επομένως, μπορούμε να συμπεράνουμε ότι τα εσωτερικά γινόμενα και τα διανυσματικά αθροίσματα που χρειάζεται η προτεινόμενη μέθοδος για να λύσει το ΥΠΕ, μέσα σε έναν αλγόριθμο περιοχής εμπιστοσύνης, είναι ανέξοδα, σε σχέση με την εφαρμογή της μεθόδου Lanczos, την οποία χρησιμοποιεί η μέθοδος LSTRS.

Βασιζόμενοι στις παραπάνω συγκρίσεις, μπορούμε να παρατηρήσουμε ότι η προτεινόμενη μέθοδος παρουσίασε λιγότερες αποτυχίες, μικρότερο αριθμό επαναλήψεων και CPU χρόνου για να λύσει τα προβλήματα, σε όλες τις διαστάσεις, σε σύγκριση με τις άλλες μεθόδους. Εμφάνισε δηλαδή την καλύτερη συμπεριφορά. Επίσης, δεδομένου ότι η προτεινόμενη μέθοδος παρουσίασε την καλύτερη επίδοση στην επίλυση του ΥΠΕ, όπως είδαμε στην προηγούμενη ενότητα, μπορούμε να πούμε ότι αυτό το γεγονός επηρέασε και την απόδοση της μεθόδου στην επίλυση προβλημάτων με τη μέθοδο της περιοχής εμπιστοσύνης. Συνεπώς, μπορούμε να συμπεράνουμε ότι η προτεινόμενη μέθοδος είναι πολλά υποσχόμενη για την επίλυση μεγάλης διάστασης προβλημάτων, όταν χρησιμοποιείται μέσα σε ένα πλαίσιο μεθόδων περιοχών εμπιστοσύνης.

6.3 Συμπεράσματα

Σε αυτό το κεφάλαιο αποτιμήσαμε την προτεινόμενη μέθοδο επίλυσης του ΥΠΕ, τόσο στην επίλυση τυχαίων στιγμιotypών, όσο και προβλημάτων βελτιστοποίησης, μέσω ενός τυπικού αλγόριθμου περιοχών εμπιστοσύνης.

Όπως είδαμε από όλα τα πειραματικά αποτελέσματα, η προτεινόμενη μέθοδος παρουσίασε την καλύτερη συμπεριφορά από όλες τις συγκρινόμενες μεθόδους, σε όλες τις ενημερώσεις. Αναφορικά με την επίλυση τυχαίων στιγμιotypών ΥΠΕ, είδαμε ότι η μέθοδος μας παρουσίασε την καλύτερη ακρίβεια λύσης, τον μικρότερο αριθμό επαναλήψεων και είχε το μεγαλύτερο ποσοστό επιτυχίας, από ότι οι υπόλοιπες μέθοδοι. Επίσης,

χρειάστηκε τον λιγότερο χρόνο, από το οποίο προκύπτει ότι η μέθοδος είναι ανέξοδη υπολογιστικά. Συνεπώς, τα εσωτερικά γινόμενα και τα διανυσματικά αθροίσματα που απαιτούνται για να υπολογιστεί το δοκιμαστικό βήμα φαίνεται ότι απαιτούν πολύ λιγότερο χρόνο από ότι η χρησιμοποίηση της μεθόδου Lanczos κατά τη διάρκεια επίλυσης του ΥΠΕ. Επίσης, φαίνεται ότι οι ποσότητες που μας ενδιαφέρει να υπολογίσουμε βρίσκονται με πολλή μεγάλη ακρίβεια, αφού η προτεινόμενη μέθοδος εμφανίζει τη μεγαλύτερη ακρίβεια λύσης και σχετικής λύσης. Επιπλέον, χειρίζεται τόσο την τυπική όσο και τη δύσκολη περίπτωση αποτελεσματικά. Μάλιστα, μπορούμε να πούμε ότι η δύσκολη περίπτωση είναι η περίπτωση όπου ο προτεινόμενος αλγόριθμος είχε τη μεγαλύτερη αποδοτικότητα.

Επίσης, παρατηρήσαμε ότι η διάσταση δεν αποτελεί πρόβλημα για την μέθοδό μας. Ακόμα και στις διαστάσεις της τάξης του 1 000 000 μεταβλητών, ο αλγόριθμος παρουσίασε 100% επιτυχία στην λύση του ΥΠΕ. Συνεπώς, η προτεινόμενη μέθοδος είναι κατάλληλη για την επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης. Επιπρόσθετα, η μέθοδος είναι αποδοτική, ανεξάρτητα από την ενημέρωση που χρησιμοποιήθηκε για την προσέγγιση του εσοιανού πίνακα. Τέλος, χειρίζεται τις περιπτώσεις τόσο των γραμμικά ανεξάρτητα διανυσμάτων όσο και των γραμμικά εξαρτημένων, με τον ίδιο αποτελεσματικό τρόπο, κάτι το οποίο δεν κάνουν οι άλλες μέθοδοι.

Όσο αναφορά την επίλυση του ΥΠΕ μέσα σε ένα πλαίσιο μεθόδων περιοχής εμπιστοσύνης, παρατηρήσαμε ότι η αποτελεσματικότητα της μεθόδου στην επίλυση μεγάλης κλίμακας υποπροβλημάτων, είχε θετικό αντίκτυπο στην απόδοση του αλγορίθμου περιοχών εμπιστοσύνης. Ο αλγόριθμος ο οποίος έλυσε μεγάλης κλίμακας προβλήματα βελτιστοποίησης χωρίς περιορισμούς με την καλύτερη συμπεριφορά, ήταν αυτός που χρησιμοποίησε την προτεινόμενη μέθοδο στην επίλυση του ΥΠΕ. Συμπερασματικά, μπορούμε να πούμε ότι η ακρίβεια της λύσης του ΥΠΕ, επηρέασε τη συνολική συμπεριφορά του αλγορίθμου βελτιστοποίησης. Ο τελευταίος είχε το μεγαλύτερο ποσοστό επιτυχίας και τις λιγότερες επαναλήψεις, σε σχέση με τους υπόλοιπους συγκρινόμενους αλγορίθμους.

Εν κατακλείδι, η προτεινόμενη μέθοδος δείχνει να είναι ανέξοδη, αποτελεσματική και ακριβής όταν χρησιμοποιείται στην επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης. Η απόδοση του αλγορίθμου παραμένει ανεπηρέαστη, όσο μεγάλη και αν είναι η διάσταση του προβλή-

ματος που επιλύεται.

Μια καμπυλόγραμμη αναζήτηση ελάχιστης μνήμης

Σε αυτό το κεφάλαιο θα αναπτύξουμε μια μέθοδο καμπυλόγραμμης αναζήτησης, για την επίλυση του προβλήματος ελαχιστοποίησης χωρίς περιορισμούς $\min\{f(x) \mid x \in \mathbb{R}^n\}$, όπου η αντικειμενική συνάρτηση $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι συνεχώς διαφορίσιμη. Η προτεινόμενη μέθοδος είναι συνδυασμός της κλασικής γραμμικής αναζήτησης και της καμπυλόγραμμης αναζήτησης των Moré & Sorensen [63]. Η μέθοδος χρησιμοποιεί την BFGS ενημέρωση ελάχιστης μνήμης για τον υπολογισμό των κατευθύνσεων μείωσης. Σκοπός μας είναι να αντλήσουμε πληροφορία από την ιδιοσύνθεση του προσεγγιστικού εσσιανού πίνακα, αποφεύγοντας οποιαδήποτε αποθήκευση ή παραγοντοποίηση πίνακα. Επιπρόσθετα, να παράγουμε μια κατάλληλη κατεύθυνση αρνητικής καμπυλότητας, έτσι ώστε η μέθοδος να μπορεί να εφαρμοστεί για την επίλυση μεγάλης κλίμακας προβλημάτων. Τα ακόλουθα αποτελούν ερευνητικά αποτελέσματα της παρούσας διατριβής, τα οποία καταγράφονται στις εργασίες των Apostolopoulou et al. [5] και [7].

7.1 Το ζεύγος κατευθύνσεων μείωσης

Όπως είδαμε στο Κεφάλαιο 2 και συγκεκριμένα στην Ενότητα 2.6, οι μέθοδοι καμπυλόγραμμης αναζήτησης απαιτούν τον υπολογισμό της πιο αρνητικής ιδιοτιμής και του αντίστοιχου ιδιοδιανύσματος, του εσσιανού ή του προσεγγιστικού εσσιανού πίνακα. Στην πρόσφατη διεθνή βιβλιογραφία, η έρευνα για την επίλυση προβλημάτων πολλών μεταβλητών, εστιάζεται στη χρησιμοποίηση τόσο της μεθόδου Lanczos, όσο και εσωτερικών επαναληπτικών διαδικασιών, για τον υπολογισμό των ανωτέρω ποσοτήτων. Στόχος μας είναι να αποφύγουμε κάθε εσωτερική επανάληψη ή αποθήκευση πί-

νακα, ώστε η μέθοδος να είναι πολύ γρήγορη και ανέξοδη υπολογιστικά, χωρίς όμως να χάνει σε αποδοτικότητα ή αξιοπιστία.

Στην προτεινόμενη μέθοδο καμπυλόγραμμης αναζήτησης, χρησιμοποιούνται δύο κατευθύνσεις μείωσης. Όταν ο πίνακας είναι θετικά ορισμένος, ο αλγόριθμος ακολουθεί μια διαδικασία γραμμικής αναζήτησης. Όταν ο πίνακας είναι μη θετικά ορισμένος, τότε η αναζήτηση του αλγορίθμου είναι καμπυλόγραμμη. Σε αυτή την περίπτωση, χρησιμοποιείται ένα ζεύγος κατευθύνσεων μείωσης, η Quasi-Newton κατεύθυνση και η κατεύθυνση αρνητικής καμπυλότητας. Πιο αναλυτικά, όπως είδαμε και στην Ενότητα 2.6 του Κεφαλαίου 2, το ζεύγος των κατευθύνσεων μείωσης ορίζεται ως εξής:

$$p_k = \begin{cases} -B_k^{-1}g_k, & \text{αν } B_k > 0; \\ -g_k, & \text{διαφορετικά.} \end{cases} \quad (7.1)$$

όπου g_k η κλίση της αντικειμενικής συνάρτησης στην επανάληψη x_k και

$$d_k = \begin{cases} 0, & \text{αν } B_k > 0; \\ -\text{sgn}(u_k^T g_k)u_k, & \text{διαφορετικά,} \end{cases} \quad (7.2)$$

όπου u_k το ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του B_k . Να υπενθυμίσουμε εδώ ότι, όπως είδαμε στο Κεφάλαιο 2, στην Υποενότητα 2.6.1, για το ζεύγος των κατευθύνσεων μείωσης θα πρέπει να ισχύουν οι Συνθήκες 2.1-2.3, δηλαδή οι συνθήκες

Συνθήκη 7.1. $g_k^T p_k = 0$ έπεται ότι $g_k = 0$ και $p_k = 0$.

Συνθήκη 7.2. $g_k^T p_k \rightarrow 0$ έπεται ότι $g_k \rightarrow 0$ και $p_k \rightarrow 0$.

Συνθήκη 7.3. $d_k^T B_k d_k \rightarrow 0$ έπεται ότι $\min(\lambda_1, 0) \rightarrow 0$ και $d_k \rightarrow 0$, όπου λ_1 είναι η μικρότερη ιδιοτιμή του B_k .

Στην προτεινόμενη μέθοδο, και οι δύο κατευθύνσεις παράγονται από την BFGS ενημέρωση ελάχιστης μνήμης,

$$B_{k+1} = \theta_{k+1}I - \theta_{k+1} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

όπου $\theta_{k+1} \in \mathbb{R}^*$. Στη συνέχεια θα υπολογίσουμε τις δύο κατευθύνσεις μείωσης, δηλαδή την Quasi-Newton κατεύθυνση p και την κατεύθυνση

αρνητικής καμπυλότητας d , όπως αυτές προκύπτουν όταν ο προσεγγιστικός εσσιανός πίνακας είναι ο BFGS πίνακας ελάχιστης μνήμης.

Όπως μπορούμε να παρατηρήσουμε από τη σχέση (7.1), η Quasi-Newton κατεύθυνση p υπολογίζεται απευθείας πολλαπλασιάζοντας τον αντίστροφο του πίνακα B με την κλίση της αντικειμενικής συνάρτησης, g . Επειδή ο πίνακας αυτός είναι γνωστός, ο υπολογισμός της συγκεκριμένης κατεύθυνσης γίνεται αρκετά εύκολα. Ενώ, όταν ο B είναι μη θετικά ορισμένος, η p είναι απλώς η μείον κλίση της αντικειμενικής συνάρτησης f .

Όσον αφορά την κατεύθυνση αρνητικής καμπυλότητας d , από τη σχέση (7.2), βλέπουμε ότι απαιτείται ο υπολογισμός του ιδιοδιανύσματος που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του πίνακα B . Ο υπολογισμός αυτός επιτυγχάνεται εφαρμόζοντας τη μέθοδο εύρεσης ιδιοδιανυσμάτων των ελάχιστης μνήμης Quasi-Newton ενημερώσεων που αναπτύξαμε στο Κεφάλαιο 3 και συγκεκριμένα στην Ενότητα 3.6 (σελ. 107), η οποία αναφέρεται στη BFGS ενημέρωση ελάχιστης μνήμης.

7.2 Η περίπτωση του θετικά ορισμένου πίνακα

Όταν ο BFGS πίνακας ελάχιστης μνήμης B_{k+1} είναι θετικά ορισμένος, η Quasi-Newton κατεύθυνση υπολογίζεται μέσω του αντιστρόφου του (σελ. 42),

$$H_{k+1} = \frac{1}{\theta_{k+1}} I + \left(1 + \frac{y_k^T y_k}{\theta_{k+1} s_k^T y_k} \right) \frac{s_k s_k^T}{s_k^T y_k} - \frac{s_k y_k^T + y_k s_k^T}{\theta_{k+1} s_k^T y_k}, \quad (7.3)$$

όπου ο αρχικός πίνακας H_0 είναι ένα πολλαπλάσιο του μοναδιαίου πίνακα, της μορφής $H_0 = (1/\theta_{k+1} I)$. Επομένως, αφού

$$p_{k+1} = -H_{k+1} g_{k+1},$$

χρησιμοποιώντας τη σχέση (7.3), έχουμε ότι

$$p_{k+1} = -\frac{g_{k+1}}{\theta_{k+1}} - \left(1 + \frac{y_k^T y_k}{\theta_{k+1} s_k^T y_k} \right) \frac{s_k^T s_k}{s_k^T y_k} s_k + \frac{y_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} s_k + \frac{s_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} y_k,$$

η οποία μπορεί να γραφεί ως εξής:

$$p_{k+1} = -\frac{g_{k+1}}{\theta_{k+1}} - \left[\left(1 + \frac{y_k^T y_k}{\theta_{k+1} s_k^T y_k} \right) \frac{s_k^T g_{k+1}}{s_k^T y_k} - \frac{y_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} \right] s_k + \frac{s_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} y_k. \quad (7.4)$$

Στην περίπτωση όπου τα διανύσματα s_k και y_k είναι γραμμικώς εξαρτημένα, δηλαδή ισχύει ότι $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, τότε ο B_{k+1} λαμβάνει την απλουστευμένη έκφραση

$$B_{k+1} = \theta_{k+1} I - (\theta_{k+1} - \kappa) \frac{s_k s_k^T}{s_k^T s_k},$$

με αποτέλεσμα ο αντίστροφός του να είναι ο πίνακας

$$H_{k+1} = \frac{1}{\theta_{k+1}} I - \frac{\kappa - \theta_{k+1}}{\theta_{k+1} \kappa} \frac{s_k s_k^T}{s_k^T s_k}.$$

Επομένως, από τον παραπάνω πίνακα προκύπτει ότι η Quasi-Newton κατεύθυνση είναι της μορφής

$$p_{k+1} = -\frac{g_{k+1}}{\theta_{k+1}} + \frac{\kappa - \theta_{k+1}}{\theta_{k+1} \kappa} \frac{s_k^T g_{k+1}}{s_k^T s_k} s_k. \quad (7.5)$$

Συνοψίζοντας τα παραπάνω, από τις σχέσεις (7.4) και (7.5), βλέπουμε ότι για τον υπολογισμό της κατεύθυνσης p ισχύουν τα εξής:

$$p_{k+1} = \begin{cases} -\frac{g_{k+1}}{\theta_{k+1}} - \left[\left(1 + \frac{y_k^T y_k}{\theta_{k+1} s_k^T y_k} \right) \frac{s_k^T g_{k+1}}{s_k^T y_k} - \frac{y_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} \right] s_k + \frac{s_k^T g_{k+1}}{\theta_{k+1} s_k^T y_k} y_k, & \text{αν } y_k \neq \kappa s_k; \\ -\frac{g_{k+1}}{\theta_{k+1}} + \frac{\kappa - \theta_{k+1}}{\theta_{k+1} \kappa} \frac{s_k^T g_{k+1}}{s_k^T s_k} s_k, & \text{αν } y_k = \kappa s_k, \end{cases}$$

όπου $\kappa \in \mathbb{R}^*$.

Προφανώς, όταν τα διανύσματα s_k και y_k είναι γραμμικά ανεξάρτητα, απαιτούνται περισσότεροι υπολογισμοί από ότι στην αντίθετη περίπτωση. Πιο συγκεκριμένα, στην πρώτη περίπτωση απαιτείται άθροισμα τριών διανυσμάτων και τεσσάρων εσωτερικών γινομένων, ενώ αν η παράμετρος θ_{k+1}

είναι η BB φασματική παράμετρος $\theta_{k+1} = s_k^T y_k / s_k^T s_k$, απαιτούνται πέντε εσωτερικά γινόμενα. Στην περίπτωση όπου τα s_k και y_k είναι γραμμικά εξαρτημένα, αθροίζονται δύο διανύσματα και δύο εσωτερικά γινόμενα, ανεξάρτητα από την επιλογή της παραμέτρου θ_{k+1} . Τα απαραίτητα διανύσματα και εσωτερικά γινόμενα για τις δύο παραπάνω περιπτώσεις συνοψίζονται στους Πίνακες 7.1 και 7.2.

Ο Πίνακας 7.1 αναφέρεται στην περίπτωση των γραμμικά ανεξάρτητων διανυσμάτων. Η πρώτη στήλη δείχνει τα διανύσματα που προστίθενται για να υπολογιστεί η p . Η δεύτερη στήλη δηλώνει τα εσωτερικά γινόμενα που απαιτούνται. Να σημειώσουμε εδώ ότι η παράμετρος θ_{k+1} μπορεί να αυξήσει ανεπαίσθητα την πολυπλοκότητα του υπολογισμού της p . Όπως είδαμε πιο πριν, αν επιλέξουμε ως θ_{k+1} την BB φασματική παράμετρο $s_k^T y_k / s_k^T s_k$, τότε θα πρέπει να υπολογίσουμε άλλο ένα εσωτερικό γινόμενο, το $s_k^T s_k$.

Πίνακας 7.1: Ποσότητες υπολογισμού της p_{k+1} όταν τα s_k, y_k είναι γραμμικά ανεξάρτητα

διανύσματα	εσωτερικά γινόμενα
g_{k+1}	$y_k^T y_k$
s_k	$s_k^T y_k$
y_k	$s_k^T g_{k+1}$
	$y_k^T g_{k+1}$
Σύνολο:	3
	4

Πίνακας 7.2: Ποσότητες υπολογισμού της p_{k+1} όταν τα s_k, y_k είναι γραμμικά εξαρτημένα

	διανύσματα	εσωτερικά γινόμενα
	g_{k+1}	$s_k^T s_k$
	s_k	$s_k^T g_{k+1}$
Σύνολο:	2	2

Ο Πίνακας 7.2 αναφέρεται στην περίπτωση των γραμμικά εξαρτημένων διανυσμάτων. Η πρώτη στήλη αναφέρεται στα διανύσματα που προστίθενται ώστε να υπολογιστεί η Quasi-Newton κατεύθυνση. Η δεύτερη στήλη δείχνει τα απαιτούμενα εσωτερικά γινόμενα, που όπως βλέπουμε είναι ανεξάρτητα από την επιλογή της παραμέτρου θ_{k+1} , λόγω του ότι τα s_k, y_k συνδέονται με τη σχέση $y_k = \kappa s_k, \kappa \in \mathbb{R}^*$.

Πίνακας 7.3: Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της p

	$s_k \neq \kappa y_k$	$s_k = \kappa y_k$
διανύσματα:	3	2
εσωτερικά γινόμενα:	4	2

Ο Πίνακας 7.3 δείχνει συγκεντρωτικά τον αριθμό των διανυσμάτων και εσωτερικών γινομένων που απαιτούνται ώστε να υπολογιστεί η p , αναφορικά με τη σχέση που συνδέει τα διανύσματα s_k και y_k . Η δεύτερη στήλη αναφέρεται στην περίπτωση όπου αυτά είναι γραμμικά ανεξάρτητα, ενώ η τρίτη στήλη, στην περίπτωση όπου είναι γραμμικά εξαρτημένα. Προφανώς στην τελευταία αυτή περίπτωση, ο συνολικός αριθμός υπολογισμών είναι κατά πολύ μικρότερος από ότι στην πρώτη περίπτωση.

Όσον αφορά την κατεύθυνση αρνητικής καμπυλότητας d , η οποία ορίζεται στη σχέση (7.2), στην περίπτωση του θετικά ορισμένου πίνακα είναι ίση με μηδέν. Συνεπώς, δεν απαιτείται κανένας απολύτως υπολογισμός

για τη συγκεκριμένη κατεύθυνση.

7.3 Η περίπτωση του αόριστου πίνακα

Όταν ο B_{k+1} είναι μη θετικά ορισμένος, η Quasi-Newton κατεύθυνση p , ορίζεται ως η μείον κλίση της αντικειμενικής συνάρτησης, δηλαδή

$$p_{k+1} = -g_{k+1}.$$

Όμως, για την κατεύθυνση αρνητικής καμπυλότητας d , απαιτείται η εύρεση του ιδιοδιανύσματος που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του πίνακα B . Η όλη διαδικασία της εύρεσης αυτού του ιδιοδιανύσματος αναλύεται στο Κεφάλαιο 3 και συγκεκριμένα στην Ενότητα 3.6, για τις ιδιοτιμές και τα ιδιοδιανύσματα της BFGS ενημέρωσης ελάχιστης μνήμης. Για να βρούμε το συγκεκριμένο ιδιοδιάνυσμα διακρίνουμε δύο περιπτώσεις:

- την περίπτωση όπου η πιο αρνητική ιδιοτιμή του πίνακα είναι διακριτή και,
- την περίπτωση όπου η πιο αρνητική ιδιοτιμή του πίνακα είναι πολλαπλή.

Να υπενθυμίσουμε εδώ ότι, σύμφωνα με το Θεώρημα 3.1, σελ. 86, η πρώτη περίπτωση εμφανίζεται όταν τα διανύσματα s_k και y_k είναι γραμμικά ανεξάρτητα. Επίσης, σύμφωνα με την ανάλυση της Ενότητας 3.2, εμφανίζεται και όταν τα s_k και y_k είναι συγγραμμικά, της μορφής $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$, και ισχύει ότι $\kappa < \theta_{k+1}$. Αντίθετα, η δεύτερη περίπτωση εμφανίζεται μόνο αν τα s_k και y_k είναι συγγραμμικά και ισχύει ότι $\kappa \geq \theta_{k+1}$.

Στην πρώτη περίπτωση, όταν τα διανύσματα s_k και y_k είναι γραμμικά ανεξάρτητα, για να βρούμε το ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή, μπορούμε να εφαρμόσουμε την Πρόταση 3.2 (σελ. 96). Σύμφωνα με αυτή, το επιθυμητό ιδιοδιάνυσμα είναι το

$$u(\lambda_1) = \frac{(\lambda_1 s_k - y_k)^T y_k}{(\lambda_1 s_k - y_k)^T s_k} s_k - y_k, \quad (7.6)$$

όπου λ_1 η μικρότερη ιδιοτιμή του B_{k+1} . Ένας άλλος τρόπος είναι να χρησιμοποιήσουμε το Πρόσχημα 4.1 (σελ. 136), όπου σύμφωνα με αυτό, το

ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή λ_1 του B είναι το κανονικοποιημένο διάνυσμα

$$u(\sigma) = \frac{c_1(\sigma)u_0 + c_2(\sigma)s_k + c_3(\sigma)y_k}{\sqrt{n(\sigma^2 - \beta_1\sigma + \beta_2)(\theta_{k+1} - \sigma)}}, \quad (7.7)$$

όπου $\sigma = \lambda_1 + \epsilon$, $\epsilon \rightarrow 0$, μια ϵ -διαταραγμένη διακριτή ιδιοτιμή του BFGS πίνακα ελάχιστης μνήμης,

$$u_0 = (1, 1, \dots, 1)^T, \quad \beta_1 = \theta_{k+1} + y_k^T y_k / s_k^T y_k, \quad \beta_2 = \theta_{k+1} s_k^T y_k / s_k^T s_k$$

και οι συντελεστές των διανυσμάτων u_0 , s_k και y_k ορίζονται από τις σχέσεις

$$\begin{aligned} c_1(\sigma) &= \sigma^2 - \beta_1\sigma + \beta_2, \\ c_2(\sigma) &= \left[(\beta_1 - \sigma) \sum_{i=1}^n s_k^i - \sum_{i=1}^n y_k^i \right] \frac{\theta_{k+1}}{s_k^T s_k}, \quad \text{και} \\ c_3(\sigma) &= \frac{\sigma}{s_k^T y_k} \sum_{i=1}^n y_k^i - \frac{\theta_{k+1}}{s_k^T s_k} \sum_{i=1}^n s_k^i. \end{aligned}$$

Όταν τα s_k και y_k είναι συγγραμμικά, με $y_k = \kappa s_k$, για να είναι η μικρότερη ιδιοτιμή διακριτή θα πρέπει να ισχύει η ανίσωση

$$\kappa < \theta_{k+1},$$

οπότε

$$\lambda_1 = \kappa \quad \text{και} \quad \lambda_2 = \lambda_3 = \dots = \lambda_n = \theta_{k+1}.$$

Τότε, σύμφωνα με την ανάλυση της Ενότητας 3.2, το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή κ είναι το κανονικοποιημένο διάνυσμα s_k . Αυτό συμβαίνει γιατί σε αυτή την περίπτωση ο πίνακας B_{k+1} είναι της μορφής

$$B_{k+1} = \theta_{k+1}I + (\kappa - \theta_{k+1}) \frac{s_k s_k^T}{s_k^T s_k},$$

επομένως, εύκολα βλέπουμε ότι ισχύει το εξής:

$$B_{k+1}s_k = \theta_{k+1}s_k + (\kappa - \theta_{k+1}) \frac{s_k^T s_k}{s_k^T s_k} s_k = \kappa s_k.$$

Στη δεύτερη περίπτωση όπου η μικρότερη ιδιοτιμή λ_1 του B_{k+1} είναι πολλαπλή, θα πρέπει τα s_k , y_k να είναι συγγραμμικά, με $y_k = \kappa s_k$ και επιπλέον να ισχύει η σχέση

$$\kappa \geq \theta_{k+1}. \quad (7.8)$$

Τότε έχουμε ότι

$$\lambda_1 = \lambda_2 = \dots = \lambda_{n-1} = \theta_{k+1} \quad \text{και} \quad \lambda_n = \kappa.$$

Προφανώς, τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή θ_{k+1} είναι γενικευμένα. Επιπλέον, αν στη σχέση (7.8) ισχύει αυστηρά η ανίσωση, δηλαδή ισχύει ότι

$$\kappa > \theta_{k+1},$$

όπως μπορούμε να παρατηρήσουμε από τη μορφή του B_{k+1} , τα ιδιοδιανύσματα που αντιστοιχούν στη θ_{k+1} θα πρέπει να είναι κάθετα στο s_k . Σε αυτή την περίπτωση, ένα τέτοιο διάνυσμα (βλέπε σελ. 95) είναι το κανονικοποιημένο διάνυσμα

$$u = \left(-\frac{s_k^n}{s_k^1}, 0, \dots, 0, 1 \right)^T. \quad (7.9)$$

Τέλος, αν στη σχέση (7.8) ισχύει μόνο η ισότητα, δηλαδή

$$\kappa = \theta_{k+1},$$

τότε $B_{k+1} = \theta_{k+1} I$, άρα τα ιδιοδιανύσματα είναι οι στήλες του μοναδιαίου πίνακα. Οπότε, ένα ιδιοδιάνυσμα που αντιστοιχεί στη θ_{k+1} είναι το

$$u = (1, 0, \dots, 0)^T. \quad (7.10)$$

Συνοψίζοντας τις παραπάνω περιπτώσεις και λαμβάνοντας υπόψη τα κανονικοποιημένα διανύσματα που ορίζονται στις σχέσεις (7.7), (7.9) και (7.10), ο υπολογισμός της κατεύθυνσης αρνητικής καμπυλότητας γίνεται

ως εξής:

$$d_{k+1} = \begin{cases} -\text{sgn}(u^T g_{k+1})u, & y_k \neq \kappa s_k; \\ -\text{sgn}(s_k^T g_{k+1})s_k, & y_k = \kappa s_k \text{ \& } \kappa < \theta_{k+1}; \\ -\text{sgn}\left(-\frac{s_k^n}{s_k^1}g_{k+1}^1 + g_{k+1}^n\right)\left(-\frac{s_k^n}{s_k^1}, 0, \dots, 1\right)^T, & y_k = \kappa s_k \text{ \& } \kappa > \theta_{k+1}; \\ -\text{sgn}(g_{k+1}^1)(1, 0, \dots, 0)^T, & y_k = \kappa s_k \text{ \& } \kappa = \theta_{k+1}, \end{cases}$$

όπου

$$u^T g_{k+1} = \frac{(\lambda_1 s_k - y_k)^T y_k}{(\lambda_1 s_k - y_k)^T s_k} s_k^T g_{k+1} - y_k^T g_{k+1},$$

ή

$$u^T g_{k+1} = \frac{c_1(\sigma) \sum_{i=1}^n g_{k+1}^i + c_2(\sigma) s_k^T g_{k+1} + c_3(\sigma) y_k^T g_{k+1}}{\sqrt{n}(\sigma^2 - \beta_1 \sigma + \beta_2)(\theta_{k+1} - \sigma)},$$

όπου $\sigma = \lambda_1 + \epsilon$ και $\kappa \in \mathbb{R}^*$.

Για καθεμία από τις παραπάνω περιπτώσεις, στους πίνακες που ακολουθούν συνοψίζονται τα απαραίτητα διανύσματα και εσωτερικά γινόμενα που πρέπει να υπολογιστούν ώστε να βρούμε την κατεύθυνση αρνητικής καρπυλότητας d_{k+1} .

Πίνακας 7.4: Ποσότητες υπολογισμού της d_{k+1} όταν τα s_k , y_k είναι γραμμικά ανεξάρτητα

διανύσματα	εσωτερικά γινόμενα
s_k	$s_k^T s_k$
y_k	$s_k^T y_k$
g_{k+1}	$y_k^T y_k$
	$s_k^T g_{k+1}$
	$y_k^T g_{k+1}$
Σύνολο:	3
	5

Ο Πίνακας 7.4 αναφέρεται στην περίπτωση των γραμμικά ανεξάρτητων διανυσμάτων. Η πρώτη στήλη δείχνει τα διανύσματα που προστίθενται για να υπολογιστεί η d . Η δεύτερη στήλη δηλώνει τα εσωτερικά γινόμενα που απαιτούνται. Όσον αφορά την παράμετρο θ_{k+1} , αυτή δεν αυξάνει την

πολυπλοκότητα των υπολογισμών, στην περίπτωση που ορίζεται ως

$$\theta_{k+1} = 1, \quad \theta_{k+1} = s_k^T y_k / s_k^T s_k \quad \text{ή} \quad \theta_{k+1} = y_k^T y_k / s_k^T y_k.$$

Βέβαια, στη γενική περίπτωση, ανάλογα με το πως θα οριστεί, μπορεί να αυξήσει τον αριθμό των πράξεων που απαιτούνται για τον υπολογισμό της d .

Ο Πίνακας 7.5 αναφέρεται στην περίπτωση των γραμμικά εξαρτημένων διανυσμάτων. Η πρώτη στήλη αναφέρεται στα διανύσματα που προστίθενται ώστε να υπολογιστεί η κατεύθυνση αρνητικής καμπυλότητας. Η δεύτερη στήλη δείχνει τα απαιτούμενα εσωτερικά γινόμενα, που όπως βλέπουμε είναι ανεξάρτητα από την επιλογή της παραμέτρου θ_{k+1} , λόγω του ότι τα s_k και y_k συνδέονται με τη σχέση $y_k = \kappa s_k$, $\kappa \in \mathbb{R}^*$.

Πίνακας 7.5: Ποσότητες υπολογισμού της d_{k+1} όταν τα s_k , y_k είναι γραμμικά εξαρτημένα

σχέση ιδιοτιμών	διανύσματα	εσωτερικά γινόμενα
$y_k = \kappa s_k$ και $\kappa < \theta_{k+1}$	g_{k+1} s_k	$s_k^T g_{k+1}$
Σύνολο:	2	1
$y_k = \kappa s_k$ και $\kappa > \theta_{k+1}$	—	$-(s_k^n / s_k^1) g_{k+1}^1 + g_{k+1}^n$
Σύνολο:	0	1
$y_k = \kappa s_k$ και $\kappa = \theta_{k+1}$	—	—
Σύνολο:	0	0

Επίσης, στις περιπτώσεις όπου ισχύουν οι σχέσεις

$$\kappa > \theta_{k+1}$$

και

$$\kappa = \theta_{k+1},$$

δεν απαιτείται ο υπολογισμός κανενός διανύσματος. Αυτό συμβαίνει διότι τα διανύσματα που χρησιμοποιούνται στην πρώτη και τη δεύτερη περίπτωση, έχουν μηδενικές συνιστώσες, εκτός από την πρώτη και την τελευταία, ή μόνο την πρώτη συνιστώσα, αντίστοιχα.

Οι Πίνακες 7.6 και 7.7 δείχνουν το συγκεντρωτικό αριθμό διανυσμάτων και εσωτερικών γινομένων που απαιτούνται, έτσι ώστε να υπολογίσουμε την κατεύθυνση αρνητικής καμπυλότητας d , όσον αφορά και τους δύο τρόπους που μπορεί να σχετίζονται τα διανύσματα s_k και y_k .

Πίνακας 7.6: Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της d όταν $y_k \neq \kappa s_k$

διανύσματα:	3
εσωτερικά γινόμενα:	5

Πίνακας 7.7: Σύνολο απαιτούμενων διανυσμάτων και εσωτερικών γινομένων για τον υπολογισμό της d_{k+1} όταν $y_k = \kappa s_k$

Ποσότητες	$\kappa < \theta_{k+1}$	$\kappa > \theta_{k+1}$	$\kappa = \theta_{k+1}$
διανύσματα:	2	—	—
εσωτερικά γινόμενα:	1	—	—

Ο Πίνακας 7.6 αναφέρεται στην περίπτωση όπου τα s_k και y_k είναι γραμμικά ανεξάρτητα, ενώ ο Πίνακας 7.7, στην περίπτωση όπου τα διανύσματα s_k , y_k είναι γραμμικά εξαρτημένα. Στην πρώτη στήλη αναγράφονται οι ποσότητες υπολογισμού. Η δεύτερη στήλη αναφέρεται στη περίπτωση όπου η πιο αρνητική ιδιοτιμή είναι διακριτή, ενώ η τρίτη στην περίπτωση όπου η πιο αρνητική ιδιοτιμή είναι πολλαπλή. Τέλος, η τελευταία στήλη δίνει τους απαιτούμενους υπολογισμούς όταν ο B_{k+1} έχει ως μοναδική ιδιοτιμή τη θ_{k+1} , πολλαπλότητας n . Προφανώς, όπως είναι α-

ναμενόμενο, η κατεύθυνση d απαιτεί πολύ λιγότερους υπολογισμούς όταν τα διανύσματα s_k, y_k είναι γραμμικά εξαρτημένα.

Το ζεύγος κατευθύνσεων μείωσης που προκύπτει χρησιμοποιώντας τη BFGS ενημέρωση ελάχιστης μνήμης, είναι ζεύγος κατευθύνσεων μείωσης. Πράγματι, όταν $B_k > 0$, τότε το ζεύγος κατευθύνσεων μείωσης $\{p, d\}$, ορίζεται ως

$$p_k = -B_k^{-1}g_k \quad \text{και} \quad d_k = 0,$$

αντίστοιχα. Συνεπώς, οι σχέσεις

$$p_k^T g_k = -g_k^T B_k^{-1} g_k < 0, \quad d_k^T g_k = 0, \quad \text{και} \quad d_k^T B_k d_k = 0,$$

ισχύουν. Όταν ο B_k είναι μη θετικά ορισμένος, τότε έχουμε ότι

$$p_k = -g_k \quad \text{και} \quad d_k = -\text{sgn}(u_k^T g_k) u_k.$$

Επομένως, για την κατεύθυνση p έχουμε ότι

$$g_k^T p_k = -\|g_k\|^2 < 0,$$

ενώ για την d ισχύουν οι παρακάτω ανισότητες:

$$d_k^T g_k = -\text{sgn}(u_k^T g_k) u_k^T g_k \leq 0, \quad \text{και} \quad d_k^T B_k d_k = \lambda_1 \|u_k\|^2 = \lambda_1 < 0.$$

Ακόμα, οι p_k και d_k ικανοποιούν τις Συνθήκες 7.1-7.3, αφού η p_k είναι μια Quasi-Newton κατεύθυνση, ενώ η d_k συνδέεται με το κανονικοποιημένο ιδιοδιάνυσμα που αντιστοιχεί στην πιο αρνητική ιδιοτιμή του B_k [42].

7.4 Ο προτεινόμενος αλγόριθμος

Θεωρούμε το ακόλουθο πρόβλημα βελτιστοποίησης

$$\min_{x \in \mathbb{R}^n} f(x),$$

όπου $f : \mathbb{R}^n \rightarrow \mathbb{R}$ είναι η αντικειμενική συνάρτηση της μεταβλητής x . Θεωρούμε ότι η αντικειμενική συνάρτηση f ικανοποιεί την παρακάτω υπόθεση:

Υπόθεση 7.1. Έστω $f : \mathbb{R}^n \rightarrow \mathbb{R}$ δύο φορές συνεχώς διαφορίσιμη συνάρτηση σε ένα ανοιχτό σύνολο $\mathbb{B}$ και υποθέτουμε ότι για κάποιο $x_0 \in \mathbb{B}$, το σύνολο

$$\Omega = \{x \in \mathbb{B} \mid f(x) \leq f(x_0)\}$$

είναι συμπαγές υποσύνολο του $\mathbb{B}$.

Για την επίλυση του παραπάνω προβλήματος, προτείνουμε έναν αλγόριθμο βελτιστοποίησης, ο οποίος χρησιμοποιεί πληροφορία, η οποία βρίσκεται στον προσεγγιστικό εσσιανό πίνακα της f . Ως προσεγγιστικό εσσιανό της αντικειμενικής συνάρτησης f , θεωρούμε τον BFGS πίνακα ελάχιστης μνήμης

$$B_{k+1} = \theta_{k+1}I - \theta_{k+1} \frac{s_k s_k^T}{s_k^T s_k} + \frac{y_k y_k^T}{s_k^T y_k},$$

όπου $\theta_{k+1} \in \mathbb{R}^*$. Για τον παραπάνω BFGS πίνακα ελάχιστης μνήμης κάνουμε την ακόλουθη υπόθεση:

Υπόθεση 7.2. Έστω B_k οι πίνακες που προκύπτουν από τη BFGS ενημέρωση ελάχιστης μνήμης. Τότε υπάρχει θετική σταθερά M , τέτοια ώστε $\|B_k\| \leq M$, για όλα τα k .

Ο ακόλουθος προτεινόμενος αλγόριθμος περιγράφει μια καμπυλόγραμμη αναζήτηση, βασισμένη σε μια διαδικασία Armijo. Το ζεύγος κατευθύνσεων μείωσης $\{p, d\}$ που χρησιμοποιείται, υπολογίζεται όπως αναλύθηκε στις προηγούμενες ενότητες, το οποίο χρησιμοποιείται από τον προτεινόμενο αλγόριθμο CMBFGS. Συγκεκριμένα, ο Αλγόριθμος 7.4 εκτελεί μια γραμμική αναζήτηση, όταν ο B είναι θετικά ορισμένος. Σε αντίθετη περίπτωση, ο αλγόριθμος ακολουθεί την καμπυλόγραμμη αναζήτηση που προτάθηκε από τους Moré & Sorensen [63]. Σε αυτή την καμπυλόγραμμη αναζήτηση, χρησιμοποιείται η κατεύθυνση αρνητικής καμπυλότητας d . Επομένως, το επαναληπτικό σχήμα είναι της μορφής

$$x_{k+1} = \begin{cases} x_k + \alpha_k p_k, & \text{όταν } B_k > 0; \\ x_k + \alpha_k^2 p_k + \alpha_k d_k, & \text{διαφορετικά,} \end{cases} \quad (7.11)$$

όπου η p_k είναι η Quasi-Newton κατεύθυνση, d_k η κατεύθυνση αρνητικής καμπυλότητας και α_k το μήκος βήματος.

Αλγόριθμος 7.1 Ο αλγόριθμος CMBFGS

Βήμα 1. Δοθέντων x_0 , $m = 1$, $0 < \sigma_1 < \sigma_2 < 1$, $0 < \mu < 1$, $0 < \eta < 1$ και $\epsilon \rightarrow 0$; θέσε $k = 0$ και $\ell = 0$; υπολόγισε τη g_0 ;

Βήμα 2. Αν $\|g_k\| \leq \epsilon$ σταμάτα; αλλιώς υπολόγισε τις ιδιοτιμές λ_i του B_k ;

Βήμα 3. Αν $\lambda_1 > 0$ τότε

α) υπολόγισε την p_k ; θέσε $d_k = 0$ και $\alpha = 1$;

β) Βρες μήκος βήματος $\alpha_k \in [\sigma_1\alpha, \sigma_2\alpha]$ τέτοιο ώστε να ισχύει η ανίσωση

$$f(x_k + \alpha p_k) - f(x_k) \leq \mu \alpha g_k^T p_k; \quad (7.12)$$

και θέσε $\alpha_k = \alpha$.

Αλλιώς

γ) θέσε $p_k = -g_k$ και υπολόγισε το ιδιοδιάνυσμα u_k ;

θέσε $d_k = -\text{sgn}(u_k^T g_k) u_k$ και $\alpha_k = 1$;

δ) Διάλεξε το μικρότερο μη αρνητικό ακέραιο i έτσι ώστε να ισχύει η ανίσωση

$$f(x_k + \alpha_k^{2i} p_k + \alpha_k^i d_k) - f(x_k) \leq \eta \alpha_k^{2i} (g_k^T p_k + \frac{1}{2} d_k^T B_k d_k) \quad (7.13)$$

Βήμα 4. Θέσε

$$x_{k+1} = \begin{cases} x_k + \alpha_k p_k, & \text{αν } \lambda_1 > 0; \\ x_k + \alpha_k^{2i} p_k + \alpha_k^i d_k, & \text{διαφορετικά.} \end{cases} \quad (7.14)$$

Βήμα 5. Υπολόγισε τα g_{k+1} , $s_k = x_{k+1} - x_k$ και $y_k = g_{k+1} - g_k$; αν

$$|s_k^T y_k| > 10^{-6} \|s_k\| \|y_k\|,$$

αποθήκευσε το διανυσματικό ζεύγος $\{s_k, y_k\}$ και θέσε $\ell = \ell + 1$;
αν $\ell > m$ διέγραψε το διανυσματικό ζεύγος $\{s_{\ell-m}, y_{\ell-m}\}$ από τη μνήμη;

Βήμα 6. Θέσε $k = k + 1$ και πήγαινε στο Βήμα 2;

Να σημειώσουμε εδώ ότι ο B_k μπορεί να είναι αόριστος. Επομένως, για

να είναι καλά ορισμένος, δεν κάνουμε ενημέρωση αν ισχύει η ανίσωση

$$|s_k^T y_k| \leq 10^{-6} \|s_k\| \|y_k\|.$$

Δοθέντος ζεύγους κατευθύνσεων μείωσης $\{p_k, d_k\}$, οι οποίες προήλθαν από τη BFGS ενημέρωση ελάχιστης μνήμης, θέλουμε να βρούμε ένα μήκος βήματος $\alpha > 0$ τέτοιο ώστε είτε

$$f(x + \alpha p) < f(x), \quad \text{όταν } B > 0,$$

ή

$$f(x + \alpha^2 p + \alpha d) < f(x), \quad \text{διαφορετικά.}$$

Το ακόλουθο λήμμα τεκμηριώνει ότι υπάρχει $\bar{\alpha} > 0$ τέτοιο ώστε $f(x + \alpha^2 p + \alpha d) < f(x)$, $\alpha \in (0, \bar{\alpha}]$.

Λήμμα 7.1. (Moré & Sorencen [63]) Έστω $\phi : \mathbb{R} \rightarrow \mathbb{R}$ δυο φορές συνεχώς διαφορίσιμη συνάρτηση σε ένα ανοιχτό διάστημα $\mathbf{I}$ το οποίο περιέχει την αρχή των αξόνων, με $\phi(\alpha) = f(x + \alpha^2 p + \alpha d)$. Υποθέτουμε ότι $c \in (0, 1)$. Τότε υπάρχει $\bar{\alpha} > 0$ στο $\mathbf{I}$, τέτοιο ώστε

$$\phi(\alpha) \leq \phi(0) + c \left[\phi'(0)\alpha + \frac{1}{2}\phi''(0)\alpha^2 \right]$$

για όλα τα $\alpha \in [0, \bar{\alpha}]$, παρέχοντας ότι είτε $\phi'(0) < 0$, ή $\phi'(0) = 0$ και $\phi''(0) < 0$.

Λήμμα 7.2. Έστω ότι οι f και B_k ικανοποιούν τις Υποθέσεις 7.1 και 7.2, αντίστοιχα. Υποθέτουμε ότι το $\{p_k, d_k\}$ είναι ένα διανυσματικό ζεύγος κατευθύνσεων μείωσης παραγόμενο από τη BFGS ενημέρωση ελάχιστης μνήμης και τον Αλγόριθμο 7.4. Τότε υπάρχει $\alpha_k > 0$ τέτοιο ώστε ικανοποιείται είτε η σχέση (7.12) ή η σχέση (7.13).

Απόδειξη. Θεωρούμε τις ακόλουθες περιπτώσεις:

(i) Ο B_k είναι θετικά ορισμένος. Οι κατευθύνσεις είναι οι $p_k = -B_k^{-1}g_k$ και $d_k = 0$, ενώ η διαδικασία της γραμμικής αναζήτησης θα πρέπει να ικανοποιεί την (7.12). Αν $\alpha_k = 1$ τότε προφανώς η συνθήκη (7.12) ικανοποιείται. Υποθέτουμε ότι η ανισότητα (7.12) δεν ικανοποιείται ποτέ. Τότε υπάρχει μια ακολουθία α_j που συγκλίνει στο 0 καθώς το $j \rightarrow \infty$, έτσι

ώστε να ισχύει η ανίσωση

$$f(x_k + \alpha_j p_k) - f(x_k) \geq \mu \alpha_j g_k^T p_k.$$

Από το θεώρημα μέσης τιμής (βλέπε σελίδα 18) υπάρχει $\delta \in (0, 1)$ τέτοιο ώστε

$$\alpha_j g(x_k + \delta \alpha_j p_k)^T p_k > \mu \alpha_j g_k^T p_k. \quad (7.15)$$

Αφού ο B_k είναι θετικά ορισμένος και ισχύει $\|B_k\| \leq M$ για όλα τα k , έχουμε ότι

$$p_k^T B_k p_k \leq M \|p\|^2. \quad (7.16)$$

Συνδυάζοντας τις (7.16) και (7.15) έχουμε ότι

$$M(\alpha_j \|p_k\|)^2 \geq \alpha_j [g(x_k + \delta \alpha_j p_k) - g(x_k)]^T p_k > (\mu - 1) \alpha_j g_k^T p_k,$$

από το οποίο έπεται ότι

$$\frac{(\mu - 1) g_k^T p_k}{M \|p_k\|^2} \leq \alpha_j. \quad (7.17)$$

Για $j \rightarrow \infty$, από τη σχέση (7.17) βρίσκουμε ότι

$$(\mu - 1) g_k^T p_k \leq 0$$

το οποίο έρχεται σε αντίθεση με το γεγονός ότι $\mu \in (0, 1)$ και $g_k^T p_k < 0$. Άρα, η ανισότητα (7.12) ικανοποιείται πάντα.

(ii) Ο B_{k+1} είναι μη θετικά ορισμένος. Τότε οι κατευθύνσεις ορίζονται ως $p_k = -g_k$ και $d_k = -\text{sgn}(u_k^T g_k) u_k$, ενώ η καμπυλόγραμμη αναζήτηση πρέπει να ικανοποιεί τη σχέση (7.13). Θέτοντας

$$\phi_k(\alpha_k) = f(x_k + \alpha_k^2 p_k + \alpha_k d_k),$$

έχουμε ότι

$$\phi_k'(0) = g_k^T d_k \leq 0 \quad \text{και} \quad \phi_k''(0) = g_k^T p_k + d_k^T B_k d_k < 0.$$

Από το Λήμμα 7.1 ξέρουμε ότι υπάρχει $a_k > 0$ τέτοιο ώστε η σχέση (7.13) να ισχύει. Επομένως, υπάρχει $a_k > 0$ τέτοιο ώστε να ισχύει είτε η σχέση (7.12) ή η σχέση (7.13). $\square$

Στη συνέχεια δίνουμε το θεώρημα σύγκλισης του προτεινόμενου Αλγορίθμου 7.4.

Θεώρημα 7.1. Έστω ότι οι f και B_k ικανοποιούν τις Υποθέσεις 7.1 και 7.2, αντίστοιχα και έστω ότι οι $\{\|p_k\|\}$ και $\{\|d_k\|\}$ είναι φραγμένες. Έστω $\{x_k\}$ η ακολουθία σημείων που παράγεται από τον Αλγόριθμο 7.4. Τότε ισχύει ότι

$$\lim_{k \rightarrow \infty} g_k = 0 \quad \text{και} \quad \lim_{k \rightarrow \infty} d_k^T B_k d_k = 0. \quad (7.18)$$

Απόδειξη. Αφού το Ω είναι συμπαγές, η ακολουθία επαναλήψεων $\{x_k\}$ επιδέχεται τουλάχιστον ένα σημείο συσσώρευσης που ανήκει στο Ω . Έστω K_p και K_{pd} σύνολα δεικτών δύο υπακολουθιών επαναλήψεων, τα οποία συγκλίνουν σε ένα σημείο έτσι ώστε

(i) για όλα τα $k \in K_p$, ισχύει η σχέση

$$f(x_k + \alpha_k p_k) - f(x_k) \leq \mu \alpha_k g_k^T p_k,$$

και

(ii) για όλα τα $k \in K_{pd}$, ισχύει η σχέση

$$f(x_k + \alpha_k^{2i} p_k + \alpha_k^i d_k) - f(x_k) \leq \eta \alpha_k^{2i} \left(g_k^T p_k + \frac{1}{2} d_k^T B_k d_k \right).$$

Το σύνολο δεικτών δεν μπορεί να είναι κενό.

Πρώτα υποθέτουμε ότι το σύνολο δεικτών K_p δεν είναι κενό. Αυτό το σύνολο δεικτών σχετίζεται με την υπακολουθία των επαναλήψεων $\{x_{K_p}\}$, για τις οποίες οι πίνακες B_{K_p} είναι θετικά ορισμένοι. Άρα ισχύει η σχέση (7.16). Από τις (7.17) και (7.12) έχουμε ότι

$$f(x_k + \alpha_k p_k) \leq f(x_k) - \frac{\mu(1-\mu)}{M} \frac{(g_k^T p_k)^2}{\|p_k\|^2}, \quad k \in K_p.$$

Αν ϑ_k είναι η γωνία μεταξύ των p_k και $-g_k$, η παραπάνω σχέση μπορεί να γραφεί ως

$$f(x_k + \alpha_k p_k) \leq f(x_k) - \frac{\mu(1-\mu)}{M} \cos^2 \vartheta_k \|g_k\|^2. \quad (7.19)$$

Αφού $s_k = -\alpha_k B_k^{-1} g_k$ και $B_k s_k = \theta_k s_k$, έχουμε ότι

$$\cos \vartheta_k = -\frac{g_k^T p_k}{\|g\| \|p_k\|} = \frac{s_k^T B_k s_k}{\|s_k\| \|B_k s_k\|} = 1.$$

Επομένως, η σχέση (7.13) γίνεται

$$f(x_k + \alpha_k p_k) \leq f(x_k) - \frac{\mu(1-\mu)}{M} \|g_k\|^2.$$

Από την υπόθεση του θεωρήματος, η ακολουθία $\{f_k\}$, $k \in K_p$ είναι κάτω φραγμένη και άρα η ακολουθία $\{f_{k+1} - f_k\}$ συγκλίνει στο μηδέν, το οποίο συνεπάγεται ότι $\lim_{k \rightarrow \infty} \|g_k\| = 0$. Επιπλέον, $d_k = 0$ για όλα τα $k \in K_p$, και προφανώς ισχύει ότι

$$\lim_{k \rightarrow \infty} d_k^T B_k d_k = 0.$$

Έστω τώρα ότι το σύνολο δεικτών K_{pd} , το οποίο σχετίζεται με τις επαναλήψεις $x_{K_{pd}}$ για τις οποίες οι πίνακες $B_{K_{pd}}$ είναι αόριστοι, δεν είναι κενό. Αν i_k είναι ο μικρότερος μη αρνητικός ακέραιος για τον οποίο να ισχύουν η σχέση

$$x_{k+1} = x_k + \alpha^{2i_k} p_k + \alpha^{i_k} d_k \in \Omega$$

και η σχέση

$$f(x_k + \alpha_k^{2i} p_k + \alpha_k^{i_k} d_k) - f(x_k) \leq \eta \alpha_k^{2i} \left(g_k^T p_k + \frac{1}{2} d_k^T B_k d_k \right),$$

έπεται από την (7.13) ότι

$$|f(x_{k+1}) - f(x_k)| \geq c \alpha^{2i_k} |g_k^T p_k + \frac{1}{2} d_k^T B_k d_k|,$$

για όλα τα $k \in K_{pd}$. Αφού $\{f_{k+1} - f_k\} \rightarrow 0$, $k \in K_{pd}$, έχουμε ότι

$$\alpha^{2i_k} |g_k^T p_k + \frac{1}{2} d_k^T B_k d_k| \rightarrow 0, \quad k \rightarrow \infty, \quad k \in K_{pd}.$$

Επομένως, λαμβάνοντας υπόψη ότι $g_k^T p_k = -\|g_k\|^2$, θα ισχύει είτε

$$\lim_{k \rightarrow \infty} g_k = 0 \quad \text{και} \quad \lim_{k \rightarrow \infty} d_k^T B_k d_k = 0, \quad k \in K_{pd},$$

ή $\alpha^{i_k} \rightarrow 0$ καθώς $k \rightarrow \infty$, $k \in K_{pd}$, δηλαδή $\lim_{k \rightarrow \infty} i_k = \infty$. Υποθέτουμε ότι $\alpha_k^{i_k} \rightarrow 0$. Τότε, από τον ορισμό των i_k έχουμε ότι

$$f(x_k + \alpha^{2i_k} p_k + \alpha^{i_k} d_k) > f(x_k) + c\alpha^{2i_k} (g_k^T p_k + \frac{1}{2} d_k^T B_k d_k). \quad (7.20)$$

Αναπτύσσοντας το αριστερό μέρος της (7.20) χρησιμοποιώντας το θεώρημα Taylor, έχουμε ότι

$$\begin{aligned} & \alpha^{2i_k} g_k^T p_k + \alpha^{i_k} g_k^T d_k + \frac{1}{2} (\alpha^{2i_k} p_k + \alpha^{i_k} d_k)^T B_k (\alpha^{2i_k} p_k + \alpha^{i_k} d_k) \\ & + o(\alpha^{2i_k}) > c\alpha^{2i_k} (g_k^T p_k + \frac{1}{2} d_k^T B_k d_k). \end{aligned}$$

Χρησιμοποιώντας το γεγονός ότι $g_k^T d_k \leq 0$, συσσωρεύοντας όλους τους όρους τάξεως μεγαλύτερης από $O(\alpha^{2i_k})$ στους όρους τάξης $o(\alpha^{2i_k})$, έχουμε ότι

$$o(\alpha^{2i_k}) > -(1-c)\alpha^{2i_k} (g_k^T p_k + \frac{1}{2} d_k^T B_k d_k). \quad (7.21)$$

Αφού

$$\lim_{k \rightarrow \infty} \frac{o(\alpha^{2i_k})}{(1-c)\alpha^{2i_k}} = 0, \quad -g_k^T p_k \geq 0, \quad -d_k^T B_k d_k \geq 0, \quad k \in K_{pd} \quad (7.22)$$

με $g_k^T p_k = -\|g_k\|^2$, το συμπέρασμα προκύπτει από τις (7.21) και 7.22). Οπότε, αφού $K_p \cup K_{pd} \neq \emptyset$, κάθε σημείο συσσώρευσης της ακολουθίας, ικανοποιεί τις σχέσεις (7.18). $\square$

7.5 Συμπεράσματα

Στο κεφάλαιο αυτό προτείναμε μια μέθοδο καμπυλόγραμμης αναζήτησης, για την επίλυση προβλημάτων ελαχιστοποίησης χωρίς περιορισμούς, στα οποία η αντικειμενική συνάρτηση είναι δύο φορές συνεχώς διαφορίσιμη και κάτω φραγμένη. Η μέθοδος χρησιμοποιεί την πληροφορία που είναι διαθέσιμη στον προσεγγιστικό εσσιανό πίνακα της f .

Το ζεύγος κατευθύνσεων μείωσης που χρησιμοποιήθηκε βασίστηκε στις ιδιότητες της BFGS ενημέρωσης ελάχιστης μνήμης. Όπως είδαμε, ο τρόπος υπολογισμού τους δεν απαιτεί μεγάλο υπολογιστικό κόστος. Αυτό έχει ως αποτέλεσμα, ο προτεινόμενος αλγόριθμος να μην αποθηκεύει ή

παραγοντοποιεί πίνακες, έτσι ώστε να μπορεί να εφαρμοστεί για την επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης. Οι κατευθύνσεις μείωσης υπολογίζονται μέσω εσωτερικών γινομένων και διανυσματικών αθροισμάτων.

Χρησιμοποιώντας την κατάλληλη πληροφορία που περιέχεται στις ιδιοτιμές και τα ιδιοδιανύσματα της BFGS ενημέρωσης ελάχιστης μνήμης, ο προτεινόμενος αλγόριθμος έχει ενσωματώσει αμφότερες μια γραμμική και μια καμπυλόγραμμη αναζήτηση. Αποδείξαμε ότι η αναζήτηση γίνεται με τέτοιο τρόπο, ώστε ο αλγόριθμος να διατηρεί τις καλές ιδιότητες σύγκλισης τόσο της γραμμικής, όσο και της καμπυλόγραμμης αναζήτησης.

Αποτίμηση του αλγορίθμου CMBFGS

Στο κεφάλαιο αυτό αποτιμούμε τον προτεινόμενο αλγόριθμο βελτιστοποίησης, που αναπτύχθηκε στο Κεφάλαιο 7. Ο αλγόριθμος δοκιμάζεται στην επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης χωρίς περιορισμούς. Επίσης αξιολογείται η επίδοσή του σε εκπαίδευση νευρωνικών δικτύων και ειδικότερα σε προβλήματα κατηγοριοποίησης βιοϊατρικών δεδομένων. Τα αριθμητικά αποτελέσματα έχουν καταγραφεί στις εργασίες των Apostolopoulou et al. [7, 11, 57].

8.1 Βελτιστοποίηση συναρτήσεων της βιβλιοθήκης CUTEr

Σε αυτό το κεφάλαιο αποτιμούμε αριθμητικά την απόδοση του Αλγορίθμου 7.4, (CMBFGS), σε μεγάλης διάστασης προβλήματα χωρίς περιορισμούς, τα οποία έχουν καθιερωθεί στη διεθνή βιβλιογραφία ως τα κατεξοχήν προβλήματα αξιολόγησης αλγορίθμων βελτιστοποίησης. Τα αποτελέσματα που παρουσιάζονται έχουν καταγραφεί στην εργασία *A curvilinear method based on minimal memory BFGS updates* [7].

Ο αλγόριθμος CMBFGS χρησιμοποιεί το ζεύγος κατευθύνσεων μείωσης $\{p, d\}$, ο υπολογισμός των οποίων βασίζεται στις ιδιότητες της BFGS ενημέρωσης ελάχιστης μνήμης. Οι κατευθύνσεις αυτές χρησιμοποιούνται σε μια γραμμική ή καμπυλόγραμμη αναζήτηση, ανάλογα αν ο πίνακας B είναι θετικά ορισμένος ή όχι, αντίστοιχα. Υλοποιήσαμε τον CMBFGS και τον δοκιμάσαμε στην επίλυση μεγάλης κλίμακας προβλημάτων ελαχιστοποίησης χωρίς περιορισμούς. Τα προβλήματα αυτά προέρχονται από τη συλλογή CUTEr, που έχει προταθεί από τους Gould et al. [47]. Πιο συγκεκριμένα, πρόκειται για μη γραμμικά προβλήματα χωρίς περιορισμούς, με αριθμό μεταβλητών από 999 έως 100 000, τα οποία συλλέχθηκαν μέσω του

shell script SELECT της CUTer. Έχουν συνεχείς δεύτερες παραγώγους και είναι κάτω φραγμένα. Συνολικά, το επιλεγόμενο σύνολο αποτελείται από πενήντα εννέα (59) συναρτήσεις. Στον Πίνακα 8.1 αναφέρονται τα ονόματα των συναρτήσεων που χρησιμοποιήθηκαν.

Πίνακας 8.1: Προβλήματα

ARWHEAD	DIXMAANG	EXTROSNB	MOREBV	SPARSINE
BROYDN7D	DIXMAANH	FLETGBV2	MSQRTALS	SPARSQUR
BRYBND	DIXMAANH	FLETGBV3	MSQRTBLS	SPMSRTLS
CHAINWOO	DIXMAANI	FLETCHBV	NONCVXU2	SROSENBR
COSINE	DIXMAANJ	FLETCHCR	NONCVXUN	TESTQUAD
CRAGGLVY	DIXMAANK	FMINSRF2	NONDIA	TOINTGSS
DIXMAANA	DIXMAANL	FMINSURF	NONDQUAR	TQUARTIC
DIXMAANB	DIXON3DQ	FREUROTH	PENALTY1	TRIDIA
DIXMAANC	DQDRTIC	GENHUMPS	POWELLSG	VARDIM
DIXMAAND	DQRTIC	GENROSE	POWER	VAREIGVL
DIXMAANE	EDENSCH	INDEF	QUARTC	WOODS
DIXMAANF	ENGVAL1	LIARWHD	SCHMVETT	

Για κάθε συνάρτηση θεωρήσαμε 3 αριθμητικά πειράματα, με αριθμό μεταβλητών 10^3 , 10^4 , και 10^5 , αντίστοιχα. Σε όλες τις περιπτώσεις χρησιμοποιήσαμε ως αρχικά σημεία, τα σημεία που παρέχονται αυτόματα από τη βιβλιοθήκη CUTer, μέσω του εργαλείου αποκωδικοποίησης Sif Decode.

Τρέξαμε όλα τα πειράματα σε Pentium 1.86 GHz προσωπικό υπολογιστή, μνήμης RAM 2GB, σε λειτουργικό σύστημα Linux. Ο κώδικας του προτεινόμενου αλγορίθμου έχει υλοποιηθεί σε FORTRAN 90, όπου και χρησιμοποιήθηκε η επιλογή ‘-O2’ του GNU compiler gfortran, έκδοσης 4.4. Στην υλοποίησή μας για τις παραμέτρους μ και η , που περιγράφονται στον Αλγόριθμο 7.4 χρησιμοποιήσαμε τις τιμές $\mu = \eta = 10^{-4}$.

Συγκρίναμε τη μέθοδό μας με τη μέθοδο γραμμικής αναζήτησης L-BFGS (αλγόριθμος LBFGS), [67] επειδή ο αλγόριθμός μας μπορεί να θεωρηθεί ως μια τροποποίηση της αντίστοιχης μεθόδου ελάχιστης μνήμης. Ο κώδικας της L-BFGS αποκτήθηκε από την ιστοσελίδα του Jorge Nocedal, ενώ χρησιμοποιήθηκε ένα διανυσματικό ζεύγος ($m=1$) για την ενημέρωση

του πίνακα.

Και στις δύο μεθόδους, η παράμετρος θ που χρησιμοποιήθηκε, ορίστηκε ως $y_k^T y_k / s_k^T y_k$. Ως κριτήριο τερματισμού τέθηκε το $\|g_k\|_\infty \leq 10^{-5}$. Θεωρήσαμε ως αποτυχία όλα τα τρεξίματα στα οποία το κριτήριο σύγκλισης δεν ικανοποιήθηκε μέσα σε συγκεκριμένο αριθμό επαναλήψεων (maxiter = 10000).

Στην πιο μικρή διάσταση, δηλαδή για $n = 1\,000$, ο αλγόριθμος CMBFGS απέτυχε να λύσει τέσσερα (4) από τα πενήντα εννέα (59) προβλήματα. Ο LBFGS αλγόριθμος απέτυχε να λύσει ένα ακόμα, συνολικά δηλαδή, πέντε (5) προβλήματα. Τα προβλήματα αυτά συνοψίζονται στον Πίνακα 8.2.

Πίνακας 8.2: Αποτυχίες για $n = 1\,000$

CMBFGS	LBFGS
FLETGBV3	FLETGBV3
FLETGBV	FLETGBV
INDEF	INDEF
NONCVXUN	NONCVXUN
	FREUROTH
Σύνολο: 4	5

Για $n = 10\,000$ ο CMBFGS απέτυχε να λύσει δεκαπέντε (15) προβλήματα, ενώ ο LBFGS, δεκάξι (16) προβλήματα. Τα ονόματα των προβλημάτων αυτών φαίνονται στον Πίνακα 8.3. Τέλος, στη μεγαλύτερη από όλες τις διαστάσεις, όπου $n = 100\,000$, και οι δύο αλγόριθμοι απέτυχαν να λύσουν δεκαεπτά (17) από τα πενήντα εννέα (59) προβλήματα που υπάρχουν συνολικά. Επιπρόσθετα, ο αλγόριθμος LBFGS απέτυχε να λύσει και άλλα δύο, σύνολο δεκαεννέα (19). Συγκεντρωτικά, τα προβλήματα αυτά καταγράφονται στον Πίνακα 8.4

Πίνακας 8.3: Αποτυχίες για $n = 10\,000$

CMBFGS		LBFGS	
ARWHEAD	GENROSE	ARWHEAD	FREUROTH
CHAINWOO	INDEF	BROYDN7D	GENROSE
CRAGGLVY	NONCVXU2	CHAINWOO	INDEF
FLETGBV3	NONCVXUN	CRAGGLVY	NONCVXU2
FLETCHBV	SPARSINE	ENGVAL1	NONCVXUN
FLETCHCR	TESTQUAD	FLETGBV3	SPARSINE
FREUROTH	VARDIM	FLETCHBV	TESTQUAD
GENHUMPS		FLETCHCR	VARDIM
Σύνολο :	15	16	

Πίνακας 8.4: Αποτυχίες για $n = 100\,000$

CMBFGS		LBFGS	
BROYDN7D	INDEF	BROYDN7D	INDEF
CHAINWOO	NONCVXU2	CHAINWOO	NONCVXU2
CRAGGLVY	NONCVXUN	CRAGGLVY	NONCVXUN
ENGVAL1	SPARSINE	ENGVAL1	SPARSINE
FLETGBV3	TESTQUAD	FLETGBV3	TESTQUAD
FLETCHBV	TRIDIA	FLETCHBV	TRIDIA
FLETCHCR	VARDIM	FLETCHCR	VARDIM
FREUROTH		FREUROTH	EDENSCH
GENHUMPS		GENHUMPS	PENALTY1
GENROSE		GENROSE	
Σύνολο :	17	19	

Πίνακας 8.5: Ολικός CPU χρόνος για τα επιτυχώς λυμένα προβλήματα και συνολικές αποτυχίες σε όλες τις διαστάσεις

	cpu		αποτυχίες	
	CMBFGS	LBFGS	CMBFGS	LBFGS
$n = 1\ 000$	29.58	26.77	4	5
$n = 10\ 000$	70.36	85.94	15	16
$n = 100\ 000$	570.71	510.99	17	19
Σύνολο	670.65	623.70	26	30

Ο Πίνακας 8.5 παρουσιάζει το συνολικό CPU χρόνο (σε δευτερόλεπτα) που χρειάστηκε καθένας από τους δύο αλγόριθμους για να λύσει τα προβλήματα. Ο χρόνος που μετρήθηκε αφορούσε μόνο τα προβλήματα που λύθηκαν επιτυχώς. Μπορούμε να παρατηρήσουμε ότι, μολονότι ο αλγόριθμος CMBFGS έλυσε περισσότερα προβλήματα από τον LBFGS, ο συνολικός CPU χρόνος που απαιτήθηκε, είναι παρόμοιος και για του δύο αλγόριθμους (διαφέρουν κατά 47 sec). Ακόμα, στον Πίνακα 8.5 συνοψίζεται και ο συνολικός αριθμός των προβλημάτων που απέτυχε να λύσει κάθε αλγόριθμος.

Τέλος, ο Πίνακας 8.6 δείχνει το συνολικό αριθμό από των κατευθύνσεων αρνητικής καμπυλότητας που υπολογίστηκαν από τον αλγόριθμο CMBFGS, στα επιτυχώς λυμένα προβλήματα.

Πίνακας 8.6: Συνολικός αριθμός υπολογιζόμενων κλίσεων αρνητικής καμπυλότητας στα επιτυχώς λυμένα προβλήματα

διάσταση	ιδιοδιανύσματα
$n = 1\ 000$	110
$n = 10\ 000$	24
$n = 100\ 000$	24
Συνολικός αριθμός ιδιοδιανυσμάτων:	158

Για να απεικονίσουμε τις επιδόσεις του κάθε αλγόριθμου, χρησιμο-

ποιήσαμε τα προφίλ απόδοσης των Dolan & Moré [30], ως προς συναρτησιακούς υπολογισμούς και υπολογισμούς κλίσεων.

(α) Απόδοση βασισμένη σε συναρτησιακούς υπολογισμούς

(β) Απόδοση βασισμένη σε υπολογισμούς κλίσεων

Σχήμα 8.1: Προφίλ απόδοσης των CMBFGS και LBFGS, σε 1000 διαστάσεις

(α) Απόδοση βασισμένη σε συναρτησιακούς υπολογισμούς

(β) Απόδοση βασισμένη σε υπολογισμούς κλίσεων

Σχήμα 8.2: Προφίλ απόδοσης των CMBFGS και LBFGS, σε 10 000 διαστάσεις

(α) Απόδοση βασισμένη σε συναρτησιακούς υπολογισμούς

(β) Απόδοση βασισμένη σε υπολογισμούς κλίσεων

Σχήμα 8.3: Προφίλ απόδοσης των CMBFGS και LBFGS, σε 100 000 διαστάσεις

Το Σχήμα 8.1 παρουσιάζει τα προφίλ απόδοσης για προβλήματα 1000 μεταβλητών. Τα επιμέρους Σχήματα 8.1(α) και 8.1(β) δείχνουν τα προφίλ απόδοσης στα οποία ως δείκτης απόδοσης ήταν οι συναρτησιακοί υπολο-

γισμοί και οι υπολογισμοί κλίσεων, αντίστοιχα. Στο Σχήμα 8.1(α') μπορούμε να παρατηρήσουμε ότι και οι δύο αλγόριθμοι συμπεριφέρθηκαν παρόμοια. Ωστόσο, είναι εμφανές ότι ο αλγόριθμος CMBFGS, φαίνεται να υπερέρχει ελάχιστα ως προς τον αριθμό των συναρτησιακών υπολογισμών. Τέλος, το Σχήμα 8.1(β') δείχνει ότι ο προτεινόμενος καμπυλόγραμμος αλγόριθμος είναι πιο αποδοτικός από τον κλασικό αλγόριθμο γραμμικής αναζήτησης, σε ότι αναφορά τους υπολογισμούς κλίσεων.

Παρόμοια συμπεράσματα μπορούν να γίνουν και από τα Σχήματα 8.2 και 8.3, τα οποία αποτυπώνουν τη σύγκριση της απόδοσης των δύο αλγορίθμων για $n = 10\,000$ και $n = 100\,000$, αντίστοιχα. Πιο αναλυτικά, και τα δύο Σχήματα 8.2(α') και 8.3(α'), δείχνουν ότι, ως προς τους συναρτησιακούς υπολογισμούς, οι CMBFGS και LBFGS έχουν σχεδόν την ίδια απόδοση, με τον CMBFGS να έχει ένα προβάδισμα. Τα Σχήματα 8.2(β') και 8.3(β') παρουσιάζουν την απόδοση ως προς τους υπολογισμούς κλίσεων. Και στα δύο σχήματα, φαίνεται ότι ο CMBFGS πέτυχε την καλύτερη απόδοση, σε όρους τόσο αποδοτικότητας, όσο και δυναμικότητας. Συγκεκριμένα, ο CMBFGS έχει 70% πιθανότητα να λύσει όλα τα προβλήματα σε λιγότερο χρόνο από ότι ο LBFGS.

Εν κατακλείδι, η καλύτερη απόδοση και ως προς τους συναρτησιακούς υπολογισμούς αλλά και ως προς τους υπολογισμούς κλίσεων, επιτεύχθηκε από τον αλγόριθμο CMBFGS, ειδικά ως προς τον τελευταίο δείκτη απόδοσης. Η γενική εικόνα του προτεινόμενου αλγορίθμου δείχνει ότι ο CMBFGS, ο οποίος βασίζεται σε πληροφορία καμπυλότητας που αποκτήθηκε από τη BFGS ενημέρωση ελάχιστης μνήμης, είναι πολύ αποδοτικός για την επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης χωρίς περιορισμούς.

8.2 Ο CMBFGS στην εκπαίδευση νευρωνικών δικτύων

Σε αυτή την ενότητα αξιολογούμε την απόδοση του προτεινόμενου αλγορίθμου σε έξι γνωστά προβλήματα που αποκτήθηκαν από τη UCI βάση δεδομένων μηχανικής μάθησης (Repository of Machine Learning Databases) [65]. Περιγράφουμε εν συντομία κάθε πρόβλημα και συγκρίνουμε τον αλγόριθμο CMBFGS με τον LBFGS.

Σε όλα τα δίκτυα εισαγάγαμε την ίδια ακολουθία εισερχόμενων προτύπων (patterns), ενώ τα αρχικά βάρη αρχικοποιήθηκαν με την καθιερωμένη μέθοδο Nguyen-Widrow [66]. Για να αξιολογήσουμε την ακρίβεια κατηγοριοποίησης, χρησιμοποιήσαμε τη διαδικασία *k-fold cross-validation* [53].

Όλες οι προσομοιώσεις έγιναν σε Pentium-IV dual core υπολογιστή (2.0MHz, 1Gbyte RAM), χρησιμοποιώντας το λειτουργικό σύστημα Linux. Για την προσομοίωση χρησιμοποιήθηκε πρωτίστως η πλατφόρμα MATLAB και εν συνεχεία ενσωματώθηκε αυτούσιος ο υλοποιημένος σε FORTRAN κώδικας, δημιουργώντας δυαδικά MEX αρχεία. Τα αποτελέσματα που παρουσιάζονται έχουν καταγραφεί στην εργασία *A memoryless BFGS neural network training algorithm* [11]

8.2.1 Το πρόβλημα του καρκίνου του μαστού

Το πρώτο πρόβλημα αφορά τη διάγνωση του κακοήγη καρκίνου του μαστού. Τα δεδομένα έχουν συλλεχθεί από εξακόσιους ογδόντα τρεις (683) ασθενείς από το Πανεπιστήμιο του Wisconsin. Κάθε ασθενής έχει εννέα (9) χαρακτηριστικά (features) και μια ετικέτα κατηγοριοποίησης (κακοήθης ή καλοήθης όγκος).

Χρησιμοποιήσαμε νευρωνικά δίκτυα με δύο (2) κρυφά στρώματα, με τέσσερις (4) και δύο (2) νευρώνες, αντίστοιχα, όπως έχει προταθεί στη συλλογή προβλημάτων PROBEN1 [75]. Ως κριτήριο τερματισμού τέθηκε η συνάρτηση σφάλματος να είναι $E \leq 0.02$, μέσα σε ένα όριο δύο χιλιάδων (2000) εποχών (epochs). Όλα τα δίκτυα ελέγχθηκαν χρησιμοποιώντας την 10-fold cross validation.

Στο Σχήμα 8.4 παρουσιάζονται τα προφίλ απόδοσης για το πρόβλημα του καρκίνου του μαστού, με σκοπό τη διερεύνηση της αποτελεσματικότητας και της δυναμικότητας της κάθε μεθόδου εκπαίδευσης. Αναφορικά με τους συναρτησιακούς υπολογισμούς, αμφότερες οι μέθοδοι είχαν την ίδια συμπεριφορά, με τον CMBFGS να επιτυγχάνει μια ελαφρώς καλύτερη αποδοτικότητα.

Το προφίλ απόδοσης ως προς τους υπολογισμούς κλίσεων (gradients) δείχνει ότι ο αλγόριθμος CMBFGS είναι πολύ πιο αποδοτικός και δυναμικός από ότι είναι ο LBFGS. Να σημειώσουμε ότι ο υπολογισμός της κλίσης μέσω του αλγορίθμου της οπισθοδρόμησης (back propagation) είναι η πιο χρονοβόρα διαδικασία κατά την εκπαίδευση νευρωνικών δικτύων.

(α) Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β) Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.4: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα του καρκίνου του μαστού.

Επομένως, είναι σαφές ότι ο CMBFGS υπερέχει κατά πολύ διότι εξοικονομεί πολλούς υπολογισμούς κλίσεων και κατά συνέπεια CPU χρόνο.

8.2.2 Το πρόβλημα του διαβήτη

Ο σκοπός αυτού του προβλήματος κατηγοριοποίησης είναι να αποφασίσει πότε οι γυναίκες της ινδιάνικης φυλής Pima (ή Akimel O'odham) είναι θετικές στον διαβήτη ή όχι. Τα δεδομένα αποτελούνται από επτακόσια εξήντα οκτώ (768) διαφορετικά πρότυπα [65, 75], καθένα εκ των οποίων έχει οχτώ (8) χαρακτηριστικά και μια ετικέτα κατηγοριοποίησης (θετικό ή μη ως προς το διαβήτη).

Χρησιμοποιήσαμε νευρωνικά δίκτυα αποτελούμενα από δύο (2) κρυφά στρώματα τεσσάρων (4) νευρώνων και ένα στρώμα εξόδου, αποτελούμενο από δύο (2) νευρώνες [75]. Το κριτήριο τερματισμού ήταν το $E < 0.14$, μέσα σε ένα όριο 2000 εποχών. Όλα τα δίκτυα ελέγχθηκαν χρησιμοποιώντας την 10-fold cross validation.

(α') Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β) Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.5: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα του διαβήτη.

Το Σχήμα 8.5 παρουσιάζει τα προφίλ αποδοτικότητας ως προς τον αριθμό συναρτησιακών υπολογισμών και τον αριθμό υπολογισμών κλίσεων, για το πρόβλημα του διαβήτη. Στο Σχήμα 8.5(α) βλέπουμε ότι και οι δύο μέθοδοι έχουν την ίδια απόδοση, με τον CMBFGS να έχει ένα προβάδισμα έναντι του LBFGS, ως προς τους συναρτησιακούς υπολογισμούς. Επιπλέον, το Σχήμα 8.5(β) δείχνει ότι η προτεινόμενη μέθοδος είναι πιο αποδοτική σε ό,τι αφορά τους υπολογισμούς κλίσεων.

8.2.3 Σύνολο SPECT δεδομένων καρδιάς

Το σύνολο δεδομένων περιέχει στιγμιότυπα εικόνων από τον αξονικό τομογράφο Proton Emission Computed Tomography (SPECT), από το Πανεπιστήμιο του Colorado [65]. Το συγκεκριμένο πρόβλημα είναι πρόβλημα κατηγοριοποίησης, στο οποίο οι ασθενείς κατηγοριοποιούνται ως προς το αν έχουν κανονική εικόνα καρδιάς ή όχι. Η κατανομή αποτελείται από πενήντα πέντε (55) στιγμιότυπα της ανώμαλης κατηγορίας (20.6%) και

(α') Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β') Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.6: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το SPECT πρόβλημα καρδιάς.

διακόσια δώδεκα (212) στιγμιότυπα της κανονικής (79.4%). Η αρχιτεκτονική του δικτύου που χρησιμοποιήθηκε αποτελείται από ένα κρυμμένο στρώμα με έξι (6) νευρώνες και ένα στρώμα εξόδου ενός νευρώνα [90]. Το κριτήριο τερματισμού ήταν το $E \leq 0.1$ μέσα σε ένα όριο χιλίων (1000) εποχών.

Το Σχήμα 8.6 παρουσιάζει το προφίλ απόδοσης για το συγκεκριμένο πρόβλημα καρδιάς. Ως προς τους συναρτησιακούς υπολογισμούς βλέπουμε ότι και οι δύο μέθοδοι συμπεριφέρονται το ίδιο, ενώ ο CMBFGS είναι πολύ πιο αποδοτικός αναφορικά με τον αριθμό υπολογισμών κλίσεων.

8.2.4 Το πρόβλημα της έγκρισης πίστωσης στην Αυστραλία

Η αυστραλιανή βάση δεδομένων για την έγκριση πίστωσης, περιέχει λεπτομέρειες αναφορικά με τις αιτήσεις για απόκτηση πιστωτικής κάρτας. Η συγκεκριμένη βάση είναι αρκετά ενδιαφέρουσα, διότι τα δεδομένα ποικίλουν και έχουν ένα μείγμα χαρακτηριστικών. Το μείγμα αυτό έχει είτε μικρό αριθμό μεταβλητών ή και μεγαλύτερο.

Χρησιμοποιήσαμε νευρωνικά δίκτυα με δύο (2) κρυφά στρώματα [75], αποτελούμενα από δεκάξι (16) και οχτώ (8) νευρώνες, αντίστοιχα. Ακόμα χρησιμοποιήσαμε ένα στρώμα εξόδου δύο (2) νευρώνων. Το κριτήριο τερματισμού ήταν το $E \leq 0.1$, μέσα σε όριο χιλίων (1000) εποχών, ενώ όλα τα δίκτυα ελέγχθηκαν χρησιμοποιώντας την 10-fold cross validation.

Το Σχήμα 8.7 δείχνει τα προφίλ απόδοσης του προβλήματος έγκρισης πίστωσης στην Αυστραλία, όπου βλέπουμε ότι καλύτερος αλγόριθμος και στους δύο δείκτες απόδοσης, φαίνεται να είναι ο CMBFGS. Πιο συγκεκριμένα, το Σχήμα 8.7(α) δείχνει ότι ο CMBFGS είναι ο πιο αποδοτικός αλγόριθμος, αφού η καμπύλη του είναι πιο πάνω από την αντίστοιχη καμπύλη του LBFGS αλγορίθμου. Επίσης, είναι ενδιαφέρον να παρατηρήσουμε ότι στο Σχήμα 8.7(β), ο προτεινόμενος αλγόριθμος έχει την καλύτερη απόδοση, υπερσχύοντας του κλασικού αλγορίθμου, αναφορικά με τον αριθμό υπολογισμών κλίσεων που απαιτούνται για να επιλυθεί το συγκεκριμένο πρόβλημα.

(α') Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β') Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.7: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα της αυστραλιανής έγκρισης πίστωσης.

8.2.5 Το πρόβλημα κολοδακτηριδίων

Αυτό το πρόβλημα βασίζεται σε δεδομένα έλλειψης ισορροπίας τριακοσίων τριάντα έξι (336) προτύπων. Πιο συγκεκριμένα, αφορά την κατηγοριοποίηση της χωροθέτησης της *E. coli* πρωτεΐνης σε οκτώ (8) τοποθεσίες.

Η αρχιτεκτονική του δικτύου αποτελείται από ένα κρυφό στρώμα δεκάξι (16) νευρώνων και ένα στρώμα εξόδου οκτώ (8) νευρώνων. Το κριτήριο τερματισμού ήταν το $E \leq 0.02$, μέσα σε όριο δύο χιλιάδων (2000) εποχών, ενώ όλα τα δίκτυα ελέγχθηκαν χρησιμοποιώντας την 4-fold cross validation [50].

Στο Σχήμα 8.8 παρουσιάζονται τα προφίλ απόδοσης ως προς τους συναρτησιακούς υπολογισμούς και τους υπολογισμούς κλίσεων. Όπως προκύπτει από το συγκεκριμένο σχήμα, ο καλύτερος αλγόριθμος φαίνεται να είναι ο CMBFGS, τόσο ως προς τους συναρτησιακούς υπολογισμούς, όσο και ως προς τον αριθμό των υπολογισμένων κλίσεων. Προφανώς, από το Σχήμα 8.8(β), μπορούμε να δούμε ότι η απόδοση των δύο συγκρινόμενων αλγορίθμων διαφέρει αισθητά, αναφορικά με το δεύτερο δείκτη απόδοσης.

(α) Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β) Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.8: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα των κολοβακτηριδίων.

8.2.6 Το πρόβλημα της μαγιάς

Αυτό το πρόβλημα είναι παρόμοιο με το προηγούμενο πρόβλημα της χωροθέτησης των κολοβακτηριδίων. Πιο συγκεκριμένα, αντιμετωπίζουμε το πρόβλημα χωροθέτησης της κυκλικής πρωτεΐνης που περιέχεται στη μαγιά.

Χρησιμοποιήσαμε νευρωνικά δίκτυα ενός κρυφού στρώματος αποτελούμενου από δεκάξι (16) νευρώνες και ενός στρώματος εξόδου, αποτελούμενο από δέκα (10) νευρώνες [2]. Το κριτήριο τερματισμού ήταν το $E \leq 0.05$, μέσα σε όριο δύο χιλιάδων (2000) εποχών, ενώ όλα τα δίκτυα ελέγχθηκαν χρησιμοποιώντας την 10-fold cross validation [50].

Στο Σχήμα 8.9, το οποίο απεικονίζει τα προφίλ απόδοσης ως προς τους δύο δείκτες απόδοσης, δείχνει ότι ο CMBFGS υπερσχύει σημαντικά έναντι του LBFGS, παρουσιάζοντας την καλύτερη αποδοτικότητα και δυναμική, τόσο ως προς τους συναρτησιακούς υπολογισμούς, όσο και ως προς τους υπολογισμούς κλίσεων.

(α) Απόδοση ως προς τους συναρτησιακούς υπολογισμούς

(β) Απόδοση ως προς τους υπολογισμούς κλίσεων

Σχήμα 8.9: Προφίλ απόδοσης αριθμού συναρτησιακών υπολογισμών και υπολογισμών κλίσεων για το πρόβλημα της μαγιάς.

8.3 Εφαρμογή σε βιοϊατρικά δεδομένα

Σε αυτή την ενότητα εκτιμούμε την απόδοση της γενίκευσης τεχνητών νευρωνικών δικτύων, που εκπαιδεύτηκαν μέσω του CMBFGS αλγορίθμου, σε μεγάλης κλίμακας δεδομένα βιοϊατρικής. Λαμβάνοντας υπόψη ότι ο CMBFGS μπορεί να εφαρμοστεί απευθείας σε μεγάλες διαστάσεις, θελήσαμε να αποτιμήσουμε παράλληλα διάφορες τεχνικές μείωσης διάστασης, που έχουν προταθεί πρόσφατα στη διεθνή βιβλιογραφία, με γνώμονα πάντα τη γενίκευση των εκπαιδευόμενων δικτύων. Συγκεκριμένα, διερευνήσαμε πειραματικά την επίπτωση της εφαρμογής μιας τεχνικής μείωσης διάστασης, ως προστάδιο της εκπαίδευσης των νευρωνικών δικτύων. Τα αποτελέσματα που παρουσιάζονται έχουν καταγραφεί στην εργασία *Classification of Large Biomedical Data using ANNs based on BFGS method* [57].

Ως προβλήματα διαλέξαμε τα ακόλουθα μεγάλης κλίμακας σύνολα δεδομένων:

- *Όγκος του ορθού (Colon Tumor)* [Δ1]: Περιλαμβάνει εξήντα δύο (62) δείγματα που συλλέχθηκαν από ασθενείς με καρκίνο του ορθού (colon-cancer) [1]. Μεταξύ αυτών, σαράντα (40) βιοψίες όγκων είναι από κακοήγη (χαρακτηρισμένες ως "αρνητικές") και είκοσι δύο (22) από υγιή μέρη του ορθού των ίδιων ασθενών ("θετικές"). Συνολικά επιλέχθηκαν δύο χιλιάδες (2000) από περίπου έξι χιλιάδες πεντακόσια (6500) γονίδια, με βάση την αξιοπιστία της βιοψίας.
- *Πρόβλεψη διάχυσης μεγάλου λεμφώματος β-κυττάρων* [Δ2] και *διάχυσης όγκων β-κυττάρων* [Δ3]: Υπάρχουν δύο είδη κατηγοριοποίησης για τη διάχυση μεγάλου λεμφώματος β-κυττάρων (DLBCL) σε αυτό το σύνολο δεδομένων [81]. Το πρώτο είναι η μορφολογία του DLBCL έναντι της μορφολογίας του λεμφώματος ωοθηκών (FL). Το σύνολο αυτό των δεδομένων περιέχει πενήντα οχτώ (58) DLBCL δείγματα και δεκαεννιά (19) δείγματα FL. Το δεύτερο είδος κατηγοριοποίησης αναφέρεται στην πρόβλεψη εάν ένας ασθενής έχει DLBCL. Μεταξύ των πενήντα οχτώ (58) δειγμάτων από ασθενείς με DLBCL, τριάντα δύο (32) από αυτά είναι από ασθενείς που ανάρρωσαν, ενώ είκοσι έξι (26) από αυτά είναι από ασθενείς με θανατηφόρο ή ανθεκτική νόσο.

- *Πνευμονικά αδενοκαρκινώματα (Michigan [Δ4]:* Το σύνολο αυτό των δεδομένων αποτελείται από ογδόντα έξι (86) δείγματα πνευμονικών αδενοκαρκινωμάτων και από δέκα (10) μη νεοπλασματικά πνευμονικά δείγματα [16]. Κάθε δείγμα αποτελείται από επτά χιλιάδες εκατόν είκοσι εννέα (7129) γονίδια.
- *Πρόβλεψη όγκων κεντρικού νευρικού συστήματος [Δ5]:* Πρόκειται για πρόβλεψη εμβρυικών όγκων στο κεντρικό νευρικό σύστημα [72]. Ως επιζώντες αναφέρονται οι ασθενείς οι οποίοι παρέμειναν στη ζωή μετά τη θεραπευτική αγωγή, ενώ αποτυχόντες είναι αυτοί οι οποίοι υπέκυψαν στην ασθένεια. Το σύνολο δεδομένων περιέχει εξήντα (60) δείγματα ασθενών, είκοσιένα (21) επιζώντες και τριάντα εννέα (39) αποτυχόντες. Συνολικά υπάρχουν επτά χιλιάδες εκατόν είκοσι εννέα (7129) γονίδια σε αυτό το σύνολο δεδομένων.
- *Πρόβλεψη προστάτη [Δ6]:* Το συγκεκριμένο σύνολο δεδομένων αναφέρεται στην πρόβλεψη του κλινικού αποτελέσματος όσον αφορά τον καρκίνο του προστάτη [82]. Είκοσι ένας (21) ασθενείς εκτιμήθηκαν ως προς τη μετεγχειρητική τους ανάρρωση. Οχτώ (8) ασθενείς υποτροπίασαν, ενώ δεκατρείς (13) ανάρρωσαν για τουλάχιστον τέσσερα (4) χρόνια.

Για όλα τα παραπάνω σύνολα δεδομένων [Δ1]-[Δ6], οι παράμετροι στον αλγόριθμο CMBFGS τέθηκαν οι εξ' ορισμού τιμές ($c_1 = c_2 = 10^{-4}$). Τα αρχικά βάρη αρχικοποιήθηκαν σύμφωνα με τη μέθοδο των Nguyen-Widrow. Για την εκτίμηση της ακρίβειας της κατηγοριοποίησης, για τα πέντε πρώτα προβλήματα, [Δ1]-[Δ5], χρησιμοποιήθηκε η 10-fold cross-validation επαναλαμβανόμενη εκατό (100) φορές. Ενώ για το τελευταίο πρόβλημα, [Δ6], χρησιμοποιήθηκε η 4-fold cross-validation.

Προκειμένου να εφαρμόσουμε τις τεχνικές μείωσης διάστασης στα αρχικά δεδομένα, απαιτείται να ορίσουμε την επιθυμητή τιμή της διάστασης (target dimensionality), η οποία αναπαριστά την διάσταση του υπόχωρου στον οποίο θέλουμε να απεικονίσουμε τα δεδομένα μας. Ο ορισμός της συγκεκριμένης αυτής τιμής γίνεται μέσω του εκτιμητή μέγιστης πιθανοφάνειας (*maximum likelihood intrinsic dimensionality estimator*) [55]. Για όλες τις τεχνικές μείωσης διάστασης χρησιμοποιήσαμε τις προεπιλεγμένες παραμέτρους όπως ορίζονται στο [86]. Όλες οι προσομοιώσεις έγιναν

σε Pentium-IV υπολογιστή διπλού πυρήνα (2.0MHz, 1Gbyte RAM), ενώ χρησιμοποιήθηκε η εργαλειοθήκη νευρωνικών δικτύων του MATLAB.

Δεδομένα	Δεδομένα πολλών μεταβλητών			Δεδομένα λίγων μεταβλητών		
	Είσοδοι	Νευρώνες σε κρυφά στρώματα	Σύνολο βαρών	Είσοδοι	Νευρώνες σε κρυφά στρώματα	Σύνολο βαρών
Δ1	2000	10-5	20077	11	6	86
Δ2	7129	10-5	71367	22	11	277
Δ3	7129	20-5-10	142787	24	12	326
Δ4	7129	5-5	35692	24	12	326
Δ5	7129	20-5-10	142787	30	15	497
Δ6	12600	10-10	126142	16	8	157

Πίνακας 8.7: Αρχιτεκτονικές νευρωνικών δικτύων

Ο Πίνακας 8.7 περιέχει πληροφορίες σχετικά με την αρχιτεκτονική των δικτύων και το συνολικό αριθμό βαρών που χρησιμοποιήθηκαν. Τα δίκτυα αυτά εκπαιδεύτηκαν σε μικρής και μεγάλης διάστασης δεδομένα για καθένα από τα προβλήματα [Δ1]-[Δ6]. Στον Πίνακα 8.7, ο αριθμός των εισόδων συμπίπτει με την εγγενή διάσταση d που λαμβάνεται από τον εκτιμητή μέγιστης πιθανοφάνειας.

Δεδομένα	None	PCA	LPP	OLPP	KPCA	LEM	LTSA
Δ1	84.5	85.4	47.6	76.9	51.5	80.1	56.8
Δ2	53.7	53.7	58.1	52.0	54.8	50.9	44.3
Δ3	84.0	84.5	66.0	84.3	66.4	82.1	58.9
Δ4	89.5	89.6	88.7	90.1	89.2	89.4	89.4
Δ5	61.3	63.9	49.1	51.2	65.0	57.3	46.0
Δ6	52.3	54.7	52.8	55.6	41.9	51.8	53.6

Πίνακας 8.8: Απόδοση γενίκευσης (%) νευρωνικών δικτύων χρησιμοποιώντας τη CMBFGS μέθοδο.

Στον Πίνακα 8.8 συνοψίζονται τα αποτελέσματα της γενίκευσης των τεχνητών νευρωνικών δικτύων, τα οποία εκπαιδεύτηκαν με τον αλγόριθμο CMBFGS, πάνω σε μεγάλης και μικρής διάστασης δεδομένα. Κάθε στήλη αναφέρει τη μέση απόδοση σε ποσοστό για κάθε σύνολο δεδομένων, χρησιμοποιώντας διαφορετικές τεχνικές μείωσης διάστασης. Η στήλη “None” δείχνει τα αποτελέσματα που προέκυψαν με τα αρχικά δεδομένα

(μεγάλης κλίμακας προβλήματα). Η καλύτερη τεχνική για ένα σύνολο δεδομένων απεικονίζεται με έντονη γραφή. Πρωτίστως, παρατηρούμε ότι η απόδοση ως προς την κατηγοριοποίηση των εκπαιδευόμενων δικτύων δε βελτιώθηκε σημαντικά από την μείωση διάστασης. Εντούτοις, βλέπουμε ότι οι γραμμικές τεχνικές (PCA, LPP, OLPP) υπερτερούν έναντι των μη γραμμικών τεχνικών (KPCA, LEM, LTSA), αφού παρουσιάζουν τα καλύτερα αποτελέσματα γενίκευσης σε πέντε σύνολα δεδομένων. Επιπρόσθετα παρατηρούμε ότι ο παραδοσιακός PCA είναι η καλύτερη τεχνική μείωσης διάστασης, η οποία παρουσίασε την καλύτερη συνολικά απόδοση.

Δεδομένα	PCA	LPP	OLPP	KPCA	LEM	LTSA
Δ1	85.4	48.5	76.7	52.9	80.6	57.3
Δ2	53.9	58.9	52.7	55.3	50.1	44.0
Δ3	84.1	66.0	84.1	65.9	82.9	57.2
Δ4	89.6	88.9	90.0	89.2	89.5	89.4
Δ5	64.5	50.3	51.0	65.8	57.9	46.0
Δ6	54.4	52.8	56.1	41.8	53.4	54.2

Πίνακας 8.9: Απόδοση γενίκευσης (%) νευρωνικών δικτύων χρησιμοποιώντας την LBFGS μέθοδο.

Ο Πίνακας 8.9 αναφέρει τα αποτελέσματα γενίκευσης των τεχνητών νευρωνικών δικτύων που εκπαιδεύτηκαν χρησιμοποιώντας τον αλγόριθμο BFGS (`trainbfg`), του toolbox του MATLAB, πάνω στα δεδομένα μικρής διάστασης, τα οποία ανακτήθηκαν από τις παρούσες τεχνικές μείωσης διάστασης. Ας σημειώσουμε ότι η διαδικασία εκπαίδευσης για τον BFGS αλγόριθμο είναι αδύνατη αν χρησιμοποιηθούν τα αρχικά δεδομένα. Οι ακριβείς αρχιτεκτονικές των δικτύων φαίνονται στο δεύτερο μισό του Πίνακα 8.7.

Συγκρίνοντας τα αποτελέσματα του Πίνακα 8.9 με τη δεύτερη στήλη (“None”) του Πίνακα 8.8, παρατηρούμε ότι και οι δύο αλγόριθμοι έχουν παρόμοια συμπεριφορά. Αυτό δείχνει ότι ο αλγόριθμος CMBFGS είναι κατάλληλος όχι μόνο για μεγάλης κλίμακας δίκτυα, αλλά μπορεί να επιτύχει εξίσου ικανοποιητικά αποτελέσματα γενίκευσης και σε μικρής διάστασης δίκτυα.

8.4 Συμπεράσματα

Σε αυτό το κεφάλαιο ελέγξαμε την αποδοτικότητα του CMBFGS αλγορίθμου ελάχιστης μνήμης. Ο προτεινόμενος αλγόριθμος δοκιμάστηκε τόσο στην επίλυση μεγάλης κλίμακας προβλημάτων βελτιστοποίησης χωρίς περιορισμούς, όσο και στην εκπαίδευση νευρωνικών δικτύων και κατηγοριοποίησης πολλών βιοϊατρικών δεδομένων.

Όπως είδαμε από όλα τα πειραματικά αποτελέσματα, ο αλγόριθμος έχει μεγάλη αποδοτικότητα, όπως και πολύ καλή ταχύτητα επίλυσης των προβλημάτων. Σε αυτό συμβάλει το γεγονός ότι ο αλγόριθμος CMBFGS εκμεταλλεύεται πληροφορία από την ιδιοσύνθεση του προσεγγιστικού εσσιανού πίνακα της αντικειμενικής συνάρτησης. Την πληροφορία αυτή την εισάγει μέσω μιας κατάλληλης αναζήτησης, είτε γραμμικής ή καμπυλόγραμμης, ανάλογα με το είδος της πληροφορίας στην εκάστοτε επανάληψη.

Ο αλγόριθμος είναι σχεδιασμένος με τέτοιο τρόπο, ώστε το ποσό της μνήμης που απαιτείται για τον υπολογισμό τόσο των μικρότερων ιδιοτιμών και των αντίστοιχων ιδιοδιανύσματος να είναι αμελητέο. Συγκεκριμένα, οι θέσεις μνήμης που απαιτούνται είναι $6n + c$, όπου n ο αριθμός των μεταβλητών και c το πλήθος των τοπικών μεταβλητών (αμελητέο). Κατά συνέπεια, ο αλγόριθμος έχει χωρική πολυπλοκότητα $\Theta(n)$, με αποτέλεσμα πολύ μεγάλα προβλήματα να μπορούν να επιλυθούν γρήγορα και αποτελεσματικά.

Βιβλιογραφία

- [1] U. Alon, N. Barkai, D.A. Notterman, K. Gishdagger, S. Ybarra-dagger, D. Mackdagger, and A.J. Levine, “Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays”, In the Proceedings of the *Proceedings of the National Academy of Sciences of the United States of America*, volume 96, 6745–6750, 1999.
- [2] A.D. Anastasiadis, G.D. Magoulas, and M.N. Vrahatis, “New globally convergent training scheme based on the resilient propagation algorithm”, *Neurocomputing*, volume 64, 253–270, 2005.
- [3] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du-Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorensen, *LAPACK User’s Guide*, SIAM, Philadelphia, PA, USA, 1994.
- [4] N. Andrei, “Scaled memoryless BFGS preconditioned conjugate gradient algorithm for unconstrained optimization”, *Optimization Methods and Software*, in press.
- [5] M.S. Apostolopoulou, D.G. Sotiropoulos, and C.A. Botsaris, “A curvilinear method for large scale optimization problems”, In the Proceedings of the *8th Hellenic European Research on Computer Mathematics and its Applications*, edited by A. Lipitakis, ISBN: 978-960-88975-3-3, LEA Publishers, 2007.
- [6] M.S. Apostolopoulou, D.G. Sotiropoulos, and C.A. Botsaris, “Ελαχιστοποίηση Τετραγωνικών Συναρτήσεων που υπόκεινται σε ένα Τετραγωνικό Περιορισμό”, In the Proceedings of the *20ο Εθνικό Συνέδριο της Ελληνικής Εταιρείας Επιχειρησιακών Ερευνών*, 827–838, 2008.

- [7] M.S. Apostolopoulou, D.G. Sotiropoulos, and C.A. Botsaris, “A curvilinear method based on minimal memory BFGS updates”, *Applied Mathematics and Computation*, volume 217, No. 2, 882–892, September 2010.
- [8] M.S. Apostolopoulou, D.G. Sotiropoulos, C.A. Botsaris, and P. Pintelas, “A nearly exact method for solving large-scale TRS”, In the Proceedings of the *7th IEEE International Conference on Industrial Informatics*, Cardif, UK, 337–342, June 2009.
- [9] M.S. Apostolopoulou, D.G. Sotiropoulos, C.A. Botsaris, and P. Pintelas, “A practical method for solving large-scale TRS”, Technical Report 09-01, Department of Mathematics, University of Patras, March 2009.
- [10] M.S. Apostolopoulou, D.G. Sotiropoulos, C.A. Botsaris, and P. Pintelas, “A practical method for solving large-scale TRS”, *Optimization Letters*, 1–21, May 2010, Published online: 14 May, 2010.
- [11] M.S. Apostolopoulou, D.G. Sotiropoulos, I.E. Livieris, and P. Pintelas, “A memoryless BFGS neural network training algorithm”, In the Proceedings of the *7th IEEE International Conference on Industrial Informatics*, Cardif, UK, 216–221, June 2009.
- [12] M.S. Apostolopoulou, D.G. Sotiropoulos, and P. Pintelas, “Solving the quadratic trust-region subproblem in a low-memory BFGS framework”, *Optimization Methods and Software*, volume 23, No. 5, 651–674, October 2008.
- [13] L. Armijo, “Minimization of functions having Lipschitz continuous first partial derivatives”, *Pacific Journal of Mathematics*, volume 16, No. 1, 1–3, 1966.
- [14] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, 1996.
- [15] J. Barzilai and J.M. Borwein, “Two point step size gradient method”, *IMA Journal of Numerical Analysis*, volume 8, 141–148, 1988.
- [16] D.G. Beer, S.L. Kardia, C.C. Huang, T.J. Giordano, A.M. Levin, D.E. Misek, L. Lin, G. Chen, T.G. Gharib, D.G. Thomas, M.L. Lizy-ness, R. Kuick, S. Hayasaka, J.M. Taylor, M.D. Iannettoni, M.B.

- Orringer, and S. Hanash, "Gene-expression profiles predict survival of patients with lung adenocarcinoma", *Nature Medicine*, volume 8, No. 8, 816–823, 2002.
- [17] E.G. Birgin and J.M. Martínez, "Structured minimal-memory inexact quasi-newton method and secant preconditioners for augmented lagrangian optimization", *Computational Optimization and Applications*, volume 39, No. 1, 1–16, 2008.
- [18] J.M. Bofill, "An conjugate-gradient algorithm with a trust region for molecular-geometry optimization", *Journal of Molecular Modeling*, volume 1, No. 1, 11–17, 1995.
- [19] P. Borwein and T. Erdélyi, *Polynomials and Polynomial Inequalities*, Springer-Verlag, New York, 1995.
- [20] C.A. Botsaris, "A class of methods for unconstrained minimization based on stable numerical integration techniques", *Journal of Mathematical Analysis and Applications*, volume 63, No. 3, 729–749, 1978.
- [21] C.A. Botsaris, "A curvilinear optimization method based upon iterative estimation of the eigensystem of the Hessian matrix", *Journal of Mathematical Analysis and Applications*, volume 63, No. 2, 396–411, 1978.
- [22] C.A. Botsaris, "Differential gradient methods", *Journal of Mathematical Analysis and Applications*, volume 63, No. 1, 177–198, 1978.
- [23] C.A. Botsaris and D.H. Jacobson, "A Newton-type curvilinear search method for optimization", *Journal of Mathematical Analysis and Applications*, volume 54, No. 1, 217–229, 1976.
- [24] C.G. Broyden, "The convergence of a class double-rank minimization algorithms", *Journal of the Institute of Mathematics and its Applications*, volume 6, 76–90, 1970.
- [25] A.R. Conn, N.I.M. Gould, and Ph.L. Toint, *Trust-Region Methods*, MPS/SIAM Series on Optimization, SIAM, Philadelphia, PA, USA, 2000.
- [26] G.B. Dantzig, *Linear Programming and extensions*, Princeton University Press, Princeton, New Jersey, 1963.

- [27] W.C. Davidon, "Variable metric method for minimization", revised ANL-5990, Argonne National Laboratory, Argonne, 1959.
- [28] W.C. Davidon, "Variable metric method for minimization", *SIAM Journal on Optimization*, volume 1, 1-17, 1991.
- [29] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall, 1983.
- [30] E. Dolan and J.J. Moré, "Benchmarking optimization software with performance profiles", *Mathematical Programming*, volume 91, 201-213, 2002.
- [31] G. Fasano and M. Roma, "Iterative computation of negative curvature directions in large scale optimization", *Computational Optimization and Applications*, volume 38, No. 1, 81-104, 2007.
- [32] W.M. Faucette, "A geometric interpretation of the solution of the general quartic polynomial", *American Mathematical Monthly*, volume 103, No. 1, 51-57, January 1996.
- [33] L. Ferreira-Mendonça, V.L.R. Lopes, and J.M. Martínez, "Quasi-newton acceleration for equality-constrained minimization", *Computational Optimization and Applications*.
- [34] M.C. Ferris, S. Lucidi, and M. Roma, "Nonmonotone curvilinear line search methods for unconstrained optimization", *Computational Optimization and Applications*, volume 6, 117-136, 1996.
- [35] A.V. Fiacco and G.P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques*, Wiley, New York, 1968.
- [36] R. Fletcher, "A new approach to variable metric algorithms", *Computer Journal*, volume 13, 317-322, 1970.
- [37] R. Fletcher, *Practical Methods of Optimization*, John Wiley and Sons, New York, 2nd edition, 1987.
- [38] R. Fletcher and M.J.D. Powell, "A rapid convergent descent method for minimization", *Computer Journal*, volume 7, 163-168, 1963.
- [39] D.M. Gay, "Computing optimal locally constrained steps", *SIAM Journal on Scientific and Statistical Computing*, volume 2, No. 2, 186-197, 1981.

- [40] J.C. Gilbert and X. Jonsson, “LIBOPT - an environment for testing solvers on heterogeneous collections of problems - version 1.0”, *CoRR*, volume abs/cs/0703025, 2007.
- [41] D. Goldfarb, “A family of variable metric methods derived by variation mean”, *Mathematics of Computation*, volume 23, 23-26, 1970.
- [42] D. Goldfarb, “Curvilinear path steplength algorithms for minimization which use directions of negative curvature”, *Mathematical Programming*, volume 18, 31-40, 1980.
- [43] S.M. Goldfeld, R.E. Quandt, and H.F. Trotter, “Maximization by quadratic hill-climbing”, *Econometrica*, volume 34, No. 2, 541-551, 1966.
- [44] A.A. Goldstein, “On steepest descent method”, *SIAM Journal on Control*, volume 3, 147-151, 1965.
- [45] N.I.M. Gould, S. Lucidi, M. Roma, and Ph.L. Toint, “Solving the trust-region subproblem using the Lanczos method”, *SIAM Journal on Optimization*, volume 9, No. 2, 504-525, 1999.
- [46] N. Gould, S. Lucidi, M. Roma, and Ph.L. Toint, “Exploiting negative curvature directions in linesearch methods for unconstrained optimization”, *Optimization Methods and Software*, volume 14, 75-98, 2000.
- [47] N. Gould, D. Orban, and Ph.L. Toint, “CUTEr (and SifDec): a constrained and unconstrained testing environment, revised”, *ACM Transactions on Mathematical Software*, volume 29, No. 4, 373-394, 2003.
- [48] W.W. Hager, “Minimizing a quadratic over a sphere”, *SIAM Journal on Optimization*, volume 12, No. 1, 188-208, 2001.
- [49] W.W. Hager and Y. Krylyuk, “Graph partitioning and continuous quadratic programming”, *SIAM Journal on Discrete Mathematics*, volume 12, No. 4, 500-523, 1999.
- [50] P. Horton and K. Nakai, “Better prediction of protein cellular localization sites with the k Nearest Neighbors classifier”, In the Proceedings of the *Intelligent Systems in Molecular Biology*, 368-383, 1997.

- [51] HSL 2007, “A collection of Fortran codes for large scale scientific computation”, 2007, Available from: www.hsl.rl.ac.uk.
- [52] I.C.F. Ipsen, “Computing an eigenvector with inverse iteration”, *SIAM Review*, volume 39, No. 2, 254–291, June 1997.
- [53] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection”, In the Proceedings of the *International Joint Conference on Artificial Intelligence*, AAAI Press and MIT Press, 223–228, 1995.
- [54] K. Levenberg, “A method for the solution of certain problems in least squares”, *Quarterly of Applied Mathematics*, volume 2, 164–168, 1944.
- [55] I.S. Lim, P.H. Ciechomski, S. Sarni, and D. Thalmann, “Planar arrangement of high-dimensional biomedical data sets by isomap coordinates”, In the Proceedings of the *16th IEEE Symposium on Computer-Based Medical Systems*, 50–55, 2003.
- [56] D.C. Liu and J. Nocedal, “On the limited memory BFGS method for large scale optimization”, *Mathematical Programming*, volume 45, No. 3, 503–528, 1989.
- [57] I.E. Livieris, M.S. Apostolopoulou, D.G. Sotiropoulos, S.A. Sioutas, and P. Pintelas, “Classification of large biomedical data using ANNs based on BFGS method”, In the Proceedings of the *13th Panhellenic Conference on Informatics*, ISBN: 978-0-7695-3788-7, Corfu, GR, 87–91, September 2009.
- [58] S. Lucidi, F. Rochetich, and M. Roma, “Curvilinear stabilization techniques for truncated Newton methods in large scale unconstrained optimization”, *SIAM Journal on Optimization*, volume 8, No. 4, 916–939, 1998.
- [59] S. Lucidi and M. Roma, “Numerical experiences with new truncated newton methods in large scale unconstrained optimization”, *Computation Optimization and Applications*, volume 7, 71–87, 1997.
- [60] D.W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters”, *SIAM Journal on Applied Mathematics*, volume 11, No. 2, 431–441, 1963.

- [61] G.P. McCormick, “A modification of Armijo’s step-size rule for negative curvature”, *Mathematical Programming*, volume 13, 111–115, 1977.
- [62] J.J. Moré, B.S. Garbow, and K.E. Hillstom, “Testing unconstrained optimization software”, *ACM Transactions on Mathematical Software*, volume 7, No. 1, 17–41, 1981.
- [63] J. Moré and D. Sorensen, “On the use of directions of negative curvature in a modified Newton method”, *Mathematical Programming*, volume 16, 1–20, 1979.
- [64] J.J. Moré and D.C. Sorensen, “Computing a trust region step”, *SIAM Journal on Scientific and Statistical Computing*, volume 4, No. 3, 553–572, 1983.
- [65] P.M. Murphy and D.W. Aha, “UCI repository of machine learning databases”, Irvine, CA: University of California, Department of Information and Computer Science, 1994.
- [66] D. Nguyen and B. Widrow, “Improving the learning speed of 2-layer neural network by choosing initial values of adaptive weights”, *Biological Cybernetics*, volume 59, 71–113, 1990.
- [67] J. Nocedal, “Updating quasi-Newton matrices with limited storage”, *Mathematics of Computation*, volume 35, No. 151, 773–782, 1980.
- [68] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer, New York, 1999.
- [69] J. Nocedal and Y. Yuan, “Combining trust region and line search techniques”, In *Advances in Nonlinear Programming*, edited by Y. Yuan, Kluwer, Dordrecht, The Netherlands, 153–175, 1998.
- [70] A. Olivares, J.M. Moguerza, and F.J. Prieto, “Nonconvex optimization using negative curvature within a modified linesearch”, *European Journal of Operational Research*, volume 189, No. 3, 706–722, September 2008.
- [71] E. Polak, *Optimization*, Springer-Verlag, New York, 1997.
- [72] S.L. Pomeroy, P. P. Tamayo, M. Gaasenbeek, L.M. Sturla, M. Angelo, M.E. McLaughlin, J.Y. Kim, L.C. Goumnerova, P.M. Black,

- C. Lau, J.C. Allen, D. Zagzag, J.M. Olson, T. Curran, C. Wetmore, J.A. Biegel, T. Poggio, S. Mukherjee, R. Rifkin, A. Califano, G. Stolovitzky, D.N. Louis, J.P. Mesirov, E.S. Lander, and T.R. Golub, "Prediction of central nervous system embryonal tumour outcome based on gene expression", *Nature*, volume 415, 436–442, 2002.
- [73] M.J.D. Powell, "A new algorithm for unconstrained optimization", In *Nonlinear Programming*, edited by J. Rosen, O. Mangasarian, and K. Ritter, Academic Press, New York, NY, 31–66, 1970.
- [74] M.J.D. Powell, "Convergence properties of a class of minimization algorithms", In *Nonlinear Programming 2*, edited by O. Mangasarian, R. Meyer, and S. Robinson, Academic Press, New York, 1–27, 1975.
- [75] L. Prechelt, "PROBEN1-A set of benchmarks and benchmarking rules for neural network training algorithms", Technical Report 21/94, Fakultt fr Informatik, University of Karlsruhe, 1994.
- [76] C. Reinsch, "Smoothing by spline functions", *Numerische Mathematik*, volume 10, 177–183, 1967.
- [77] F. Rendl and H. Wolkowicz, "A semidefinite framework for trust region subproblems with applications to large scale minimization", *Mathematical Programming*, volume 77, No. 2, 273–299, 1997.
- [78] M. Rojas, S.A. Santos, and D.C. Sorensen, "A new matrix-free algorithm for the large-scale trust-region subproblem", *SIAM Journal on Optimization*, volume 11, No. 3, 611–646, 2000.
- [79] M. Rojas, S.A. Santos, and D.C. Sorensen, "Algorithm 873: LSTRS: MATLAB software for large-scale trust-region subproblems and regularization", *ACM Transactions on Mathematical Software*, volume 34, No. 2, 1–28, 2008.
- [80] D.F. Shanno, "Conditioning of quasi-Newton method for function minimization", *Mathematics of Computation*, volume 24, 27–30, 1970.
- [81] M.A. Shipp, K.N. Ross, P. Tamayo, A.P. Weng, J.L. Kutok, R.C. Aguiar, M. Gaasenbeek, M. Angelo, M. Reich, G.S. Pinkus, T.S. Ray, M.A. Koval, K.W. Last, A. Norton, T.A. Lister, J. Mesirov, D.S. Neuberg, E.S. Lander, J.C. Aster, and T.R. Golub, "Diffuse large b-cell

- lymphoma outcome prediction by gene-expression profiling and supervised machine learning”, *Nature Medicine*, volume 8, No. 1, 68–74, 2002.
- [82] D. Singh, P.G. Febbo, K. Ross, D.G. Jackson, J. Manola, C. Ladd, P. Tamayo, A.A. Renshaw, A.V. D’Amico, J.P. Richie, E.S. Lander, M. Loda, P.W. Kantoff, T.R. Golub, and W.R. Sellers, “Gene expression correlates of clinical prostate cancer behavior”, *Cancer Cell*, volume 1, 203–209, 2002.
- [83] T. Steihaug, “The conjugate gradient method and trust regions in large scale optimization”, *SIAM Journal on Numerical Analysis*, volume 20, No. 3, 626–637, 1983.
- [84] W. Sun and Y.X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, 2006.
- [85] Ph.L. Toint, “Towards an efficient sparsity exploiting newton method for minimization”, In *Sparse Matrices and Their Uses*, edited by I. Duff, Academic Press, Inc., New York, NY, 57–87, 1981.
- [86] L.J.P. Van der Maaten and H.J. Van den Herik, “Dimensionality reduction: A comparative review”, Technical Report TiCC-TR 2009-005, Tilburg University, 2009.
- [87] H. Wielandt, *Mathematische Werke/Mathematical works*, Walter de Gruyter & Co., Berlin, 1996.
- [88] J.H. Wilkinson, *The algebraic eigenvalue problem*, Oxford University Press, London, 1965.
- [89] P. Wolfe, “Convergence conditions for ascent methods”, *SIAM Review*, volume 11, 226–235, 1969.
- [90] M. Zhang and P. Wong, “Genetic programming for medical classification: a program simplification approach”, *Genetic Programming and Evolvable Machines*, volume 9, 229–255, 2008.

Ευρετήριο

1BFGS ενημέρωση ελάχιστης μνή- μης, 164, 177, 223	39, 40, 42
2BFGS ενημέρωση μικρής μνή- μης, 178, 223	ανάπτυγμα Taylor, 52, 262
BFGS ενημέρωση, 41, 80	δύσκολη περίπτωση, 62, 72, 154, 170, 171, 187, 224
BFGS ενημέρωση ελάχιστης μνή- μης, 81, 107, 161, 174, 209, 249, 255, 262, 265	ελάχιστο πολυώνυμο, 13, 93, 94, 106, 134, 146, 147
BFGS ενημέρωση μικρής μνήμης, 125, 154, 229	εοσιανός πίνακας, 18, 53, 126, 243, 256, 262
Broyden γενική ενημέρωση πρώ- της τάξης, 36, 38	εξίσωση τέμνουσας, 35, 37, 39, 41, 44, 97, 116
DFP ενημέρωση, 39, 80	γενική PSB ενημέρωση, 36, 37, 79
DFP ενημέρωση ελάχιστης μνή- μης, 82, 112, 162, 175	γενική PSB ενημέρωση ελάχιστης μνήμης, 81, 119, 156, 171
L-BFGS μέθοδος, 44, 266, 273	γενικευμένος αντίστροφος, 4, 170
PSB ενημέρωση, 42	γραμμικώς ανεξάρτητα διανύσμα- τα, 8, 87, 90, 96, 108, 120, 132, 141, 164, 246, 249
PSB ενημέρωση ελάχιστης μνή- μης, 82, 163, 176	γραμμικώς εξαρτημένα διανύσμα- τα, 8, 94, 95, 132, 135, 141, 162, 163, 165, 166, 172, 246, 247, 249
Quasi-Newton ενημέρωση ελάχι- στης μνήμης, 79, 80, 86, 92, 95, 154, 160, 187, 209	ιδιοδιάνυσμα, 11, 52, 79, 95, 108, 113, 120, 135, 148, 170, 172–177, 249, 250
Quasi-Newton ενημερώσεις, 79	
Quasi-Newton εξίσωση, 35, 36	
Quasi-Newton μέθοδοι, 44	
SR1 ενημέρωση, 38, 80	
SR1 ενημέρωση ελάχιστης μνή- μης, 81, 160, 173	
Sherman-Morrison-Woodbury, 9,	

- ιδιοτιμή, 11, 53, 79, 86–88, 90, 95, 126, 130, 132, 135, 154, 170, 172, 244, 249, 250
- θεώρημα Caley-Hamilton, 93, 94, 106, 134
- θεώρημα μέσης τιμής, 18, 259
- θεώρημα συνδιαπλοκής ιδιοτιμών, 14, 87–89, 130, 139
- κατεύθυνση Newton, 32
- κατεύθυνση αρνητικής καμπυλότητας, 20, 50, 52, 170, 243–245, 248, 249, 252, 254, 256
- κατεύθυνση μείωσης, 52, 244, 255
- κρίσιμο σημείο δεύτερης τάξης, 54
- μέθοδος Lanczos, 51, 154, 155, 189, 210, 225, 243
- μέθοδος Newton, 31, 67, 155, 156, 159–163, 165, 181
- μέθοδος αντίστροφης δύναμης, 14, 79, 97, 98, 108, 113, 120, 135, 136, 148, 155
- μέθοδος γραμμικής αναζήτησης, 48, 49, 244, 256, 263, 266
- μέθοδος περιοχής εμπιστοσύνης, 181, 229
- μέθοδος συζυγών κλίσεων, 51
- ολική σύγκλιση, 181
- ορίζουσα, 13, 83, 104, 108, 112
- πίνακας άνω τριγωνικός, 5
- πίνακας αόριστος, 5
- πίνακας αρνητικά ημιορισμένος, 5
- πίνακας αρνητικά ορισμένος, 5
- πίνακας θετικά ημιορισμένος, 5, 50
- πίνακας θετικά ημιορισμένος, 50
- πίνακας θετικά ορισμένος, 5, 87, 88, 139, 153, 181, 244, 256
- πίνακας κάτω τριγωνικός, 5
- πίνακας μη θετικά ορισμένος, 170
- πίνακας ορθογώνιος, 5
- πίνακας συμμετρικός, 5, 153
- πίνακας τριδιαγώνιος, 5
- παραγοντοποίηση Cholesky, 10, 69, 154, 155, 189
- πολλαπλασιαστής Lagrange, 26, 154, 155, 159, 160, 165, 170, 180
- ποσότητα Rayleigh, 12, 126
- χαρακτηριστικό πολυώνυμο, 13, 79, 83, 86, 88–90, 119, 130, 133, 140, 142, 145, 147, 171, 173–178
- ρυθμός σύγκλισης, 8
- συνάρτηση δύο φορές συνεχώς διαφορίσιμη, 17, 52
- συνάρτηση διαφορίσιμη, 17
- συνάρτηση παραγωγίσιμη, 17
- συνάρτηση συνεχώς διαφορίσιμη, 17, 262
- συνθήκες Armijo, 47, 256
- συνθήκες Goldstein, 47

- συνθήκες Karysh-Kuhn-Tucker,
27, 154
- συνθήκες Wolfe, 47, 48, 55
- συνθήκες βελτιστότητας, 3, 21, 154
- συνθήκη καμπυλότητας, 35
- τεχνικές ολικής σύγκλισης, 31
- τετραγωνική σύγκλιση, 8
- τοπικές μέθοδοι, 31
- τυπική περίπτωση, 62, 63, 154,
155, 187, 211, 224
- υποπρόβλημα περιοχής εμπιστο-
σύνης, 57, 60, 153, 187

Κατάλογος Δημοσιεύσεων

Δημοσιεύσεις σε διεθνή περιοδικά

1. M.S. Apostolopoulou, D.G. Sotiropoulos, C. A. Botsaris and P. Pintelas. *A practical method for solving large-scale TRS*. Optimization Letters, published online: 14 May 2010. (Impact Factor: 0.926)
2. M.S. Apostolopoulou, D.G. Sotiropoulos and C. A. Botsaris. *A curvilinear method based on minimal memory BFGS updates*. Applied Mathematics and Computation, Volume 217, No. 2, pp. 882-892, September 2010. (Impact Factor: 1.124)
3. M.S. Apostolopoulou, D.G. Sotiropoulos and P. Pintelas. *Solving the quadratic trust-region subproblem in a low-memory BFGS framework*, Optimization Methods and Software, Volume 23, No. 5, pp. 651-674, October 2008 (Impact Factor: 0.969)

Δημοσιεύσεις σε πρακτικά συνεδρίων

1. M.S. Apostolopoulou, D.G. Sotiropoulos, C.A. Botsaris and P. Pintelas. *A nearly exact method for solving large-scale TRS*. In: Proceedings of the 7th IEEE International Conference on Industrial Informatics (INDIN 2009), 24-26th June 2009, Cardiff UK, pp. 337-342, ISBN 978-1-4244-3759-7. Published by IEEE Computer Society, 2009
2. M.S. Apostolopoulou, D.G. Sotiropoulos, I.E. Livieris and P. Pintelas. *A memoryless BFGS neural network training algorithm*. In: Proceedings of the 7th IEEE International Conference on Industrial Informatics (INDIN 2009), 24-26th June 2009, Cardiff UK,

- pp. 216-221, ISBN 978-1-4244-3759-7. Published by IEEE Computer Society, 2009.
3. I.E. Livieris, M.S. Apostolopoulou, D.G. Sotiropoulos, S.A. Siou-tas and P. Pintelas. *Classification of Large Biomedical Data using ANNs based on BFGS method*. In: Proceedings of the 13th Panhellenic Conference on Informatics (PCI 2009), 10-12th September 2009, Corfu, Greece, pp. 87-91, ISBN 978-0-7695-3788-7. Published by IEEE Computer Society, 2009
 4. M.S. Apostolopoulou, D.G. Sotiropoulos and C.A. Botsaris. *Ε-λαχιστοποίηση Τετραγωνικών Συναρτήσεων που υπόκεινται σε ένα Τετραγωνικό Περιορισμό*. Πρακτικά του 20ου Εθνικού Συνεδρίου της Ελληνικής Εταιρείας Επιχειρησιακών Ερευνών (20ΕΕΕΕ 2008), pp.827-838, 2008.
 5. M.S. Apostolopoulou, D.G. Sotiropoulos and C.A. Botsaris. *A Curvilinear Method for Large Scale Optimization Problems*. In: Proceedings of the 8th Hellenic European Research on Computer Mathematics and its Applications (HERCMA 2007), A.E. Lipitakis ed., ISBN: 978-960-88975-3-3, LEA Publishers, 2007.

Οι δημοσιεύσεις της υποψήφιας είναι διαθέσιμες σε ηλεκτρονική μορφή στην ιστοσελίδα :

<http://www.math.upatras.gr/~msa/>

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ,
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ,
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ