ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΜΑΘΗΜΑΤΙΚΑ ΤΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΤΩΝ ΑΠΟΦΑΣΕΩΝ»

Προσεγγίσεις για μοντέλα Γραμμικού Στοχαστικού Προγραμματισμού

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνα Φ. Μπασέτα

Επιβλέπων : Νικόλαος Τσάντας

Καθηγητής Πανεπιστημίου Πατρών

Πάτρα, Απρίλιος 2015

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΜΑΘΗΜΑΤΙΚΑ ΤΩΝ ΥΠΟΛΟΓΙΣΤΩΝ ΚΑΙ ΤΩΝ ΑΠΟΦΑΣΕΩΝ»

Προσεγγίσεις για μοντέλα Γραμμικού Στοχαστικού Προγραμματισμού

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κωνσταντίνα Φ. Μπασέτα

Επιβλέπων : Νικόλαος Τσάντας

Καθηγητής Πανεπιστημίου Πατρών

Εγκρίθηκε από την τριμελής εξεταστική επιτροπή την $30_{η}$ Απριλίου 2015 :

.....	.....	.....
Ν. Τσάντας	Κ. Πετρόπουλος	Ι. Δημητρίου
Καθηγητής Πανεπιστημίου Πατρών	Επίκουρος Καθηγητής Πανεπιστημίου Πατρών	Λέκτορας Πανεπιστημίου Πατρών

Πάτρα, Απρίλιος 2015

……

Κωνσταντίνα Φ. Μπασέτα

Πτυχιούχος Μαθηματικός Πανεπιστημίου Πατρών

Copyright © Κωνσταντίνα Φ. Μπασέτα, 2015
Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέ-ρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγρα-φέα και δεν πρέπει να ερμηνευτεί ότι εκφράζουν τις επίσημες θέσεις του Πανεπιστημίου Πατρών.

ΠΕΡΙΛΗΨΗ

Πολλά είναι τα προβλήματα που καλούμαστε να αντιμετωπίσουμε στην καθημερινότητά μας, και που χρίζουν την ανάγκη προσδιορισμού αυτών, μέσω του Γραμμικού Στοχαστικού Προγραμματισμού. Βασικό εργαλείο των προβλημάτων του Γραμμικού Στοχαστικού Προγραμματισμού που χρησιμοποιείται για την υπολογιστική τους επίλυση είναι οι μέθοδοι του Γραμμικού και του Μη Γραμμικού Προγραμματισμού.

Στο 1^{ο} κεφάλαιο της παρούσας Διπλωματικής Εργασίας, υπενθυμίζονται οι βασικές ιδιότητες και μέθοδοι επίλυσης των Γραμμικών και Μη Γραμμικών προβλημάτων, όπως αυτές χρησιμοποιούνται από τον Στοχαστικό Προγραμματισμό.

Στο 2^{ο} κεφάλαιο, παρουσιάζεται μια σειρά από Γραμμικά μοντέλα Στοχαστικού Γραμμικού Προγραμματισμού ενός σταδίου συζητώντας τις θεωρητικές τους ιδιότητες, σχετικές με την υπολογιστική τους δυνατότητα, μία από τις οποίες αποτελεί η κυρτότητά τους.

Στο 3^{ο}, και τελευταίο κεφάλαιο, ακολουθείται μια αντίστοιχη παρουσίαση των Γραμμικών Στοχαστικών μοντέλων πολλαπλών σταδίων, τονίζοντας τις ιδιότητες αυτές που επιτρέπουν την κατασκευή προσεγγιστικών μεθόδων επίλυσης.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ

Στοχαστικός γραμμικός προγραμματισμός, γραμμικό μοντέλο, μέθοδοι επίλυσης, κυρτότητα.

ABSTRACT

There are various special problem formulations to be dealt with in our daily life, and our conclusion is that a basic toolkit of Linear and Nonlinear Programming methods cannot be waived if we want to deal with the computational solution of Stochastic Linear Programming problems.

In chapter 1, basic properties of Linear Problems and Non Linear Problems, as well as their solution methods, are reminded, as they are used in the Stochastic Linear Programming.

In chapter 2, various Single-stage Stochastic Linear Programming (SLP) models are presented and their theoretical properties are discussed, which are relevant for their computational tractability, as convexity statements.

Conclusions are presented in chapter 3, followed by an analogous discussion of Multi-stage SLP models, focussed, among others, on properties allowing for the construction of particular approximation methods for computing solutions.

KEY WORDS

Stochastic Linear Programming, linear model, solution methods, convexity.

Ευχαριστίες

Πρώτα απ' όλα, θέλω να ευχαριστήσω τον επιβλέποντα της διπλωματικής εργασίας μου, Καθηγητή κ. Ν. Τσάντα, για την πολύτιμη βοήθεια και καθοδήγησή του κατά τη διάρκεια της προετοιμασίας της εκπόνησης της εργασίας μου. Επίσης, είμαι ευγνώμων στα υπόλοιπα μέλη της εξεταστικής επιτροπής της διπλωματικής εργασίας μου, κ. Κ. Πετρόπουλο και κ. Ι. Δημητρίου για την προσεκτική ανάγνωση της εργασίας μου και για τις πολύτιμες υποδείξεις τους.

Σε αυτό το σημείο θέλω να αναφέρω ανθρώπους, εκτός του στενού ακαδημαϊκού περιβάλλοντος, που υπήρξαν σημαντικοί πόλοι στη ζωή μου. Ευχαριστώ πολύ τις καλές μου φίλες για την σωστή καθοδήγησή τους κατά την τελική διεκπεραίωση της εργασίας. Επίσης, ευχαριστώ την οικογένειά μου για την αμέριστη συμπαράσταση του χρόνου των σπουδών μου. Πιο συγκεκριμένα, ευχαριστώ την αδερφή μου Ασπασία για την κατανόησή της, ιδιαίτερα κατά τη διάρκεια των τελευταίων μηνών της προσπάθειάς μου. Πάνω απ' όλα όμως , είμαι ευγνώμων στους γονείς μου, Φώτιο και Ολυμπία Μπασέτα για την ολόψυχη αγάπη και υποστήριξή τους όλα αυτά τα χρόνια. Αφιερώνω αυτή την εργασία στον πατέρα μου.

Κωνσταντίνα Μπασέτα

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ 1
ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ 3
ΠΕΡΙΛΗΨΗ 5
ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ 5
ABSTRACT 6
KEY WORDS 6
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ΣΧΗΜΑΤΩΝ 9
ΚΕΦΑΛΑΙΟ 1^{Ο} «ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ» 11
1.1. Στοχαστικά Προγράμματα 11
1.2. Ιδιότητες Γραμμικού Προγραμματισμού 14
1.2.1. Αλγεβρικές έννοιες και ιδιότητες 14
1.2.2. Γεωμετρική ερμηνεία 16
1.2.3. Η Δυική μέθοδος 20
1.2.4. Η μέθοδος Simplex 23
1.2.5. Η Δυική μέθοδος Simplex 24
1.2.6. Δυική Ανάλυση 24
1.2.7. Αλυσιδωτή Ανάλυση 27
1.2.8. Ανάλυση Κανονικότητας 34
1.2.9. Μέθοδος Εσωτερικού Σημείου 36
1.3. Ιδιότητες Μη Γραμμικού Προγραμματισμού 40
1.3.1. Συνθήκες Βελτιστοποίησης 42
1.3.2. Μέθοδοι Επίλυσης 45
ΚΕΦΑΛΑΙΟ 2^{ο} «ΓΡΑΜΜΙΚΑ ΣΤΟΧΑΣΤΙΚΑ ΠΡΟΓΡΑΜΜΑΤΑ ΕΝΟΣ ΣΤΑΔΙΟΥ» 50
2.1. Εισαγωγή 50
2.2. Στοχαστικά μοντέλα βελτιστοποίησης με συναρτήσεις πιθανότητας 62
2.2.1 Βασικές ιδιότητες 65
2.2.2 Μελέτη Στοχαστικών Μοντέλων Προγραμματισμού 68
2.2.3. Ανεξαρτησία 79
2.2.4 Από κοινού περιορισμοί 80
2.3 Συναρτήσεις Ποσοστημορίων - Επίπεδο Σημαντικότητας 85
2.4. Μοντέλα βασισμένα στην εκτίμηση 86

ΚΕΦΑΛΑΙΟ 3^{ο} «ΠΟΛΥΔΙΑΣΤΑΤΑ ΜΟΝΤΕΛΑ ΓΡΑΜΜΙΚΟΥ ΣΤΟΧΑΣΤΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ» 88
3.1. Γενικευμένα προβλήματα γραμμικού προγραμματισμού με κατανομή πόρων 88
3.2. Προβλήματα Γραμμικού Στοχαστικού Προγραμματισμού δύο σταδίων 91
3.2.1. Ολοκληρωμένη Σταθερή Κατανομή πόρων 92
3.3 Πολυμεταβλητά Προβλήματα Γραμμικού Στοχαστικού Προγραμματισμού 99
ΒΙΒΛΙΟΓΡΑΦΙΑ.....101

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ΣΧΗΜΑΤΩΝ

Εικόνα 1: Δομή Δυικής χαλαρότητας 13
Εικόνα 2: Κυρτό πολύεδρο 17
Εικόνα 3 : Simplex Tableau 21
Εικόνα 4 : Cutting plane method 47
Εικόνα 5 : Slater points 48
Εικόνα 6 : Τομή Επιπέδων 48
Εικόνα 7 : Scenario tree για προβλήματα 4 σταδίων στοχαστικής διαδικασίας 99
Εικόνα 8: Basic scenario tree: First split. 100

ΚΕΦΑΛΑΙΟ 1^{Ο} «ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ»

Τα Γραμμικά Προγράμματα αποτελούν μοντέλα πολλών πρακτικών προβλημάτων με αποτέλεσμα να αποτελούν αντικείμενο μελέτης για πολλά χρόνια και για πολλούς τομείς.

Κάθε γραμμικό μοντέλο είναι της μορφής :

υπό τους περιορισμούς $\left.\begin{array}{c} \min c^{\mathrm{T}} x \\ Ax = b \\ x \geq 0 \end{array}\right\}$

όπου c, x , A πίνακες-διανύσματα.

Σκοπός είναι να βρεθεί ένας βέλτιστος συνδυασμός των μεταβλητών απόφασης $x = \{x_1, x_2, \ldots, x_n\}$ ικανοποιώντας τους παραπάνω περιορισμούς.

Ωστόσο, στην πραγματική ζωή οι συναρτήσεις που καθορίζουν το πρόβλημα δεν είναι πάντα γραμμικές ως προς τις μεταβλητές απόφασης, όπως είναι η ζήτηση, η παραγωγή, οι τιμές κ. ά οι οποίες ως τυχαίες μεταβλητές δεν παίρνουν δεδομένες τιμές. Αυτές μπορούν να καθοριστούν από συναρτήσεις πιθανότητας της μορφής $G(\mathrm{x}) := \mathrm{P}(\mathrm{Tx} \geq \mathrm{h})$ η οποία είναι φανερό πως είναι μη γραμμική ως προς το x.

Η μορφή αυτή του προβλήματος με την παρουσία τυχαίων μεταβλητών που ακολουθούν μια συνεχή συνάρτηση κατανομής είναι γνωστή ως πρόβλημα Γραμμικού Στοχαστικού Προγραμματισμού.

1.1. Στοχαστικά Προγράμματα

Για μια πιο γενική εικόνα του Στοχαστικού Προγραμματισμού θα μπορούσαμε να πούμε πως αποτελεί μια προσέγγιση για την μοντελοποίηση των προβλημάτων βελτιστοποίησης που εμπεριέχουν αβεβαιότητα. Σε αντίθεση με τα ντετερμινιστικά προβλήματα βελτιστοποίησης που διατυπώνονται με γνωστές παραμέτρους, τα

προβλήματα που συναντάμε στην καθημερινή ζωή περιλαμβάνουν παραμέτρους οι οποίες είναι άγνωστες κατά την λήψη μιας απόφασης. Για συγκεκριμένο σύνολο δυνατών τιμών τους, όμως, μπορεί να αναζητηθεί μια λύση που θα είναι εφικτή για όλες τις πιθανές παραμέτρους βελτιστοποιώντας μια δοσμένη αντικειμενική συνάρτηση.

Τα μοντέλα στοχαστικού προγραμματισμού διακρίνονται από το γεγονός ότι οι κατανομές πιθανότητας που διέπουν τα στοιχεία είναι γνωστές ή εκτιμώμενες. Τέτοια μοντέλα εφαρμόζονται συχνά σε περιπτώσεις όπου είτε οι αποφάσεις λαμβάνονται κατ' επανάληψη κάτω από τις ίδιες συνθήκες με στόχο την λήψη μιας απόφασης η οποία θα έχει καλή απόδοση κατά μέσο όρο, είτε όταν η λήψη γίνεται ακαριαία.

Τα πιο διαδεδομένα και θεωρητικά ανεπτυγμένα μοντέλα στοχαστικού προγραμματισμού είναι τα γραμμικά προγράμματα δύο σταδίων. Σε αυτά ο λήπτης, σε πρώτη φάση, προχωρά σε κάποιες ενέργειες, μετά τις οποίες συμβαίνει ένα τυχαίο γεγονός που επηρεάζει το αποτέλεσμα της πρώτης απόφασης.

Παρά το γεγονός ότι τα γραμμικά στοχαστικά προγράμματα δύο σταδίων θεωρούνται ως κλασικό παράδειγμα μοντελοποίησης στοχαστικού προγραμματισμού, η ισχύς τους έχει αναπτυχθεί και διευρυνθεί προκειμένου να περιλαμβάνει ένα ευρύτερο φάσμα μοντέλων και προσεγγίσεων λύσεων. Οι εφαρμογές τους εκτείνονται από την οικονομία μέχρι και την διαχείριση έργων.

Γενίκευση του μοντέλου δύο σταδίων είναι το μοντέλο πολλαπλών σταδίων. Στην περίπτωση αυτή κάθε στάδιο αποτελείται από μία απόφαση που ακολουθείται από μια σειρά εμφάνισης αβέβαιων παραμέτρων κατά την πάροδο του χρόνου. Εδώ ο στοχαστικός προγραμματισμός χρησιμοποιεί μαθηματικά εργαλεία, όπως την ανάλυση αποφάσεων, την βελτιστοποίηση διακριτών περιπτώσεων, την στοχαστική θεωρία ελέγχου, της διαδικασίες λήψης αποφάσεων Markov καθώς και τον δυναμικό προγραμματισμό.

Συνοψίζοντας, θα λέγαμε ότι η μέθοδος συνδυάζει τη δύναμη του μαθηματικού προγραμματισμού με προηγμένες τεχνικές πιθανότητας προς την επίλυση προβλημάτων βελτιστοποίησης που εμπεριέχουν την αβεβαιότητα. Για παράδειγμα, τα εργαλεία της κυρτής ανάλυσης και η δυαδική θεωρία μπορούν να εφαρμοστούν για να αποφέρουν σημαντικά στοιχεία και να αναπτύξουν τεχνικές επίλυσης βασισμένες στην τμηματοποίηση μεγάλων προβλημάτων για καλύτερη μελέτη.

Η προσέγγιση της λύσης των μοντέλων στοχαστικού προγραμματισμού καθορίζεται από τον τύπο των κατανομών πιθανότητας των τυχαίων μεταβλητών. Η συνηθέστερη προσέγγιση για την αντιμετώπιση της αβεβαιότητας είναι ο καθορισμός μικρού πλήθους σεναρίων για μελλοντικές καταστάσεις. Στην περίπτωση αυτή είναι εφικτό να υπολογιστεί μια λύση στο πρόβλημα στοχαστικού προγραμματισμού μέσω ενός ισοδύναμου ντετερμινιστικού γραμμικού προγράμματος.

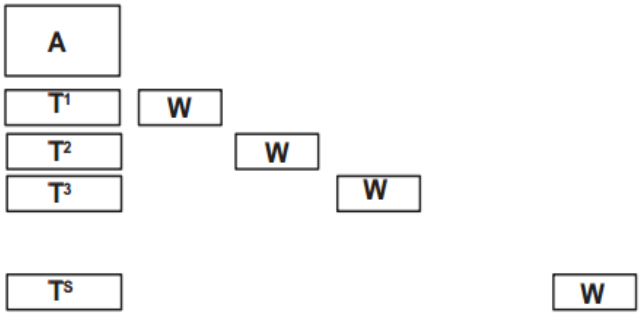
Συνεπώς, όταν οι κατανομές πιθανότητας των τυχαίων παραμέτρων είναι συνεχείς εμφανίζεται το πρόβλημα κατασκευής σεναρίων για την προσέγγιση της αβεβαιότητας. Μια προσέγγιση του προβλήματος κατασκευάζει δύο διαφορετικά ντετερμινιστικά ισοδύναμα προβλήματα, οι βέλτιστες λύσεις των οποίων αποτελούν το άνω και κάτω όριο της βέλτιστης τιμής $\hat{\gamma}$ του αρχικού προβλήματος.

Απ' την άλλη μεριά, αν οι τυχαίες μεταβλητές ακολουθούν από κοινού πεπερασμένη διακριτή κατανομή πιθανότητας $p_j, j = 1,...,S$ τότε το πρόβλημα μπορεί να επιλυθεί ως πρόβλημα Γραμμικού Προγραμματισμού, έχοντας την δομή Δυικής χαλαρότητας, όπως φαίνεται παρακάτω :

$$\min c^T x + \sum_{j=1}^{S} p_j q^T y^j$$

υ. π.

$$\begin{aligned} &\mathrm{A}x = b \\ &T^j x + W y^j = h^j,\ j = 1,...S \\ &x \geq 0 \\ &y^j \geq 0 \end{aligned}$$



Εικόνα 1: Δομή Δυικής χαλαρότητας

Παρ' όλα αυτά τα βασικά εργαλεία και οι μέθοδοι επίλυσης των γραμμικών και μη γραμμικών προγραμμάτων δεν επαρκούν για την υπολογιστική επίλυση ενός στοχαστικού γραμμικού προγράμματος.

Παρακάτω θα παρουσιαστούν κάποιες από τις ιδιότητες των μεθόδων επίλυσης του γραμμικού και μη γραμμικού μοντέλου οι οποίες αναφέρονται σε στοχαστικό γραμμικό μοντέλο.

1.2. Ιδιότητες Γραμμικού Προγραμματισμού

Για την δεδομένη μορφή του Γραμμικού Προγραμματισμού :

υ. π

$$\left.\begin{array}{l} \min c^T x \\ \mathrm{A}x = b \\ x \geq 0 \end{array}\right\}$$

όπου Α $(m \times n)-$ πίνακας και b, c διανύσματα-στήλες ίδιας διάστασης.

Τότε το πρόβλημα $Ax = b$ είναι επιλύσιμο αν και μόνο αν:

- $rank(\mathrm{A}) = \mathrm{m}$ ή
- θεωρώντας ότι οι πρώτες m_1 στήλες του $\mathrm{A}_1, \mathrm{A}_2, ...\mathrm{A}_{m_1}$ ($\tilde{A}$) είναι γραμμικώς ανεξάρτητες, κάθε γραμμικός συνδυασμός $\tilde{A}x = \tilde{b}$ έχει εφικτή λύση.

1.2.1. Αλγεβρικές έννοιες και ιδιότητες

Η επίλυση ενός Γραμμικού Προγράμματος και συγκεκριμένα η εύρεση της εφικτής εκείνης λύσης η οποία θα ικανοποιεί τους περιορισμούς του προβλήματος μπορεί να ερμηνευτεί και αλγεβρικά.

Πρόταση 1.1 Κάθε εφικτή λύση λέγεται βασική εφικτή λύση αν οι στήλες του πίνακα Α με θετικές βασικές μεταβλητές είναι γραμμικά ανεξάρτητες.

Έτσι, για κάθε εφικτή λύση ισχύει :

$\hat{x}_i > 0$ για κάθε $i \in I(\hat{x})$, $\hat{x}_j = 0$ για κάθε $j \notin I(\hat{x})$ και $\sum_{i \in I(x)} A_i \hat{x}_i = b$.

Αν ωστόσο, το σύνολο των στοιχείων του $I(\hat{x})$ είναι μικρότερο από τον βαθμό του πίνακα Α ($I(\hat{x})$<m), τότε η βασική λύση $\hat{x}$ περιέχει λιγότερο από m βασικές μεταβλητές και ονομάζεται μη εκφυλλισμένη. Οπότε, υπάρχει άλλο ένα σύνολο $I_B(\hat{x})$, με υπερσύνολο του $I(\hat{x})$, όπου $B = \left(A_i, i \in I_B(\hat{x}) \right)$ αντιστρέψιμος πίνακας.

Τότε μπορούμε να ορίσουμε τις βασικές και τις μη βασικές μεταβλητές,

$$x_k^{\{B\}} = x_{i_k}, \mathrm{i}_k \in I_B(\hat{x}) \text{ για k=1,…,m}$$

$$x_l^{\{N\}} = x_{j_l}, \mathrm{j}_l \in I_N(\hat{x}) \text{ για l=1,…,n-m.}$$

Σύμφωνα με τα παραπάνω το αρχικό σύστημα *Ax=b* αναδιαμορφώνεται και ισοδυναμεί με το

$$Bx^{\{B\}} + Nx^{\{N\}} = b$$

και το πρόβλημα παίρνει την μορφή

$$\min c^{\{B\}^T} x^{\{B\}} + c^{\{N\}^T} x^{\{N\}}$$

υ. π

$$Bx^{\{B\}} + Nx^{\{N\}} = b$$

$$x^{\{B\}} \geq 0$$

$$x^{\{N\}} \geq 0$$

Με βασική εφικτή λύση $\hat{x}$ όπου $\hat{x}^{\{N\}} = 0$ (x βασική λύση) και $\hat{x}^{\{B\}} = B^{-1}b \geq 0$ (x εφικτή λύση) που αντιστοιχεί σε αντικειμενική τιμή $\gamma_B := c^{\{B\}}B^{-1}b$.

Πρόταση 1.2: Η ύπαρξη μιας βέλτιστης τιμής $\hat{\gamma}$ συνεπάγεται την ύπαρξη βασικού εφικτού συνόλου για το οποίο ισχύει : $c^{\{\hat{B}\}}\hat{B}^{-1}b = \hat{\gamma}$.

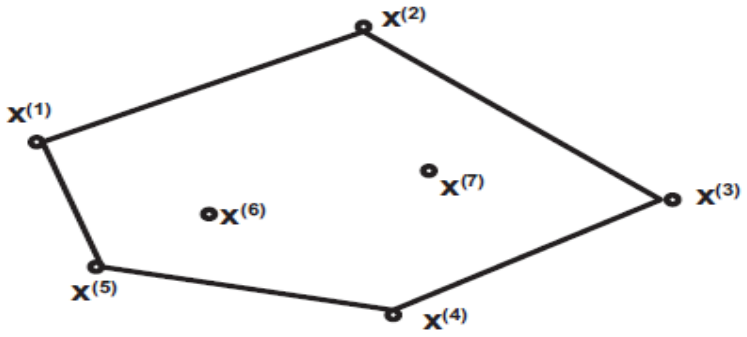
Ύπαρξη εκφυλισμένης λύσης Η βασική λύση $\hat{x}$ ονομάζεται εκφυλισμένη εάν το πλήθος των αυστηρά θετικών βασικών μεταβλητών είναι μικρότερο του m.

Κριτήριο Άριστης λύσης (Simplex) Βασικό χαρακτηριστικό της μεθόδου Simplex είναι η δυνατότητα να αναγνωρίζει αν μια βασική εφικτή λύση είναι η άριστη. Έτσι η βασική λύση του παραπάνω προβλήματος είναι η άριστη λύση του προβλήματος γραμμικού προγραμματισμού αν $d^T \leq 0$, όπου $d^T = c^{\{B\}^T}D - c^{\{N\}^T}$ και $D = B^{-1}N$.

1.2.2. Γεωμετρική ερμηνεία

Πριν προχωρήσουμε στην παρουσίαση των βασικών μεθοδολογιών για την επίλυση ενός προβλήματος γραμμικού προγραμματισμού θα ήταν χρήσιμο να μελετήσουμε την γενική εικόνα αυτών των προβλημάτων. Έτσι, παρατίθενται οι παρακάτω ορισμοί.

Ορισμός 1.1 Έστω $x^{(1)},...,x^{(r)} \in \mathbb{R}^n$, τότε οι κυρτή θήκη αυτών είναι $P = conv\{x^{(1)},...,x^{(r)}\} := \{x \mid x = \sum_{j=1}^{r}\lambda_j x^{(j)} \quad \text{με} \quad \sum_{j=1}^{r}\lambda_j = 1, \quad \lambda_j \geq 0 \quad \forall j\}$ και ονομάζεται κυρτό πολύεδρο.



Εικόνα 2: Κυρτό πολύεδρο

Ορισμός 2.2 Το σύνολο $C = \{y \mid Ay = 0,\ y \geq 0\}$ λέγεται κυρτός πολυεδρικός κώνος, ο οποίος είναι είτε τετριμμένος με $C = \{0\}$,

είτε για $y \in C$ και $y \neq 0$ προέρχεται από το σύνολο των βασικών εφικτών λύσεων του συστήματος

$$\begin{aligned} &Ay = 0 \\ &e^{T} y = 1 \quad \text{με } e^{T} = (1,\ldots,1) \\ &y \geq 0 \end{aligned}$$

Ορισμός 1.3 Για κάθε μη κενό υποσύνολο M του $\mathbb{R}^n$, το σύνολο $M^P := \{z \in \mathbb{R}^n \mid z^T x \leq 0 \forall x \in M\}$ καλείται πολικός κώνος του M.

Πρόταση 1.3 Για κάθε μη κενό υποσύνολο $M \subset \mathbb{R}^n$ ισχύει ότι ο πολικός του κώνος $M^P \subset \mathbb{R}^n$ είναι κλειστός κυρτός κώνος, συνεπώς, το $M^P \neq 0$ είναι κλειστό σύνολο έτσι ώστε για κάθε $z^{(i)} \in M^P$, i=1, 2, ισχύει $\lambda_1 z^{(1)} + \lambda_2 z^{(2)} \in M^P \ \forall \lambda_i \geq 0$.

Συγκεκριμένα, για κάθε κυρτό πολυεδρικό κώνο C ο κυρτός του κώνος C^P είναι επίσης κλειστός πολυεδρικός κώνος.

Απόδειξη : Ισχύει ότι $0 \in M^P$, οπότε ο πολικός κώνος δεν είναι κενός $M^P \neq \varnothing$ και συνεπώς είναι κυρτός κώνος.

Για $\{z^{(v)} \in M^P, v \in \mathbb{N}\}$ που συγκλίνουν στο $\hat{z}$ ισχύει ότι $z^{(v)^T} \hat{x} \leq 0, \forall v \in \mathbb{N}$ για τυχαίο $\hat{x} \in M$, το οποίο συνεπάγεται $\hat{z}^T \hat{x} = \lim_{v \to \infty} z^{(v)^T} \hat{x} \leq 0$ με $\hat{z} \in M^P$, και κατά συνέπεια M^P κλειστός.

Αν C είναι κυρτός πολυεδρικός κώνος που παράγεται από $\{d^{(1)}, ..., d^{(r)}\}$, τα οποία αποτελούν τον πίνακα D=($d^{(1)}, ..., d^{(r)}$), ο πολικός του κώνος δίνεται από το σύνολο $C^P = \{z \,|\, D^T z \leq 0\}$ το οποίο αποτελεί κυρτό πολυεδρικό κώνο.

(Πρόταση : Κάθε σύνολο της μορφής $C = \{y \,|\, Ay \geq 0\}$ είναι κλειστός πολυεδρικός κώνος ο οποίος είτε παράγεται από {0} είτε από $y \in C$, $y \neq 0$ και είναι λύσεις του συστήματος $\begin{matrix} Ay = 0 \\ e^T y = 1 \\ y \geq 0 \end{matrix}$ με $e^T = (1, ..., 1)$).▲

Πρόταση 1.4 Έστω το σύνολο των εφικτών λύσεων $B \neq \varnothing$,τότε για $P = conv\{x^{(1)}, ..., x^{(r)}\}$ και $C = \{y \,|\, Ay = 0,\ y \geq 0\}$, το σύνολο $B = P + C = \{z \,|\, z = x + y$ με $x \in P$ και $y \in C\}$ είναι το νέο σύνολο των εφικτών λύσεων.

Σύμφωνα με την παραπάνω *πρόταση* και χρησιμοποιώντας το σύνολο των βασικών εφικτών λύσεων $\{x^{(1)}, ..., x^{(r)}\}$, δηλαδή τις κορυφές του πολύεδρου P, και το εκφυλισμένο σύνολο $\{y^{(1)}, ..., y^{(s)}\}$ του C, τότε το πρόβλημα γραμμικού προγραμματισμού μπορεί να πάρει την μορφή :

$$\left.\begin{array}{l} \min \sum_{i=1}^{r} \lambda_i c^T x^{(\mathrm{i})} + \sum_{j=1}^{s} \mu_j c^T y^{(\mathrm{j})} \\ \text{υ. π .} \sum_{i=1}^{r} \lambda_i = 1 \\ \lambda_i \geq 0 \forall i \\ \mu_j \geq 0 \forall j \end{array}\right\} \quad (1)$$

Το παραπάνω πρόβλημα είναι επιλύσιμο αν και μόνον αν $c^T y \geq 0 \quad \forall y \in C$, με βέλτιστη λύση μία από τις κορυφές $x^{(\mathrm{i}_0)}$ του συνόλου Β για την οποία ισχύει $c^T x^{(\mathrm{i}_0)} = \min_{i \in \{1,\ldots,\mathrm{r}\}} c^T x^{(\mathrm{i})}$.

Πρόταση 1.5 Με την υπόθεση $B \neq 0$, το γραμμικό πρόβλημα $\begin{cases} \min c^T x \\ Ax = b \\ x \geq 0 \end{cases}$

Είναι επιλύσιμο αν και μόνο αν $c^T y \geq 0, \forall y \in C$, δηλαδή $-c \in C^P$. Στην περίπτωση αυτή, σαν βέλτιστη λύση μπορεί να θεωρηθεί μία κορυφή $x^{(\mathrm{i}_0)}$ του συνόλου Β τέτοια ώστε $c^T x^{(\mathrm{i}_0)} = \min_{i \in \{1,\ldots,\mathrm{r}\}} c^T x^{(\mathrm{i})}$.

Απόδειξη : Η υπόθεση αυτή $c^T y \geq 0, \forall y \in C$ ισοδυναμεί με την απαίτηση να ισχύει $c^T y \geq 0, \forall y \in C$.

Αν η συνθήκη αυτή παραβιαστεί για τουλάχιστον ένα $y^{(\mathrm{j})}$ (έστω το $y^{(\mathrm{j}_1)}$) τότε για $\mu_{j_1} \to \infty \implies z \to -\infty$ έτσι ώστε το πρόβλημα γραμμικού προγραμματισμού να είναι επιλύσιμο.

Αν από την άλλη μεριά, ικανοποιείται η συνθήκη, τότε επιλύοντας το πρόβλημα (1), επιλέγουμε $\mu_j = 0, \forall j$, το οποίο επαληθεύει την υπόθεση.▲

Πρόταση 1.6 Αν $B \neq 0$ και $c^T x \geq \gamma \quad \forall x \in B, \forall \gamma \in \mathbb{R}$ τότε το γραμμικό πρόβλημα $\min\{\mathrm{c}^T x \,|\, \mathrm{x} \in \mathrm{B}\}$ είναι επιλύσιμο.

***Απόδειξη*:** Για κάθε σταθερό σημείο $\hat{x} \in B$ και για τυχαίο $y \in C$ ισχύει ότι $\hat{x} + \mu y \in B$, $\forall \mu > 0$. Από υπόθεση και την προηγούμενη πρόταση έχουμε $c^T x + \mu c^T y \geq \gamma \Rightarrow c^T y \geq 0, \forall y \in C$.▲

1.2.3. Η Δυική μέθοδος

Το αρχικό μας πρόβλημα ελαχιστοποίησης γραμμικού προγραμματισμού

$$\text{υ. π.} \left.\begin{array}{l} \min c^T x \\ Ax = b \\ x \geq 0 \end{array}\right\}$$

μπορεί να μετατραπεί σε ισοδύναμο πρόβλημα μεγιστοποίησης παίρνοντας την μορφή:

$$\text{υ. π.} \left.\begin{array}{l} \max b^T u \\ A^T u \leq c \end{array}\right\}$$

και έτσι ορίζεται το δυικό του, όπου ο περιορισμός για τις θετικές μεταβλητές έχει αντιστοιχηθεί με την ανισότητα $A^T u \leq c$.

Είναι σημαντικό να αναφέρουμε την ισχυρή σχέση μεταξύ της δυικής βασικής εφικτής λύσης και του Simplex κριτηρίου.

Το πρωτεύον πρόβλημα, είτε εφικτό είτε όχι, έχει διαιρεθεί σύμφωνα με τον Simplex πίνακα

$$\begin{array}{|c|c|}\hline \zeta & d^{\mathrm{T}} \\ \hline \beta & D \\ \hline \end{array} = \begin{array}{|c|c|}\hline \zeta & \delta_1 \quad \cdots \quad \delta_{n-m} \\ \hline \beta_1 & \alpha_{11} \quad \cdots \quad \alpha_{1n-m} \\ \vdots & \vdots \qquad\quad \vdots \\ \beta_m & \alpha_{m1} \quad \cdots \quad \alpha_{mn-m} \\ \hline \end{array}$$

Εικόνα 3 : Simplex Tableau

και αν ικανοποιεί το simplex κριτήριο $d^T = (\mathrm{c}^{\{\mathrm{B}\}^T}\mathrm{B}^{-1}\mathrm{N} - \mathrm{c}^{\{\mathrm{N}\}^T}) \le 0$, θα καλείται Δυικός Εφικτός πίνακας.

Προτάσεις-ιδιότητες πρωτεύοντος και δυικού προβλήματος

- Η μετατροπή του δυικού προβλήματος ενός πρωτεύοντος προβλήματος ισοδυναμεί με το πρωτεύον πρόβλημα.
- Για κάθε ζευγάρι εφικτών λύσεων του αρχικού και του δυικού προβλήματος $\tilde{x}$, $\tilde{u}$, ισχύει $b^T \tilde{u} \le c^T \tilde{x}$ (**Θεώρημα Δυικής Χαλαρότητας**).

Απόδειξη : Σύμφωνα με τις παραδοχές $A\tilde{x} = b, \tilde{x} \ge 0$ και $A^T u \le c$ ισχύει ότι

$b^T \tilde{u} = (\mathrm{A}\tilde{\mathrm{x}})^T \tilde{u} = \tilde{x}^T (\mathrm{A}^T \tilde{\mathrm{u}}) \le \tilde{\mathrm{x}}^T c$.▲

- Αν καθένα από τα δύο προβλήματα είναι εφικτά είναι και επιλύσιμα.

Απόδειξη : Έστω $\hat{u}$ η βέλτιστη λύση του δυικού προβλήματος $\left.\begin{array}{l} \max b^T u \\ A^T u \le c \end{array}\right\}$. Τότε από το θεώρημα δυικής χαλαρότητας έχουμε $c^T x \ge b^T \hat{u}, \forall x \in B$. Έτσι από την **"πρόταση 1.6"** προκύπτει η επιλυσιμότητα του πρωτεύοντος προβλήματος, αναλόγως και του δυικού του.▲

- Αν το αρχικό πρόβλημα είναι επιλύσιμο τότε και το αντίστοιχο δυικό του είναι επίσης επιλύσιμο με ίσες βέλτιστες τιμές των αντικειμενικών τους συναρτήσεων. (**Ισχυρός Νόμος Δυικότητας**)

Απόδειξη : Από το κριτήριο Simplex το πρόβλημα $\left.\begin{array}{l}\min c^T x \\ Ax = b \\ x \geq 0\end{array}\right\}$ είναι επιλύσιμο αν και μόνο αν υπάρχει μια βέλτιστη λύση του B τέτοια ώστε να ικανοποιεί την

$$d^T = (\mathrm{c}^{\{\mathrm{B}\}^T}\mathrm{B}^{-1}\mathrm{N} - \mathrm{c}^{\{\mathrm{N}\}^T}) = (\mathrm{c}^{\{\mathrm{B}\}^T}\mathrm{D} - \mathrm{c}^{\{\mathrm{N}\}^T}) \leq 0$$

Τότε από την αναδιάταξη των στηλών του πίνακα A, έχουμε $(\mathrm{B}, \mathrm{N}) = \mathrm{A}$ και για $\hat{u} = B^{-1^T} c^{\{\mathrm{B}\}}$ ισχύει ότι $\begin{array}{l} B^T \hat{u} = c^{\{\mathrm{B}\}} \\ N^T \hat{u} \leq c^{\{\mathrm{N}\}} \end{array}$.

Έτσι, το $\hat{u}$ είναι εφικτή για το δυικό πρόγραμμα με αντικειμενική τιμή

$$b^T \hat{u} = b^T B^{-1^T} c^{\{\mathrm{B}\}} = c^{\{\mathrm{B}\}^T} B^{-1} b$$

η οποία και συμπίπτει με αυτήν του πρωτεύοντος προγράμματος.▲

- x^*, u^* αποτελούν τις βέλτιστες λύσεις του αρχικού και δυικού προβλήματος αντίστοιχα αν και μόνο αν ισχύει (***Συνθήκη Συμπληρωματικής Χαλαρότητας***):

$$x^{*T}(\mathrm{A}^T u^* - c) = 0$$

Απόδειξη : Αν το πρωτεύον και το δυικό πρόβλημα είναι επιλύσιμα, τότε υπάρχουν βέλτιστες εφικτές λύσεις x^*, u^* όπου από την εφικτότητά τους και την ισχυρή δυικότητα ισχύει

$$0 = c^T x^* - b^T u^* = c^T x^* - x^{*T} A^T u^* = x^{*T}(\mathrm{c} - \mathrm{A}^T \mathrm{u}^*).$$

Από την άλλη μεριά, λόγω της εφικτότητας, της συμπληρωματικότητας και την δυική χαλαρότητα, ισχύει

$$0 \leq c^T x^* - b^T u^* = x^{*T}(\mathrm{c} - \mathrm{A}^T \mathrm{u}^*) = 0$$

Δηλαδή η βελτιστότητα των λύσεων.▲

Πρόταση-Λήμμα Farkas Ισχύει $\{x\left|\mathrm{Ax}=\mathrm{b},\mathrm{x}\geq 0\right.\}\neq\varnothing$ αν και μόνο αν $A^{T}u\leq 0$ $\Rightarrow b^{T}u\leq 0$.

Απόδειξη : Έστω ότι ισχύει $A\hat{x}=b$ για κάποιο $\hat{x}\geq 0$. Τότε για κάθε $\hat{u}$ με $A^{T}\hat{u}\leq 0$

συνεπάγεται $B^{T}\tilde{u}=(A\hat{x})^{T}\tilde{u}=(\hat{\mathrm{x}}^{T}\ \mathrm{A}^{T})\tilde{\mathrm{u}}=\hat{x}^{T}\ (\mathrm{A}^{T}\ \tilde{\mathrm{u}})\leq 0$.

Από την άλλη πλευρά, θεωρώντας ότι $A^{T}u\leq 0$ συνεπάγεται $b^{T}u\leq 0$. Για τυχαίο $\hat{u}\neq 0$, ορίζουμε $c:=A^{T}\hat{u}$. Από την "**πρόταση 1.5"** έχουμε ότι το γραμμικό πρόγραμμα $\max\{b^{T}u\left|\mathrm{A}^{T}u\leq c\right.\}$ είναι επιλύσιμο. Τότε από τον ισχυρό νόμο της δυικότητας, και το δυικό πρόγραμμα $\min\{c^{T}x\left|\mathrm{Ax}=\mathrm{b},\mathrm{x}\geq 0\right.\}$ είναι επιλύσιμο και συνεπώς εφικτό.

1.2.4. Η μέθοδος Simplex

Η βασική ιδέα της μεθόδου Simplex για ένα πρόβλημα γραμμικού προγραμματισμού που έχει βέλτιστη λύση, είναι να ψάχνει μεταξύ των εφικτών λύσεων αυτήν που βελτιώνει την τιμή της αντικειμενικής συνάρτησης. Η οποία συνοψίζεται στην παρακάτω πρόταση :

Πρόταση 1.7 Το πρώτο βήμα της μεθόδου μας παρέχει το σύνολο των εφικτών λύσεων, οι οποίες είναι μη εκφυλισμένες. Τότε η μέθοδος, αφού ελέγξει όλες αυτές τις λύσεις, τερματίζει σε πεπερασμένο πλήθος επαναλήψεων δίνοντας την βέλτιστη εφικτή λύση ή δίνοντας την πληροφορία ότι $\inf\{\mathrm{c}^{T}\mathrm{x}\left|x\in B\right.\}=-\infty$.

Αν ωστόσο, υπάρχουν εκφυλισμένες βασικές εφικτές λύσεις τότε δεν θα υπάρξει καμία αλλαγή στην τιμή της αντικειμενικής συνάρτησης και θα πάρουμε την ίδια εφικτή λύση με την οποία ξεκινήσαμε.

Το γεγονός αυτό μπορεί να επαναλαμβάνετε απείρως χωρίς να έχουμε καμία βελτίωση. Η διαδικασία αυτή μπορεί να αποφευχθεί εάν προσθέσουμε κάποιους κανόνες κατά την εφαρμογή του αλγορίθμου Simplex, τους κανόνες Bland ή λεξικογραφικοί κανόνες, όπου θα βοηθούν στην επιλογή των μεταβλητών όταν αυτές δεν είναι μοναδικά ορισμένες. Τότε εμποδίζεται η ατέρμονη συμπεριφορά της Simplex και επιτρέπεται στην μέθοδο να τερματίσει και να επιστρέψει και πάλι ένα από τα παραπάνω αποτελέσματα.

1.2.5. Η Δυική μέθοδος Simplex

Η Δυική Simplex αποτελεί μια παραλλαγή της μεθόδου Simplex η οποία συμπεριφέρεται σαν αλγόριθμος που κινείται στο σύνολο των εφικτών λύσεων του πρωτεύοντος προβλήματος αναζητώντας μια εφικτή λύση του δυικού. Εναλλακτικά, μπορεί να ξεκινήσει από μια εφικτή λύση του Δυικού αναζητώντας μια εφικτή λύση του πρωτεύοντος.

Σύγκριση Δυικής και Απλής Simplex

Συγκριτικά με την απλή μέθοδο Simplex, η Δυική έχει παρατηρηθεί ότι απαιτεί μικρότερο πλήθος πινάκων που πρέπει να επαναπροσδιορίζονται σε κάθε βήμα. Επίσης, έχει παρατηρηθεί αρκετές φορές να είναι πιο αποτελεσματική έναντι της απλής Simplex.

1.2.6. Δυική Ανάλυση

Στην εισαγωγή αναφέραμε στοχαστικά προγράμματα με τις μεταβλητές να ακολουθούν κάποια διακριτή κατανομή πιθανότητας $p_j, j = 1,...,S$, τα οποία μπορούν να επιλυθούν ως προβλήματα με δομή Δυικής χαλαρότητας. Θεωρώντας την απλή περίπτωση με μία παρατηρούμενη τιμή (S=1), το πρόβλημα έχει την μορφή:

$$\left.\begin{array}{l} \min\{\mathbf{c}^T\mathbf{x}+\mathbf{q}^T\mathbf{y}\} \\ \mathbf{A}\mathbf{x}=\mathbf{b} \\ \mathbf{T}\mathbf{x}+\boldsymbol{W}\boldsymbol{y}=\boldsymbol{h} \\ \boldsymbol{x}\geq\mathbf{0} \\ \boldsymbol{y}\geq\mathbf{0} \end{array}\right\}$$

Υποθέτοντας ότι το πρόβλημα είναι επιλύσιμο και ότι το σύνολο $\{x|\mathrm{Ax}=\mathrm{b},\mathrm{x}\geq 0\}$ είναι φραγμένο, το πρόβλημα μπορεί να γραφεί ισοδύναμα ως εξής :

$$\min\{\mathrm{c}^T+f(\mathrm{x})\}$$

όταν

$$Ax=b$$
$$x\geq 0$$

με $f(\mathrm{x}):=\min\{\mathrm{q}^T\ y|\mathrm{Wy}=\mathrm{h}-\mathrm{Tx},\mathrm{y}\geq 0\}$.

Πρόταση 1.8 Η συνάρτηση f(x), που καθορίζεται από το φραγμένο σύνολο $\{x|\mathrm{Ax}=\mathrm{b},\mathrm{Tx}+\mathrm{Wy}=\mathrm{h},\mathrm{x}\geq 0,\mathrm{y}\geq 0\}\neq\varnothing$, είναι μερικώς γραμμική, κυρτή και φραγμένη από κάτω.

Απόδειξη : Από τις υποθέσεις μας, για το σύνολο $\mathrm{B}_1=\{x|\mathrm{Ax}=\mathrm{b},\mathrm{x}\geq 0\}$ συνεπάγεται ότι το σύνολο $B:=\mathrm{B}_1\cap\{x|\exists\,\mathrm{y}\geq 0:\mathrm{Wy}=\mathrm{h}-\mathrm{Tx}\}\neq\varnothing$ είναι φραγμένο. Τότε το $\{x|\exists\,\mathrm{y}\geq 0:\mathrm{Wy}=\mathrm{h}-\mathrm{Tx}\}\neq\varnothing$ είναι η προβολή του κυρτού πολυεδρικού συνόλου $\{(x,y)|\mathrm{Tx}+\mathrm{Wy}=\mathrm{h},\mathrm{y}\geq 0\}$ στον χώρο (x,y). Τότε το σύνολο B ως η τομή του κυρτού πολυέδρου με ένα κυρτό πολυεδρικό σύνολο είναι κυρτό πολύεδρο και για $x\in B$ ισχύει

$$f(\mathrm{x})=\mathrm{q}^{\{\mathrm{Bw}\}^T}\ B_W^{-1}(\mathrm{h}-\mathrm{Tx})\ \text{αν}\ B_W^{-1}(\mathrm{h}-\mathrm{Tx})\geq 0$$

με B_W να είναι μία κατάλληλη βέλτιστη βάση που επιλέχθηκε ανάμεσα από πολλές εφικτές βάσεις του W. Έτσι η f(x) είναι μερικώς γραμμική και έχει κάτω φράγμα στο

$x \in B$. Τέλος, για $x^1 \in B$ και $x^2 \in B$ τέτοια ώστε η $f(\mathrm{x}^i), \mathrm{i} = 1,2$ να είναι πεπερασμένη, σε αντιστοιχία με τις λύσεις y^1, y^2 ισχύει

$$f(\mathrm{x}^i) = \mathrm{q}^T y^i, i = 1,2 \text{ και } Wy^i = h - Tx^i, y^i \geq 0, i = 1,2$$

Για τυχαίο $\lambda \in (0,1)$ και $\tilde{x} = \lambda x^1 + (1-\lambda)\,\mathrm{x}^2$ ισχύει ότι

$$\lambda y^1 + (1-\lambda)\,\mathrm{y}^2 \in \{y \,|\, \mathrm{Wy} = \mathrm{h} - \mathrm{T}\,\tilde{\mathrm{x}}, \mathrm{y} \geq 0\}$$

οπότε $f(\tilde{\mathrm{x}}) \begin{cases} = \min\{\mathrm{q}^T y \,|\, \mathrm{Wy} = \mathrm{h} - \mathrm{T}\,\tilde{\mathrm{x}}, \mathrm{y} \geq 0\} \\ \leq q^T(\lambda\,\mathrm{y}^1 + (1-\lambda)\,\mathrm{y}^2) = \lambda f(\mathrm{x}^1) + (1-\lambda) f(\mathrm{x}^2) \end{cases}$

αποδεικνύοντας έτσι την κυρτότητα της f(x) στην αποτελεσματική της περιοχή (dom f=B).

Επαναδιατυπώνοντας,

$$\min\{\mathrm{c}^T \mathrm{x} + \theta\}$$

όταν

$$\begin{aligned} &Ax = b \\ &\theta - f(\mathrm{x}) \geq 0 \\ &\mathrm{x} \geq 0 \end{aligned}$$

Στην περίπτωση όμως αυτή απαιτείται η εκ των προτέρων γνώση της f. Ως εκ τούτου, μπορούμε να προσθέσουμε κάποιους γραμμικούς περιορισμούς $\psi_1(\mathrm{x}), \ldots, \psi_K(\mathrm{x})$ προκειμένου να μειώσουμε το σύνολο των εφικτών λύσεων *B* και να πετύχουμε εφικτότητα του προβλήματος.

Ωστόσο, και πάλι δεν είναι γνωστοί οι γραμμικοί περιορισμοί εκ των προτέρων και γι' αυτό η βασική ιδέα της μεθόδου της Δυικής χαλαρότητας είναι να δημιουργούνται διαδοχικά οι επιπρόσθετοι περιορισμοί που χρειάζονται για την προσέγγιση της αρχικής λύσης.▲

Πρόταση 1.9 : Αν ισχύουν οι παραπάνω προϋποθέσεις, δηλαδή το πρόβλημα να είναι επιλύσιμο και το σύνολο $\{x \mid \mathrm{Ax} = \mathrm{b}, \mathrm{x} \geq 0\}$ να είναι φραγμένο, τότε η δυική μέθοδος χαλαρότητας παράγει μια βέλτιστη λύση x^* μετά από πεπερασμένου πλήθους βήματα.

Απόδειξη : Από την προηγούμενη πρόταση υπάρχει ένα κάτω όριο θ_0 κατά το πρώτο βήμα της μεθόδου, το οποίο μπορεί να θεωρηθεί ότι είναι η ποσότητα $b^T w + h^T u$.

Λόγω της επιλυσιμότητας του προβλήματος, οι δυικοί περιορισμοί $W^T u \leq q$ είναι εφικτοί και ανεξάρτητοι του x. Συνεπώς, στο δεύτερο βήμα η δυική μορφή της $f(\hat{\mathrm{x}})$ θα είναι πάντα εφικτή, το οποίο σημαίνει ότι η $f(\hat{\mathrm{x}})$ είναι είτε πεπερασμένη είτε άπειρη ($+\infty$ $\Leftrightarrow$ μη αρχική εφικτότητα).

Αν $f(\hat{\mathrm{x}}) = +\infty$, δηλαδή $\hat{\mathrm{x}}$ μη εφικτό, τότε υπάρχει $\hat{u} : W^T \tilde{u} \leq 0$ και $(h - T\hat{x})^T \tilde{u} > 0$. ▲

1.2.7. Αλυσιδωτή Ανάλυση

Τα παραπάνω εφαρμόστηκαν για πρόβλημα με μία παρατηρούμενη τιμή, αν ωστόσο είχαμε Τ παρατηρούμενες τιμές επιτρέποντας την εμφάνιση τυχαίων παραμέτρων, το πρόβλημα θα είχε την μορφή :

$$\left.\begin{array}{l} \min\{\mathrm{c}_1^T \mathrm{x}_1 + E_{\zeta_t}[\sum_{t=2}^{T} c_t^T(\zeta_t)\,\mathrm{x}_t(\zeta_t)]\} \\ \mathrm{W}_1 x_1 = b_1 \\ T_t(\zeta_t)\,\mathrm{x}_{t-1}(\zeta_{t-1}) + W_t(\zeta_t) x_t(\zeta_t) = b_t(\zeta_t) \\ \mathrm{x}_1, x_t(\zeta_t) \geq 0 \end{array}\right\} \forall t$$

με $x_t(\hat{\zeta}_t)$ να είναι η απόφαση του t σταδίου ικανοποιώντας όλους τους προηγούμενους περιορισμούς.

Η παραπάνω διαδικασία μπορεί να γίνει και μέσω ενός δεντροδιαγράμματος (σενάριο) όπου οι διαφορετικές τιμές ζ_t αντιστοιχούν στα διάφορα μονοπάτια από την ρίζα στα φύλλα του δέντρου. Έτσι, συμβολίζοντας με N το σύνολο των κόμβων του σεναρίου-δέντρο, N_t το σύνολο των κόμβων του t σταδίου και $G(\mathrm{n})$ το σύνολο των κόμβων κάθε υποδέντρου, το αρχικό πρόβλημα μπορεί να γραφεί :

όταν

$$\left.\begin{aligned} & F_1 = \min c_1^T x_1 + \sum_{\nu \in N-\{1\}} p_\nu c_\nu^T x_\nu \\ & W_1 x_1 = b_1 \\ & T_\nu x_{h_\nu} + W_\nu x_\nu = b_\nu, \nu \in \mathrm{N} - \{1\} \\ & x_\nu \geq 0, \nu \in N \end{aligned}\right\}$$

Υπόθεση Το μοντέλο έχει άνω όριο τέτοιο ώστε $x_n \leq U_n, U_n \geq 0, \forall n \in N$. Υποθέτουμε πως αυτό το άνω όριο αποτελείται από γραμμές της μορφής (W_n, b_n) τα οποία αντιστοιχούν στις γραμμές των T_n αποτελούμενες από μηδενικά.

Σύμφωνα με την παραπάνω υπόθεση, το πρόβλημα μπορεί να είναι είτε μη εφικτό, είτε να έχει μία βέλτιστη λύση, ενώ το δυικό του είναι εφικτό.

Απόδειξη: Θεωρώντας το αρχικό και το δυικό πρόβλημα γραμμικού προγραμματισμού,

$$\left\{\begin{array}{ll} \min c^T x & \\ & Ax = b \\ & Ix + y = d \\ s.t & x \geq 0 \\ & y \geq 0 \end{array}\right\}, \left\{\begin{array}{ll} \max b^T u + d^T v & \\ & A^T u + v \leq c \\ s.t & v \leq 0 \end{array}\right\}$$

Είναι φανερό πως για αυθαίρετη τιμή του u υπάρχει τιμή v τέτοια ώστε το σύνολο (u,v) να αποτελεί την βασική εφικτή λύση του δυικού.

Έτσι, το παραπάνω πρόβλημα με κόμβους πλήθους n, έχει συνάρτηση $F_n(x_{h_n})$ γραμμική κατά τμήματα και κυρτή στις x_{h_n}, $\forall n \in N_t$.▲

Ορισμός προβλημάτων

- Σε αυτό το σημείο μπορούμε να ορίσουμε το $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ (Φθίνων Πρόβλημα) το οποίο έχει την μορφή :

όταν

$$\left.\begin{aligned} & F_n(\mathrm{x}_{h_n}) = \mathrm{minc}_n^T + \sum_{v \in G(\mathrm{n})-\{\mathrm{n}\}} \frac{p_v}{p_n} c_v^T x_v \\ & W_n x_n = b_n - T_n x_{h_n} \\ & T_v x_{h_v} + W_v x_v = b_v, v \in G(\mathrm{n}) - \{\mathrm{n}\} \\ & \mathrm{x}_n \geq 0, v \in G(\mathrm{n}) \end{aligned}\right\}$$

Το $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ είτε είναι μη εφικτό είτε έχει μία βέλτιστη λύση, ενώ, το δυικό του είναι εφικτό για κάθε $n \in N$ και για όλες τις μεταβλητές x_{h_n} .

- Θέτοντας ένα άνω όριο ϑ στον πρόσθετο όρο $\sum_{m \in C(\mathrm{n})} \frac{p_m}{p_n} F_m(\mathrm{x}_n)$, με ϑ να ικανοποιεί τον γραμμικό περιορισμό $d_n^T x_n + \vartheta_n \geq \delta_{nk}, k = 1, ..., \mathrm{S}_n$ και επιπλέον να ισχύει $a_{n_j}^T x_n \geq a_{n_j}, j = 1, ..., \mathrm{R}_n$ εξασφαλίζοντας έτσι την εφικτότητα του προβλήματος, μπορούμε να ορίσουμε το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ με μορφή :

$$\left.\begin{aligned}
&\tilde{F}_n(\mathrm{x}_{h_n}) = \mathrm{minc}_n^T\, x_n + \vartheta_n \\
&W_n x_n = b_n - T_n x_{h_n} \\
&a_{n_j}^T x_n \geq a_{n_j},\ j = 1,\ldots,\mathrm{r}_n \\
&d_{n_k}^T x_n + \vartheta_n \geq \delta_{n_k},\ k = 1,\ldots,s_n \\
&x_n \geq 0
\end{aligned}\right\}$$

το οποίο σε κάθε επόμενο στάδιο θεωρείται το αρχικό πρόβλημα και μπορεί να είναι μη εφικτό ή να έχει μία βέλτιστη λύση με το δυικό του να είναι εφικτό.

Οι περιορισμοί $a_{n_j}^T x_n \geq a_{n_j},\ j = 1,\ldots,r_n$ και $d_{n_k}^T x_n + \vartheta_n \geq \delta_{n_k}, \mathrm{k} = 1,\ldots,\mathrm{s}_n$ αποτελούν τα διαστήματα εφικτότητας και βελτιστότητας αντίστοιχα.

Για $r_n = 0$ και $\mathrm{s}_n = 0$ έχουμε την απουσία των περιορισμών αυτών.

Για $s_n > 0$ ο τρίτος περιορισμός μπορεί να γραφεί ως $\vartheta_n \geq \max_{1\leq k\leq s_n}(\delta_{n_k} - \mathrm{d}_{n_k}^T \mathrm{x}_n)$.

Ανάλυση Ευαισθησίας

Θεωρώντας, ότι το πρόβλημα $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ έχει λύση την $(\hat{x}_n, \hat{\vartheta}_n)$, τότε για έναν επόμενο κόμβο του δέντρου $m \in C(\mathrm{n})$, το πρόβλημα $LPMast(\mathrm{m}, \hat{x}_n)$ ή θα είναι μη εφικτό ή θα έχει βέλτιστη λύση.

Ορισμός 1.4 Στο πρόβλημα $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ το διάστημα εφικτότητας $a_{n_j}^T x_n \geq a_{n_j},\ j = 1,\ldots,r_n$ ισχύει όταν για κάθε εφικτή λύση $(\tilde{\mathrm{x}}_\nu, \nu \in G(\mathrm{n}))$ του προβλήματος $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ ισχύει η ανίσωση $a_{n_j}^T \tilde{x}_n \geq a_{n_j}$.

Επιπλέον, αν το πρόβλημα $LPMast(\mathrm{m}, \hat{x}_n)$ είναι επιλύσιμο για κάθε $m \in C(\mathrm{n})$ τότε θεωρούμε ένα αντίστοιχο διάστημα βελτιστότητας του αρχικού προβλήματος $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$. Για $(\hat{\mathrm{x}}_m, \vartheta_m)$ η βέλτιστη λύση του $LPMast(\mathrm{m}, \hat{x}_n)$ και

$(\hat{u}_m, \hat{v}_m, \hat{w}_m)$ η αντίστοιχη βέλτιστη του δυικού του, έχουμε την μορφή του διαστήματος βελτιστότητας : $d_{n_k}^T x_n + \vartheta_n \geq \delta_{n_k}, k = s_n + 1$.

Ορισμός 1.5 Στο πρόβλημα $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ το διάστημα βελτιστότητας $d_{n_k}^T x_n + \vartheta_n \geq \delta_{n_k}$ ισχύει αν ισχύει η ανισότητα $\delta_{n_k} - d_{n_k}^T x_n \leq \sum_{m \in C(\mathrm{n})} \frac{p_m}{p_n} F_m(\mathrm{x}_n)$ για κάθε αυθαίρετη μεταβλητή x_n .

Πρόταση 1.10

i. Έστω ο τυχαίος κόμβος $n \in N$. Για $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ θα ισχύει ότι είτε $r_n = 0$ είτε όλες οι τομές εφικτότητας είναι έγκυρες, τότε κάθε εφικτή λύση $(\bar{\mathrm{x}}_\nu, \nu \in G(\mathrm{n}))$ του $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$, η $\bar{x}_n$ θα είναι η εφικτή λύση του $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$.

ii. Έστω οι κόμβοι $n \in N$ και $m \in C(\mathrm{n})$ οι οποίοι παράγουν την τομή εφικτότητας. Υποθέτοντας για το $LPMast(\mathrm{m}, \hat{\mathrm{x}}_n)$ είτε ότι $r_m = 0$ είτε ότι όλες οι τομές εφικτότητας είναι έγκυρες, η νέα αυτή τομή θα είναι ισχύς εφικτή τομή για το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$.

Απόδειξη :

i. Αν $r_n = 0$ τότε ισχύει η πρόταση. Διαφορετικά, η υπόθεση είναι άμεση συνέπεια του ορισμού της εγκυρότητας.

ii. Υποθέτουμε ότι η $(\bar{x}_\nu, \nu \in \mathrm{G(n)})$ είναι εφικτή λύση του $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$.

Πιο συγκεκριμένα, για τον κόμβο $m \in C(\mathrm{n})$, ο οποίος έχει δημιουργήσει την τομή, έχουμε

$$W_m \bar{x}_m = b_m - T_m \bar{x}_n$$

$$a_{mj}^T \bar{x}_m \geq a_{mj}, \; j = 1, \ldots, \mathrm{r}_m$$

όπου η ανισότητα ισχύει για τον λόγο ότι : η $(\bar{x}_\mu, \mu \in \mathrm{G(m)})$ είναι εφικτή λύση του $LPDesc(\mathrm{m}, \bar{\mathrm{x}}_n)$ και τότε η ανισότητα είναι λογικό επακόλουθο της υπόθεσης για την τομή εφικτότητας του $LPMast(\mathrm{m}, \hat{\mathrm{x}}_n)$. Πολλαπλασιάζοντας τις παραπάνω σχέσεις με $\tilde{u}_m, \tilde{v}_m$ αντίστοιχα και προστίθοντας κατά μέλη έχουμε

$$(\mathrm{b}_m - \mathrm{T}_m \bar{x}_n)^T \tilde{u}_m + \sum_{j=1}^{r_m} a_{mj} \tilde{v}_{mj} \le (\mathrm{W}_m^T \tilde{u}_m + \sum_{j=1}^{r_m} a_{mj} \tilde{v}_{mj}) \bar{x}_m \le 0$$

με $\bar{x}_m \ge 0$.▲

Πρόταση 1.11

i. Έστω ένας τυχαίος κόμβος $n \in N$, θεωρούμε ότι όλες οι τομές εφικτότητας ισχύουν για το πρόβλημα $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$. Αν $n \in N_T$ ή για το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ ισχύει $s_n > 0$ και ισχύουν όλες οι τομές βελτιστότητας, τότε ισχύει $\tilde{F}(\mathrm{x}_{h_n}) \le \mathrm{F}_n(\mathrm{x}_{h_n})$ για κάθε x_{h_n}.

ii. Έστω ο κόμβος $n \in N$ στον οποίο οφείλονται οι τομές βελτιστότητας. Αν $n \in N_{T-1}$ ή αν για όλα τα $m \in C(\mathrm{n})$ ισχύει $s_m > 0$ και οι τομές εφικτότητας και βελτιστότητας ισχύουν για το $LPMast(\mathrm{m}, \hat{x}_n)$, τότε η νέα τομή θα είναι επίσης τομή βελτιστότητας για το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$.

Απόδειξη:

i. Στην περίπτωση που $n \in N_T$, τα $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$, $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ είναι ταυτόσημα προβλήματα και ισχύει $\tilde{F}_n(\mathrm{x}_{h_n}) = \mathrm{F}_n(\mathrm{x}_{h_n})$ για όλα τα x_{h_n}.

Υποθέτουμε ότι $n \in N_t$ με $t \le T-1$. Η υπόθεση αυτή συνεπάγεται την εξής ανισότητα :

$$c_n^T x_n + \max_{1\le k\le s_n}(\delta_{nk} - \mathrm{d}_{nk}^T \mathrm{x}_n) \le c_n^T x_n + \sum_{m\in C(\mathrm{n})} \frac{p_m}{p_n} F_m(\mathrm{x}_n) \text{ για τυχαίο } \mathrm{x}_n .$$

Θεωρούμε τις μορφές

$$\left.\begin{array}{l} F_n(\mathrm{x}_{h_n}) = \min \mathrm{c}_n^T x_n + \sum_{m\in C(\mathrm{n})} \frac{p_m}{p_n} F_m(x_n) \\ W_n x_n = b_n - T_n x_{h_n} \\ \mathrm{x}_n \ge 0 \end{array}\right\} \;\&\; \left\{\begin{array}{l} \tilde{F}_n(\mathrm{x}_{h_n}) = \min_{1\le k\le s_n}[\mathrm{c}_n^T x_n + \max(\delta_{nk} - \mathrm{d}_{nk}^T \mathrm{x}_n) \\ W_n x_n = b_n - T_n x_{h_n} \\ a_{nj}^T x_n \ge a_{nj},\ j = 1,\ldots,r_n \\ \mathrm{x}_n \ge 0 \end{array}\right\}$$

των $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$, $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ αντίστοιχα.

- Αν το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ είναι μη εφικτό, δηλαδή $\tilde{\mathrm{F}}_n(\mathrm{x}_{h_n}) = +\infty$, τότε από την προηγούμενη πρόταση (i) το $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ είναι επίσης μη εφικτό και κατά συνέπεια ισχύει και $\tilde{\mathrm{F}}_n(\mathrm{x}_{h_n}) = +\infty$. Οπότε, $\tilde{\mathrm{F}}_n(\mathrm{x}_{h_n}) = F_n(\mathrm{x}_{h_n}) = +\infty$.
- Αν το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$ είναι εφικτό, τότε για κάθε εφικτή λύση του $LPDesc(\mathrm{n}, \mathrm{x}_{h_n})$ η οποία είναι μη εφικτή για το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$, ισχύει ένας περιορισμός τομής εφικτότητας για το $LPMast(\mathrm{n}, \mathrm{x}_{h_n})$. Αυτό συνεπάγεται ότι το δεύτερο μέρος της ανισότητας είναι ίσο με $+\infty$. Έτσι, δεν λαμβάνουμε υπόψη τα αποτελέσματα πάνω από την εφικτή περιοχή που ορίζει η ανισότητα.

ii. Από το (i) συνεπάγεται ότι $\tilde{\mathrm{F}}_m(\mathrm{x}_n) = F_m(\mathrm{x}_n)$ για κάθε $m \in C(\mathrm{n})$ και τυχαίο x_n.

Έτσι, ισχύει $\delta_{nk} - d_{nk}^T x_n \le \sum_{m\in C(\mathrm{n})} \frac{p_m}{p_n} \tilde{F}_m(\mathrm{x}_n) \le \sum_{m\in C(\mathrm{n})} \frac{p_m}{p_n} F_m(\mathrm{x}_n)$.▲

1.2.8. Ανάλυση Κανονικότητας

Προκειμένου να πετύχουμε την ελάττωση των διάφορων συμβολισμών που απαιτούνται για το εκάστοτε αρχικό πρόβλημα (master problem), μπορούμε να γράψουμε το πρόβλημα ως :

$$\min\{\mathrm{c}^T\mathrm{x}+\sum_{j=1}^{S} p_j\vartheta_j \,|(\mathrm{x},\vartheta_1,...,\vartheta_S)\in \mathrm{D}_k\}$$

Τότε μπορούμε να ελαχιστοποιήσουμε την γραμμική συνάρτηση

$$\hat{\Phi}_k(\mathrm{x}) = \mathrm{c}^T\mathrm{x}+\min_{\vartheta}\{\sum_{j=1}^{S} p_j\vartheta_j \,|(\mathrm{x},\vartheta_1,...,\vartheta_S)\in \mathrm{D}_k\}\,.$$

όπου για την συνάρτηση $\hat{\Phi}_k$ θέτουμε ένα τετραγωνικό όριο προκειμένου να αποφύγουμε την απομάκρυνση από την βέλτιστη λύση κατά τις επαναλήψεις.

Έτσι, για $z^{(\mathrm{k})}$ η προσεγγιστική εφικτή λύση, το αρχικό πρόβλημα παίρνει την μορφή:

$$\min\{\frac{1}{2p}\left\|x-z^{(\mathrm{k})}\right\|^2+\mathrm{c}^T\mathrm{x}+\sum_{j=1}^{S} p_j\vartheta_j \,|(\mathrm{x},\vartheta_1,...,\vartheta_S)\in \mathrm{D}_k\}$$

ψάχνοντας το επόμενο δυνατό (προσεγγιστικό) σημείο το οποίο σύμφωνα με τα παραπάνω κριτήρια ή θα το αποδεχτούμε ή θα συνεχίσουμε με το ήδη υπάρχων.

Παρακάτω δίνεται ο αλγόριθμος που εξασφαλίζει την βέλτιστη λύση σε δεδομένες επαναλήψεις.

Αλγόριθμος QDECOM

Βήμα 1 : Καθορισμός πρώτης προσεγγιστικής λύσης $z^{(1)}$ και της εφικτής περιοχής.

$G_1 := \{\mathrm{Ax}=\mathrm{b}\}\bigcup$ {όλα τα διαστήματα βελτιστότητας του $z^{(1)}$}

Βήμα 2 : Επίλυση της $\min\{\frac{1}{2p}\left\|x-z^{(\mathrm{k})}\right\|^2+\mathrm{c}^T\mathrm{x}+\sum_{j=1}^{S} p_j\vartheta_j \,|(\mathrm{x},\vartheta_1,...,\vartheta_S)\in \mathrm{D}_k\}$

ως προς $x^{(k)}$,το προσεγγιστικό σημείο για το πρώτο στάδιο με $\vartheta^{(k)} = (\vartheta_1^{(k)},...,\vartheta_S^{(k)})^T$.

Αν $\Phi(z^{(k)}) = \hat{\Phi}(x^{(k)})$, τότε $z^{(k)}$ είναι η βέλτιστη λύση του πρώτου σταδίου και σταματά ο αλγόριθμος, αλλιώς συνέχισε.

Βήμα 3 : Διαγραφή από το σύνολο των περιορισμών G_k τους ανενεργούς περιορισμούς της μορφής $(x^{(k)}, \vartheta^{(k)})$, έτσι ώστε το σύνολο να μην ξεπερνά τους n+S περιορισμούς.

Βήμα 4 : Αν η $x^{(k)}$ ικανοποιεί όλους τους περιορισμούς του πρώτου σταδίου, πήγαινε στο "Βήμα 5", διαφορετικά,

πρόσθεσε στο σύνολο των περιορισμών G_k τους περιορισμούς που παραβιάζονται στο πρώτο στάδιο πλήθους το πολύ S, αντικαθιστώντας το με G_{k+1}, θέσε

$z^{(k+1)} := z^{(k)}$
$k := k+1$ και επέστρεψε στο "Βήμα 2" .

Βήμα 5 : Ορισμός της $f_j(x^{(k)})$, j=1,…,S.

Αν $f_j(x^{(k)}) = +\infty$, τότε πρόσθεσε στο σύνολο G_k το διάστημα εφικτότητας,

διαφορετικά, αν $f_j(x^{(k)}) > \vartheta_j^{(k)}$, τότε πρόσθεσε το σύνολο βελτιστότητας στο G_k .

Βήμα 6 : Αν $f_j(x^{(k)}) = +\infty$ για ένα τουλάχιστον j τότε θέσε $z^{(k+1)} := z^{(k)}$ και πήγαινε στο "Βήμα 8", διαφορετικά πήγαινε στο "Βήμα 7".

Βήμα 7 : Αν $\begin{cases} \Phi(x^{(k)}) = \hat{\Phi}(x^{(k)}) \\ \Phi(x^{(k)}) \le \mu\Phi(z^{(k)}) + (1-\mu)\hat{\Phi}(x^{(k)}) \end{cases}$ ή

για $\mu \in (0,1)$ και με περιορισμούς τους n+S ενεργούς στο $(x^{(k)}, \vartheta^{(k)})$,

τότε θέσε $z^{(k+1)} := x^{(k)}$, διαφορετικά θέσε $z^{(k+1)} := z^{(k)}$.

Βήμα 8 : Για G_{k+1} το τελικό σύνολο περιορισμών, ύστερα από προσθήκη και αφαίρεση περιορισμών από το σύνολο G_k , και D_{k+1} η αντίστοιχη εφικτή περιοχή, θέσε $k := k+1$ και επέστρεψε στο "Βήμα 2".

Σύμφωνα με τον αλγόριθμο όλες οι προσεγγιστικές τιμές είναι εφικτές είτε αν η γραμμική συνάρτηση $\hat{\Phi}$ συμπίπτει με την Φ για την μεταβλητή $x^{(k)}$, όπως και για την $z^{(k)}$, και οδηγεί στην ανισότητα $\hat{\Phi}(\mathrm{x}^{(k)}) \le \hat{\Phi}(\mathrm{z}^{(k)}) \Leftrightarrow \Phi(\mathrm{x}^{(k)}) < +\infty$

εξασφαλίζοντας έτσι την εφικτότητα των $\mathrm{x}^{(k)}$, είτε αν το $(\mathrm{x}^{(k)}, \vartheta^{(k)})$ αποτελεί κορυφή του συνόλου D_k και ισχύει $\Phi(\mathrm{x}^{(k)}) - \Phi(\mathrm{z}^{(k)}) < 0$, τότε $\Phi(\mathrm{x}^{(k)}) < +\infty$ οπότε πάλι ενισχύεται η υπόθεση για εφικτότητα των $\mathrm{x}^{(k)}$.

1.2.9. Μέθοδος Εσωτερικού Σημείου

Έστω το αρχικό γραμμικό πρόβλημα και το δυικό του,

$$\text{όταν } \left.\begin{array}{l} \min c_x^T \\ Ax = b \\ x \ge 0 \end{array}\right\}, \text{ όταν } \left.\begin{array}{r} \max b^T u \\ A^T u \le 0 \end{array}\right\}.$$

Προσθέτοντας μια χαλαρή μεταβλητή s_i στο δυικό , για τις λύσεις των δύο προβλημάτων ισχύει :

$$\begin{aligned} Ax &= b \\ A^T u + s &= c \\ x &\ge 0 \\ s &\ge 0 \\ x^T s &= 0 \end{aligned}$$

ή μπορούμε να έχουμε τον επόμενο πίνακα

$$F(\mathrm{x},\mathrm{u},\mathrm{s}) := \begin{pmatrix} A^T u + Is - c \\ Ax - b \\ XSe \end{pmatrix} = 0$$

$$\text{για } \begin{matrix} X := diag(\mathrm{x}_i) \\ \mathrm{S} := \mathrm{diag}(\mathrm{s}_i) \\ \mathrm{e} = (1,\ldots,1)^T \end{matrix} \quad \text{και } x \geq 0,\ s \geq 0 .$$

Τότε το εφικτό ζευγάρι $(\hat{\mathrm{x}}, \hat{u}, \hat{s})$ ικανοποιεί την συνθήκη

$$Ax = b, A^T u + s = c, x > 0, s > 0$$

η οποία λέγεται συνθήκη ισχυρής εφικτότητας ή εσωτερικού σημείου

και λύνοντας την ισότητα $F(\hat{x},\hat{u},\hat{s}) + \Im(\hat{x},\hat{u},\hat{s}) \begin{pmatrix} \Delta x \\ \Delta u \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ (**Newton Method**)

(με $\Im(\mathrm{x},\mathrm{u},\mathrm{s}) = \begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix}$ η Ιακωβιανή ορίζουσα)

καταλήγουμε στην λύση του αρχικού συστήματος.

Πρόταση 1.12 Δεδομένης της συνθήκης ισχυρής εφικτότητας ισχύει, για κάθε $w \in \Re^n$ **:** $w_i > 0 \ \ \forall i$, υπάρχει μονοσήμαντος προσδιορισμός των **x,u,s** έτσι ώστε να ικανοποιούν την συνθήκη :

$$Ax = b,\ x \geq 0\ ,\ A^T u + s = c,\ s \geq 0\ ,\ \text{και } x_i s_i = w_i,\ i = 1,\ldots,\mathrm{n}\ .$$

Από την πρόταση αυτή είναι φανερή η σημασία της κεντρικής διαδρομής για την μέθοδο εσωτερικού σημείου, έτσι δίνεται ο ορισμός αυτής.

Ορισμός 1.6 Για μ>0 το κεντρικό μονοπάτι ορίζεται ως εξής :

$$C := \left\{ (x^{\{\mu\}}, u^{\{\mu\}}, s^{\{\mu\}}) \middle| F(x^{\{\mu\}}, u^{\{\mu\}}, s^{\{\mu\}}) = \begin{pmatrix} 0 \\ 0 \\ \mu e \end{pmatrix}, (x^{\{\mu\}}, s^{\{\mu\}}) > 0 \right\}.$$

Όπου για $\mu \to 0$, το όριο $\lim_{\mu \to 0} (\mathrm{x}^{\{\mu\}}, \mathrm{u}^{\{\mu\}}, \mathrm{s}^{\{\mu\}})$ δίνει το ζεύγος λύσεων του αρχικού και δυικού προβλήματος $\mathrm{x}^*, \mathrm{u}^*, \mathrm{s}^*$.

Αντικαθιστώντας στο παραπάνω σύστημα που προέκυψε από την μέθοδο Newton και δεδομένης της συνθήκης ισχυρής εφικτότητας, για $\hat{\mu} = \frac{\hat{x}^T \hat{s}}{n}$ έχουμε :

$$F(\hat{x}, \hat{u}, \hat{s}) + \Im(\hat{x}, \hat{u}, \hat{s}) \begin{pmatrix} \Delta x \\ \Delta u \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \hat{\mu} e \end{pmatrix}.$$

Τέλος, αν θέσουμε το $\hat{\mu}$ με μια παράμετρο $\sigma \in [0,1]$, δηλαδή σ $\hat{\mu}$, τότε για σ=0 έχουμε το αρχικό σύστημα της μεθόδου Newton, ενώ για σ=1 έχουμε το κεντρικό μονοπάτι. Έτσι, το σύστημα παίρνει την τελική του μορφή

$$F(\hat{x}, \hat{u}, \hat{s}) + \Im(\hat{x}, \hat{u}, \hat{s}) \begin{pmatrix} \Delta x \\ \Delta u \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sigma \hat{\mu} e \end{pmatrix}$$

ή

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S} & 0 & \hat{X} \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta u \\ \Delta s \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ -\hat{X}\hat{S}e+\sigma\hat{\mu}e \end{pmatrix}.$$

Αλγόριθμος Εσωτερικού Σημείου

Βήμα 1 : Βρες μια αρχική λύση $(\mathrm{x}^0,\mathrm{u}^0,\mathrm{s}^0)$ που να ικανοποιεί την συνθήκη εσωτερικού σημείου και θέσε k:=0.

Βήμα 2 : Για κάποια τιμή του σ_k στο διάστημα [0,1] και για $\mu_k=\dfrac{x^{k^T}s^k}{n}$

λύσε το σύστημα $\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S^k & 0 & X^k \end{pmatrix}\begin{pmatrix} \Delta x^k \\ \Delta u^k \\ \Delta s^k \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ -X^kS^ke+\sigma_k\mu_ke \end{pmatrix}.$

Βήμα 3 : Θέσε $(x^{k+1},u^{k+1},s^{k+1}):=(\mathrm{x}^k,u^k,s^k)+a_k(\Delta\mathrm{x}_{k,}\Delta u_{k,}\Delta s_k)$ με a_k τέτοιο ώστε $(x^{k+1},s^{k+1})>0$.

- Αν $X^{k+1}S^{k+1}e<\varepsilon e$, όπου ε σταθερή μεταβλητή, τότε σταμάτα, αλλιώς πήγαινε στο "Βήμα 2" με $k:=k+1$.
- Από την άλλη μεριά, αν η αρχική λύση είναι μη εφικτή ικανοποιώντας τους περιορισμούς, για παράδειγμα για $\begin{cases} \hat{w}_p:=A\hat{x}-b\neq 0 \\ \hat{w}_d:=A^T\hat{u}+I\hat{s}-c\neq 0 \end{cases}$ το σύστημα στο "Βήμα 2" παίρνει την μορφή :

$$\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ \hat{S} & 0 & \hat{X} \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta u \\ \Delta s \end{pmatrix}=\begin{pmatrix} -\hat{w}_d \\ -\hat{w}_p \\ -\hat{X}\hat{S}e+\sigma\hat{\mu}e \end{pmatrix}$$

το οποίο θα ισοδυναμεί με το αρχικό όταν η αρχική λύση γίνει εφικτή.

Αυτή η τροπολογία είναι γνωστή ως *μη εφικτή μέθοδος εσωτερικού σημείου*.

1.3. Ιδιότητες Μη Γραμμικού Προγραμματισμού

Ορισμός 1.7 Ένα σύνολο $C \in \mathbb{R}^n$ καλείται *κυρτό* αν για τυχαία σημεία του $x, y \in C$ και για κάθε $\lambda \in [0,1]$ ισχύει $\lambda x + (1-\lambda)\,\mathrm{y} \in \mathrm{C}$.

Τότε η συνάρτηση $\varphi : C \to \mathbb{R}$ είναι *κυρτή* όταν

$$\varphi(\lambda\,\mathrm{x} + (1-\lambda)\,\mathrm{y}) \le \lambda\varphi(\mathrm{x}) + (1-\lambda)\varphi(\mathrm{y}) \quad \forall x, y \in C \quad , \forall \lambda \in [0,1] .$$

Από τον παραπάνω ορισμό προκύπτει η επόμενη πρόταση.

Πρόταση 1.13 Αν μια συνάρτηση $\varphi : \mathbb{R}^n \to \mathbb{R}$ είναι κυρτή, τότε είναι και συνεχής.

Ορισμός 1.8 Ένα διάνυσμα $g \in \mathbb{R}^n$ είναι μία κλίση της κυρτής συνάρτησης φ στο σημείο x όταν ισχύει $g^T(\mathrm{z}-\mathrm{x}) \le \varphi(\mathrm{z}) - \varphi(\mathrm{x}) \quad \forall \mathrm{z}$. Το σύνολο όλων των υποκλίσεων της φ στο σημείο x είναι η μερική παράγωγος της φ και συμβολίζεται με $\partial\varphi(\mathrm{x})$. Αν $\partial\varphi(\mathrm{x}) \neq 0$, φ καλείται μερικώς παραγωγίσιμη στο x.

Πρόταση 1.14 Έστω η κυρτή συνάρτηση $\varphi : \mathbb{R}^n \to \mathbb{R}$, τότε για κάθε $x \in \mathbb{R}^n$ το σύνολο των μερικών παραγώγων της φ, $\partial\varphi(\mathrm{x})$ είναι διάφορο του κενού, κυρτό, κλειστό και φραγμένο.Επιπλέον, η φ είναι παραγωγίσιμη στο x με κλίση $\hat{g} = \nabla\varphi(\mathrm{x})$ αν και μόνο αν είναι μοναδική, $\partial\varphi(\mathrm{x}) = \{\hat{g}\}$.

Πρόβλημα Μη Γραμμικού Προγραμματισμού

Έστω οι κυρτές συναρτήσεις $\begin{cases} f : \mathbb{R}^n \to \mathbb{R} \\ g_i : \mathbb{R}^n \to \mathbb{R} \end{cases}$, το πρόβλημα

$$\left.\begin{array}{l}\min f(\mathrm{x}) \\ \mathrm{g}_i(\mathrm{x}) \le 0, \mathrm{i} = 1,...,\mathrm{m}\end{array}\right\}$$

αποτελεί *πρόβλημα Μη Γραμμικού Προγραμματισμού*.

Λόγω της κυρτότητας της συνάρτησης το εφικτό σύνολο

$$B := \{x \,|\, \mathrm{g}_i(\mathrm{x}) \le 0, \mathrm{i} = 0,...,\mathrm{m}\}$$

είναι κυρτό και κατά συνέπεια κάθε τοπικό ελάχιστο $\hat{x} \in B$ είναι και ολικό , δηλαδή $f(\hat{x}) = \min_{x \in B} f(\mathrm{x})$.

Από τα παραπάνω προκύπτει και η ακόλουθη πρόταση ,

Πρόταση 1.15 Η συνάρτηση φ είναι κυρτή αν και μόνο αν

$$(\mathrm{y}-\mathrm{x})^T \nabla\varphi(\mathrm{x}) \le \varphi(\mathrm{y}) - \varphi(\mathrm{x}) \quad \forall x, y \in \mathbb{R}^n .$$

Απόδειξη : Έστω ότι ισχύει η $(\mathrm{y}-\mathrm{x})^T \nabla\varphi(\mathrm{x}) \le \varphi(\mathrm{y}) - \varphi(\mathrm{x})$, τότε για τυχαία $y, z \in \mathbb{R}^n$, $\lambda \in (0,1)$ και $x = \lambda y + (1-\lambda)\,\mathrm{z}$ ισχύει

$$(\mathrm{y}-\mathrm{x})^T \nabla\varphi(\mathrm{x}) \le \varphi(\mathrm{y}) - \varphi(\mathrm{x})$$
$$(\mathrm{z}-\mathrm{x})^T \nabla\varphi(\mathrm{x}) \le \varphi(\mathrm{z}) - \varphi(\mathrm{x})$$

Το οποίο συνεπάγεται

$$\underbrace{(\lambda\,\mathrm{y}-(1-\lambda)\,\mathrm{z}-\mathrm{x})^T}_{=0} \nabla\varphi(\mathrm{x}) \le \lambda\varphi(\mathrm{y}) + (1-\lambda)\varphi(\mathrm{z}) - \varphi(\mathrm{x})$$

δηλαδή η κυρτότητα της φ.

Αν υποθέσουμε ότι η φ είναι κυρτή, τότε για τυχαία $x, y \in \mathbb{R}^n$ και $\lambda \in (0,1)$, από το θεώρημα μέσης τιμής με $\theta_\lambda \in (0,1)$ έχουμε

$$\varphi(\mathrm{y})-\varphi(\mathrm{x})\geq\frac{1}{1-\lambda}\{\varphi(\mathrm{x}+(1-\lambda)(\mathrm{y}-\mathrm{x}))-\varphi(\mathrm{x})\}$$

$$=\frac{1}{1-\lambda}\{(1-\lambda)(\mathrm{y}-\mathrm{x})^{T}\nabla\varphi(\mathrm{x}+\theta_{\lambda}(1-\lambda)(\mathrm{y}-\mathrm{x}))\}$$
$$=(\mathrm{y}-\mathrm{x})^{T}\nabla\varphi(\mathrm{x}+\theta_{\lambda}(1-\lambda)(\mathrm{y}-\mathrm{x}))$$

για $\lambda\to 1$.▲

Προκειμένου να ορίσουμε τις συνθήκες Βελτιστοποίησης και για το Μη Γραμμικό πρόβλημα, θεωρούμε την βέλτιστη λύση $\hat{x}\in B$ για την οποία θα ισχύει

- $g_i(\hat{x}+\alpha z)\leq 0$ για $i\notin I(\mathrm{x}):=\{i\,|\,\mathrm{g}_i(\mathrm{x})=0\}$, $z\in\mathbb{R}^n$ και α = πολύ μικρή θετική τιμή
- $\forall i\in I(\mathrm{x})\ \exists\bar{\alpha}>0$ τ. ώ. $\forall\alpha\in(0,\tilde{\alpha})$, $g_i(\hat{x}+\alpha\tilde{z})\leq 0$ όπου $\tilde{z}\in\mathbb{R}^n$, $\tilde{z}^{T}\nabla g_i(\mathrm{x})<0$.

Θεωρώντας ότι το $\hat{x}$ είναι το ελάχιστο της f προκύπτει ότι $f(\hat{x}+\alpha\tilde{z})-f(\hat{x})\geq 0\forall\alpha\in(0,\tilde{\alpha})$.

Έτσι προκύπτουν οι παρακάτω συνθήκες.

1.3.1. Συνθήκες Βελτιστοποίησης

Συνθήκη Κανονικότητας

Για κάθε βέλτιστη λύση του Μη γραμμικού προβλήματος ισχύει ότι

$z^{T}\nabla g_i(\hat{x})\leq 0\ \ \forall i\in I(\hat{x})$ το οποίο συνεπάγεται $z^{T}\nabla f(\hat{x})\geq 0$.

Παρατήρηση

Η παραπάνω συνθήκη ικανοποιείται για κάθε γραμμικό περιορισμό, ωστόσο, για προβλήματα μη γραμμικών περιπτώσεων δεν ισχύει.

Επίσης, η αδυναμία να ελεγχθεί η ισχύς της συνθήκης αυτής λόγω της απαραίτητης εκ των προτέρων γνώσης της βέλτιστης λύσης, την καθιστά δύσκολη στην εφαρμογή.

Οπότε, για κυρτά, κυρίως, προβλήματα με μη γραμμικούς περιορισμούς έχει καθιερωθεί η χρήση μιας άλλης συνθήκης, της συνθήκης Slater.

Συνθήκη Slater

Για το μη γραμμικό πρόβλημα

$$\left.\begin{array}{l}\min f(\mathrm{x}) \\ \text{όταν } \mathrm{g}_i(\mathrm{x}) \le 0, \mathrm{i}=1,\ldots,\mathrm{m}\end{array}\right\}$$

υπάρχει μία εφικτή λύση $\tilde{x}$ για την οποία ισχύει $g_i(\tilde{x}) < 0 \quad \forall i$.

*Συνθήκες Karush-Kuhn-Tucker (KKT***)**

Πρόταση 1.16 Θεωρώντας ότι η συνθήκη Slater ικανοποιείται για το κυρτό μη γραμμικό πρόγραμμα, ισχύει ότι η $\hat{x} \in B$ επαληθεύει το πρόβλημα μη γραμμικού προγραμματισμού αν και μόνο αν υπάρχει $\hat{u} \in \mathbb{R}^m$ για το οποίο ισχύουν οι παρακάτω συνθήκες :

$$\left.\begin{aligned} &i)\nabla f(\hat{x}) + \sum_{i=1}^{m} \hat{u}_i \nabla g_i(\hat{x}) = 0 \\ &ii) g_i(\hat{x}) \le 0 \forall i \\ &iii) \hat{u}_i\, g_i(\hat{x}) = 0 \forall i \\ &iv) \hat{u} \ge 0 \end{aligned}\right\}$$

Αν και οι συνθήκες αυτές προσαρμόζονται ανάλογα με την μορφή του μη γραμμικού προβλήματος παρουσιάζει μεγάλο ενδιαφέρον η ερμηνεία τους για την απλή μορφή.

Εφαρμόζοντας στο πρόβλημα την συνάρτηση Lagrange :

$$L(\mathrm{x},\mathrm{u}) := \mathrm{f}(\mathrm{x}) + \sum_{i=1}^{m} u_i g_i(\mathrm{x})$$

προκύπτει ότι

- για σταθερό $\tilde{u} \ge 0$ η $\mathrm{L}(\cdot, \tilde{u})$ είναι κυρτή στο x και επιπλέον, από την (i) συνθήκη έχουμε ότι $\mathrm{L}(\hat{x}, \hat{u}) = \min_{x \in \mathbb{R}^n} \mathrm{L}(\mathrm{x}, \hat{u})$.
- για σταθερό $\tilde{x}$ η συνάρτηση $L(\tilde{x}, \cdot)$ είναι γραμμικά συναφής και κατά συνέπεια κοίλη στο u. Από τις συνθήκες (ii) & (iii) προκύπτει ότι $L(\hat{x}, \mathrm{u}) = \max_{u \ge 0} L(\hat{x}, u)$.

Το σημείο $(\mathrm{x}, \hat{u} \ge 0)$ που ικανοποιεί τις συνθήκες ΚΚΤ αποτελεί το σαγματικό σημείο για την συνάρτηση Lagrange :

$$\mathrm{L}(\hat{\mathrm{x}}, \mathrm{u}) \le L(\mathrm{x}, \hat{u}) \le L(\mathrm{x}, \hat{u}) \quad \forall u \ge 0,\ \forall x \in \mathbb{R}^n .$$

Θεώρημα Σαγματικού Σημείου (*Δυικό Θεώρημα για προβλήματα Μη Γραμμικού Προγραμματισμού*) Αν για την συνάρτηση Lagrange στο μη γραμμικό πρόγραμμα υπάρχει σημείο καμπής $(\hat{x}, \hat{u} \geq 0)$, τότε

$$\max_{u \geq 0} \inf_{x \in \mathbb{R}^n} L(x, u) = \min_{x \in \mathbb{R}^n} \sup_{u \geq 0} L(x, u)$$

και $(\hat{x}, \hat{u})$ αποτελεί λύση και του αρχικού προβλήματος $\min_{x} \sup_{u \geq 0} L(x, u)$ αλλά και του δυικού του $\max_{u \geq 0} \inf_{x} L(x, u)$.

1.3.2. Μέθοδοι Επίλυσης

Πολλές προσεγγιστικές μέθοδοι έχουν αναπτυχθεί για την επίλυση μη γραμμικών προγραμμάτων. Μερικές από τις κατηγορίες που μπορούμε να τις συνοψίσουμε είναι οι εξής :

- Μέθοδοι Τεμνόμενων Επιπέδων
- Μέθοδοι Εφικτών Κατευθύνσεων
- Μέθοδοι του Lagrange

Προκειμένου να εφαρμόσουμε την μέθοδο που χρησιμοποιεί την τομή των επιπέδων στον στοχαστικό γραμμικό προγραμματισμό, θεωρούμε αρχικά το μη γραμμικό πρόβλημα

$$\min_{x \in \mathbb{R}^n} \{ f(x) \,|\, g_i(x) \leq 0, i = 1, \ldots, m \}$$

με $B = \{ x \,|\, g_i(x) \leq 0, i = 1, \ldots, m \}$.

Ή γενικότερα, $\min\{ c^T x \,|\, x \in B \}$ με x εσωτερικό σημείο του B. Τότε, υπάρχει ένα κυρτό πολύγωνο P το οποίο περιέχει το σύνολο B, $P \supset B$.

Πιο συγκεκριμένα, υπάρχουν γραμμικοί περιορισμοί που ορίζουν το εφικτό σύνολο P και ισχύει : $\min_{x\in P} c^T x \le \min_{x\in B} c^T x$.

Αλγόριθμος Τομής Επιπέδων

Βήμα 1 : Βρες ένα εσωτερικό σημείο $\tilde{x} \in B$, το κυρτό πολύεδρο $P_0 \supset B$ και θέσε k:=0.

Βήμα 2 : Λύσε το γραμμικό πρόβλημα $\min\{c^T x \,|\, x \in P_k\}$ συμβολίζοντας την λύση με $\hat{x}^{(k)}$.

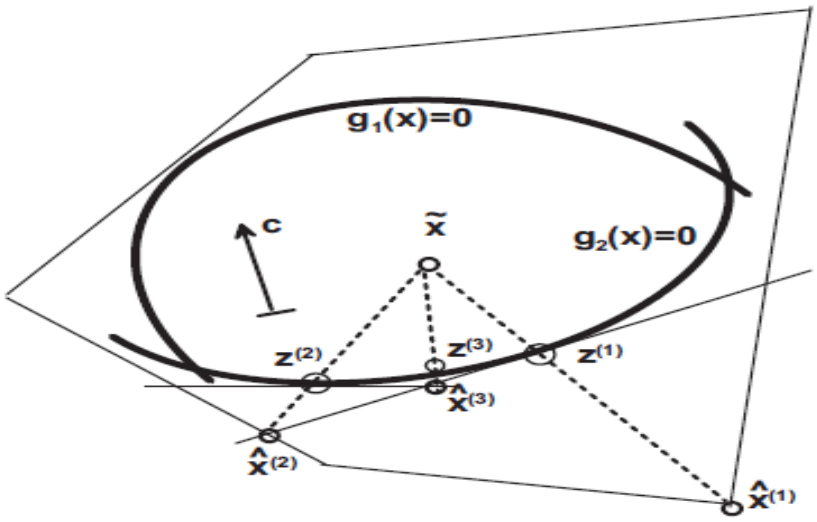
Αν $\hat{x}^{(k)} \in B$ τότε σταμάτα και η $\hat{x}^{(k)}$ είναι η λύση του μη γραμμικού προβλήματος.

Αλλιώς, βρες το σημείο τομής $z^{(k)} \in [\hat{x}^{(k)}, \tilde{x}] \cap \mathrm{bd}B$ της ευθείας που ενώνει τα $\hat{x}^{(k)}, \tilde{x}$ με το σύνορο του B.

Βήμα 3 : Προσδιόρισε το εφαπτόμενο υπερεπίπεδο H_k του B, δηλαδή να βρεθεί ένα $\alpha^{(k)} \in \mathbb{R}^n$ όπου $\alpha_k = \alpha^{(k)^T} z^{(k)}$ τέτοιο ώστε :

$$H_k := \{x \,|\, \alpha^{(k)^T} x = \alpha_k\} \text{ και } \alpha^{(k)^T} \hat{x}^{(\kappa)} > \alpha_k \ge \alpha^{(k)^T} x \quad \forall x \in B.$$

Βρες $P_{k+1} := P_k \cap \{x \,|\, \alpha^{(k)^T} x \le \alpha_k\}$ και θέσε k:=k+1 και επέστρεψε στο "Βήμα 2" .



Εικόνα 4 : Cutting plane method

Πρόταση 1.17 Σύμφωνα με τα παραπάνω, τα σημεία συσσώρευσης $\{\hat{x}^{(k)}\}$ και $\{z^{(k)}\}$ επαληθεύουν το μη γραμμικό πρόβλημα $\min\{c^T x \mid x \in B\}$. Επιπλέον, οι τιμές της αντικειμενικής συνάρτησης, $\{c^T \hat{x}^{(k)}\}, \{c^T z^{(k)}\}$, συγκλίνουν στο $\min\{c^T x \mid x \in B\}$. Τέλος, για κάθε επανάληψη ορίζουμε έναν εκτιμητή σφάλματος για την βέλτιστη λύση δ, ως

$$\Delta_k = \min_{l=1,\ldots,k} c^T z^{(l)} - c^T \hat{x}^{(k)} .$$

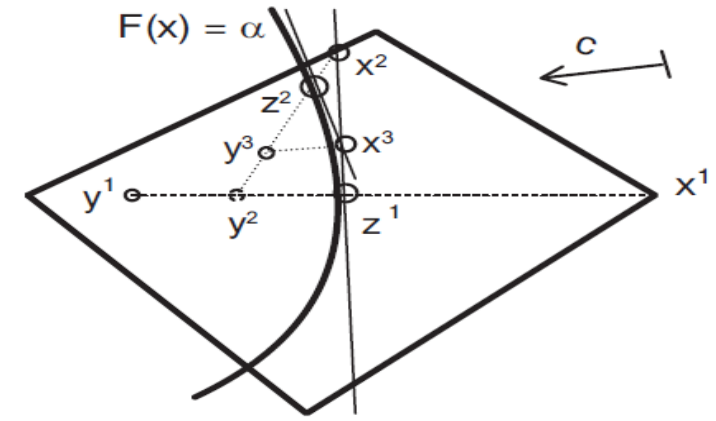
Η παραπάνω έμμεση επίλυση του αρχικού προβλήματος αποτελεί μια περίπτωση της μεθόδου των Τεμνόμενων Επιπέδων. Στην συνέχεια θα παρουσιάσουμε περιγραφικά και άλλες περιπτώσεις που επιλύουν προβλήματα πιθανοτικών (τυχαίων) περιορισμών όπως του παρακάτω :

$$\min c^T x$$

όταν $\alpha^{(i)^T} x \geq b_i \quad i = 1,\ldots,m$

$$F(x) \geq \alpha$$

Μέθοδος Τομής Επιπέδων με μεταβλητό σαγματικό σημείο



7 Outer linearization with moving Slater points.

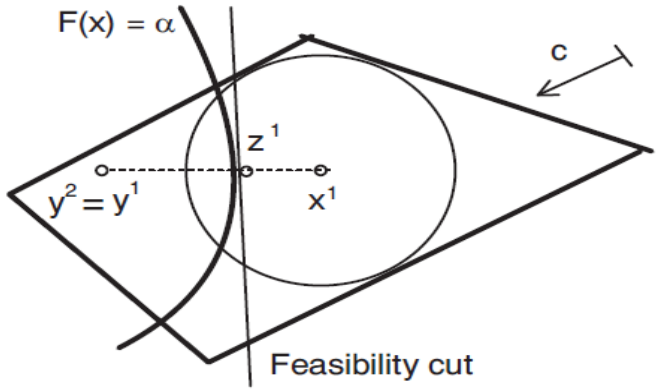
Εικόνα 5 : Slater points

Παρατήρηση

Στην περίπτωση αυτή το εσωτερικό σημείο ποικίλει σε κάθε επανάληψη σύμφωνα με το σύνολο $\{x \mid \mathrm{F}(\mathrm{x}) \geq \alpha\}$ και την παραπάνω εικόνα. Έτσι για κάθε κυρτό σύνολο D , $y \in \operatorname{int} D$ και για $z \in bdD$ ισχύει $\lambda z + (1-\lambda)\,\mathrm{y} \in (0,1)$.

Επίσης, στην περίπτωση αυτή βελτιώνεται ο βαθμός σύγκλισης του αλγορίθμου.

Μέθοδος κεντρικής τομής επιπέδων



8 Central cutting plane method: Outer cut.

Εικόνα 6 : Τομή Επιπέδων

Παρατήρηση

Η βασική ιδέα της περίπτωσης αυτής της μεθόδου σχετίζεται με την μορφή της ισότητας $d^T x = \rho$ που χαρακτηρίζεται ως κανονική όταν $\|d\| = 1$ (Ευκλείδεια Απόσταση). Τότε θα ισχύει $\sigma = d^T y - \rho$ με $|\sigma|$ η Ευκλείδεια απόσταση του y από το υπερεπίπεδο $\{x | \mathrm{d}^T \mathrm{x} = \rho\}$. Έτσι κάθε τυχαίο $\alpha^T \mathrm{x} = b$ με $\alpha \neq 0$ παίρνει την κανονική μορφή $\frac{\alpha^{\mathrm{T}}}{\|\alpha\|} x = \frac{b}{\|\alpha\|}$.

Τότε κάθε γραμμικό πρόβλημα της μορφής $\max\{\eta | \mathrm{d}^{(\mathrm{i})^T} \mathrm{x} - \|d^{(\mathrm{i})}\| \eta \geq \rho_i, \mathrm{i} \in \mathrm{I}\}$ (όπου η η ακτίνα του εγγεγραμμένου κύκλου στο πολύεδρο) θα επιλύεται και θα δίνει το κέντρο $\hat{x}$ του πολυέδρου $\{x | \mathrm{d}^{(\mathrm{i})^T} \geq \rho_i, \mathrm{i} \in \mathrm{I}\}$.

ΚΕΦΑΛΑΙΟ 2ο «ΓΡΑΜΜΙΚΑ ΣΤΟΧΑΣΤΙΚΑ ΠΡΟΓΡΑΜΜΑΤΑ ΕΝΟΣ ΣΤΑΔΙΟΥ»

Σε αυτό το κεφάλαιο θα ασχοληθούμε με ποικίλα προβλήματα γραμμικού στοχαστικού προγραμματισμού ενός σταδίου και θα συζητήσουμε τις θεωρητικές τους ιδιότητες καθώς και την υπολογιστική τους επιλυσιμότητα.

2.1. Εισαγωγή

Τα προβλήματα ενός σταδίου αναφέρονται στα προβλήματα όπου η απόφαση παίρνεται ακαριαία και που καμία διόρθωση στις τυχαίες μεταβλητές δεν λαμβάνεται υπόψη.

Τα προβλήματα αυτά αποτελούνται από τυχαίες μεταβλητές, είτε στους περιορισμούς είτε και στην αντικειμενική συνάρτηση, οι οποίες είναι της μορφής :

$$\zeta(\mathrm{x},\xi) := \mathrm{T}(\xi)\,\mathrm{x} - \mathrm{h}(\xi)$$

Όπου $\xi : \Omega \to \mathbb{R}^r$ τυχαίο διάνυσμα του χώρου πιθανότητας $(\Omega, \mathrm{F}, \mathrm{P})$.

$T(\xi)$ τυχαίος πίνακας μεγέθους $s \times n$

$h(\xi) \in \mathbb{R}^s$ τυχαίο διάνυσμα

Το ξ καθορίζεται ως το μικρότερο κλειστό σύνολο του $\Xi \subset \mathbb{R}^r$ που έχει την ιδιότητα $P(\xi \in \Xi) = 1$.

Η εξάρτηση των $T(\xi)$, $h(\xi)$ από το ξ είναι τελείως γραμμικές , όπως φαίνεται και παρακάτω :

$$T(\xi) = \mathrm{T} + \sum_{j=1}^{r} T_j \xi_j$$

$$h(\xi) = \mathrm{h} + \sum_{j=1}^{r} h_j \xi_j$$

όπου T , $T_j \in \mathbb{R}^{s \times n}$ πίνακες και $h, h_j \in \mathbb{R}^s$ διανύσματα με $j = 1, \ldots, r$.

Αυτό που θα μας είναι χρήσιμο παρακάτω όμως είναι η από κοινού συνάρτηση πιθανότητας $(\mathrm{T}(\xi), \mathrm{h}(\xi))$ η οποία όπως φαίνεται δεν εξαρτάται από το x, συνεπώς η απόφασή μας δεν επηρεάζει την τιμή της πιθανότητας.

Θεωρώντας ότι η $\zeta(\mathrm{x}, \xi)$ έχει τον ρόλο της χαλαρής μεταβλητής σε περιορισμούς με μορφή γραμμικής ανίσωσης, όπως

$$\mathrm{T}(\xi)\,\mathrm{x} \geq h(\xi) \Rightarrow \zeta(\mathrm{x}, \xi) \geq 0$$

το πρόβλημά μας μπορεί να γραφεί είτε σε μορφή γραμμής

$$\zeta_i(\mathrm{x}; \xi) := \mathrm{t}_i^T(\xi)\,\mathrm{x} - \mathrm{h}_i(\xi), \mathrm{i} = 1, \ldots, \mathrm{s}$$

είτε στήλης $\zeta(\mathrm{x}; \xi) = \sum_{j=1}^{n} T_j(\xi)\,\mathrm{x}_j - \mathrm{h}(\xi), \mathrm{j} = 1, \ldots, \mathrm{n}$.

Συνεπώς, η $\zeta(\mathrm{x}, \xi)$ μπορεί να θεωρηθεί ως ένας γραμμικός συνδυασμός τυχαίων διανυσμάτων. Στον γραμμικό μας όμως συνδυασμό οι συντελεστές είναι οι μεταβλητές απόφασης x_j , οπότε η συνάρτηση πιθανότητας εξαρτάται από τη απόφασή μας.

Στον Στοχαστικό Προγραμματισμό όμως το βασικό σημείο είναι η δοσμένη από κοινού κατανομή ($T_j(\xi), \mathrm{j} = 1, \ldots, \mathrm{n}; \mathrm{h}(\xi)$) με στόχο την δημιουργία μιας κατανομής πιθανότητας της $\zeta(\mathrm{x}, \xi)$ (x αιτιοκρατική μεταβλητή) με ωφέλιμες ιδιότητες εξαρτώμενες από την ποσοτική τους σημασία. Αυτό επιτυγχάνεται κατατάσσοντας τα διάφορα μοντέλα Γραμμικού Στοχαστικού Προγραμματισμού ως ακολούθως :

- Αρχικά θεωρούμε μια συνάρτηση $\rho : \mathrm{Y} \to \mathbb{R}^1$, με Y μερικώς γραμμικός χώρος τυχαίου διανύσματος διάστασης s, που καθορίζεται από τον χώρο πιθανοτήτων $(\Omega, \mathrm{F}, \mathrm{P})$, η οποία θα αξιολογεί τα τυχαία διανύσματα. Έτσι, για τυχαίο διάνυσμα $\vartheta \in Y$, η $\rho(\vartheta)$ ερμηνεύεται ως το μέτρο ποιότητας της προσέγγισης του αντίστοιχου μοντέλου. Γενικότερα, η ρ είναι το μέτρο ρίσκου.
- Ομοίως αξιολογείται και η μεταβλητή απόφασης x, θεωρούμε την συνάρτηση $V : \mathbb{R}^n \to \mathbb{R}^1$ αντικαθιστώντας την $\zeta(\mathrm{x}, \xi)$ με την ρ. Έτσι, $V(\mathrm{x}) := \rho(\zeta(\mathrm{x}, \xi))$ για $\zeta(\mathrm{x}, \xi) \in \mathrm{Y}$ για κάθε x η οποία αποτελεί ποιοτικό μέτρο, χρησιμοποιείται για την δημιουργία του μοντέλου και καλείται από κοινού συνάρτηση αξιολόγησης.
- Εναλλακτικά, εάν πρόκειται για τους περιορισμούς, πρέπει να ορίσουμε τα μέτρα ποιότητας των στοιχείων της $\zeta(\mathrm{x}, \xi)$ μεμονωμένα. Έτσι, η $V_i(\mathrm{x}) := \rho(\zeta_i(\mathrm{x}, \xi))$ αξιολογεί την x στο i-οστό στοιχείο της $\zeta(\mathrm{x}, \xi)$, με i=1,…,s και καλείται διαχωρίσιμη συνάρτηση αξιολόγησης. Για s=1 και οι δύο έννοιες «από κοινού» και «διαχωρίσιμη» ισχύουν.

Συνεπώς, επιλέγοντας το κατάλληλο μέτρο ρ θα παράγεται ένα διαφορετικής κατηγορίας μοντέλο Γ.Σ.Π.

Από την άλλη πλευρά, θεωρώντας $\zeta(\mathrm{x}, \xi)$ την απόκλιση μεταξύ των $\mathrm{T}(\xi)\,\mathrm{x}$ και $h(\xi)$, όπου από την Στατιστική προσεγγίζουμε την έννοια των ελαχίστων τετραγώνων. Έτσι, η $\zeta(\mathrm{x}, \xi)$ αποτελεί το σφάλμα. Οπότε, κινούμενοι όπως προηγουμένως θέλοντας ένα μέτρο ποιότητας, δηλαδή τώρα το μέτρο απόκλισης, μπορούμε να κατασκευάσουμε τα διάφορα πρωτότυπα μοντέλα Στοχαστικού Γραμμικού Προγραμματισμού χρησιμοποιώντας την συνάρτηση αξιολόγησης V ανάλογα με το μέτρο ρ.

Τα μοντέλα έχουν ως εξής :

$$\left.\begin{array}{l} \max c^T x \\ V(\mathrm{x}) \geq \mathrm{k} \\ x \in B \end{array}\right\} \quad \text{και} \quad \left.\begin{array}{l} \max c^T x \\ V_i(\mathrm{x}) \geq \mathrm{k}_i, i = 1, \ldots, s \\ x \in B \end{array}\right\}$$

όπου k, k_i είναι προκαθορισμένες τιμές και B το σύνολο πολυέδρων :

$$B = \{\mathrm{x} \,|\, Ax \propto b, l \le x \le u\}$$

με A ένας πίνακας διάστασης $m \times n$ και x, b, l, u αντίστοιχων διαστάσεων.

❖ Η έννοια του συμβόλου $\propto$ υποδηλώνει τη σχέση ισότητας/ανισότητας.

Όταν όμως η συνάρτηση αξιολόγησης είναι στον στόχο, το πρωτότυπο μοντέλο είναι της μορφής :

$$\begin{aligned} &\max c^T x + V(\mathrm{x}) \\ &\mathrm{x} \in \mathrm{B} \end{aligned}$$

Η μορφή αυτή εξυπηρετεί την εφαρμογή των σύνθετων μοντέλων με εύκολο τρόπο.

Είναι φανερό πως η αντικειμενική συνάρτηση αποτελείται από ένα άθροισμα δύο όρων διαβαθμίζοντας την βαρύτητά τους. Θέτοντας λ το βάρος (λ≥0), έχουμε :

$$\left.\begin{aligned} &\max c^T x + \lambda V(\mathrm{x}) \\ &\mathrm{x} \in \mathrm{B} \end{aligned}\right\}$$

Η διαβάθμιση αυτή μπορεί εύκολα να ερμηνευτεί με την δυικότητα, δηλαδή :

$$\left.\begin{aligned} &\mathrm{v}(\lambda) := \max c^T x + \lambda(V(\mathrm{x}) - \mathrm{k}) \\ &\mathrm{x} \in \mathrm{B} \end{aligned}\right\}$$

με ‘λk’ να αποτελεί μια μετατόπιση από την βέλτιστη αντικειμενική τιμή.

Το παραπάνω πρόβλημα ονομάζεται χαλάρωση Lagrangian του προβλήματος

$$\left.\begin{aligned} &\max c^T x \\ &V(\mathrm{x}) \ge \mathrm{k} \\ &x \in B \end{aligned}\right\}$$

ενώ το αντίστοιχο δυικό του, δυικό κατά Lagrange $\min\{\upsilon(\lambda) \,|\, \lambda \ge 0\}$.

Ωστόσο, ο απλούστερος τρόπος να αποδώσεις ένα μέτρο ποιότητας στην $\zeta(\mathrm{x},\xi)$ είναι να πάρεις τον εκτιμητή της. Έτσι, για $p_{\mathrm{E}}(\vartheta) := \mathrm{E}(\vartheta), \vartheta \in L_1^1,\ k_i = 0 \forall i$ και θεωρώντας ότι υπάρχουν οι εκτιμητές των $\mathrm{T}(\xi), \mathrm{h}(\xi)$ με

$$\bar{t}_i := E_\xi[\mathrm{t}_i(\xi)], \bar{h}_i := E_\xi[\mathrm{h}_i(\xi)],$$

το μοντέλο παίρνει την μορφή :

$$\left.\begin{array}{l} \max c^T x \\ \bar{t}_i^T x \geq \bar{h}_i, i = 1,...,n \\ x \in B \end{array}\right\}$$

όπου τα στοιχεία του $\mathrm{t}_i(\xi)$ είναι στοιχεία της i-οστής γραμμής του $\mathrm{T}(\xi)$.

Το παραπάνω πρόβλημα είναι γνωστό ως πρόβλημα αναμενόμενης αξίας.

Εν συνεχεία, αν συμπτύξουμε όλους τους περιορισμούς ανισώσεων του προβλήματος $\left.\begin{array}{l} \max c^T x \\ V_i(\mathrm{x}) \geq \mathrm{k}_i, i = 1,...,s \\ x \in B \end{array}\right\}$ στην μορφή $\mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)$, έχουμε το Στοχαστικό Γραμμικό Πρόβλημα :

$$\begin{array}{l} \max c^T x \\ T(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi), \xi \in \Xi \\ x \in \mathrm{B} \end{array}$$

του οποίου η λύση ονομάστηκε από τον Madansky ως “fat solution” με αντίστοιχο μέτρο ποιότητας $\rho_{fat}(\vartheta) := \min_{\hat{\vartheta} \in \Theta} \min_{1 \leq i \leq s} \hat{\vartheta}_i, \vartheta \in L_s^\infty$.

Η εφικτή περιοχή αποτελείται από τομές κυρτών συνόλων, συνεπώς αποτελεί επίσης ένα κυρτό σύνολο, $D = \bigcap_{\xi \in \Xi} \{x \big| \mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi), \mathrm{x} \in \mathrm{B}\}$.

Το παραπάνω πρόβλημα μειώνεται σε γραμμικό πρόβλημα όταν το σύνολο Ξ είναι πεπερασμένο, δηλαδή στην περίπτωση πεπερασμένης διακριτής κατανομής.

Προκειμένου να γενικεύσουμε το παραπάνω μοντέλο, θεωρούμε το x ως εφικτή λύση που ικανοποιεί όλους τους ανισοτικούς περιορισμούς των υποσυνόλων του Ξ δεδομένου ότι αυτά παίρνουν συγκεκριμένες τιμές πιθανότητας.

Τότε το επίπεδο ποιότητας είναι $\rho_P(\vartheta) := \mathrm{P}(\vartheta \geq 0), \vartheta \in \mathrm{L}_s^0$

και ορίζεται σε όλα τα τυχαία διανύσματα του χώρου πιθανοτήτων $(\Omega, \mathrm{F}, \mathrm{P})$.

Για το συγκεκριμένο μοντέλο η συνάρτηση υπολογισμού V(x) μπορεί να αντικατασταθεί από την συνάρτηση πιθανοτήτων $G(\mathrm{x}) := P_\xi(\mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi))$ με $G(\mathrm{x}) \geq \alpha$, όπου α η μέγιστη τιμή πιθανότητας (≈ 0.99). Το μοντέλο πλέον παίρνει την μορφή :

$$\left.\begin{array}{l} \max c^T x \\ P_\xi(\mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)) \geq \alpha \\ x \in B \end{array}\right\}$$

Ωστόσο, όταν ισχύει $G(\mathrm{x}) \leq \beta$ με $\beta \approx 0.01$, έχουμε να αντιμετωπίσουμε ένα πρόβλημα με πιθανότητα ζημίας το οποίο θέλουμε να το ελαχιστοποιήσουμε.

Συνεχίζοντας με τα μοντέλα με υψηλές τιμές του α, θα αναφέρουμε κάποια σημαντικά χαρακτηριστικά τους. Από την μία πλευρά, τα μοντέλα αυτά δίνουν πολύ μικρή πιθανότητα $(=1-\alpha)$ εμφάνισης ζημίας, ενώ από την άλλη πλευρά, στην περίπτωση εμφάνισης αυτής, το μοντέλο στέκεται αδύναμο στην μοντελοποίηση του μεγέθους της. Είναι σαφές ότι το δεύτερο χαρακτηριστικό του μοντέλου αποτελεί και μειονέκτημα αυτού. Στα παραπάνω χαρακτηριστικά στηρίζεται και η διάκριση των μοντέλων σε ποιοτικά μοντέλα, δηλαδή αυτών που βασίζονται σε πιθανοτικούς περιορισμούς και σε ποσοτικά, αυτών δηλαδή που υπολογίζουν το μέγεθος της ζημίας.

Στην πορεία, θα συμβολίζουμε την ζημία με τις αρνητικές τιμές της τυχαίας μεταβλητής $\zeta(\mathrm{x},\xi)$, δηλαδή

$$\zeta^{-}(\mathrm{x},\xi) := (\mathrm{t}^{T}(\xi)\,\mathrm{x} - \mathrm{h}(\xi))^{-}$$

με $t(\xi)$ το διάνυσμα γραμμή του $\mathrm{T}(\xi)$, $h(\xi)$ τυχαία μεταβλητή και όπου $z^{-} = \max\{0, -\mathrm{z}\}$ τα αρνητικά στοιχεία του $z \in \mathbb{R}$,

ενώ οι θετικές τιμές θα συμβολίζουν τα κέρδη.

Έτσι, ο περιορισμός $G(\mathrm{x}) \geq \alpha$ εκτιμώμενος από την αναμενόμενη τιμή του μπορεί να γραφεί ως

$$G(\mathrm{x}) \geq \alpha \Leftrightarrow E_{\xi}[\chi(\zeta^{-}(\mathrm{x},\xi))] \leq 1-\alpha$$

όπου χ η δείκτρια συνάρτηση $\chi(\mathrm{z}) = \begin{cases} 0, & z \leq 0 \\ 1, & z > 0 \end{cases}$ αν.

Η συνάρτηση όμως αυτή προάγει ισότητες μεταξύ διαφορετικών ζημιών. Για να ξεπεραστεί αυτή η αδυναμία θεωρούμε το μέτρο

$$\rho_{sic}^{-}(\vartheta) := \mathrm{E}[\vartheta^{-}], \vartheta \in \mathrm{L}_1^1$$

Το οποίο οδηγεί σε μια νέα μορφή της συνάρτησης υπολογισμού, αυτή της αναμενόμενης τιμής των τυχαίων μεταβλητών ζημίας $H(\mathrm{x}) := \mathrm{E}_{\xi}[\zeta^{-}(\mathrm{x},\xi)]$.

Θέτοντας γ την μέγιστη τιμή ανεκτής απώλειας, οι περιορισμοί θα έγκεινται στην μορφή $H(\mathrm{x}) \leq \gamma$ και έτσι μπορεί να πάρει το μοντέλο μας την μορφή ελαχιστοποίησης

$$\begin{aligned} &\min c^{T} x \\ &E_{\xi}[\zeta^{-}(\mathrm{x},\xi)] \leq \gamma \\ &x \in B \end{aligned}$$

Στην συνέχεια, θα εξετάσουμε την περίπτωση όπου οι θετικές τιμές της μεταβλητής $\zeta(\mathrm{x},\xi)$ θα συμβολίζουν τις ζημίες και οι αρνητικές τα κέρδη. Η χρήση των ποσοστιμορίων φάνηκε σαν λύση για τον υπολογισμό του μέτρου πιθανότητας της ζημίας. Έτσι, για δεδομένη τιμή του α $(0<\alpha<1)$, το μέτρο ποιότητας θα είναι

$$\rho_{\mathrm{var}}^{\alpha}(\vartheta) := \nu(\vartheta,\alpha) := \min\{\mathrm{z}\,|\,F_{\vartheta}(\mathrm{z}) \geq \alpha\}\vartheta \in \mathrm{L}_1^0$$

το οποίο και αποτελεί το αριστερό άκρο του κλειστού διαστήματος των α-ποσοστιμορίων του ϑ και είναι γνωστό ως Επίπεδο Κινδύνου (Value at Risk).

Τότε η συνάρτηση υπολογισμού είναι:

$$\upsilon(\mathrm{x},\alpha) := \min\{\mathrm{z}\,|\,\Psi(\mathrm{x},\mathrm{z}) \geq \alpha\}$$

με $\Psi(\mathrm{x},\bullet)$ η συνάρτηση πιθανότητας για κάθε σταθερό x και $\alpha = 0.95$.

Η συνάρτηση αυτή γενικά δεν είναι κυρτή και γι' αυτό είναι αδύνατη η κατασκευή ενός μοντέλου αριθμητικά προσεγγίσιμο που βασίζεται στο επίπεδο α .

Ωστόσο, ένας συνδυασμός του επίπεδου κινδύνου και της δεσμευμένης αναμενόμενης τιμής θα φάνταζε μια ιδανική προσέγγιση της παραπάνω κατασκευής. Το μέτρο ποιότητας θα ήταν

$$\rho_{CVaR}^{\alpha}(\vartheta) := \min_{z}[\mathrm{z} + \frac{1}{1-\alpha}\mathrm{E}[(\vartheta-\mathrm{z})^{+}]], \vartheta \in \mathrm{L}_1^1$$

το οποίο εξυπηρετεί δύο σκοπούς, από την μία πλευρά από το θεώρημα πιθανότητας το σύνολο λύσεων του παραπάνω προβλήματος συμπίπτει με το σύνολο των α-ποσοστιμορίων της συνάρτησης του ϑ και από την άλλη πλευρά δεδομένης της συνεχούς συνάρτησης του ϑ έχουμε :

$$\rho_{CVaR}^{a}(\vartheta) = E[\vartheta\,|\,\vartheta \geq \nu(\vartheta,a)], \vartheta \in L_1^1$$

όπου $\nu(\vartheta,a)$ είναι το VaR. Συνεπώς, το μέτρο αυτό αποτελεί μια εκτίμηση της ζημίας δεδομένου ότι η ζημία υπερβαίνει το VaR.

Έτσι, ορίζουμε την συνάρτηση υπολογισμού με κυρτές ιδιότητες

$$\upsilon_c(\mathrm{x},\alpha) := \rho^a_{CVaR}(\zeta(\mathrm{x},\xi))$$

Συνεπώς το αρχικό πρόβλημα θα περιλαμβάνει ανισοτικούς περιορισμούς της μορφής $\upsilon_c(\mathrm{x},\alpha) \le \gamma$ δηλαδή η συνάρτηση περιορισμού θα ελαχιστοποιείται. Επιπλέον, δεδομένης μιας πεπερασμένης διακριτής συνάρτησης, ένα πρόβλημα βελτιστοποίησης μειώνεται σε γραμμικό πρόβλημα.

Στην συνέχεια θα ασχοληθούμε με το ερώτημα που έχει γεννηθεί έχοντας υπολογίσει τόσα μέτρα ποιότητας, δηλαδή με το ποια η αριθμητική απόδοση της συμπεριφοράς των στοχαστικών προβλημάτων βελτιστοποίησης βασισμένα σε αυτά τα μέτρα.

Για την εύρεση μιας αποτελεσματικής αριθμητικής λύσης βασική ιδιότητα ενός μη γραμμικού προβλήματος βελτιστοποίησης είναι να αποτελεί κυρτό γραμμικό πρόβλημα. Τότε μπορούμε να κατασκευάσουμε ένα πρωτότυπο πρόβλημα των προβλημάτων

$$\begin{array}{l} \max c^T x \\ V_i(\mathrm{x}) \ge \mathrm{k}_i, i = 1,...,s \\ x \in B \end{array} \quad \& \quad \begin{array}{l} \max c^T x + V(\mathrm{x}) \\ x \in B \end{array}$$

τα οποία μπορούν να θεωρηθούν κυρτά γραμμικά προβλήματα όταν οι V και V_i αποτελούν κοίλες συναρτήσεις.

Παρακάτω δίνονται κάποιοι ορισμοί κοίλων συναρτήσεων που θα χρειαστούν στην πορεία.

Ορισμός 2.1 Έστω η συνάρτηση $f := C \to \mathbb{R}$ που ορίζεται στο κυρτό σύνολο C.

- η f ονομάζεται <u>σχεδόν κοίλη</u>, αν ισχύει

$$f(\lambda\,\mathrm{x} + (1-\lambda)\,\mathrm{y}) \ge \min\{\mathrm{f(x)}, \mathrm{f(y)}\}$$

για κάθε $x \in C, y \in C, \lambda \in [0,1]$.

- η f ονομάζεται <u>σχεδόν κυρτή</u>, αν f είναι σχεδόν κοίλη.

- Η συνάρτηση η οποία είναι και σχεδόν κοίλη και σχεδόν κυρτή θα ονομάζεται σχεδόν γραμμική.
- Επιπλέον, είναι φανερό πως μια συνάρτηση f είναι σχεδόν κοίλη όταν το άνω όριο $U_\gamma := \{x | \mathrm{f(x)} \geq \gamma\}$ είναι κυρτό σύνολο για όλες τις τιμές των $\gamma \in \mathbb{R}$. Τότε προκειμένου να έχουμε μια εφικτή περιοχή του προβλήματος θεωρούμε ότι η V είναι σχεδόν κοίλη.

Ορισμός 2.2 Έστω η $f := C \to \mathbb{R}$ συνεχής διαφορίσιμη συνάρτηση που ορίζεται σε ένα ανοιχτό κυρτό σύνολο C.

- η f λέγεται ψευδοκοίλη εάν ισχύει

  $$\nabla^{\mathrm{T}} f(\mathrm{x})(\mathrm{y}-\mathrm{x}) \leq 0 \Rightarrow f(\mathrm{y}) - f(\mathrm{x}) \quad \forall x, y \in C$$

- η f ονομάζεται ψευδοκυρτή, αν f είναι ψευδοκοίλη.

- Οι ψευδοκοίλες συναρτήσεις είναι επίσης σχεδόν κοίλες.

Προκειμένου να μεγιστοποιήσουμε προβλήματα με σχεδόν κοίλους περιορισμούς και ψευδοκοίλη αντικειμενική συνάρτηση απαιτούνται οι εξής ιδιότητες :

- Οι τοπικές λύσεις να είναι και ολικές λύσεις.
- Οι συνθήκες βελτιστοποίησης των Kuhn-Tucker να είναι επαρκείς.

Συνεπώς και η συνάρτηση V του προβλήματος $\max \mathrm{c}^T x + V(\mathrm{x}) \; \mathrm{x} \in \mathrm{B}$ αποτελεί μια ψευδοκοίλη συνάρτηση. Ωστόσο η απαίτηση μόνο της σχεδόν κοιλότητας της V συνεπάγεται αποτελέσματα μόνο σε μη κυρτά προβλήματα βελτιστοποίσης.

Επίσης, μολονότι η απαίτηση της σχεδόν κοιλότητας στους περιορισμούς της V στο πρόβλημα $\max \mathrm{c}^T x \; V_i(\mathrm{x}) \geq \mathrm{k}_i, i = 1, \ldots, s \; \mathrm{x} \in \mathrm{B}$ διασφαλίζει την κυρτότητα της εφικτής περιοχής, αλγοριθμικά στέκει αδύναμη να την διασφαλίσει διότι είναι αδύνατος ο έλεγχος των

συνθηκών κανονικότητας. Αυτό συνεπάγεται την μη διασφάλιση των συνθηκών Kuhn-Tucker, και επισημαίνεται η ανάγκη της χρήσης των συνθηκών κανονικότητας του Slater.

Η παρακάτω πρόταση αφορά συναρτήσεις σε μορφή κλάσματος.

Πρόταση 2.1 Έστω C ένα ανοιχτό κυρτό σύνολο και f,g δύο αυστηρά θετικές συναρτήσεις ορισμένες στο C. Αν f και g συνεχείς διαφορίσιμες συναρτήσεις με f να είναι κοίλη και g κυρτή, τότε η $h(\mathrm{x}) := \dfrac{f(\mathrm{x})}{g(\mathrm{x})}$ είναι ψευδοκοίλη στο C.

Απόδειξη : Αν $x \in C, y \in C$ και έστω ότι ισχύει $\nabla^{\mathrm{T}} h(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$. Τότε κάνοντας τους υπολογισμούς έχουμε :

$$g(\mathrm{x})\nabla^{\mathrm{T}}\,\mathrm{f}(\mathrm{x})(\mathrm{y}-\mathrm{x}) - \mathrm{f}(\mathrm{x})\nabla^{\mathrm{T}}\,\mathrm{g}(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$$

Χρησιμοποιώντας τις παραπάνω υποθέσεις καταλήγουμε στην ανισότητα :

$$g(\mathrm{x})\,\mathrm{f}(\mathrm{y}) - \mathrm{f}(\mathrm{x})\,\mathrm{g}(\mathrm{y}) \le 0$$

συνεπώς $h(\mathrm{y}) - \mathrm{h}(\mathrm{x}) \le 0$.▲

Οι μετασχηματισμοί των ψευδοκοίλων συναρτήσεων δημιουργούν την ανάγκη χρήσης της επόμενης πρότασης.

Πρόταση 2.2 Έστω C ένα ανοιχτό κυρτό σύνολο και g μία συνεχής διαφορίσιμη ψευδοκοίλη ή ψευδοκυρτή συνάρτηση ορισμένη στο C. Αν $f : \mathbb{R} \to \mathbb{R}$ μία συνεχής διαφορίσιμη συνάρτηση αυστηρά αύξουσα με $f'(\mathrm{x}) \neq 0$ για κάθε $x \in \mathbb{R}$. Τότε η $h(\mathrm{x}) := \mathrm{f}(\mathrm{g}(\mathrm{x}))$ είναι αντίστοιχα ψευδοκοίλη ή ψευδοκυρτή στο C.

Απόδειξη : Η κλίση της h συνεπάγεται την σχέση $\nabla h(\mathrm{x}) = \mathrm{f}'((\mathrm{gx}))\nabla\,\mathrm{g}(\mathrm{x})$.

- Για g ψευδοκοίλη , έστω $x \in C, y \in C$ και $\nabla^{T} h(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$. Χρησιμοποιώντας τις υποθέσεις μας συνεπάγεται ότι $\nabla^{T}\,\mathrm{g}(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$ και δεδομένου την ψευδοκοιλότητα της g συνεπάγεται ότι $g(\mathrm{y}) \le \mathrm{g}(\mathrm{x})$ όπου

από την μονοτονία της f έχουμε την τελική σχέση $\mathrm{h(y)} \le h(\mathrm{x})$ δηλαδή h ψευδοκοίλη.

- Για g ψευδοκυρτή η απόδειξη είναι όμοια.▲

Σημείωση: Το άθροισμα γραμμικών και ψευδοκοίλων συναρτήσεων δεν αποτελεί ψευδοκοίλη συνάρτηση.

Στην συνέχεια θα δούμε μια νέα κατηγορία κοίλων συναρτήσεων, τις αποκαλούμενες λογαριθμικά κοίλες συναρτήσεις.

Ορισμός 2.3 Έστω $f := C \to \mathbb{R}$ μη αρνητική συνάρτηση ορισμένη στο κυρτό σύνολο C.

- η f ονομάζεται λογαριθμικά κοίλη (log-concave) αν ισχύει η ανίσωση $f(\lambda \mathrm{x} + (1-\lambda)\,\mathrm{y}) \ge [f(x)]^{\lambda}[f(\mathrm{y})]^{(1-\lambda)}$ για κάθε $x \in C, y \in C, \lambda \in (0,1)$.
- η f ονομάζεται λογαριθμικά κυρτή (log-convex) αν ισχύει η αντίστροφη παραπάνω ανισότητα $f(\lambda \mathrm{x} + (1-\lambda)\,\mathrm{y}) \le [f(x)]^{\lambda}[f(\mathrm{y})]^{(1-\lambda)}$.

Συνεπώς, για μία λογαριθμικά κοίλη συνάρτηση το σύνολο

$$C^{+} := \{x \,|\, \mathrm{f(x)} > 0, \mathrm{x} \in \mathrm{C}\}$$

είναι κυρτό. Επίσης, η ανισότητα από τον παραπάνω ορισμό ισχύει είτε για $x \notin C^{+}$ είτε για $y \notin C^{+}$. Τότε δημιουργείται ένα νέο χαρακτηριστικό για τις λογαριθμικά κοίλες συναρτήσεις :

Πρόταση 2.3 Μία μη αρνητική συνάρτηση f είναι λογαριθμικά κοίλη στο κυρτό σύνολο C αν και μόνο αν το σύνολο $C^{+} := \{x \,|\, \mathrm{f(x)} > 0, \mathrm{x} \in \mathrm{C}\}$ είναι κυρτό και ο λογάριθμος της f (logf) είναι κοίλη συνάρτηση στο C^{+}.

Σημείωση : Το γινόμενο λογαριθμικά κοίλων συναρτήσεων $f(\mathrm{x}) = \prod_{i=1}^{r} f_i(\mathrm{x})$ αποτελεί επίσης λογαριθμικά κοίλη συνάρτηση. Επιπλέον, οι μετατοπίσεις λογαριθμικά κοίλων συναρτήσεων $\mathrm{g(x)} := \mathrm{f(x+y)}$ με $y \in \mathbb{R}^n$ σταθερό αποτελούν επίσης λογαριθμικά κοίλες συναρτήσεις στον $\mathbb{R}^n$.

Πρόταση 2.4 Έστω f συνεχής διαφορίσιμη, αυστηρά θετική και λογαριθμικά κοίλη συνάρτηση στο ανοιχτό κυρτό σύνολο C. Τότε f είναι ψευδοκοίλη στο C.

Απόδειξη: Έστω $x \in C, y \in C, \lambda \in (0,1)$ και θεωρούμε ότι ισχύει $\nabla^{\mathrm{T}} f(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$.

Έτσι ισχύει η σχέση $\nabla^{\mathrm{T}} \log f(\mathrm{x})(\mathrm{y}-\mathrm{x}) = \frac{1}{f(\mathrm{x})} \nabla^{T} f(\mathrm{x})(\mathrm{y}-\mathrm{x}) \le 0$. Επίσης επειδή f(x) είναι κοίλη συνάρτηση συνεπάγεται ότι $f(\mathrm{y}) \le \log \mathrm{f}(\mathrm{x})$ άρα η f αποτελεί και ψευδοκοίλη συνάρτηση.▲

Τέλος, χρησιμοποιώντας μια ισοδύναμη μορφή του προβλήματος

$$\begin{array}{l} \max \mathrm{c}^{T}\, x + V(\mathrm{x}) \\ \mathrm{x} \in \mathrm{B} \end{array} \Leftrightarrow \begin{array}{l} \max c^{T} + z \\ V(\mathrm{x}) - \mathrm{z} \ge 0 \\ \mathrm{x} \in \mathrm{B} \end{array}$$

για f κοίλη συνάρτηση, η νέα μορφή αποτελεί ένα κυρτό πρόβλημα προγραμματισμού. Έτσι κάθε πρόβλημα βελτιστοποίησης μπορεί να καλείται κυρτό πρόβλημα προγραμματισμού όταν η V είναι ψευδοκοίλη ή σχεδόν κοίλη συνάρτηση.

2.2. Στοχαστικά μοντέλα βελτιστοποίησης με συναρτήσεις πιθανότητας

Σκοπός της ενότητας αυτής είναι να εισάγει την έννοια της πιθανότητας ως μέτρο ποιότητας.

Προκειμένου να ορίσουμε το διάνυσμα απόφασης x ορίζουμε ως μέτρο ποιότητας το $\rho_{\mathrm{P}}(\vartheta) := P(\vartheta \ge 0)$, $\vartheta \in L_s^0$ ορισμένο στο σύνολο όλων των τυχαίων διανυσμάτων $(\Omega, \mathrm{F}, \mathrm{P})$ και αντιστοιχίζουμε σε αυτό την συνάρτηση αξιολόγησης $G(\mathrm{x}) := \rho_P(\zeta(\mathrm{x}, \xi)) := \mathrm{P}_{\xi}(\zeta(\mathrm{x}, \xi) \ge 0)$. Τότε η συνάρτηση G καλείται συνάρτηση πιθανότητας και παίρνει την μορφή :

$$G(\mathrm{x}) := \mathrm{P}_{\xi}(\mathrm{T}(\xi)\,\mathrm{x} - \mathrm{h}(\xi) \geq 0)$$

Έστω αυθαίρετο σταθερό $x \in \mathbb{R}^n$ και $S(\mathrm{x}) := \{\mathrm{z} \in \mathbb{R}^n \,|\, T(\mathrm{z})\,\mathrm{x} - \mathrm{h}(\mathrm{z}) \geq 0\}$ με $T(\bullet), \mathrm{h}(\bullet)$ συσχετισμένες γραμμικές συναρτήσεις, τότε σύμφωνα με τις υποθέσεις μας το σύνολο $S(\mathrm{x}) \subset \mathbb{R}^r$ αποτελεί ένα σύνολο πολυέδρων, έτσι ισχύει

$$G(\mathrm{x}) = \mathrm{P}(\xi \in \mathrm{S}(\mathrm{x}))$$

Τότε μπορούμε να ορίσουμε τα παρακάτω νέα προβλήματα :

$$\left.\begin{array}{l} \max \mathrm{c}^T x \\ P_{\xi}(\mathrm{T}(\xi)\,\mathrm{x} - \mathrm{h}(\xi) \geq 0) \geq \alpha \\ x \in B \end{array}\right\} \text{ και } \left\{\begin{array}{l} \max P_{\xi}(\mathrm{T}(\xi)\,\mathrm{x} - \mathrm{h}(\xi) \geq 0) \\ x \in B \end{array}\right\}$$

Όπου Β γνωστό φραγμένο, διάφορο του κενού, πολύεδρο.

Γενικότερα, τα παραπάνω προβλήματα αποτελούν μη κυρτά προβλήματα βελτιστοποίησης. Σε αυτή την ενότητα θα ασχοληθούμε με το να προσδιορίσουμε εκείνες τις υποκλάσεις όπου θα κατατάσσουν τα παραπάνω προβλήματα σε κυρτά προβλήματα βελτιστοποίησης ορίζοντας αρχικά τις βασικές ιδιότητες των μοντέλων αυτών καθώς επίσης και τα αντίστοιχα αποτελέσματα για τα αντίστροφα αυτών προβλήματα.

Στα προβλήματα αυτά οι περιορισμοί οι οποίοι περιέχουν συνάρτηση πιθανότητας ονομάζονται πιθανοτικοί περιορισμοί.

- Αν $s = 1$ ο περιορισμός καλείται διαχωρίσιμος.
- Αν $s > 1$ ο περιορισμός καλείται από κοινού.

Με ανάλογο τρόπο καλείται και η συνάρτηση πιθανότητας ως διαχωρίσιμη ή από κοινού και το μοντέλο, μοντέλο πιθανοτικών περιορισμών.

Η συνάρτηση πιθανότητας για τα μοντέλα που χρησιμοποιούν την συνάρτηση αυτή ως μέτρο ποιότητας, θα παίρνει την μορφή

$$P_{\xi}(\mathrm{f}(\mathrm{x}, \xi) \geq 0)$$

όπου $f(\mathrm{x},\bullet)$ είναι το μέτρο Borel για κάθε x και συγκεκριμένα επιλέγουμε $f(\mathrm{x},\xi)=\mathrm{T}(\xi)\,\mathrm{x}-\mathrm{h}(\xi)$.

Σύμφωνα με αυτά, τα παραπάνω μοντέλα καλούνται Μοντέλα Γραμμικού Στοχαστικού Προγραμματισμού με Συνάρτηση Πιθανότητας.

Στην συνέχεια, προκειμένου να σχεδιάσουμε τους περιορισμούς των προβλημάτων με αντίστροφη φορά των $G(\mathrm{x})\geq\alpha$, έχουμε :

$$P_\xi(\zeta(\mathrm{x},\xi)\geq 0)\geq\alpha \Leftrightarrow P_\xi([\min_{1\leq i\leq s}\zeta_i(\mathrm{x},\xi)]\geq 0)\geq\alpha$$

$$\Leftrightarrow P_\xi([\min_{1\leq i\leq s}\zeta_i(\mathrm{x},\xi)]<0)\leq 1-\alpha$$

$$\Leftrightarrow P_\xi([\max_{1\leq i\leq s}\zeta_i(\mathrm{x},\xi)^-]>0)\leq 1-\alpha$$

όπου για κάθε πραγματικό αριθμό z, το $z^-:=\max\{0,-z\}$ δηλώνει το αρνητικό μέρος του z.

Ωστόσο, οι παραπάνω υπολογισμοί είναι δύσκολο να γίνουν. Για τον λόγο αυτό περιοριζόμαστε στην περίπτωση με s=1, οπότε έχουμε :

$$P_\xi(\zeta(\mathrm{x},\xi)\geq 0)\geq\alpha \Leftrightarrow P_\xi(\zeta(\mathrm{x},\xi)^->0)\leq 1-\alpha \Leftrightarrow P_\xi(\mathrm{t}^T(\xi)\,\mathrm{x}-\mathrm{h}(\xi)<0)\leq 1-\alpha$$

με t(ξ) τα στοιχεία της γραμμής του T(ξ). Στην περίπτωση συνεχούς συνάρτησης η ανισότητα '<' μπορεί να αντικατασταθεί με '$\leq$'.

Ορίζοντας τώρα τον περιορισμό μας με βάση την εκτίμησή του, παίρνουμε την εξής μορφή :

$$P_\xi(\zeta(\mathrm{x},\xi)\geq 0)\geq\alpha \Leftrightarrow E_\xi[\chi(\max_{1\leq i\leq s}\zeta_i(\mathrm{x},\xi)^-)]\leq 1-\alpha$$

με $\chi(\mathrm{z})=\begin{cases}0 & \mathrm{z}\leq 0\\ 1 & \mathrm{z}>0\end{cases}$ όταν δείκτρια συνάρτηση.

Επιπλέον, το σύνολο των διανυσμάτων περιλαμβάνοντας πιθανοτικούς περιορισμούς μπορεί να πάρει την μορφή

$$B(\alpha) = \{x \,|\, \mathrm{G(x)} \geq \alpha\}$$

τότε μπορούμε να ανακατασκευάσουμε τα παραπάνω μοντέλα μας μέσω της συνάρτησης πιθανότητας :

$$\left.\begin{array}{l} \max c^T x \\ G(\mathrm{x}) \geq \alpha \\ x \in B \end{array}\right\} \text{ και } \left\{\begin{array}{l} \max G(\mathrm{x}) \\ \mathrm{x} \in \mathrm{B} \end{array}\right\}.$$

Τα προβλήματα αυτά θεωρούνται κυρτά προβλήματα προγραμματισμού αν G είναι σχεδόν κοίλη ή ψευδοκοίλη αντίστοιχα. Ωστόσο, και στην περίπτωση που η G δεν είναι σχεδόν κοίλη το πρόβλημα θεωρείται και πάλι κυρτό πρόβλημα προγραμματισμού διότι από τον "*ορισμό 2.1"* αρκεί το άνω φράγμα της συνάρτησης να είναι κυρτό (το οποίο συνεπάγεται ότι είναι σχεδόν κοίλη η συνάρτηση), και αυτό συμβαίνει αφού η εφικτή περιοχή $B(\alpha)$ έχει αντίστοιχο άνω όριο το α. Έτσι, κάθε τέτοιο πρόβλημα εξαρτάται από το προβλεπόμενο επίπεδο α, το οποίο όταν αυξάνεται αρκετά εξασφαλίζει την κυρτότητα του $B(\alpha)$.

2.2.1 Βασικές ιδιότητες

Στην ενότητα αυτή θα δούμε μερικά αποτελέσματα σχετικά με την συνάρτηση κατανομής της μεταβλητής ξ.

Αρχικά θα ορίσουμε την συνάρτηση πιθανότητας και στην συνέχεια το εφτικτό σύνολο.

Έστω η συνάρτηση πιθανότητας η οποία εμφανίζεται και στον περιορισμό του προβλήματος :

$$G(\mathrm{x}) = \mathrm{P}_{\xi}(\mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)) \ , \ G(\mathrm{x}) \geq \alpha$$

Τότε το αντίστοιχο εφικτό σύνολο συμβολίζεται ως :

$$\mathrm{B}(\alpha) = \{x \,|\, G(\mathrm{x}) \geq \alpha\} = \{x \,|\, \mathrm{S(x)} \in G_{\alpha}\}$$

όπου

$$S(\mathrm{x}) := \{\xi \,|\, \mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)\} \in \mathrm{B}^r \subseteq G_\alpha = \{\mathrm{A} \in \mathrm{B}^r \,|\, P_\xi(\mathrm{A}) \geq \alpha\}$$

Στην συνέχεια το σύνολο B(α) μπορεί να πάρει την ακόλουθη μορφή :

$$B(\alpha) = \bigcup_{\mathrm{A} \in G_\alpha} \bigcap_{\xi \in \mathrm{A}} \{x \,|\, \mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)\}$$

Για το σύνολο αυτό ισχύουν τα παρακάτω **θεωρήματα** :

2.1 Το σύνολο $B(\alpha)$ είναι ένα κλειστό σύνολο.

2.2 Το σύνολο $B(\alpha)$ είναι κυρτό σύνολο για α=0 και α=1.

Απόδειξη Κυρτότητας

Για α=0 ισχύει $G_\alpha = \mathrm{B}^r$ και συνεπώς $\mathrm{B}(\alpha) = \mathbb{R}^n$.

Για α=1 συνεπάγεται ότι $\mathrm{A} \in G_1$, $B \in G_1$ οπότε και $\mathrm{A} \cap B \in G_1$.

Αν $x \in B(1), \mathrm{y} \in \mathrm{B}(1)$ και $\lambda \in [0,1]$ με $z = \lambda x + (1-\lambda)\,\mathrm{y}$ τότε $S(\mathrm{x}) \in \mathrm{G}_1$, $S(\mathrm{y}) \in \mathrm{G}_1 \Rightarrow S(\mathrm{x}) \cap S(\mathrm{y}) \in \mathrm{G}_1$. Για κάθε σταθερό $\xi \in \mathbb{R}^r$ από τις σχέσεις $\begin{cases} \mathrm{T}(\xi)\,\mathrm{x} \geq h(\xi) \\ \mathrm{T}(\xi)\,\mathrm{y} \geq h(\xi) \end{cases}$ προκύπτει η εξής ανίσωση $\mathrm{T}(\xi)\,\mathrm{z} \geq \mathrm{h}(\xi)$.

Συνεπώς, $S(\mathrm{x}) \cap \mathrm{S}(\mathrm{y}) \subset \mathrm{S}(\mathrm{z})$ το οποίο συνεπάγεται ότι $\mathrm{S}(\mathrm{z}) \in \mathrm{G}_1$.

Επιπλέον, μπορούμε να μελετήσουμε και την αντίστροφη ανισότητα, δηλαδή την μορφή $G(\mathrm{x}) \leq \beta$. Σκεπτόμενοι αναλόγως προκύπτει το εφικτό σύνολο :

$$B(\beta) = \bigcup_{A \in H_\beta} \bigcap_{\xi \in A} \{x \,|\, \mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi)\}$$

με $H_\beta = \{\mathrm{A} \in \mathrm{B}^r \,|\, P_\xi(\mathrm{A}) \leq \beta\}$. Για β=0 και β=1 τα σύνολα B(0) και B(1) είναι κυρτά και μη κυρτά αντίστοιχα.

- Στην περίπτωση όπου α=0 και β=0 πρόκειται για πλεονάζων περιορισμό

Από την άλλη πλευρά θεωρώντας ότι η ξ ακολουθεί πεπερασμένη διακριτή κατανομή που δίνεται από τον παρακάτω πίνακα

$$\begin{pmatrix} p_1 & \dots & p_N \\ \hat{\xi}^1 & \dots & \hat{\xi}^{\mathrm{N}} \end{pmatrix} \text{ με } p_i > 0 \text{ και } \sum_{i=1}^{N} p_i = 1 .$$

Το αρχικό μας πρόβλημά τότε διαμορφώνεται ως εξής : $\left.\begin{matrix} \max c^T x \\ x \in B(\alpha) \cap \mathrm{B} \end{matrix}\right\}$

όπου στην διακριτή περίπτωση το σύνολο $\mathrm{B}(\alpha) = \{x \big| \mathrm{G(x)} \geq \alpha\}$ παίρνει την μορφή

$$\mathrm{B}(\alpha) = \bigcup_{\substack{J \subset I \\ \sum_{j \in J} p_j \geq \alpha}} \bigcap_{j \in J} \{x \big| \mathrm{T}(\hat{\xi}^j)\mathrm{x} \geq h(\hat{\xi}^j)\} \text{ για } \mathrm{I} = \{1, \dots, \mathrm{N}\} .$$

Ή ορίζοντας το σύνολο κυρτών πολυέδρων :

$$K_j = \{x \big| \mathrm{T}(\hat{\xi}^j)\,\mathrm{x} \geq \mathrm{h}(\hat{\xi}^j)\},\ \mathrm{j} = 1, \dots, \mathrm{N}$$

Το Β(α) παίρνει την μορφή : $B(\alpha) = \bigcup_{\substack{j \subset I \\ \sum_{j \in J} p_j \geq \alpha}} \bigcap_{j \in J} K_j$.

Από την μορφή αυτή διαπιστώνουμε ότι το πρόβλημά μας περιλαμβάνει την μεγιστοποίηση μιας γραμμικής συνάρτησης στο σύνολο της τομής κυρτών πολυέδρων. Αυτό δημιουργεί μια νέα οικογένεια προβλημάτων βελτιστοποίησης, αυτή των συζευγμένων προβλημάτων προγραμματισμού.

Και η μορφή αυτών :

$$\left.\begin{array}{l} \min c^T x \\ T(\hat{\xi}^k)\,\mathrm{x} + \mathrm{M}(1-\mathrm{z}_k) \geq h(\hat{\xi}^k), \mathrm{k} = 1,\ldots,\mathrm{N} \\ \sum_{i=1}^{N} p_i z_i \geq \alpha \\ z_k \in \{0,1\}, \mathrm{k} = 1,\ldots,\mathrm{N} \\ \mathrm{x} \in \mathrm{B} \end{array}\right\}$$

όπου z_j δυαδική μεταβλητή και M σταθερά με μεγάλη τιμή που μπορεί να υπολογιστεί επιλύοντας το γραμμικό πρόβλημα

$M_k = \max\{\gamma \,|\, \gamma + \mathrm{T}(\hat{\xi}^k)\,\mathrm{x} \geq h(\hat{\xi}^k), \mathrm{x} \in \mathrm{B}\}$ για $k \notin I_0$

και το M να παίρνει την μέγιστη τιμή αυτού.

Θεώρημα 2.3 Έστω $p_{i0} = \min_{i \in I} p_i$. Τότε το B(α) είναι κυρτό σύνολο για $\alpha > 1 - p_{i0}$.

Θεώρημα 2.4 Έστω $p_{i0} = \min_{i \in I} p_i$ και θεωρούμε ότι το p_{i0} είναι μοναδικά ορισμένο.

Έστω $p_{i1} = \min_{i \in I - \{\mathrm{i}_0\}} p_i$. Τότε το σύνολο B(α) είναι κυρτό όταν $\alpha > 1 - p_{i1}$.

2.2.2 Μελέτη Στοχαστικών Μοντέλων Προγραμματισμού

Στόχος της ενότητας αυτής είναι να προσδιορίσουμε στοχαστικά μοντέλα προγραμματισμού τα οποία περιέχουν διαχωρίσιμη συνάρτηση πιθανότητας τα οποία οδηγούν σε κυρτά προβλήματα.

Η μορφή του μοντέλου μας είναι η ίδια με αυτή που μας έχει απασχολήσει μέχρι τώρα με την διαφορά ότι η μεταβλητή $\zeta(\mathrm{x}, \xi)$ είναι τυχαία.

Θεωρούμε την διαχωριστική συνάρτηση $G(\mathrm{x}) = \mathrm{P}_{\xi}(x|\,\mathrm{t}(\xi)^{\mathrm{T}}\,\mathrm{x} \geq \mathrm{h}(\xi))$.

Θέτοντας $\eta := t(\xi)$ και $\xi := h(\xi)$, έχουμε την απλούστερη μορφή :

$$G(\mathrm{x}) = \mathrm{P}_{\xi}(x|\,\eta^{T}\,\mathrm{x} - \xi \geq 0)$$

Και η τυχαία μας μεταβλητή γίνεται $\zeta(\mathrm{x}, \eta, \xi) := \eta^{\mathrm{T}} x - \xi$.

Στην συνέχεια θα ασχοληθούμε με την ειδική περίπτωση όπου μόνο τα δεξιά μέλη είναι στοχαστικά. Έπειτα στην περίπτωση όπου οι μεταβλητές (η, ξ) ακολουθούν Πολυμεταβλητή Κανονική Κατανομή. Και τέλος, θα γίνει μια γενίκευση στις σταθερές κατανομές καθώς και στην απουσία αυτών.

Στοχαστικά δεξιά μέλη

- Θεωρούμε ότι το ξ ακολουθεί μια συνεχής κατανομή και ότι $\eta \equiv t$, με t ντετερμινιστική μεταβλητή. Τότε η συνάρτηση πιθανότητας θα έχει την μορφή

$$G(\mathrm{x}) = \mathrm{P}_{\xi}(x|\,\mathrm{t}^{T}\mathrm{x} \geq \xi) = \mathrm{F}_{\xi}(\mathrm{t}^{T}\mathrm{x})$$

όπου F_{ξ} η συνάρτηση κατανομής πιθανότητας για την τυχαία μεταβλητή είναι αύξουσα και ταυτόχρονα σχεδόν-κυρτή και σχεδόν-κοίλη, συνεπώς είναι σχεδόν-γραμμική. Αυτό συνεπάγεται την σχεδόν κοιλότητα και κυρτότητα της G(x), δηλαδή καθένα από τα σύνολα $\{x|\,\mathrm{G}(\mathrm{x}) \geq \alpha\}, \{x|\,\mathrm{G}(\mathrm{x}) \leq \beta\}$ είναι κυρτά.

Ή επιλύοντας τις παραπάνω σχέσεις αλγεβρικά έχουμε τα εξής συμπεράσματα :

$$\begin{cases} \mathrm{G}(\mathrm{x}) \geq \alpha \Leftrightarrow \mathrm{F}_{\xi}(\mathrm{t}^{T}\mathrm{x}) \geq \alpha \Leftrightarrow t^{T}x \geq Q_{\xi}^{-}(\alpha) \\ G(\mathrm{x}) \leq \beta \Leftrightarrow \mathrm{F}_{\xi}(\mathrm{t}^{T}\mathrm{x}) \leq \beta \Leftrightarrow t^{T}x \leq Q_{\xi}^{+}(\beta) \end{cases}$$

με $Q_{\xi}^{-}(\alpha), Q_{\xi}^{+}(\beta)$ τα αριστερά και δεξιά μέλη του κλειστού διαστήματος των α-, β-ποσοστημορίων αντίστοιχα.

- Αν ωστόσο η συνάρτηση δεν είναι συνεχής (π. χ. μια πεπερασμένη διακριτή κατανομή) τότε προκύπτει μία αυστηρά θετική / αρνητική ανίσωση που δηλώνει έτσι την ασυμμετρία μεταξύ των περιορισμών $\mathrm{G}(\mathrm{x}) \geq \alpha$, $G(\mathrm{x}) \leq \beta$.

Πολυμεταβλητή Κανονική Κατανομή

Θεωρούμε ότι οι $(\eta^{\mathrm{T}}, \xi)^{\mathrm{T}}$ ακολουθούν την από κοινού πολυμεταβλητή Κανονική κατανομή με η : στοχαστική μεταβλητή, δηλαδή ότι δεν υπάρχει $d \in \mathbb{R}^n : \eta = d$.

Ορισμός 2.4 Το τυχαίο διάνυσμα ζ διάστασης r ακολουθεί πολυμεταβλητή Κανονική κατανομή αν υπάρχει πίνακας B διάστασης $r \times s$ και $\mu \in \mathbb{R}^n$ τέτοια ώστε να ισχύει η σχέση $\zeta = \mathrm{B}\tilde{\zeta} + \mu$

όπου $\tilde{\zeta}$ τυχαίο διάνυσμα διάστασης s με στοιχεία $\tilde{\zeta}_i$ στοχαστικά ανεξάρτητα ακολουθώντας Τυπική Κανονική κατανομή για κάθε i.

Από τον ορισμό προκύπτει ότι :

- $E[\zeta] = \mu$, η αναμενόμενη τιμή
- $\Sigma = \mathrm{B}\mathrm{B}^{\mathrm{T}}$, ο πίνακας συνδιασποράς του ζ, θετικά ημιορισμένος.

Στην περίπτωση αυτή η κατανομή ονομάζεται εκφυλισμένη ή μοναδική, ενώ διαφορετικά, αν ο Σ είναι θετικά ορισμένος, ονομάζεται μη εκφυλισμένη.

Έτσι, μπορούμε να συμβολίσουμε την κατανομή του ζ ως : $\zeta \sim \mathrm{N}(\mu, \Sigma)$.

Η κατανομή αυτή είναι συνεχής με συνάρτηση πυκνότητας πιθανότητας

$$f(\mathrm{y}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathrm{y}-\mu)^{\mathrm{T}} \Sigma^{-1} (\mathrm{y}-\mu)}$$

όπου $|\Sigma|$ η ορίζουσα του πίνακα Σ.

Αν το διάνυσμα αναμενόμενης τιμής είναι μηδενικό και οι τυπικές αποκλίσεις των στοιχείων του ζ είναι 1, τότε μιλάμε για την ειδική περίπτωση της πολυμεταβλητής Κανονικής κατανομής αυτή της Τυπικής Πολυμεταβλητής Κανονικής κατανομής. Η συνάρτηση πυκνότητας πιθανότητας αυτής είναι :

$$\varphi(\mathrm{y};\mathrm{R}) = \frac{1}{(2\pi)^{\frac{n}{2}} \left|R\right|^{\frac{1}{2}}} e^{-\frac{1}{2}\mathrm{y}^{\mathrm{T}} R^{-1} \mathrm{y}}$$

με R να είναι ο πίνακας συσχέτισης του ζ που διαμορφώνεται από τις τυπικές αποκλίσεις σ_i, σ_j των ζ_i , ζ_j αντίστοιχα ως το πηλίκο : $R_{i,j} = \frac{\Sigma_{i,j}}{\sigma_i \sigma_j}, \forall i, j$.

Σημείωση: Κάθε γραμμικός μετασχηματισμός τυχαίων διανυσμάτων πολυμεταβλητών κανονικών κατανομών ακολουθεί επίσης πολυμεταβλητή κανονική κατανομή με

Αναμενόμενη τιμή : $E[\zeta(\mathrm{x},\eta,\xi)] = \mu^{\mathrm{T}} x - \mu_{n+1}$

Συνδιασπορά : $Var[\zeta(\mathrm{x},\eta,\xi)] = \left\|D^T x - d\right\|^2 = x^T D D^T x - 2(Dd)^T x + \left\|d\right\|^2$

Αν $Var[\zeta(\mathrm{x},\eta,\xi)] = 0$ τότε $\zeta(\mathrm{x},\eta,\xi) = \mathrm{E}[\zeta(\mathrm{x},\eta,\xi)]$ και έτσι κατάλήγουμε στην τυπική κανονική κατανομή.

Διαφορετικά αν $\left\|D^T x - d\right\|^2 > 0$, από την τυποποίηση έχουμε :

$$G(\mathrm{x}) = \mathrm{P}(\zeta(\mathrm{x},\eta,\xi) \geq 0) = 1 - P(\zeta(\mathrm{x},\eta,\xi) \leq 0)$$

$$= 1 - P\left(\frac{\zeta(\mathrm{x},\eta,\xi) - \mathrm{E}[\zeta(\mathrm{x},\eta,\xi)]}{\left\|D^T x - d\right\|} \leq \frac{-\mu^{\mathrm{T}} x + \mu_{n+1}}{\left\|D^T x - d\right\|}\right)$$

$$= 1 - \Phi\left(\frac{-\mu^{\mathrm{T}} x + \mu_{n+1}}{\left\|D^T x - d\right\|}\right) = \Phi\left(\frac{\mu^{\mathrm{T}} x - \mu_{n+1}}{\left\|D^T x - d\right\|}\right)$$

με $\Phi(\mathrm{x}) = 1 - \Phi(-\mathrm{x})$ λόγω συμμετρίας της τυπικής κανονικής.

Τελικά για την G(x) έχουμε :

$$G(\mathrm{x})=\begin{cases}1 & D^T x-d=0 \quad \mu^{\mathrm{T}}x-\mu_{n+1}\geq 0\\ 0 & \text{αν } D^T x-d=0 \text{ και } \mu^{\mathrm{T}}x-\mu_{n+1}<0\\ \Phi\left(\dfrac{\mu^{\mathrm{T}}x-\mu_{n+1}}{\left\|D^T x-d\right\|}\right) & D^T x-d\neq 0\end{cases}$$ **(1)** ▲

Θεώρημα 2.5 Έστω το τυχαίο διάνυσμα $\zeta^{\mathrm{T}}=(\eta^{\mathrm{T}},\xi)^{\mathrm{T}}$ διάστασης n+1 που ακολουθεί Πολυμεταβλητή Κανονική κατανομή. Για $\alpha\geq\frac{1}{2}$ το σύνολο $\mathrm{B}(\alpha)=\{x|\mathrm{G}(\mathrm{x})\geq\alpha\}$ είναι κυρτό.

Ενώ για την περίπτωση όπου $\alpha<\frac{1}{2}$, ισχύει :

Θεώρημα 2.6 Έστω n>1 και θεωρούμε ότι το τυχαίο διάνυσμα $\zeta^{\mathrm{T}}=(\eta^{\mathrm{T}},\xi)^{\mathrm{T}}$ διάστασης n+1 ακολουθεί μη εκφυλισμένη Πολυμεταβλητή Κανονική κατανομή. Για $\alpha<\frac{1}{2}$ το σύνολο Β(α) είτε είναι μη κυρτό είτε $\mathrm{B}(\alpha)=\mathbb{R}^n$.

Απόδειξη : Έστω $\hat{x}\in\mathbb{R}^n$ και $\hat{x}\notin\mathrm{B}(\alpha)$.

Θα δείξουμε ότι υπάρχουν $x^{(1)},x^{(2)}\in B(\alpha)$ διαφορετικά μεταξύ τους

($x^{(1)}\neq x^{(2)}$) τέτοια ώστε $\hat{x}=\frac{1}{2}(\mathrm{x}^{(1)}+\mathrm{x}^{(2)})$.

Για n>1 προκύπτει ότι υπάρχει $v\in\mathbb{R}^n$, διάφορο του μηδενός για το οποίο ισχύει $\mu^{\mathrm{T}}v=0$. Τότε από τον περιορισμό $G(\mathrm{x})\geq\alpha$ με $D^T x-d\neq 0$ και για $x(\lambda)=\hat{x}+\lambda v$, $\lambda\in\mathbb{R}$ προκύπτει ότι

$$\Phi^{-1}(\alpha)\left\|D^T x(\lambda)-\mathrm{d}\right\|-\mu^{\mathrm{T}}\hat{x}\leq-\mu_{n+1}$$

όπου $\mu^{\mathrm{T}}x(\lambda)=\mu^{\mathrm{T}}\hat{x},\forall\lambda\in\mathbb{R}$.

Τότε έχουμε

$$\left\|D^T x(\lambda)-\mathrm{d}\right\| \geq \left\|D^T x(\lambda)\right\| - \|d\| \Rightarrow$$

$$\left\|D^T x(\lambda)\right\|^2 = \lambda^2 v^T DD^T v + 2\lambda v^T DD^T \hat{x} + \hat{x}^T DD^T \hat{x}$$

Επιπλέον ισχύει ότι $\lim_{\lambda\to\pm\infty}\left\|D^T x(\lambda)-\mathrm{d}\right\| = \infty$. Χρησιμοποιώντας την σχέση $\Phi^{-1}(\alpha)<0$, συνεπάγεται ότι $\exists\lambda_0\in\mathbb{R}$ τέτοιο ώστε $x(\lambda_0)\in B(\alpha)$. Συνεπώς, ισχύουν $x(\lambda_0)\neq x(-\lambda_0)$ και $\hat{x}=\frac{1}{2}(x(\lambda_0)+x(-\lambda_0))$.▲

Αντίστοιχα, και στην περίπτωση της συνάρτησης πιθανότητας με αντίθετη φορά ανισότητας, $\hat{G}(\mathrm{x}) := \mathrm{P}(\eta^{\mathrm{T}}\,\mathrm{x}\leq\xi) = \mathrm{P}(\zeta(\mathrm{x},\eta,\xi)\leq 0)$, έχουμε :

$$\hat{G}(\mathrm{x}) = \begin{cases} 1 & D^T x - d = 0 \quad \mu^{\mathrm{T}} x - \mu_{n+1} \leq 0 \\ 0 & \text{αν } D^T x - d = 0 \text{ και } \mu^{\mathrm{T}} x - \mu_{n+1} > 0 \\ \Phi\left(\dfrac{-\mu^{\mathrm{T}} x + \mu_{n+1}}{\left\|D^T x - d\right\|}\right) & D^T x - d \neq 0 \end{cases} \quad \textbf{(2)}$$

Από τις (1) , (2) έχουμε :

$$P(\eta^{\mathrm{T}}\,\mathrm{x}\geq\xi)\geq\alpha \Leftrightarrow \Phi^{-1}(\alpha)\left\|D^T x-d\right\| - \mu^{\mathrm{T}} x \leq -\mu_{n+1}$$

$$P(\eta^{\mathrm{T}}\,\mathrm{x}\leq\xi)\geq\alpha \Leftrightarrow \Phi^{-1}(\alpha)\left\|D^T x-d\right\| + \mu^{\mathrm{T}} x \leq \mu_{n+1}$$

Οι οποίες ανισώσεις καθορίζουν την εφικτή περιοχή ως κυρτή.

Ενώ, όταν έχουμε περιορισμούς με αντίθετη ανίσωση, $G(\mathrm{x})\leq\beta, \hat{G}(\mathrm{x})\leq\beta$, οι σχέσεις μας γίνονται:

$$P(\eta^{\mathrm{T}}\,\mathrm{x}\geq\xi)\leq\beta \Leftrightarrow \Phi^{-1}(\beta)\left\|D^T x-d\right\| - \mu^{\mathrm{T}} x \geq -\mu_{n+1}$$

$$P(\eta^{\mathrm{T}}\,\mathrm{x}\leq\xi)\leq\beta \Leftrightarrow \Phi^{-1}(\beta)\left\|D^T x-d\right\| + \mu^{\mathrm{T}} x \geq \mu_{n+1}$$

οι οποίες καθορίζουν την εφικτή περιοχή ως κυρτή.

Από την άλλη πλευρά, όσον αφορά μοντέλα που περιέχουν στον στόχο συνάρτηση πιθανότητας περιοριζόμαστε στην περίπτωση που η $\zeta^{\mathrm{T}} = (\eta^{\mathrm{T}}, \xi)^{\mathrm{T}}$ ακολουθεί μη εκφυλισμένη Πολυμεταβλητή Κανονική κατανομή, δηλαδή όταν ο πίνακας $\begin{pmatrix} D \\ d^T \end{pmatrix}$ έχει βαθμό όσο το πλήθος των γραμμών. Τότε ισχύει ότι $D^T x - d \neq 0$ για κάθε $x \in \mathbb{R}^n$ οπότε η συνάρτηση πιθανότητας έχει την μορφή :

$$G(\mathrm{x}) = \Phi\left(\frac{\mu^{\mathrm{T}} x - \mu_{n+1}}{\left\| D^T x - d \right\|} \right) \ \forall x \in \mathbb{R}^n .$$

Σχετικά με τα προβλήματα μεγιστοποίησης η συνάρτηση $G(\mathrm{x})$ δεν μπορεί να είναι ψευδοκοίλη εφόσον δεν αποτελεί ούτε σχεδόν κοίλη συνάρτηση. Θεωρώντας την συνάρτηση $g(\mathrm{x}) = \dfrac{\mu^{\mathrm{T}} x - \mu_{n+1}}{\left\| D^T x - d \right\|}$, ισχύει ότι $G(\mathrm{x}) = \Phi(\mathrm{g}(\mathrm{x}))$ και έχουμε την παρακάτω πρόταση :

Πρόταση 2.5 Αν $\zeta^{\mathrm{T}} = (\eta^{\mathrm{T}}, \xi)^{\mathrm{T}}$ ακολουθεί μία μη εκφυλισμένη Πολυμεταβλητή Κανονική κατανομή τότε και οι συναρτήσεις $G(\mathrm{x}), \mathrm{g}(\mathrm{x})$ είναι :

a) ψευδοκοίλη στον ανοιχτό υποχώρο $\{x | \mu^{\mathrm{T}} \mathrm{x} > \mu_{n+1}\}$

b) ψευδοκυρτή στον ανοιχτό υποχώρο $\{x | \mu^{\mathrm{T}} \mathrm{x} < \mu_{n+1}\}$.

Απόδειξη: Από την υπόθεση της μη εκφυλισμένης κατανομής ισχύει ότι $\left\| D^T x - d \right\| > 0 \ \forall x \in \mathbb{R}^n$. Τότε η συνάρτηση g(x) σε ένα ανοιχτό κυρτό σύνολο, όπως αυτό της υπόθεσης, είναι

- ψευδοκοίλη όταν $\mu^{\mathrm{T}} x - \mu_{n+1} > 0$ και
- ψευδοκυρτή όταν $\mu^{\mathrm{T}} x - \mu_{n+1} < 0$.

Επίσης, δεδομένου ότι η Φ είναι αύξουσα, διαφορίσιμη συνάρτηση με $\Phi'(\mathrm{x}) \neq 0, \forall\, \mathrm{x} \in \mathbb{R}$ συνεπάγεται ότι ισχύουν τα ίδια αποτελέσματα και για την $G(\mathrm{x}) = \Phi(\mathrm{g}(\mathrm{x}))$.

Τέλος, όταν το ξ είναι τυχαίο σε συνδυασμό με την μη εκφυλισμένη κατανομή προκύπτει ότι όλα τα στοιχεία του η ακολουθούν μη εκφυλισμένη περιθωριακή κατανομή μιας διάστασης, το οποίο συνεπάγεται την τυχαιότητα και των δεξιών μελών.

Έστω $\xi \equiv \mu_{n+1} := h$ με $h \in \mathbb{R}$ τυχαίο. Το πρόβλημα ελαχιστοποίησης

$$\left.\begin{array}{l} \min \dfrac{-\mu^T x + h}{\left\| D^T x \right\|} \\ x \in B \end{array}\right\} \text{ όταν}$$

είναι κυρτό πρόβλημα προγραμματισμού όταν ισχύει $\mu^{\mathrm{T}} x < h \quad \forall x \in B$.▲

Σταθερή Κατανομή

Στην συνέχεια θα ασχοληθούμε με κάποιες βασικές ιδιότητες της Κανονικής κατανομής ορίζοντας μια νέα ταξινόμηση των συναρτήσεων κατανομών.

Έστω δύο μονοδιάστατες συναρτήσεις κατανομών F,H. Τις δύο συναρτήσεις συνδέει η σχέση:

$$F \circ H \Leftrightarrow \exists \alpha > 0, b : H\left(\frac{x-b}{\alpha}\right) = F(\mathrm{x}) \forall\, \mathrm{x} \in \mathbb{R}^1$$

η οποία είναι αυτοπαθής, συμμετρική και μεταβατική.

Θεωρώντας ως μια κλάση D την παραπάνω σειρά συναρτήσεων με $F \in D$ και H_0 μία Τυπική Κανονική στο D, ισχύει :

$\exists \alpha > 0,$ b τέτοιο ώστε $F(\mathrm{x}) = \mathrm{H}_0\left(\dfrac{x-b}{\alpha}\right), \forall x \in \mathbb{R}^1$

Η κλάση αυτή ονομάζεται κλίμακας-θέσης αφού α είναι η παράμετρος κλίμακας και b η παράμετρος θέσης της F.

Αν ζ τυχαία μεταβλητή, μπορούμε να τυποποιήσουμε την μεταβλητή αυτή θέτοντας $x = \frac{\zeta - b}{\alpha}$ και τότε η νέα μεταβλητή x θα ακολουθεί Τυπική Κανονική κατανομή καθώς και κάθε γραμμικός συνδυασμός της ζ.

Επίσης, από την ιδιότητα της συμμετρίας $f(\mathrm{x}) = \mathrm{f}(-\mathrm{x}) \forall\, \mathrm{x} \in \mathbb{R}$, ορίζουμε μια νέα κλάση που ονομάζεται κλάση συμμετρικής κλίμακας :

$$F \circ H \Leftrightarrow \exists \alpha > 0 : \mathrm{H}(\mathrm{x}) = \mathrm{F}(\alpha\, \mathrm{x}) \forall\, \mathrm{x} \in \mathbb{R}^1$$

Μια τυποποιημένη μεταβλητή επίσης ανήκει στο σύνολο κλάσεων των συμμετρικών κατανομών.

Επιπλέον, μία συνάρτηση κατανομής με την αντίστοιχη συνάρτηση πιθανότητας και μια τυχαία μεταβλητή λέγονται σταθερές όταν για κάθε πραγματικό αριθμό $s_1 > 0, m_1, s_2 > 0, m_2$ υπάρχει πραγματικός αριθμός $s > 0, m$ τέτοιοι ώστε :

$$F\left(\frac{x - m_1}{s_1}\right) * F\left(\frac{x - m_2}{s_2}\right) = F\left(\frac{x - m}{s}\right) \ \forall x \in \mathbb{R}^1$$

Τότε η κλάση αυτή ονομάζεται κλάση σταθερών κατανομών.

Ενώ την κλάση $F\left(\frac{x}{s_1}\right) * F\left(\frac{x -}{s_2}\right) = F\left(\frac{x}{s}\right) \ \forall x \in \mathbb{R}^1$

την ονομάζουμε αυστηρά σταθερή κλάση.

Η Τυπική Κανονική κατανομή αποτελεί συμμετρική και αυστηρά σταθερή κατανομή.

Ορισμός 2.5 Το τυχαίο διάνυσμα ζ διάστασης r ακολουθεί μία μη εκφυλισμένη πολυμεταβλητή Cauchy κατανομή αν υπάρχει πίνακας B διάστασης $r \times s$ με βαθμό όσο το πλήθος των γραμμών, ένα τουλάχιστον μη μηδενικό στοιχείο σε κάθε στήλη του και για $m \in \mathbb{R}^r$ να ισχύει : $\zeta = \mathrm{B}\tilde{\zeta} + m$

όπου $\tilde{\zeta}$ τυχαίο διάνυσμα διάστασης s με στοιχεία $\tilde{\zeta}_i$ στοχαστικά ανεξάρτητα που ακολουθούν Τυπική Cauchy κατανομή για κάθε i.

Για τυχαίο διάνυσμα $\zeta^{\mathrm{T}} = (\eta^{\mathrm{T}}, \xi)^{\mathrm{T}}$ διάστασης n+1 που ακολουθεί μη εκφυλισμένη πολυμεταβλητή Cauchy κατανομή ισχύει ότι η

$$\zeta(\mathrm{x}, \eta, \xi) = \tilde{\zeta}^{\mathrm{T}}(\mathrm{D}^T \mathrm{x} - \mathrm{d}) + \mathrm{mx} - \mathrm{m}_{n+1}$$

ακολουθεί επίσης κατανομή Cauchy, δηλαδή :

$$\zeta(\mathrm{x}, \eta, \xi) \sim C(\mathrm{m}^T \mathrm{x} - \mathrm{m}_{n+1}, \left\| \mathrm{D}^T \mathrm{x} - \mathrm{d} \right\|_1)$$

Τυποποιώντας και κάνοντας χρήση της συμμετρίας της τυποποιημένης Cauchy κατανομής έχουμε για την G(x):

$$G(\mathrm{x}) = \Psi\left(\frac{m^T x - m_{n+1}}{\left\| D^T x - d \right\|_1} \right)$$

Έτσι ο πιθανοτικός περιορισμός είναι

$$G(\mathrm{x}) \geq \alpha \Leftrightarrow \Psi^{-1}(\alpha) \left\| D^T x - d \right\|_1 - m^T x \leq -m_{n+1}$$

Λόγω της κυρτότητας της νόρμας έχουμε το παρακάτω θεώρημα.

Θεώρημα 2.7 Για τυχαίο διάνυσμα $\zeta^{\mathrm{T}} = (\eta^{\mathrm{T}}, \xi)^{\mathrm{T}}$ που ακολουθεί πολυμεταβλητή Cauchy κατανομή και για $\alpha \geq \frac{1}{2}$ το σύνολο $\mathrm{B}(\alpha) = \{x \,|\, \mathrm{G}(\mathrm{x}) \geq \alpha\}$ είναι κυρτό.

Ωστόσο, λόγω της δυσκολίας που επιφέρει η μη παραγωγισιμότητα της νόρμας $\left\| \bullet \right\|_1$, καταφεύγουμε στην προσθήκη μεταβλητών στους περιορισμούς ώστε αυτοί να πάρουν την μορφή γραμμικών περιορισμών. Έτσι έχουμε το παρακάτω σύστημα γραμμικών περιορισμών :

$$\begin{aligned} -m^T x + \Psi^{-1}(\alpha) \sum_{i=1}^{s} y_i &\leq -m_{n+1} \\ D_i^T x - y_k &\leq d_k, k = 1, \ldots, s \\ D_i^T x + y_k &\geq d_k, k = 1, \ldots, s \end{aligned}$$

Όπου αν $\tilde{x}$ εφικτή λύση του πιθανοτικού περιορισμού μας συνεπάγεται ότι για $\tilde{y_k} = \left| D_k^T x - d_k \right|, \forall k$ το σύνολο $(\tilde{x}, y_k, k = \tilde{1}, ..., s)$ είναι επίσης εφικτή λύση και αντίστροφα.

Από την άλλη πλευρά θεωρώντας ότι τα (η,ξ) είναι στοχαστικά ανεξάρτητα και ακολουθούν κατανομή Cauchy με $\eta_i \sim C(\mathrm{m}_i, \mathrm{t}_i), \mathrm{i} = 1, ..., \mathrm{n}$ και $\xi \sim C(\mathrm{m}_{n+1}, \mathrm{t}_{n+1})$ ισχύει ότι ο πίνακας $\begin{pmatrix} D \\ d^T \end{pmatrix}$ είναι διαγώνιος διάστασης $((\mathrm{n}+1)\times(\mathrm{n}+1))$ με διαγώνια στοιχεία τα t_i. Οπότε η νόρμα μπορεί να γραφεί σαν γραμμικός συνδυασμός $\left\| D^T x - d \right\|_1 = \sum_{i=1}^{n} t_i x_i + t_{n+1}$ χωρίς τελικά την προσθήκη νέων μεταβλητών.

Έλλειψη κατανομής

Αν θεωρήσουμε ότι υπάρχουν οι ροπές δεύτερης τάξης της μεταβλητής (η^{T}, ξ), για την μεταβλητή $\zeta(\mathrm{x}, \eta, \xi) = \eta^{\mathrm{T}}\,\mathrm{x} - \xi$ ισχύει ότι έχει :

$$E[\zeta(\mathrm{x}, \eta, \xi)] = \mu^{\mathrm{T}}\,\mathrm{x} - \mu_{n+1}$$
$$Var[\zeta(\mathrm{x}, \eta, \xi)] = \left\| D^T x - d \right\|^2$$

Τότε για το άνω φράγμα της συνάρτησης πιθανότητας $G(\mathrm{x}) = \mathrm{P}(\eta^{\mathrm{T}}\,\mathrm{x} - \xi \geq 0)$ συνεπάγεται από την ανισότητα Chebyshev ότι :

$$G(\mathrm{x}) \leq \frac{\left\| D^T x - d \right\|^2}{(-\mu^{\mathrm{T}}\,\mathrm{x} + \mu_{n+1})^2}$$

Συνεπώς για $G(\mathrm{x}) \leq \beta \Leftrightarrow \frac{\left\| D^T x - d \right\|^2}{(-\mu^{\mathrm{T}}\,\mathrm{x} + \mu_{n+1})^2} \leq \beta$.

Υποθέτοντας ότι ο παρανομαστής $-\mu^{\mathrm{T}}\,\mathrm{x}+\mu_{n+1}>\varepsilon$ είναι μεγαλύτερος από κάποια θετική σταθερά η παραπάνω ανίσωση μπορεί να πάρει την μορφή :

$$\beta^{-\frac{1}{2}}\left\|D^{T}x-d\right\|+\mu^{\mathrm{T}}x\le\mu_{n+1}$$

Το οποίο αποτελεί κυρτό σύνολο.

Η ανισότητα αυτή μπορεί να μην έχει νόημα στην περίπτωση ύπαρξης κατανομής αλλά στην περίπτωση που η κατανομή είναι ελλιπής οι εκτιμητές της μέσης τιμής και της διασποράς μέσω της ανισότητας δίνουν ένα εναλλακτικό μοντέλο.

Αναλόγως συμπεριφερόμαστε και στην περίπτωση ελαχιστοποίησης ενός στοχαστικού γραμμικού μοντέλου $\min_{\mathrm{x}\in\mathrm{B}} G(\mathrm{x})$. Χρησιμοποιώντας το άνω φράγμα της G(x) παίρνουμε ένα υποκατάστατο μοντέλο του αρχικού:

$$\min_{x\in B}\frac{\left\|D^{T}x-d\right\|^{2}}{(-\mu^{\mathrm{T}}\,\mathrm{x}+\mu_{n+1})^{2}}\overset{-\mu^{\mathrm{T}}\,\mathrm{x}+\mu_{n+1}>0}{\Leftrightarrow}\min_{x\in B}\frac{\mu^{\mathrm{T}}\,\mathrm{x}-\mu_{n+1}}{\left\|D^{T}x-d\right\|}$$

του οποίου η αντικειμενική συνάρτηση είναι ψευδοκυρτή στο B και συνεπώς το πρόβλημα αποτελεί κυρτό πρόβλημα προγραμματισμού.

Τέλος, η βέλτιστη τιμή του υποκατάστατου μοντέλου δίνει ένα άνω όριο της τιμής του αρχικού προβλήματος χωρίς αυτό να σημαίνει ότι αποτελούν ισοδύναμα προβλήματα.

2.2.3. Ανεξαρτησία

Θεωρώντας τα τυχαία διανύσματα $(\mathrm{t}_i^{T}(\xi),\mathrm{h}_i(\xi)),\mathrm{i}=1,\ldots,s$ ως στοχαστικά ανεξάρτητα η από κοινού συνάρτηση πιθανότητας μπορεί να γραφεί :

$$G(\mathrm{x})=\mathrm{P}(\zeta(\mathrm{x},\xi)\ge 0)=\prod_{i=1}^{s}P((\mathrm{t}_i^{T}(\xi)\,\mathrm{x}\ge\mathrm{h}_i(\xi))$$

Η μονοτονία της συνάρτησης κατανομής αποτελεί σχεδόν κοίλη συνάρτηση, ωστόσο το γινόμενό τους δεν συνεπάγεται και την κοιλότητα αυτών. Έτσι λογαριθμώντας τα δύο μέλη έχουμε :

$$\log G(\mathrm{x}) = \sum_{i=1}^{s} \log F_{\xi_i}(\mathrm{t}_i^T \mathrm{x})$$

η οποία είναι κοίλη συνάρτηση μόνο όταν οι F_{ξ_i} είναι λογαριθμικά κοίλες συναρτήσεις.

Συνεπώς και ο πιθανοτικός περιορισμός $G(\mathrm{x}) \geq \alpha$ αποτελεί κυρτό σύνολο $\forall \alpha \in [0,1]$ και μπορεί να πάρει την μορφή $\sum_{i=1}^{s} \log F_{\xi_i}(\mathrm{t}_i^T \mathrm{x}) \geq \log \alpha$.

2.2.4 Από κοινού περιορισμοί

Αν θεωρήσουμε ότι ο πίνακας Τ είναι αιτιοκρατικός , ο πιθανοτικός περιορισμός παίρνει την μορφή $G(\mathrm{x}) := \mathrm{P}_{\xi}(\mathrm{Tx} \geq \xi) \geq \alpha$ ή για $G(\mathrm{x}) = F_{\xi}(\mathrm{T}x)$ έχουμε :

$$F_{\xi}(\mathrm{y}) \geq \alpha$$
$$y - Tx = 0 \text{ .}$$

Το οποίο συνεπάγεται την κυρτότητα της εφικτής περιοχής $\mathrm{B}(\alpha) = \{x | \mathrm{G}(\mathrm{x}) \geq \alpha\}$

όταν οι συναρτήσεις πιθανότητας κατανομής F_{ξ} είναι σχεδόν κοίλες.

Γενικευμένα κοίλα μέτρα πιθανότητας

Ορίζουμε τον γενικευμένο μέσο $\mathrm{M}_{\gamma}^{\lambda}(\alpha, \mathrm{b})$ για $\alpha \geq 0, b \geq 0$ και $\lambda \in [0,1]$.

Τότε έχουμε τον ακόλουθο ορισμό :

Ορισμός 2.6 Μία μη αρνητική συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}_+$ λέγεται γ-κοίλη αν για $x, y \in \mathbb{R}^n$

και $\lambda \in [0,1]$ ισχύει : $\mathrm{f}(\lambda \mathrm{x}+(1-\lambda)\mathrm{y}) \geq \mathrm{M}_{\gamma}^{\lambda}(\mathrm{f}(\mathrm{x}),\mathrm{f}(\mathrm{y}))$.

Ορίζοντας το σύνολο $C^{+} := \{x \mid \mathrm{f}(\mathrm{x}) > 0\}$ μιας γ-κοίλης συνάρτησης f, ισχύει ότι το σύνολο είναι κυρτό. Τότε διάφορες τιμές του γ στο σύνολο C^{+} δίνουν διαφορετικές ερμηνείες της συνάρτησης. (π. χ. σταθερής (γ= $+\infty$), λογαριθμικά κοίλη (γ=0) κλπ).

Το σύνολο της κάθε συνάρτησης έχει ιδιαίτερη σημασία καθώς ενώ γενικά ισχύει ότι γραμμικοί συνδυασμοί και γενικότερα παράγωγα κυρτών συνόλων αποτελούν επίσης κυρτά σύνολα, δεν συμβαίνει όμως το ίδιο με σύνολα όπως το Borel.

Ορισμός 2.7 Το μέτρο πιθανότητας P στο σύνολο Borel B^{r} λέγεται γ-κοίλο αν για κάθε κυρτό, μετρήσιμο σύνολο A,B και κάθε $\lambda \in [0,1]$ όπου το σύνολο $\lambda \mathrm{A}+(1-\lambda)\mathrm{B}$ είναι μέτρο Borel ισχύει η ανισότητα :

$$P(\lambda \mathrm{A}+(1-\lambda)\mathrm{B}) \geq M_{\gamma}^{\lambda}(\mathrm{P}(\mathrm{A}),\mathrm{P}(\mathrm{B}))$$

Πρόταση 2.6 Για κάθε κυρτό, κλειστό σύνολο A στον $\mathbb{R}^{r}$ με P_{ξ} μέτρο γ-κοίλης, η συνάρτηση $\Gamma_{\mathrm{A}}(\mathrm{y}) = P_{\xi}(\mathrm{A}+\{\mathrm{y}\})$ είναι γ-κοίλη συνάρτηση.

Απόδειξη : Έστω $x, y \in \mathbb{R}^{r}, \lambda \in [0,1]$. Τότε

$\Gamma_{\mathrm{A}}(\lambda \mathrm{x}+(1-\lambda)\mathrm{y}) = P_{\xi}(\mathrm{A}+\{\lambda \mathrm{x}+(1-\lambda)\mathrm{y}\})$

$= P_{\xi}([\lambda+(1-\lambda)]\mathrm{A}+\{\lambda \mathrm{x}+(1-\lambda)\mathrm{y}\})$

$= P_{\xi}(\lambda[\mathrm{A}+\{\mathrm{x}\}]+(1-\lambda)[\mathrm{A}+\{\mathrm{y}\}])$

$\geq M_{\gamma}^{\lambda}(\mathrm{P}(\mathrm{A}+\{\mathrm{x}\}),\mathrm{P}(\mathrm{A}+\{\mathrm{y}\}))$

$= M_{\gamma}^{\lambda}(\Gamma_{\mathrm{A}}(\mathrm{x}),\Gamma_{A}(\mathrm{y}))$ ▲

Θεώρημα 2.8 Έστω f η γ-κοίλη συνάρτηση πυκνότητας πιθανότητας για την συνάρτηση πιθανότητας της τυχαίας μεταβλητής ξ διάστασης r. Έστω, επίσης, $-\frac{1}{r} \leq \gamma \leq \infty$. Τότε το P_{ξ} είναι κοίλο μέτρο πιθανότητας.

Απόδειξη : Έστω $\lambda \in [0,1]$, A,B κυρτά σύνολα και $\lambda \mathrm{A} + (1-\lambda)\mathrm{B}$ ένα μέτρο Borel.

Η γ-κοιλότητα της f είναι :

$$P_{\xi}(\lambda \mathrm{A} + (1-\lambda)\mathrm{B}) = \int\limits_{\lambda \mathrm{A}+(1-\lambda)\mathrm{B}} f(\mathrm{z})\,\mathrm{dz}$$

$$\geq \int\limits_{\lambda \mathrm{A}+(1-\lambda)\mathrm{B}} (\sup_{\lambda x+(1-\lambda)\mathrm{y}=\mathrm{z}} M_{\gamma}^{\lambda}(f(\mathrm{x}), f(\mathrm{y})))\,\mathrm{dz}$$

$$\geq M_{\frac{\gamma}{1+r\gamma}}^{\lambda}(\int\limits_{\mathrm{A}} f(\mathrm{x})\,\mathrm{dx}, \int\limits_{B} f(\mathrm{y})\,\mathrm{dy})$$

$$= M_{\frac{\gamma}{1+r\gamma}}^{\lambda}(P_{\xi}(\mathrm{A}), \mathrm{P}_{\xi}(\mathrm{B}))$$ ▲

Η παραπάνω απόδειξη στηρίζεται σε μία σημαντική ιδιότητα των λογαριθμικών συναρτήσεων : "έστω συνάρτηση $f : \mathbb{R}^{n+m} \to \mathbb{R}_{+}$ λογαριθμικά κοίλη, τότε η $g(\mathrm{x}) := \int\limits_{\mathbb{R}^m} f(\mathrm{x}, \mathrm{y})dy$ είναι λογαριθμικά κοίλη συνάρτηση στον $\mathbb{R}^n$". Επίσης η ιδιότητα αυτή συνεπάγεται και την λογαριθμική κοιλότητα και των περιθωριακών συναρτήσεων πυκνότητας.

Πολλές εφαρμογές του παραπάνω θεωρήματος για πολυμεταβλητές κατανομές πιθανότητας σε κυρτά σύνολα με συνάρτηση πυνότητας να είναι σχεδόν κοίλη ή λογαριθμικά κοίλη συμπεραίνουν ότι και η συνάρτηση πιθανότητας αποτελεί αντίστοιχα σχεδόν κοίλη ή λογαριθμικά κοίλη. (π. χ. ομοιόμορφη κατανομή, κατανομή του Dirichlet κ. ά.).

Γενικευμένες κοίλες συναρτήσεις κατανομών

Στην συνέχεια θα ασχοληθούμε με πιο σύνθετες συναρτήσεις προκειμένου να δείξουμε ότι ισχύουν και για αυτές κάποιες ιδιότητες κοιλότητας.

Θεώρημα 2.9 Έστω η συνάρτηση πιθανότητας

$$\mathrm{M}(\mathrm{x}) = \mathrm{P}_{\xi}\{\mathrm{g}_i(\mathrm{x}, \xi) \geq 0, i = 1, \ldots, \mathrm{s}\} = \mathrm{P}_{\xi}\{\mathrm{g}(\mathrm{x}, \xi) \geq 0\}$$

με $g_i : \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}, i-1,..,s; g^T(\mathrm{x},\xi) = (\mathrm{g}_1^T(\mathrm{x},\xi),...,\mathrm{g}_s^T(\mathrm{x},\xi))$ να είναι σχεδόν κοίλες συναρτήσεις, δηλαδή η $g_i(\cdot,\cdot)$ είναι από κοινού σχεδόν κοίλη και για τις δύο παραμέτρους. Επίσης, g συνεχής συνάρτηση.

Έστω ξ με γ-κοίλη συνάρτηση κατανομής. Τότε η M(x) είναι μία γ-κοίλη συνάρτηση.

Απόδειξη : Έστω το κλειστό, κυρτό σύνολο $H(\mathrm{x}) := \{z \,|\, \mathrm{g}(\mathrm{x},\mathrm{z}) \geq 0\} \subset \mathbb{R}^s \;\; \forall x$ και $\mathrm{M}(\mathrm{x}) = P_\xi(\mathrm{H}(\mathrm{x}))$.

Ισχύει η εξής σχέση υποσυνόλου : $H(\lambda\,\mathrm{x} + (1-\lambda)\,\mathrm{y}) \supset \lambda H(\mathrm{x}) + (1-\lambda)\,\mathrm{H}(\mathrm{y})$

Βάσει αυτής και της γ- κοιλότητας του μέτρου πιθανότητας έχουμε :

$$\mathrm{M}(\lambda\,\mathrm{x} + (1-\lambda)\,\mathrm{y}) = \mathrm{P}_\xi(H(\lambda\,\mathrm{x} + (1-\lambda)\,\mathrm{y}))$$

$$\geq P_\xi(\lambda\,\mathrm{H}(\mathrm{x}) + (1-\lambda)\,\mathrm{H}(\mathrm{y}))$$
$$\geq \mathrm{M}_\gamma^\lambda(\mathrm{P}_\xi(\mathrm{H}(\mathrm{x})), \mathrm{P}_\xi(\mathrm{H}(\mathrm{y}))) \;\blacktriangle$$
$$= \mathrm{M}_\gamma^\lambda(\mathrm{M}(\mathrm{x}), \mathrm{M}(\mathrm{y}))$$

Ωστόσο, στην περίπτωση που μεταβούμε σε γραμμικό μετασχηματισμό των γ-κοίλων τυχαίων μεταβλητών ισχύει το παρακάτω θεώρημα :

Θεώρημα 2.10 Έστω ξ ένα τυχαίο διάνυσμα διάστασης s, D ένας πίνακας διάσταση $(\mathrm{r}\times\mathrm{s})$, και $\zeta = D\xi + d$. Αν ξ έχει γ-κοίλη κατανομή τότε ομοίως και η ζ έχει γ-κοίλη κατανομή.

Απόδειξη : Έστω $\lambda \in [0,1]$ και A, B, $C_\lambda := \lambda\mathrm{A} + (1-\lambda)\mathrm{B}$ κυρτά σύνολα στον $\mathbb{R}^r$ μέτρου Borel. Τότε οι αντίστροφες εικόνες αυτών στον $\mathbb{R}^s$ μέσω των γραμμικών μετασχηματισμών των D, d είναι :

$$\bar{A} := \{x \,|\, \mathrm{Dx} + \mathrm{d} \in \mathrm{A}\},$$
$$\bar{B} := \{x \,|\, \mathrm{Dx} + \mathrm{d} \in B\}, \quad \text{και}$$
$$\bar{C}_\lambda := \{x \,|\, \mathrm{Dx} + \mathrm{d} \in C_\lambda\}$$

οι οποίες επίσης αποτελούν κυρτά σύνολα μέτρου Borel στον $\mathbb{R}^s$.▲

Ισχύει επίσης η ιδιότητα υποσυνόλου : $\bar{C}_\lambda \supset \lambda \bar{\mathrm{A}} + (1-\lambda)\bar{\mathrm{B}}$.

Βάσει αυτής έχουμε :

$$P_\zeta(\lambda \mathrm{A} + (1-\lambda)\mathrm{B}) = P_\xi(\bar{C}_\lambda)$$

$$\geq P_\xi(\lambda \bar{\mathrm{A}} + (1-\lambda)\bar{\mathrm{B}})$$
$$\geq \mathrm{M}_\gamma^\lambda(\mathrm{P}_\xi(\bar{\mathrm{A}}), \mathrm{P}_\xi(\bar{\mathrm{B}}))$$
$$= \mathrm{M}_\gamma^\lambda(\mathrm{P}_\zeta(\mathrm{A}), \mathrm{P}_\zeta(\mathrm{B}))$$

Συνεπώς, σύμφωνα με το παραπάνω θεώρημα κάθε συνδυασμός στοχαστικών τυχαίων μεταβλητών με γ-κοίλες (ή λογαριθμικά κοίλες) συναρτήσεις πυκνότητας έχει επίσης γ-κοίλη (ή λογαριθμικά κοίλη) κατανομή.

Μεγιστοποίηση της από κοινού συνάρτησης πιθανότητας

Στην περίπτωση που αντικειμενική συνάρτηση περιέχει την συνάρτηση πιθανότητας το μοντέλο μας παίρνει την μορφή

$$\max G(\mathrm{x})$$
$$\mathrm{x} \in \mathrm{B}$$

με $G(\mathrm{x}) = \mathrm{P}_\xi(\mathrm{Tx} \geq \xi)$ και B το σύνολο πολυέδρων $B = \{x \,|\, \mathrm{Ax} \propto \mathrm{b}, \mathrm{l} \leq \mathrm{x} \leq \mathrm{u}\}$.

Όταν η G είναι λογαριθμικά κοίλη συνάρτηση τότε το πρόβλημα είναι κυρτό πρόβλημα βελτιστοποίησης γραμμικών περιορισμών.

Τυχαίος πίνακας

Για την συνάρτηση πιθανότητας

$$G(\mathrm{x}) = P_\xi(x \,|\, \mathrm{T}(\xi)\,\mathrm{x} \geq \mathrm{h}(\xi))$$

θεωρούμε ότι T(ξ) είναι στοχαστικός πίνακας.

Για να συμπεράνουμε την κυρτότητα των $\begin{cases} \mathrm{B}(\alpha) := \{x \mid \mathrm{P}(\zeta(\mathrm{x}) \geq 0) \geq \alpha\} \\ \hat{\mathrm{B}}(\alpha) := \{x \mid \mathrm{P}(\zeta(\mathrm{x}) \leq 0) \geq \alpha\} \end{cases}$ όταν s>1

υποθέτουμε ότι $(\mathrm{T}_1(\xi), ..., \mathrm{T}_1(\xi), \mathrm{h}(\xi))$ ακολουθούν από κοινού Πολυμεταβλητή Κανονική κατανομή με πίνακα συνδιασποράς

$$\Sigma(\mathrm{x}) = \mathrm{Cov}[\zeta(\mathrm{x}), \zeta(\mathrm{x})]$$

$$= \sum_{i,j}^{n} x_i x_j C^{(\mathrm{i,j})} + \sum_{j=1}^{n} x_j C^{(\mathrm{n+1},j)} + C^{(\mathrm{n+1,n+1})}$$

$$= \sum_{i,j=1}^{n+1} \hat{x}_i \hat{x}_j C^{(\mathrm{i,j})} \quad .$$

2.3 Συναρτήσεις Ποσοστημορίων - Επίπεδο Σημαντικότητας

Προκειμένου να συμπεριλάβουμε σε ένα γραμμικό στοχαστικό πρόβλημα το μέγεθος και την πιθανότητα ζημίας χρησιμοποιούμε τα ποσοστημόρια.

Το α-ποσοστημόριο μιας μεταβλητής ϑ με συνάρτηση κατανομής F_ϑ ορίζεται από τις ανισότητες : $\begin{matrix} P(\vartheta \leq \mathrm{z}) \geq \alpha \\ P(\vartheta \geq \mathrm{z}) \geq 1-\alpha \end{matrix}$ για $0 < \alpha < 1,\ z \in \mathbb{R}$

Για τυχαία μεταβλητή ϑ το επίπεδο σημαντικότητας μπορεί να οριστεί με την σχέση :

$$\rho_{VaR}^{\alpha}(\vartheta) := \nu(\vartheta, \alpha) := \min\{\mathrm{z} \mid \mathrm{F}_\vartheta(\mathrm{z}) \geq \alpha\}, \vartheta \in \mathrm{L}_1^0$$

Με $0 < \alpha < 1$ γνωστό και $\nu(\vartheta, \alpha)$ το αριστερό άκρο του κλειστού διαστήματος των α-ποσοστημορίων του ϑ.

Για τυχαία μεταβλητή $\zeta(\mathrm{x}, \eta, \xi)$ με διαχωρίσιμη συνάρτηση πιθανότητας, η συνάρτηση αξιολόγησης είναι :

$$\mathrm{u}(\mathrm{x},\alpha) := \min\{\mathrm{z} \,|\, \Psi(\mathrm{x},\mathrm{z}) \geq \alpha\}$$

με $\Psi(\mathrm{x},\cdot)$ η συνάρτηση κατανομής πιθανότητας της $\zeta(\mathrm{x},\eta,\xi)$.

Τότε η $\mathrm{u}(\mathrm{x},\alpha)$ λέγεται συνάρτηση ποσοστημορίων – Value at Risk και ερμηνεύει το μέγεθος ζημίας το οποίο δεν θα υπερβαίνει την τιμή $\mathrm{u}(\mathrm{x},\alpha)$ με πιθανότητα τουλάχιστον α.

Ελαχιστοποιώντας το άθροισμα μιας γραμμικής συνάρτησης και το VaR και ορίζοντας την ελάχιστη τιμή $\mathrm{u}(\mathrm{x},\alpha)$ ως z, έχουμε την παρακάτω μορφή γραμμικού στοχαστικού μοντέλου:

Το οποίο γενικά αποτελεί ένα μη κυρτό πρόβλημα βελτιστοποίησης.

Αν ωστόσο το VaR εμφανίζεται στους περιορισμούς του προβλήματός μας, αυτό παίρνει την μορφή :

$$\begin{aligned} &\min_{x,z} c^T x \\ &P(\eta^T \mathrm{x} - \mathrm{z} \leq \xi) \geq \alpha \\ &z \leq \kappa \\ &x \in B \end{aligned}$$

Το οποίο αποτελεί πρόβλημα γραμμικού στοχαστικού προγραμματισμού με διαχωρίσιμη συνάρτηση πιθανότητας.

2.4. Μοντέλα βασισμένα στην εκτίμηση

Ένα γραμμικό στοχαστικό πρόβλημα μπορεί να περιέχει εκτίμηση χρησιμοποιώντας το επίπεδο σημαντικότητας $\rho_{\mathrm{E}}(\vartheta) := \mathrm{E}[\vartheta], \vartheta \in \mathrm{L}_1^1$ για τυχαίες μεταβλητές με πεπερασμένη μέση τιμή.

Για την τυχαία μεταβλητή $\zeta(\mathrm{x},\xi)=\mathrm{t}^{T}(\xi)\,\mathrm{x}-\mathrm{h}(\xi)$, με $\mathrm{t}(\xi)$ τυχαιο διάνυσμα διάστασης n και $\mathrm{h}(\xi)$ τυχαία μεταβλητή, θεωρούμε ότι υπάρχει η μέση τιμή των $(\mathrm{T}(\xi),\mathrm{h}(\xi))$, τότε η συνάρτηση αξιολόγησης του x είναι :

$$E[\zeta(\mathrm{x},\xi)]=\bar{t}^{T}\,x-\bar{h}$$

με $\bar{t}=\mathrm{E}[\mathrm{t}(\xi)]$, $\bar{h}=\mathrm{E}[\mathrm{h}(\xi)]$.

Τέτοια προβλήματα ονομάζονται προβλήματα αναμενόμενης τιμής γραμμικού στοχαστικού προγραμματισμού.

Ωστόσο, τα προβλήματα αυτά συνήθως χρησιμοποιούνται ως υποκατάστατα των γραμμικών στοχαστικών προβλημάτων και η χρήση τους θα πρέπει να είναι ιδιαίτερα προσεκτική καθώς η λύση που θα καταλήξει θα είναι αρκετά ριψοκίνδυνη.

ΚΕΦΑΛΑΙΟ 3ο «ΠΟΛΥΔΙΑΣΤΑΤΑ ΜΟΝΤΕΛΑ ΓΡΑΜΜΙΚΟΥ ΣΤΟΧΑΣΤΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ»

3.1. Γενικευμένα προβλήματα γραμμικού προγραμματισμού με κατανομή πόρων

Έστω ο πιθανοτικός χώρος $(\Omega, \mathrm{G}, \mathrm{P})$, το τυχαίο διάνυσμα $\xi_t : \Omega \to \mathbb{R}^{r_t}$ και η πιθανότητα κατανομής που ορίζεται από τα $\xi = (\xi_2^{\mathrm{T}}, ..., \xi_{\mathrm{T}}^{\mathrm{T}})^{\mathrm{T}} : \Omega \to \mathbb{R}^R$ με $R = r_2 + ... + r_T$ στο σ-πεδίο Borel του $\mathbb{R}^R$. Αν $\zeta_t = (\xi_2^{\mathrm{T}}, ..., \xi_{\mathrm{T}}^{\mathrm{T}})^{\mathrm{T}}$ η μεταβλητή του t σταδίου, τότε το πολυδιάστατο στοχαστικό γραμμικό πρόβλημα ορίζεται ως :

$$\min\{\mathrm{c}_1^T \mathrm{x}_1 + \mathrm{E}\sum_{t=2}^{T} c_t^T(\zeta_t)\,\mathrm{x}_t(\zeta_t)\}$$
$$A_{11}x_1 = b_1$$
$$A_{t1}(\zeta_t) + \sum_{r=2}^{t} A_{tr}(\zeta_t)\,\mathrm{x}_r(\zeta_r) = \mathrm{b}_t(\zeta_t), \mathrm{t} = 2, ..., \mathrm{T}$$
$$\mathrm{x}_1 \geq 0, \mathrm{x}_t(\zeta_t) \geq 0, t = 2, ..., T$$

όπου $x_t : \mathbb{R}^{r_2+...+r_t} \to \mathbb{R}^{n_t}$ πρέπει να είναι μέτρο Borel το οποίο θα συνεπάγεται ότι η $x_t(\zeta_t(\cdot)) : \Omega \to \mathbb{R}^{n_t}$ είναι μέτρο F_t ,

όπου $F_t = \sigma(\zeta_t) \subset \mathrm{G}$ η σ-άλγεβρα στο Ω που γεννάται στο στάδιο t από την ζ_t .

Υπόθεση

Έστω

- $\xi_t \in L^2(\Omega, \mathbb{R}^{r_t}) \forall\, \mathrm{t}$
- $A_{tr}(\cdot), \mathrm{b}_t(\cdot), \mathrm{c}_t(\cdot)$ γραμμικά συσχετισμένα στο ζ_t μέτρου F_t, όπου $A_{tr}(\cdot)$ πίνακας διάστασης $m_t \times n_r$.

Η παραπάνω υπόθεση συνεπάγεται την εξής μορφή του παραπάνω προβλήματος :

$$\left.\begin{array}{l} \min E\{\sum_{t=1}^{T} c_t^T x_t\} \\ \sum_{r=1}^{t} A_{tr} x_r = b_t \\ \left.\begin{array}{l} x_t \geq 0 \\ x_t \in X_t \cap M_t \end{array}\right\} t = 1,...,T \end{array}\right\}$$

Για δοσμένα :

- Ο χώρος πιθανότητας (Ω, G, P)
- σ-άλγεβρα $F_t, t = 1,...,T$ τέτοια ώστε $F_t \subset G \quad \forall t$ και $F_t \subset F_{t+1}$ για t=1,…,T-1
- $F := \{F_1,...,F_T\}$
- X_t γραμμικός υποχώρος του $L^2(\Omega, \mathbb{R}^{n_t})$
- M_t το σύνολο συναρτήσεων ορισμένες στο $\Omega \to \mathbb{R}^{n_t}$ μέτρου F_t, με $X_t \cap M_t$ κλειστό γραμμικό υποασύνολο του X_t.

Χρησιμοποιώντας διάφορες μορφές των παραπάνω συνόλων μπορούμε να έχουμε μια ποικιλία τέτοιων προβλημάτων , όπως :

- Προβλήματα αθροιστικής απόφασης $P(\hat{\mathrm{F}}, \mathrm{F})$
- Αθροιστικών περιορισμών $P(\mathrm{F}, \hat{\mathrm{F}})$, όπου οι περιορισμοί παρουσιάζονται σαν εκτίμηση του $\hat{\mathrm{F}}_t$
- Ολικού αθροίσματος $P(\hat{\mathrm{F}}, \hat{\mathrm{F}})$

Των οποίων

- τα εφικτά σύνολα είναι $\mathrm{B}(\mathrm{F}, \mathrm{F})$, $\mathrm{B}(\hat{\mathrm{F}}, \mathrm{F})$, $\mathrm{B}(\mathrm{F}, \hat{\mathrm{F}})$ και $\mathrm{B}(\hat{\mathrm{F}}, \hat{\mathrm{F}})$ και

- οι βέλτιστες λύσεις είναι $\inf(\mathrm{P}(\mathrm{F},\mathrm{F}))$, $\inf(\mathrm{P}(\hat{\mathrm{F}},\mathrm{F}))$, $\inf(\mathrm{P}(\mathrm{F},\hat{\mathrm{F}}))$, $\inf(\mathrm{P}(\hat{\mathrm{F}},\hat{\mathrm{F}}))$

αντίστοιχα.

Πρόταση 3.1 Για τα παραπάνω εφικτά σύνολα ισχύει :

$$B(\mathrm{F},\mathrm{F}) \supseteq \mathrm{B}(\hat{\mathrm{F}},\mathrm{F}) \quad B(\mathrm{F},\mathrm{F}) \subseteq \mathrm{B}(\mathrm{F},\hat{\mathrm{F}})$$
$$\mathrm{B}(\hat{\mathrm{F}},\mathrm{F}) \subseteq \mathrm{B}(\hat{\mathrm{F}},\hat{\mathrm{F}}) \quad \mathrm{B}(\mathrm{F},\hat{\mathrm{F}}) \supseteq \mathrm{B}(\hat{\mathrm{F}},\hat{\mathrm{F}})$$

Και αντίστοιχα για τις εφικτές λύσεις :

$$\inf(\mathrm{P}(\mathrm{F},\hat{\mathrm{F}})) \le \inf(P(\mathrm{F},\mathrm{F})) \le \inf(P(\hat{\mathrm{F}},\mathrm{F}))$$
$$\inf(\mathrm{P}(\mathrm{F},\hat{\mathrm{F}})) \le \inf(P(\hat{\mathrm{F}},\hat{\mathrm{F}})) \le \inf(P(\hat{\mathrm{F}},\mathrm{F}))$$

Παρατήρηση

για τα προβλήματα ολικού αθροίσματος ισχύουν :

- για δοσμένο άπειρο σύνολο F , δηλαδή κάποιο ζ_t δεν ακολουθεί πεπερασμένη διακριτή κατανομή, και $\hat{F}$ πεπερασμένο, τότε η $P(\hat{F},\hat{F})$ με πεπερασμένους περιορισμούς και μεταβλητές είναι φανερά πιο αποδοτική από την $P(\mathrm{F},\mathrm{F})$.
- Με κατάλληλες βελτιώσεις των μέτρων $\{\hat{\mathrm{F}}^{\nu}\}$ μπορούμε να αναμένουμε την σύγκλιση των βέλτιστων λύσεων του προβλήματος ολικού αθροίσματος σε αυτές του αρχικού.
- Στην ουσία η σχέση μεταξύ αυτών των βέλτιστων λύσεων είναι ότι οι λύσεις του προβλήματος ολικού αθροίσματος ($P(\hat{\mathrm{F}},\hat{\mathrm{F}})$) αποτελούν ένα κάτω όριο γι'αυτές του αρχικού ($P(\mathrm{F},\mathrm{F})$).

3.2. Προβλήματα Γραμμικού Στοχαστικού Προγραμματισμού δύο σταδίων

Περιορίζοντας τα προβλήματα Τ-σταδίων σε Τ=2, θεωρούμε τον χώρο πιθανότητας (Ω,F,P) και το τυχαίο διάνυσμα $\xi : \Omega \to \mathbb{R}^r$ τέτοιο ώστε $F = \sigma(\xi)$. Τότε ισχύει $P_\xi(\mathrm{B}) = P(\xi^{-1}[\mathrm{B}]) \ \forall \mathrm{B} \in \mathrm{B}^r$, όπου B^r το σύνολο Borel σ-άλγεβρας στον $\mathbb{R}^r$.

Τότε η γενική μορφή του προβλήματος δύο σταδίων τυχαίας κατανομής πόρων είναι:

$$\begin{aligned}
&\min \mathrm{E}_\xi \{\mathrm{c}^T x + q^T(\xi)\,\mathrm{y}(\xi)\} \\
&Ax = b \\
&T(\xi)\,\mathrm{x} + \mathrm{W}(\xi)\,\mathrm{y}(\xi) = h(\xi) \\
&\mathrm{x} \geq 0 \\
&\mathrm{y}(\xi) \geq 0 \\
&\mathrm{y}(\cdot) \in \mathrm{Y} \cap \mathrm{M}
\end{aligned}$$

όπου

- $A \in \mathbb{R}^{m_1 \times n_1}, b \in \mathbb{R}^{m_1}, c \in \mathbb{R}^{n_1}$ ντετερμινιστικοί πίνακες
- Y = $L^2(\Omega, \mathbb{R}^{n_2})$
- M το σύνολο των συναρτήσεων μέτρου F ορισμένες στο $\Omega \to \mathbb{R}^{n_2}$ και $\mathrm{Y} \subset \mathrm{M}$ (συνεπώς $y(\cdot) \in \mathrm{Y}$).
- $\mathrm{T}(\xi) \in \mathbb{R}^{m_2 \times n_1}, \mathrm{W}(\xi) \in \mathbb{R}^{m_2 \times n_2}, \mathrm{h}(\xi) \in \mathbb{R}^{m_2}, \mathrm{q}(\xi) \in \mathbb{R}^{n_2}$

με

- $\mathrm{T}(\xi) = \mathrm{T} + \sum_{j=1}^{r} T^j \xi_j$, $T, T^j \in \mathbb{R}^{m_2 \times n_1}$
- $\mathrm{W}(\xi) = W + \sum_{j=1}^{r} W^j \xi_j$, $\mathrm{W}, \mathrm{W}^j \in \mathbb{R}^{m_2 \times n_2}$
- $\mathrm{h}(\xi) = h + \sum_{j=1}^{r} h^j \xi_j$, $\mathrm{h}, \mathrm{h}^j \in \mathbb{R}^{m_2}$

- $\mathrm{q}(\xi) = q + \sum_{j=1}^{r} q^j \xi_j \ , \ \mathrm{q}, \mathrm{q}^j \in \mathbb{R}^{n_2}$

Στις εφαρμογές του μοντέλου αυτού, το να επιλέξεις την εφικτή απόφαση-λύση $\hat{x}$ για το πρώτο στάδιο απόφασης συνεπάγεται και την ικανοποίηση των περιορισμών του δεύτερου σταδίου. Ωστόσο, η επιλογή του $\hat{x}$ δεν πρέπει να έχει ούτε άπειρη ούτε αόριστη αντικειμενική τιμή αλλά να αντιστοιχεί στην τιμή που θα ελαχιστοποιεί την αντικειμενική συνάρτηση του μοντέλου.

3.2.1. Ολοκληρωμένη Σταθερή Κατανομή πόρων

Γενικότερα, επιλέγεται η ολοκληρωμένη σταθερή κατανομή πόρων παρά η σχετικά ολοκληρωμένη κατανομή, κάτω από την συνθήκη για $W(\xi) \equiv \mathrm{W}$:

$$\{\mathrm{z} \,|\, \mathrm{z} = \mathrm{Wy}, \mathrm{y} \geq 0\} = \mathbb{R}^{m_2} .$$

Αν ισχύει η παραπάνω συνθήκη, τότε κάθε εφικτή λύση $\hat{x}$ ενός αυθαίρετου συνόλου περιορισμών του πρώτου σταδίου και για κάθε τιμή $\hat{\xi}$ του τυχαίου διανύσματος ξ συνεπάγεται η εφικτότητα του συνόλου περιορισμών του επόμενου σταδίου.

Η ισχύς της συνθήκης αυτής κατοχυρώνεται από το ακόλουθο Λήμμα :

Λήμμα 3.1 Ένας πίνακας $W \in \mathbb{R}^{m_2 \times m_2}$ ικανοποιεί την συνθήκη ολοκληρωμένης κατανομής πόρων αν και μόνο αν :

- $rankW = m_2$ και
- για το αυθαίρετο σύνολο των γραμμικών ανεξάρτητων στηλών του W $\{\mathrm{W}_{i_1}, \mathrm{W}_{i_2} ..., \mathrm{W}_{i_{m_2}}\}$, οι γραμμικοί περιορισμοί :

$$\begin{aligned} &Wy = 0 \\ &y_{i_k} \geq 1, k = 1,\ldots, m_2 \quad (1) \\ &y \geq 0 \end{aligned}$$

είναι εφικτοί.

Απόδειξη : Θεωρούμε ότι ο W είναι ένας πίνακας ολοκληρωμένης κατανομής πόρων, τότε από την συνθήκη συνεπάγεται ότι $rank(W) = m_2$.

Επιπλέον, θέτουμε $\hat{z} = -\sum_{k=1}^{m_2} W_{i_k}$

όπου $\{W_{i_1}, W_{i_2} \ldots, W_{i_{m_2}}\}$ γραμμικά ανεξάρτητες του W.

Τότε για τον W έχουμε ότι $\{y \mid Wy = \hat{z}, y \geq 0\} \neq \varnothing$.

Επιλέγουμε τα δείκτρια σύνολα $\{j_1,\ldots, j_{n_2-m_2}\}$ τέτοια ώστε

$$\begin{cases} \{i_1, i_2 \ldots, i_{m_2}\} \cap \{j_1,\ldots, j_{n_2-m_2}\} = \varnothing \\ \{i_1, i_2 \ldots, i_{m_2}\} \cup \{j_1,\ldots, j_{n_2-m_2}\} = \{1,\ldots, n_2\} \end{cases}$$

το οποίο συνεπάγεται την ύπαρξη εφικτής λύσης $\hat{y}$ του

$$\sum_{k=1}^{m_2} W_{i_k} \hat{y}_{i_k} + \sum_{l=1}^{n_2-m_2} W_{j_l} \hat{y}_{j_l} = \hat{z}$$

$$= -\sum_{k=1}^{m_2} W_{i_k}$$

$$\hat{y}_i \geq 0, i = 1,\ldots, n_2$$

Έτσι για $y_v = \begin{cases} \hat{y}_v + 1, v = i_1, i_2, ..., i_{m_2} \\ \hat{y}_v, \mathrm{v} = \mathrm{j}_1, i_2, ..., i_{n_2 - m_2} \end{cases}$ ισχύουν οι περιορισμοί (1).

Αντίστροφα,

Έστω ότι ισχύουν οι συνθήκες του Λήμματος και έστω αυθαίρετο $\bar{z} \in \mathbb{R}^{m_2}$. Τότε η γραμμική ισότητα $\sum_{k=1}^{m_2} W_{i_k y_{i_k}} = \bar{z}$ έχει μοναδική λύση την $\{\bar{y}_{i_1}, ..., \bar{y}_{i_{m_2}}\}$.

Αν για $k = 1, ..., m_2$ ισχύει $\bar{y}_{i_k} \geq 0$, τότε υπάρχει εφικτή λύση της ισότητας κατανομής $Wy = \bar{z}$.

Διαφορετικά, το σύνολο $\gamma := \min\{\bar{y}_{i_1}, ..., \bar{y}_{i_{m_2}}\} < 0$. Αν $\tilde{y}$ η εφικτή λύση του προβλήματος των περιορισμών (1), τότε για $\hat{y}_v = \begin{cases} \hat{y}_v - \gamma \tilde{y}_v, v = i_1, i_2, ..., i_{m_2} \\ -\gamma \tilde{y}_v, \mathrm{v} = \mathrm{j}_1, i_2, ..., i_{n_2 - m_2} \end{cases}$

συνεπάγεται ότι

$$W \hat{y} = \sum_{k=1}^{m_2} W_{i_k} \hat{y}_{i_k} + \sum_{l=1}^{n_2 - m_2} W_{j_l} \hat{y}_{j_l}$$

$$= \sum_{k=1}^{m_2} W_{i_k} \underbrace{(\bar{y}_{i_k} - \gamma \tilde{y}_{i_k})}_{\geq 0} + \sum_{l=1}^{n_2 - m_2} W_{j_l} \underbrace{(-\gamma \tilde{y}_{j_l})}_{\geq 0}$$

$$= \bar{z} - \gamma \underbrace{\sum_{r=1}^{n_2} W_r \tilde{y}_r}_{=0}$$

Συνεπώς, η $\hat{y}$ είναι η εφικτή λύση του $Wy = \bar{z}, y \geq 0$.

Υποθέσεις :

- Ο πίνακας κατανομής W ικανοποιεί την συνθήκη της ολοκληρωμένα σταθερής κατανομής πόρων.
- Ο W σε συνδυασμό με το q(.) ικανοποιούν την σχέση :

$$\{u \mid \mathrm{W}^T \mathrm{u} \leq \mathrm{q}(\xi)\} \neq \varnothing \forall \xi \in \Xi \blacktriangle$$

Λήμμα 3.2 Σύμφωνα με τις παραπάνω υποθέσεις, για κάθε $x \in \mathbb{R}^{n_1}$ υπάρχει μία βέλτιστη τιμή $y(\cdot) \in \mathrm{Y}$ τέτοια ώστε $Q(\mathrm{x}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi)) = q^T(\xi)\, \mathrm{y}(\xi)$.

Θεώρημα 3.1 Έστω ότι ικανοποιούνται οι παραπάνω υποθέσεις. Τότε η συνάρτηση κατανομής πόρων $Q(\mathrm{x}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi))$ είναι :

a) Πεπερασμένα αξιολογήσιμη $\forall x \in B_1, \xi \in \Xi$

b) Κυρτή στο x $\forall \xi \in \Xi$

c) Συνεχής κατά Lipschitz στο x $\forall \xi \in \Xi$ με σταθερά Lipschitz $D(\xi) \in \mathrm{L}^1$.

Απόδειξη

a) Το γραμμικό πρόβλημα που προσδιορίζει ως συνάρτηση κατανομής πόρων την $Q(\mathrm{x}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi))$ είναι

$$\begin{aligned} &\min q^T(\xi)\, \mathrm{y} \\ &\mathrm{Wy} = h(\xi) - \mathrm{T}(\xi)\, \mathrm{x} \\ &\mathrm{y} \geq 0 \end{aligned}$$

Ή

$$\min\{q^T(\xi)\, y \mid \mathrm{Wy} = h(\xi) - \mathrm{T}(\xi)\, \mathrm{x}, \mathrm{y} \geq 0\}$$

το οποίο έχει αρχική εφικτή λύση για τυχαίο $x \in \mathbb{R}^{n_1}$ και $\xi \in \mathbb{R}^r$ και από την 2η υπόθεση συνεπάγεται ότι είναι και δυικά εφικτή $\forall \xi \in \Xi$.

Συνεπώς, είναι επιλύσιμη για όλα τα $x \in B_1$ και για όλα τα $\xi \in \Xi$, έτσι:

η $Q(\mathrm{x}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi))$ είναι πεπερασμένα αξιολόγισιμη για κάθε $x \in B_1$ και $\xi \in \Xi$.

b) Συνεπώς, για τυχαίο $\hat{\xi} \in \Xi$ και για $x^{(1)}, x^{(2)} \in B_1$ υπάρχει $y^{(\mathrm{i})}$, i=1,2

τέτοιο ώστε $Q(\mathrm{x}^{(\mathrm{i})}; \mathrm{T}(\hat{\xi}), \mathrm{h}(\hat{\xi}), \mathrm{q}(\hat{\xi})) = \mathrm{q}^T(\hat{\xi})\, \mathrm{y}^{(\mathrm{i})}$

με $Wy^{(\mathrm{i})} = \mathrm{h}(\hat{\xi}) - \mathrm{T}(\hat{\xi}) x^{(\mathrm{i})}, \mathrm{y}^{(\mathrm{i})} \geq 0$.

Τότε για $\tilde{x} = \lambda x^{(1)} + (1-\lambda)\, \mathrm{x}^{(2)}$ με $\lambda \in (0,1)$ συνεπάγεται και $\tilde{y} = \lambda y^{(1)} + (1-\lambda)\, \mathrm{y}^{(2)}$ είναι εφικτό για $Wy = \mathrm{h}(\hat{\xi}) - \mathrm{T}(\hat{\xi})\, \tilde{x}, \mathrm{y} \geq 0$.

Οπότε,

$$Q(\tilde{\mathrm{x}}; \mathrm{T}(\hat{\xi}), \mathrm{h}(\hat{\xi}), \mathrm{q}(\hat{\xi})) \leq q^T(\hat{\xi})\, \tilde{y} = \lambda q^T(\hat{\xi})\, \mathrm{y}^{(1)} + (1-\lambda) q^T(\hat{\xi})\, \mathrm{y}^{(2)}$$

Το οποίο συνεπάγεται την κυρτότητα του $Q(\mathrm{x}; \mathrm{T}(\hat{\xi}), \mathrm{h}(\hat{\xi}), \mathrm{q}(\hat{\xi}))$ στο x.

c) Για $x^{(1)} \neq x^{(2)}$ και κάθε $\xi \in \Xi$ υπάρχει $v_i \in \{1, \ldots, \mathrm{K}\}, \mathrm{i} = 1, 2$ τέτοιο ώστε

$$Q(\mathrm{x}^{(\mathrm{i})}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi)) = \alpha_{v_i}(\xi) + \mathrm{d}^{(v_i)T}(\xi)\, \mathrm{x}^{(\mathrm{i})}$$

και ισχύει

$$\begin{aligned}
&[\alpha_{v_1}(\xi) + \mathrm{d}^{(v_1)T}(\xi)\, \mathrm{x}^{(2)}] - [\alpha_{v_1}(\xi) + \mathrm{d}^{(v_1)T}(\xi)\, \mathrm{x}^{(1)}] \\
&= \mathrm{d}^{(v_1)T}(\xi)(\mathrm{x}^{(2)} - \mathrm{x}^{(1)}) \\
&\leq Q(\mathrm{x}^{(2)}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi)) - Q(\mathrm{x}^{(1)}; \mathrm{T}(\xi), \mathrm{h}(\xi), \mathrm{q}(\xi)) \\
&\leq [\alpha_{v_2}(\xi) + \mathrm{d}^{(v_2)^T}(\xi)\, \mathrm{x}^{(2)}] - [\alpha_{v_2}(\xi) + \mathrm{d}^{(v_2)^T}(\xi)\, \mathrm{x}^{(1)}] \\
&= \mathrm{d}^{(v_2)^T}(\xi)(\mathrm{x}^{(2)} - \mathrm{x}^{(1)})
\end{aligned}$$

τέτοιο ώστε

$$\left|Q(\mathrm{x}^{(2)};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi))-Q(\mathrm{x}^{(1)};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi))\right|$$
$$\leq \max_{i\in\{1,2\}}\left|\mathrm{d}^{(\mathrm{v}_i)^T}(\xi)(\mathrm{x}^{(2)}-\mathrm{x}^{(1)})\right|$$
$$\leq \max_{i\in\{1,2\}}\left\|\mathrm{d}^{(\mathrm{v}_i)^T}(\xi)\right\|\left\|(\mathrm{x}^{(2)}-\mathrm{x}^{(1)})\right\|$$

με $D(\xi)=\max_{i\in\{1,2\}}\left\|\mathrm{d}^{(\mathrm{v}_i)^T}(\xi)\right\|\in L^1$.▲

Λήμμα 3.3 Έστω το επιλύσιμο γραμμικό πρόβλημα $\min\{\mathrm{c}^T x\,|\,\mathrm{Ax}=\mathrm{b},\mathrm{x}\geq 0\}$ $\forall b\in\mathbb{R}^m$. Η βέλτιστη λύση του $\varphi(\mathrm{b})$ είναι κυρτή στο b και υποπαραγωγίσμη σε κάθε b και δίνεται από $\partial\varphi(\mathrm{b})=\arg\max\{\mathrm{b}^T u\,|\,\mathrm{A}^T\mathrm{u}\leq\mathrm{c}\}$ δηλαδή από το σύνολο των βέλτιστων δυικών λύσεων στο b.

Θεώρημα 3.2 Από τις υποθέσεις ισχύει ότι η συνάρτηση κατανομής πόρων $Q(\mathrm{x};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi))$ είναι υποπαραγωγίσιμη στο x για κάθε $\xi\in\Xi$. Για κάθε $\hat{x}$ ισχύει :

$$Q(\hat{\mathrm{x}};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi))=\{-\mathrm{T}^T(\xi)\hat{u}\,|\,\hat{u}\in\arg\max\{(\mathrm{h}(\xi)-\mathrm{T}(\xi)\hat{\mathrm{x}})^{\mathrm{T}}\mathrm{u}\,|\,W^T u\leq q(\xi)\}\}$$

$\forall\xi\in\Xi$.

Απόδειξη : Για αυθαίρετη τιμή $\xi\in\Xi$ ορίζουμε

$b(\mathrm{x};\xi):=\mathrm{h}(\xi)-\mathrm{T}(\xi)\,\mathrm{x}$ και

$\Psi(b(\mathrm{x};\xi);\xi):=Q(\mathrm{x};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi))$

$=\min\{\mathrm{q}^T(\xi)\,y\,|\,\mathrm{Wy}=\mathrm{b}(\mathrm{x};\xi),\mathrm{y}\geq 0\}$

Τότε από το “Λήμμα 3.3” η υποπαράγωγος της $\Psi(\cdot;\xi)$ στο $\mathrm{b}(\hat{x};\xi)$ είναι :

$$\partial_b\Psi(\mathrm{b}(\hat{x};\xi);\xi)=\arg\max\{\mathrm{b}^T(\hat{\mathrm{x}};\xi)u\,|\,\mathrm{W}^T u\leq q(\xi)\}$$

Έτσι έχουμε,

$$\partial_x Q(\mathrm{x};\mathrm{T}(\xi),\mathrm{h}(\xi),\mathrm{q}(\xi)) = -T^T(\xi)\partial_b \Psi(\mathrm{b}(\hat{x};\xi);\xi)$$

$$= -\mathrm{T}^{\mathrm{T}} \arg\max\{\mathrm{b}^T(\hat{\mathrm{x}};\xi)\,u \,|\, \mathrm{W}^T u \le q(\xi)\}\,.\blacktriangle$$

Θεώρημα 3.3 Για $\xi \in L^2(\Omega,\mathbb{R}^r)$, η αναμενόμενη τιμή Q(x) είναι :

a) Πεπερασμένα υπολογίσιμη $\forall x \in B_1$ και
b) Κυρτή και συνεχής συνάρτηση κατά Lipschitz στο x.

Πόρισμα : Το "θεώρημα 3.3" ισχύει ακόμα και για τυχαίες καταχωρήσεις q(x) και στοχαστικές ανεξάρτητες (h(ξ),T(ξ)).

Παρατήρηση : $Q(\cdot)$ είναι παραγωγίσιμη στο $\hat{x}$ αν και μόνο αν το σύνολο $\partial_x Q(\hat{x}, T(\cdot), \mathrm{h}(\cdot), \mathrm{q}(\cdot))$ είναι μονοσύνολο.

Πρόταση 3.2 Από τις υποθέσεις ισχύει ότι η $Q(\mathrm{x};T(\cdot),\mathrm{h}(\cdot))$ είναι κυρτή συνάρτηση στο ξ για κάθε $x \in \mathbb{R}^{n_1}$.

Απόδειξη : Για κάθε σταθερό $x \in \mathbb{R}^{n_1}$ έχουμε για το γραμμικό πρόβλημα :

$$Q(\mathrm{x};\mathrm{T}(\xi),\mathrm{h}(\xi)) := \min\{q^T\, y \,|\, \mathrm{W}\mathrm{y} = h(\xi) - \mathrm{T}(\xi)\,\mathrm{x},\, \mathrm{y} \ge 0\}$$

με το δεξί μέλος να είναι γραμμικό ως προς ξ, το οποίο συνεπάγεται την κυρτότητα▲.

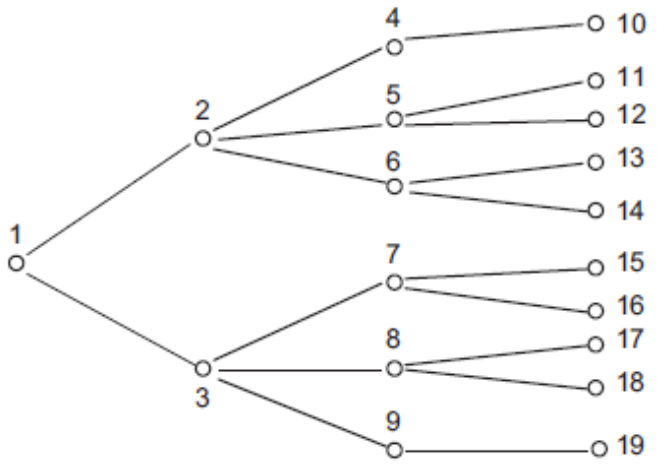
Θεώρημα 3.4 Για $\bar{\xi} = \mathrm{E}_{\xi}[\xi]$, η αναμενόμενη τιμή $Q(\mathrm{x}) = \mathrm{E}_{\xi}[\mathrm{Q}(\mathrm{x};\mathrm{T}(\xi),h(\xi)]$ είναι κάτω φραγμένη με $\mathrm{Q}(\mathrm{x};\mathrm{T}(\bar{\xi}),h(\bar{\xi})) \le \mathrm{Q}(\mathrm{x})$.

3.3 Πολυμεταβλητά Προβλήματα Γραμμικού Στοχαστικού Προγραμματισμού

Διακριτή Κατανομή

Τα προβλήματα που συναντάμε στις περισσότερες εφαρμογές είναι αυτά των οποίων το τυχαίο διάνυσμά μας ξ ακολουθεί διακριτή κατανομή.

Τότε κατά την διαδικασία αυτής της διακριτής στοχαστοποίησης ακολουθείται ένα συγκεκριμένο σενάριο όπου η δομή του φαίνεται παρακάτω :



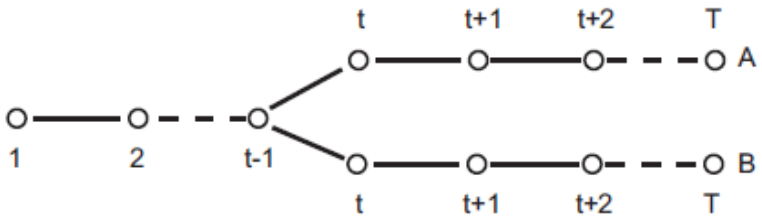
Εικόνα 7 : Scenario tree για προβλήματα 4 σταδίων στοχαστικής διαδικασίας

Όπου ο κόμβος 1, η ρίζα δηλαδή του δέντρου, αναφέρεται στο στάδιο 1. Αυτός συνδέεται με καθέναν από τους κόμβους του 2^{ου} σταδίου και η διαδικασία συνεχίζεται για όλους τους κόμβους και στάδια.

Με το παραπάνω σενάριο, περιγράφοντας γραφικά τις πιθανές εξελίξεις της στοχαστικής διαδικασίας $\{\zeta_t; \mathrm{t} = 2,\ldots \mathrm{T}\}$ μπορούμε να συνδυάσουμε πιθανοτικές πληροφορίες για να πάρουμε μια ολοκληρωμένη περιγραφή της διαδικασίας.

Μη Διακριτή Κατανομή

Η βασική ιδέα στην περίπτωση επίλυσης προβλημάτων με μη διακριτή κατανομή είναι η προσπάθεια μιας επιτυχημένης διακριτοποίηση αυτής. Αυτό επιτυγχάνεται διασπάζοντας τους κόμβους σε άλλους δύο υποκόμβους (διαδικασία splitting).



Εικόνα 8: Basic scenario tree: First split.

«ΒΙΒΛΙΟΓΡΑΦΙΑ»

[1] Peter Kall, Janos Mayer, (2005). Springer *Stochastic Linear Programming, Models, Theory, and Computation,* (7-255)

[2] King A. J., S. W. Wallace, (2012). Springer Modeling with Stochastic Programming, (1-164)

[3] Shapiro A., D. Dentcheva and A. Ruszczynski, (2009). Lectures on stochastic programming: Modeling and theory. MPS/SIAM Series on Optimization vol.9.

[4] Ruszcynski A., A. Shapiro, (2003). Stochastic Programming. Handbooks in Operations Research and Management Science Vol.10.

[5] http://www.neos-server.org/neos/ (*Numerical Optimization*).

[6] The Stochastic Programming Society (SPS) website: http://stoprog.org/

[7] Olsen P. (1976). When Is a Multistage Stochastic Programming Problem Well Defined? SIAM J. Contr. Opt. 14: 518-527.

[8] Williams A. C. (1996). Approximation formulas for stochastic linear programming. SIAM J. Appl. Math. 14: 668-877.