ΣΗΜΕΙΩΣΕΙΣ ΜΑΘΗΜΑΤΟΣ

ΒΕΛΤΙΣΤΟΣ ΣΧΕΔΙΑΣΜΟΣ ΤΩΝ ΚΑΤΑΣΚΕΥΩΝ

Β. ΚΟΥΜΟΥΣΗΣ

ΚΑΘΗΓΗΤΗΣ

ΑΘΗΝΑ

ΜΑΡΤΙΟΣ 1998

ΠΕΡΙΕΧΟΜΕΝΑ **ΣΕΛ.**

ΤΟ ΠΡΟΒΛΗΜΑ ΒΕΛΤΙΣΤΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΤΑΣΚΕΥΩΝ 5

1.1 ΕΙΣΑΓΩΓΗ 5

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ 10

2.1 ΜΕΤΑΒΛΗΤΕΣ ΣΧΕΔΙΑΣΜΟΥ (design variables) 10
2.1.1 Συνεχείς και Διακριτές μεταβλητές σχεδιασμού 11
2.1.2 Όρια των μεταβλητών σχεδιασμού 11

2.2 ΠΕΡΙΟΡΙΣΜΟΙ 12
2.2.1 Κανονικοποίηση των περιορισμών 13

2.3 ΑΝΤΙΚΕΙΜΕΝΙΚΗ ΣΥΝΑΡΤΗΣΗ (objective function) 13

2.4 Εφαρμογή 14

2.5 ΜΑΘΗΜΑΤΙΚΗ ΔΙΑΤΥΠΩΣΗ 15

ΓΡΑΜΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ 17

3.1 ΕΙΣΑΓΩΓΗ 17

3.2 ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ 17
3.2.1 Διατύπωση προβλήματος με μορφή μητρώων 18
Ανισοτικοί περιορισμοί 18
Βασικές Λύσεις 20
Εφαρμογή 21

3.3 Θεμελιώδες Θεώρημα του Γραμμικού Προγραμματισμού 24
3.3.1 Μετάβαση από Βασική-Αποδεκτή λύση σε άλλη 25

3.4 Κανονική Μορφή προβλήματος Γραμμικού Προγραμματισμού 26

3.5 Μέθοδος Simplex 27

3.6 Εκφυλισμένες λύσεις 30

3.7 Μητρωική Μορφή της Μεθόδου Simplex 31

3.8 Διερεύνηση λύσεων του προβλήματος 32
3.8.1 Μοναδική λύση 33
3.8.2 Μη μοναδική λύση 33
3.8.3 Απεριόριστη λύση 34
3.8.4 Μη επιτρεπτή λύση 35

3.9 Μέθοδοι Simplex δύο φάσεων 40

3.10 Αλγόριθμοι δύο φάσεων 44

3.11 Ο ΔΥΪΣΜΟΣ ΣΤΟ ΓΡΑΜΜΙΚΟ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ 45
3.11.1 Το Γενικό Δυικό Πρόβλημα 47
3.11.2 Προσδιορισμός πρωτογενούς λύσης από δυϊκή λύση 49

3.12 Λύση του πρωτογενούς προβλήματος με τη βοήθεια του δυϊκού πίνακα 54

3.13 Επίλυση Προβλημάτων Γραμμικού Προγραμματισμού στον Υπολογιστή 58
3.13.1 Χρήση του προγράμματος Excel της Microsoft 58
3.13.2 Χρήση προγραμμάτων συμβολικού προγραμματισμού 58

ΕΦΑΡΜΟΓΕΣ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ 60

4.1 Γενικά 60

4.2 Κάτω Όριο Οριακού Φορτίου Επιπέδου Δικτυώματος 60

4.3 Βέλτιστη χάραξη τένοντα προεντεταμένης δοκού 63

ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΜΕΣΩ ΕΠΑΝΑΛΗΠΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ 65

5.1 ΥΠΑΡΞΗ ΚΑΙ ΜΟΝΑΔΙΚΟΤΗΤΑ ΤΗΣ ΒΕΛΤΙΣΤΗΣ ΛΥΣΗΣ 67

5.2 Προβλήματα χωρίς περιορισμούς 68

5.3 Προβλήματα με περιορισμούς 69

5.4 Συνθήκες Kuhn-Tucker 72

ΠΡΟΣΕΓΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ 75

6.1. ΕΙΣΑΓΩΓΗ 75

6.2 ΓΕΝΙΚΕΣ ΕΝΝΟΙΕΣ 76
6.2.1 Ανάπτυξη σε σειρές Taylor 76
6.2.2 Αλγόριθμος άμεσων μεθόδων βελτιστοποίησης (direct methods) 77
6.2.3 Γραμμικοποίηση μη γραμμικών προβλημάτων βελτιστοποίησης 78

6.3 ΜΕΘΟΔΟΣ ΔΙΑΔΟΧΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (SLP) ... 84
6.3.1 Βασική ιδέα της μεθόδου 84
6.3.2 Ο αλγόριθμος της μεθόδου SLP 85

6.4 ΜΕΘΟΔΟΣ ΔΙΑΔΟΧΙΚΟΥ ΔΕΥΤΕΡΟΒΑΘΜΙΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ .. 92
6.4.1 Υποπρόβλημα δευτεροβάθμιου προγραμματισμού (QP) 93
6.4.3 Λύση προβλημάτων δευτεροβάθμιου προγραμματισμού (QP) 97
Διατύπωση συνθηκών Kuhn-Tucker για το πρόβλημα QP 97
Η μέθοδος simplex για την επίλυση του προβλήματος QP 99

6.5 ΜΕΘΟΔΟΣ ΕΠΙΤΡΕΠΤΩΝ ΔΙΕΥΘΥΝΣΕΩΝ (MFD) 105
6.5.1 Γενική διατύπωση 105

6.5.2 Επιλογή της διεύθυνσης αναζήτησης. 107
6.5.3 Επιλογή του βήματος α..... 110

6.6 Προσέγγιση του προβλήματος κατά Zoutendijk 112

6.7 Παρατηρήσεις επί της μεθόδου μέγιστης κλίσεως 115
6.7.1 Εισαγωγή 115
6.7.2 Η μέθοδος μέγιστης κλίσεως..... 115
6.7.3 Η μέθοδος συζυγών διευθύνσεων 116

6.8 Η ΜΕΘΟΔΟΣ ΤΩΝ ΚΕΝΤΡΩΝ 117

6.9 ΧΡΗΣΗ ΥΠΟΡΟΥΤΙΝΩΝ IMSL στην γλώσσα Fortran 122
6.9.1 Επίλυση Προβλήματος Γραμμικού Προγραμματισμού με βάση την αναθεωρημένη μέθοδο Simplex 124
6.9.2 Επίλυση Δευτοροβάθμιου προβλήματος με γραμμικούς περιορισμούς..... 127
6.9.3 Ελαχιστοποίηση οποιασδήποτε αντικειμενικής συνάρτησης που υπόκειται σε γραμμικούς περιορισμούς..... 129
6.9.4 Επίλυση γενικού μη γραμμικού προβλήματος (SQP) αριθμητικές ευαισθησίες 134
6.9.5 Επίλυση Γενικού μη γραμμικού προβλήματος (SQP) αναλυτικές ευαισθησίες 141

ΑΝΑΛΥΣΗ ΕΥΑΙΣΘΗΣΙΑΣ 146

7.1 ΕΥΑΙΣΘΗΣΙΑ 146

7.2 Ανάλυση Ευαισθησίας 147
Εφαρμογή 1 149
Εφαρμογή 2 155
Εφαρμογή 3 159
Άσκηση 163

7.3 ΣΥΖΥΓΗΣ ΜΕΘΟΔΟΣ ΥΠΟΛΟΓΙΣΜΟΥ ΕΥΑΙΣΘΗΣΙΩΝ 164
Εφαρμογή 4 165
Άμεση Μέθοδος..... 166
Συζυγής Μέθοδος (adjoint Method)..... 168

ΜΕΘΟΔΟΙ ΕΣΩΤΕΡΙΚΟΥ ΣΗΜΕΙΟΥ - Η ΚΕΝΤΡΙΚΗ ΔΙΑΔΡΟΜΗ 170

8.1 Εκτίμηση της κατάλληλης τιμής της παραμέτρου μ 176
8.1.1 Επιλογή της παραμέτρου μήκους βήματος..... 177

8.2 Αλγόριθμος Εσωτερικής διαδρομής 178

8.3 Εξισώσεις Karush-Kuhn-Tucker – KKT 179

8.4 Κανονικές Εξισώσεις KKT 180

8.5 Δευτεροβάθμιος Προγραμματισμός 181

1

ΤΟ ΠΡΟΒΛΗΜΑ ΒΕΛΤΙΣΤΟΥ ΣΧΕΔΙΑΣΜΟΥ ΚΑΤΑΣΚΕΥΩΝ

1.1 ΕΙΣΑΓΩΓΗ

Ο σχεδιασμός των κατασκευών πραγματοποιείται κατά φάσεις. Συνήθως η μελέτη ξεκινά με μία ή περισσότερες βασικές ιδέες που αποτελούν την προκαταρκτική μελέτη (conceptual design), οι οποίες στη συνέχεια υποβάλλονται στη δοκιμασία της συμμόρφωσης με τις λειτουργικές, αισθητικές κ.α απαιτήσεις της κατασκευής για να καταλήξει σε αυτό που συνήθως ονομάζουμε προμελέτη (preliminary design). Με βάση την προμελέτη καταρτίζεται στη συνέχεια η οριστική μελέτη (detailed design) που περιλαμβάνει τις απαντήσεις για όλα τα επιμέρους θέματα που οριστικοποιούν την κατασκευή του σχεδιασμού. Οι φάσεις αυτές μελέτης ενός σύνθετου εν γένει αντικειμένου ή συνόλου αντικειμένων είναι αρκετά γενικές και μπορούν να εφαρμοστούν για τεχνικά έργα, έπιπλα, σκεύη ακόμη και έργα τέχνης. Το σύνολο των λειτουργικών, αισθητικών απαιτήσεων, ανθεκτικότητας κλπ. ποικίλουν από έργο σε έργο το οποίο σε τελική φάση παράγεται και διατίθεται με τον ένα ή άλλο τρόπο στην ευρύτερη αγορά όπου και αποτιμάται.

Στην σημερινή εποχή δεν υπάρχει επεξεργασμένη θεωρία σχεδιασμού γενική ή εξειδικευμένη. Διάφορες προσπάθειες βρίσκονται σε εξέλιξη που στοχεύουν στην αποκωδικοποίηση του τρόπου σύλληψης των στοιχείων ενός έργου είτε αυτό αφορά ένα αρχιτεκτόνημα, είτε τον φέροντα οργανισμό μιας γέφυρας, είτε ένα γλυπτό. Ερωτήματα που απασχολούν τις έρευνες αυτές προσπαθούν να αναδείξουν το κατά πόσο η σύλληψη ενός έργου προκύπτει ως όλον ή συντίθεται από επιμέρους στοιχεία, ή αν το κυρίαρχο στοιχείο αποτελεί η δομή της σύνθεσης κλπ.

Για τον μελετητή Πολιτικό Μηχανικό δομοστατικής κατεύθυνσης αντικείμενο δημιουργίας αποτελεί η μόρφωση του φέροντα οργανισμού μιας κατασκευής. Για κατασκευές με ιδιαίτερες λειτουργικές και αισθητικές απαιτήσεις, στις σημερινές συνθήκες άσκησης των τεχνικών επαγγελμάτων, χρειάζεται η συνεργασία αρχιτέκτονα και πολιτικού μηχανικού. Η συνεργασία αυτή θα πρέπει να δομείται με πρωτεύοντα το ρόλο του αρχιτέκτονα στις λειτουργικές και αισθητικές απαιτήσεις του δε πολιτικού μηχανικού στη μόρφωση και επάρκεια σχεδιασμού του φέροντα οργανισμού. Η δημιουργική συνεργασία αυτή πρέπει να υπάρχει σε όλες τις φάσεις και να είναι στενότερη κατά τις αρχικές φάσεις του σχεδιασμού. Ακόμη όμως και στη περίπτωση, που συχνά απαντάται, όπου ο αρχιτέκτονας έχει προβεί στην οριστικοποίηση της μορφής μιας κατασκευής, ο ρόλος του δομοστατικού πολιτικού μηχανικού είναι ακόμη πιο σύνθετος. Ακολουθώντας την κεντρική ιδέα του αρχιτεκτονήματος θα πρέπει να μορφώσει τον φέροντα οργανισμό κατά τρόπο που να διατάσσει τις αμυντικές δυνάμεις του κατά ορθολογικό τρόπο σύμφωνα με τις γενικές και ειδικότερες απαιτήσεις που καθορίζουν οι επιστημονικές και τεχνικές γνώσεις. Συχνά κατά την αναζήτηση αυτή προκύπτουν λύσεις σημαντικά καλύτερες που προσκρούουν όμως στις επιλογές της αρχιτεκτονικής λύσης. Προκύπτουν όμως και τα επιχειρήματα που συνηγορούν για κάποια μερική ή γενικότερη τροποποίηση-αναθεώρηση η οποία σε συνεργασία με τον αρχιτέκτονα οδηγεί τον σχεδιασμό σε καλύτερους δρόμους.

Είναι σημαντικό να τονιστεί ότι με τα σημερινά δεδομένα διαχωρισμού των επαγγελματικών κατευθύνσεων η εκπαίδευση των αρχιτεκτόνων σωστά διακατέχεται από την τάση της συνθετικής δημιουργίας, ενώ για τους πολιτικούς μηχανικούς κυριαρχεί μια αναλυτική τάση στη προσπάθεια ενσωμάτωσης των θεωριών και μεθόδων των τελευταίων ετών με αποτέλεσμα την μηδαμινή εκπαίδευση τους σε συνθετικές προσπάθειες. Η χαρακτηριστική αυτή διαφορά των δύο τεχνικών ειδικοτήτων καθιστά στην πράξη δύσκολη τη συνεργασία που θα πρέπει να γεφυρωθεί με τον εμπλουτισμό της εκπαίδευσης των αρχιτεκτόνων και πολιτικών μηχανικών με στοιχεία επικάλυψης.

Με την υποστήριξη της αναλυτικής δουλειάς του πολιτικού μηχανικού από τους υπολογιστές και τα προγράμματα που έχουν αναπτυχθεί το όλο θέμα έχει ισορροπίσει σε μία νέα κατάσταση. Τα προγράμματα που διατίθενται ακολουθούν τη λογική «δίδεται ο φορέας και η φόρτιση, να επιλυθεί και να διαστασιολογηθεί ο φορέας σύμφωνα με τους ισχύοντες κανονισμούς». Έτσι ο μηχανικός μορφώνει τον φέροντα οργανισμό και τον εισάγει στο πρόγραμμα ως δεδομένο. Στην περίπτωση αστοχίας μιας κατάστασης αντοχής ή λειτουργικότητας επεμβαίνει συνήθως τοπικά με την τροποποίηση των διαστάσεων των μελών που αστόχησαν ώστε σε νέα επίλυση να ικανοποιούνται όλοι οι έλεγχοι.

Με δεδομένη την υπολογιστική ισχύ των Η/Υ που εξασφαλίζουν σε ελάχιστο χρόνο τη ανάλυση και διαστασιολόγηση ενός δεδομένου φορέα τίθεται το θέμα της δημιουργίας μιας νέας γενεάς προγραμμάτων που θα διευρύνουν τις δυνατότητες σχεδιασμού των κατασκευών προσφέροντας καλύτερες λύσεις σε σχέση με κάποια κριτήρια. Όπως λοιπόν με τη τρέχουσα διαδικασία ο μελετητής τροποποιεί τις διαστάσεις των μελών που αστόχησαν για να επαναλάβει ένα κύκλο ανάλυσης και διαστασιολόγησης μπορεί και ένας αλγόριθμος που εκτελείται από ένα πρόγραμμα να τροποποιεί αυτόματα τις διαστάσεις αυτές με κάποιο κανόνα ώστε να δημιουργείται το επόμενο βήμα στην επαναληπτική διαδικασία ανάλυσης και σχεδιασμού. Η κατεύθυνση των αλλαγών αυτών θα πρέπει να αφορά ένα σύνολο παραμέτρων οι οποίες να θεωρούνται ως μεταβλητές κατα την επαναληπτική αυτή διαδικασία. Το σύνολο των παραμέτρων ενός προβλήματος που θεωρούνται μεταβαλλόμενες για τις ανάγκες της επαναληπτικής διαδικασίας σχεδιασμού καλούνται μεταβλητές σχεδιασμού. Οι μεταβλητές αυτές μπορεί να ανοίκουν στους πραγματικούς αριθμούς οπότε καλούνται *συνεχείς*, μπορεί να είναι όμως ακέραιοι αριθμοί ή γενικότερες τιμές από ένα κατάλογο τιμών όπως για παράδειγμα οι πρότυπες διατομές των μεταλλικών στοιχείων οπότε καλούνται *διακριτές* μεταβλητές σχεδιασμού. Το εύρος διακύμανσης των μεταβλητών σχεδιασμού συνθέτουν το χώρο των λύσεων του προβλήματος που καλείται χώρος σχεδιασμού (design space). Μέσα στον χώρο αυτό υπάρχουν λύσεις που ικανοποιούν τους περιορισμούς του προβλήματος οπότε ορίζουν το υποσύνολο των αποδεκτών λύσεων (feasible solutions). Όλες οι αποδεκτές λύσεις κατά μία γενική θεώρηση είναι και λύσεις του προβλήματος. Ενδέχεται όμως να ενδιαφερόμαστε για μία βαθμονόμηση των λύσεων ως προς κάποιο κριτήριο ή κριτήρια που διατυπώνονται μαθηματικά από μία αντικειμενική συνάρτηση (objective function), όπως για παράδειγμα τον σχεδιασμό ελάχιστου βάρους ή ελάχιστου κόστους για δεδομένες τιμές μονάδος των υλικών. Άν καθοριστεί ένα τέτοιο κριτήριο αμέσως οι λύσεις κατατάσσονται ως προς το κριτήριο αυτό που αποτελεί πλέον τον αντικειμενικό στόχο οπότε έχουμε μία ή περισσότερες αποδεκτές λύσεις που αποτελούν το βέλτιστο σχεδιασμό (optimal design) σε σχέση με την επιλεγμένη αντικειμενική συνάρτηση του προβλήματος.

Μια τέτοια καθοδηγούμενη προσπάθεια συστηματικής παραμετρικής διερεύνησης ενός προβλήματος σχεδιασμού που είναι δυνατή με τα διαθέσιμα υπολογιστικά μέσα είναι και ο βέλτιστος σχεδιασμός των κατασκευών. Από την φύση του δεν μπορεί να λειτουργήσει ως θεωρία σχεδιασμού διότι υπολείπεται σημαντικά κυρίως στη φάση του προκαταρκτικού σχεδιασμού (conceptual design). Είναι βέλτιστος σε σχέση με κάποιο προκαθορισμένο κριτήριο ή κριτήρια, η δε διερεύνηση των λύσεων που προσφέρει περιορίζεται από την, εν πολλοίς, αυθαίρετη επιλογή των μεταβλητών σχεδιασμού και από το εύρος διακύμανσής τους. Έτσι η επιλογή των μεταβλητών σχεδιασμού αποκτά καθοριστική σημασία, ενώ επίσης η επιλογή της αντικειμενικής συνάρτησης επηρεάζει σημαντικά τον βέλτιστο σχεδιασμό. Η δυνατότητα αναγωγής του προβλήματος σε μαθηματικό πρόβλημα επίσης περιορίζει την διερεύνηση σε οντότητες που μπορούν να αποτιμηθούν μέσω μιας συνεχούς ή διακριτής μεταβλητής που δεν μπορούν να αποδώσουν για παράδειγμα την αισθητικά αρμονική σύνθεση κάποιων οντοτήτων.

Οι μέθοδοι βελτιστοποίησης αναπτύχθηκαν αρχικά για οικονομικά συστήματα και επεκτάθηκαν στη διοίκηση επιχειρήσεων και διαχείριση διεργασιών (operations research). Στις κατασκευές βρίσκουν σημαντική εφαρμογή στην αυτοκινητοβιομηχανία και στην αεροπορική βιομηχανία. Στις εφαρμογές αυτές η βελτιστοποίηση από πλευράς κόστους για παράδειγμα ενός βραχίονα ανάρτησης ενός τροχού αυτοκινήτου που παράγεται σε εκατοντάδες χιλιάδες τεμάχια έχει σημαντική επίδραση στο κόστος εν γένει τυποποιημένων κατασκευών. Για τις κατασκευές πολιτικού μηχανικού συχνά διατυπώνεται ο αντίλογος ότι λόγω της έλλειψης της τυποποίησης αλλά κυρίως για παράδειγμα, του μικρού σχετικά κόστους του φέροντα οργανισμού (15%-25%) του συνολικού κόστους ενός κτιριακού έργου, η αναζήτηση κάποιας βέλτιστης λύσης (ελάχιστο βάρος ή ελάχιστο κόστος) στερείται αντικειμένου. Επιπλέον δε η εξοικείωση του πολιτικού μηχανικού με τέτοια εργαλεία τον απομακρύνει από την συνθετική αναζήτηση λύσεων που πληρούν ευρύτερα κριτήρια.

Παρόλο που η κριτική αυτή εστιάζει σε υπαρκτά θέματα, θα αντέτεινε κανείς ότι η διερεύνηση με τον υπολογιστή εναλλακτικών λύσεων στη πορεία αναζήτησης της βέλτιστης λύσης με βάση κάποιο κριτήριο μπορεί να είναι χρήσιμη και προς όφελος ενός καλύτερου σχεδιασμού στην κατεύθυνση των επιλογών του μελετητή, χωρίς σε καμία περίπτωση να φιλοδοξεί να υποκαταστήσει τον δομοστατικό σχεδιασμό. Επιπλέον η διαδικασία του βέλτιστου σχεδιασμού μπορεί να τερματίζεται όχι μόνο στη βέλτιστη, από μαθηματικής πλευράς, λύση αλλά και άλλες «περίπου βέλτιστες αποδεκτές λύσεις» (near optimal feasible solutions).

Με βάση τις μεταβλητές σχεδιασμού που μπορούν να επιλεγούν καθορίζονται τρείς μεγάλες ενότητες προβλημάτων βέλτιστου σχεδιασμού των κατασκευών. Η πρώτη ενότητα αφορά στη βέλτιστη διαστασιολόγηση με βάση κάποια αντικειμενική συνάρτηση, δηλ. την επιλογή των διατάσεων των διατομών των μελών ενός φορέα (sizing optimization). Μια δεύτερη ενότητα αφορά το σχήμα του φορέα, δηλ. τις συντεταγμένες που ορίζουν τους κόμβους ενός ραβδωτού ή επιφανειακού φορέα (shape optimization), και η τρίτη ενότητα αφορά την συνδεσμολογία ενός ραβδωτού φορέα, ή το κατά πόσο σε μία περιοχή ενός επιφανειακού φορέα θα τοποθετηθεί υλικό ή όχι (topology optimization). Συνήθως οι μεταβλητές σχεδιασμού ανήκουν στις δύο πρώτες κατηγορίες ενώ η βέλτιστη διάταξη της συνδεσμολογίας δηλ. ο τοπολογικός βέλτιστος σχεδιασμός είναι ο πλέον δυσχερής. Είναι επίσης φανερό ότι η βέλτιστη διάταξη του υλικού αποτελεί τη σημαντικότερη επιλογή σε σχέση με τη βέλτιστη διαστασιολόγηση και το βέλτιστο σχήμα της κατασκευής, προβλήματα στα οποία κυρίως δίδονται σήμερα απαντήσεις από την θεωρία του βέλτιστου σχεδιασμού.

Ιδιαίτερη αναφορά θα πρέπει να γίνει με βάση την επιλογή της αντικειμενικής συνάρτησης των προβλημάτων βέλτιστου σχεδιασμού των κατασκευών. Παραδοσιακά το ελάχιστο βάρος ή το ελάχιστο κόστος χρησιμοποιούνται για το σκοπό αυτό. Συμπληρωματικά διατυπώνεται ως αντικειμενικός στόχος η μέγιστη δυσκαμψία ως προς κάποιες μετακινήσεις για δεδομένο όγκο υλικού. Είναι φανερό ότι ειδικότερα για τις κατασκευές πολιτικού μηχανικού η διερεύνηση εναλλακτικών αντικειμενικών συναρτήσεων που απορρέουν από σύγχρονες απαιτήσεις σχεδιασμού των κατασκευών αποκτά ιδιαίτερη σημασία. Για παράδειγμα η εξασφάλιση της όσο το δυνατόν καθυστερημένης κατάρρευσης ενός φορέα για δεδομένο όγκο υλικού μπορεί να έχει ιδιαίτερη σημασία για τον σχεδιασμό κατασκευών πολιτικού μηχανικού.

2

ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

Ο βέλτιστος σχεδιασμός των κατασκευών χρησιμοποιεί τις μαθηματικές θεωρίες και αλγορίθμους που αναπτύχθηκαν αρχικά σε προβλήματα διαχείρισης (operations research) και σε οικονομικά προβλήματα.

2.1 ΜΕΤΑΒΛΗΤΕΣ ΣΧΕΔΙΑΣΜΟΥ (design variables)

Ένα πρόβλημα σχεδιασμού μιας κατασκευής περιλαμβάνει τον φορέα, την καταπόνηση και την συμπεριφορά του. Κάθε μία από τις ενότητες αυτές περιγράφεται με πλήθος παραμέτρων που καθορίζουν το υλικό και τις μηχανικές του ιδιότητες, την συνδεσμολογία των μελών του φορέα, την γεωμετρία του φορέα και τις διατομές των μελών του, τις φορτίσεις και τους συνδυασμούς φορτίσεων που ορίζουν οι σχετικοί κανονισμοί και τους ελέγχους αντοχής και λειτουργικότητας που εξασφαλίζουν την συμπεριφορά του.

Αν από το σύνολο των παραμέτρων που ορίζουν το πρόβλημα επιλεγούν ορισμένες ως άγνωστες μεταβλητές που μπορούν να μεταβάλλονται εντός προκαθορισμένων ορίων είναι φανερό ότι ορίζεται μια πολυπαραμετρική οικογένεια κατασκευών που ορίζει τον χώρο των λύσεων του προβλήματος. Οι μεταβλητές αυτές καλούνται μεταβλητές σχεδιασμού του προβλήματος $x_i,\ \ i = 1, 2, \dots, n$ οι οποίες ορίζουν το διάνυσμα των μεταβλητών σχεδιασμού $\mathbf{x}$ του προβλήματος. Ο καθορισμός των μεταβλητών σχεδιασμού αποτελεί το πρώτο βήμα της διατύπωσης του προβλήματος βέλτιστου σχεδιασμού. Οι μεταβλητές σχεδιασμού θα πρέπει να είναι κατά το δυνατόν ανεξάρτητες μεταξύ τους ώστε το πρόβλημα να μην περιπλέκεται άσκοπα χρησιμοποιώντας περιττές μεταβλητές οι οποίες εκ των υστέρων θα πρέπει να δεσμεύονται κατάλληλα ώστε να εξασφαλίζεται το φυσικό νόημα της περιγραφής του προβλήματος.

Για παράδειγμα, αναζητώντας τις μεταβλητές σχεδιασμού μιας κοίλης κυκλικής διατομής, από τα τρία μεγέθη d_{in}, d_{ex}, t (εσωτερική, εξωτερική διάμετρος, πάχος διατομής), είναι φανερό όμως ότι οι δύο είναι ανεξάρτητες και η τρίτη εξαρτημένη καθόσον π.χ. το πάχος t μπορεί να εκφραστεί συναρτήσει των διαμέτρων d_{in}, d_{ex} ως $t = 0.5(d_{ex} - d_{in})$

2.1.1 Συνεχείς και Διακριτές μεταβλητές σχεδιασμού

Οι μεταβλητές σχεδιασμού διακρίνονται σε συνεχείς (continuous) και διακριτές (discrete), ανάλογα με το πεδίων τιμών στο οποίο δύνανται να μεταβάλλονται. Αν π.χ η εξωτερική διάμετρος της κοίλης κυκλικής διατομής που αναφέρθηκε μπορεί να λαμβάνει τιμές στο διάστημα $\lfloor 10,\, 20 \rfloor cm$ τότε χαρακτηρίζεται ως συνεχής. Αν όμως επιθυμούμε να επιλύσουμε το πρόβλημα και να χρησιμοποιήσουμε μία από τις πρότυπες διατομές του εμπορίου τότε η εξωτερική διάμετρος θα πρέπει να λαμβάνει διακριτές τιμές στο ίδιο ή άλλο διάστημα όπως αυτές ορίζονται από σχετικούς πίνακες. Στη περίπτωση αυτή η μεταβλητή σχεδιασμού χαρακτηρίζεται ως διακριτή.

Αντίστοιχα αν κατά τον σχεδιασμό μιας μεταλλικής δοκού επιθυμούμε να χρησιμοποιήσουμε μια πρότυπη διατομή π.χ IPE 200,220,..., 400 η σχετική μεταβλητή σχεδιασμού ανήκει στην κατηγορία των διακριτών μεταβλητών.

Οι μέθοδοι επίλυσης προβλημάτων με διακριτές μεταβλητές συχνά απαιτούν πρόσθετη υπολογιστική προσπάθεια και τη χρήση άλλων μεθόδων υπολογισμού της βέλτιστης λύσης για τις οποίες δεν ορίζονται παράγωγοι των εμπλεκόμενων συναρτήσεων. Συνήθως, το πρόβλημα λύνεται θεωρώντας συνεχείς μεταβλητές και στη συνέχεια επιλέγονται οι πλησιέστερες διακριτές τιμές από τους πίνακες και ο σχεδιασμός που προκύπτει ελέγχεται για το αν είναι εφικτός. Πάντως αποό μαθηματικής πλευράς δεν είναι πάντοτε βέβαιο ότι ο σχεδιασμός που προκύπτει είναι ο βέλτιστος.

Στην πράξη πολλές φορές ορισμένες μεταβλητές του προβλήματος είναι συνεχείς και ορισμένες διακριτές οπότε το διάνυσμα των μεταβλητών σχεδιασμού είναι μικτό.

2.1.2 Όρια των μεταβλητών σχεδιασμού

Για την αποτελεσματική θεώρηση του προβλήματος οι μεταβλητές σχεδιασμού περιορίζονται εντός συγκεκριμένων ορίων εντός των οποίων ενδιαφερόμαστε για την βέλτιστη λύση. Η επιλογή των ορίων αυτών δεν είναι πάντοτε εύκολη και γίνεται σε συνδυασμό με την επιλογή των μεταβλητών σχεδιασμού κατά τρόπο που να εξασφαλίζεται το φυσικό τους νόημα. Για παράδειγμα αν για την κοίλη διατομή επιλεγούν η εξωτερική και εσωτερική διάμετρος ως συνεχείς μεταβλητές του προβλήματος και επιλεγούν για αυτές τα διαστήματα $\lfloor 10,\, 20 \rfloor cm$ και $\lfloor 8,\, 18 \rfloor cm$ για την εξωτερική και εσωτερική διάμετρο αντίστοιχα ενδέχεται κατά την διαδικασία βελτιστοποίησης να επιλεγεί τιμή της εσωτερικής διαμέτρου μεγαλύτερη αυτής της εξωτερικής οπότε η διατομή δεν ορίζεται και το εμβαδόν της διατομής προκύπτει αρνητικό. Έτσι ο καθορισμός των ορίων των μεταβλητών σχεδιασμού απαιτεί ιδιαίτερη προσοχή και μπορεί να οδηγεί και στην αναθεώρηση της επιλογής των ανεξάρτητων μεταβλητών σχεδιασμού όπως για παράδειγμα είναι προτιμότερη η επιλογή της εξωτερικής διαμέτρου και του πάχους ως ανεξάρτητων μεταβλητών και ο καθορισμός ορίων που διασφαλίζουν θετικό εμβαδόν για την διατομή.

2.2 ΠΕΡΙΟΡΙΣΜΟΙ

Εκτός από τα όρια μεταβολής των μεταβλητών σχεδιασμού, για την ορθή διατύπωση ενός προβλήματος βέλτιστου σχεδιασμού απαιτείται και η διατύπωση ενός συνόλου περιορισμών που θα εξασφαλίζει αποδεκτές λύσεις δηλ. λύσεις που θα ανταποκρίνονται στο περιγραφόμενο πρόβλημα. Οι περιορισμοί αυτοί εν γένει αποτελούν εκφράσεις μερικών ή και όλων των μεταβλητών σχεδιασμού του προβλήματος σε άμεση ή πεπλεγμένη μορφή. Διατυπώνονται με την μορφή ανισοτήτων ή ισοτήτων και αφορούν στον περιορισμό της περιγραφής του προβλήματος και στην οριοθέτηση της συμπεριφοράς του. Για παράδειγμα μπορεί να περιγράφουν τον περιορισμό του βέλους κάμψεως μίας αμφιέρειστης δοκού κοίλης κυκλικής διατομής με ομοιόμορφο φορτίο στο μέσον της ώστε να ικανοποιείται ο σχετικός έλεγχος λειτουργικότητας.

$$\delta = \frac{5}{384}\frac{ql^4}{EI(d_{ex},\, t)} \le 0.02m \tag{1.1}$$

όπου οι μεταβλητές σχεδιασμού εμφανίζονται άμεσα στη ροπή αδρανείας της διατομής.

Οι περιορισμοί διακρίνονται σε <u>περιορισμούς ισότητας</u> (equality constraints) και <u>ανισοτικούς περιορισμούς </u>(inequality contraints).

Οι περιορισμοί ανισότητας συνήθως αφορούν τις παραμορφώσεις και τις αναπτυσσόμενες τάσεις που δεν πρέπει να ξεπερνούν τις επιτρεπόμενες τιμές $(\delta \le \delta_{\varepsilon\pi})$ και $(\sigma \le \sigma_{\varepsilon\pi})$. Σε ένα υπερστατικό φορέα οι τάσεις εν γένει εξαρτώνται από τα αδρανειακά χαρακτηριστικά των μελών του φορέα οπότε οι περιορισμοί των τάσεων αποτελούν πεπλεγμένες εκφράσεις των μεταβλητών σχεδιασμού που ανάλογα με την μέθοδο ανάλυσης ενδέχεται να εκφράζονται αναλυτικά συναρτήσει των μεταβλητών σχεδιασμού ή και όχι γεγονός που περιπλέκει το πρόβλημα όπως θα δούμε σε επόμενα κεφάλαια.

Το σύνολο των περιορισμών του προβλήματος θα πρέπει να ικανοποιείται ώστε ο σχεδιασμός να θεωρείται επιτρεπτός. Το σύνολο των ανισοτικών περιορισμών που ικανοποιείται ως ισότητα για συγκεκριμένες τιμές των μεταβλητών σχεδιασμού αποτελεί το σύνολο των ενεργών περιορισμών (active constraints) για τον συγκεκριμένο σχεδιασμό. Είναι φανερό ότι ενδέχεται να μην υπάρχει σχεδιασμός για τον οποίο όλοι οι ανισοτικοί περιορισμοί να είναι ενεργοί. Όπως επίσης το σύνολο των περιορισμών που καθορίζουμε για ένα πρόβλημα άθελά μας να ορίζει ένα κενό χώρο λύσεων για το πρόβλημα.

2.2.1 Κανονικοποίηση των περιορισμών

Για τη σωστή εφαρμογή των μεθόδων επίλυσης των προβλημάτων βέλτιστου σχεδιασμού, είναι επιθυμητό να κανονικοποιούνται όλες οι συναρτήσεις που εκφράζουν τους περιορισμούς. Αυτό γίνεται διότι διαφορετικοί περιορισμοί εμπλέκουν μεγέθη με διαφορετικές μονάδες μέτρησης, οπότε δεν μπορούν να αντιμετωπισθούν αριθμητικά κατά ισοδύναμο τρόπο με αποτέλεσμα την συσσώρευση λαθών στρογγύλευσης και αποπροσανατολισμό των αλγορίθμων.

Ας θεωρήσουμε έναν περιορισμό τάσεων

$$\sigma \leq \sigma_\alpha \Rightarrow \sigma - \sigma_\alpha \leq 0$$

και έναν περιορισμό μετακινήσεων

$$\delta \leq \delta_\alpha \Rightarrow \delta - \delta_\alpha \leq 0$$

Παρατηρούμε ότι οι δυο περιορισμοί εμπλέκουν διαφορετικά μεγέθη, την τάση με μονάδες π.χ N/m^2 και την μετακίνηση με μονάδες cm. Αν θεωρήσουμε σ_α=250MPα και δ_α=2cm, γίνεται φανερό ότι οι περιορισμοί αναφέρονται σε διαφορετικές τάξεις μεγέθους.

Η κανονικοποίηση γίνεται διαιρώντας τους περιορισμούς με τις αντίστοιχες επιτρεπόμενες τιμές. Έτσι προκύπτουν:

$$R - 1.0 \leq 0 \quad , \quad \text{όπου } R=\sigma/\sigma_\alpha \quad \text{ή} \quad R=\delta/\delta_\alpha$$

οι δε λόγοι R είναι συγκρίσιμοι με την μονάδα και το δεύτερο μέλος πάντοτε μηδέν χωρίς να βλάπτεται η γενικότητα.

2.3 ΑΝΤΙΚΕΙΜΕΝΙΚΗ ΣΥΝΑΡΤΗΣΗ (objective function)

Όλες οι λύσεις στο χώρο των λύσεων ορίζουν επιτρεπτούς σχεδιασμούς δηλ. σχεδιασμούς που ικανοποιούν τους περιορισμούς του προβλήματος. Για να καθοριστεί μεταξύ αυτών ο βέλτιστος σχεδιασμός απαιτείται ένα κριτήριο το οποίο να τους ιεραρχήσει. Το κριτήριο αυτό καλείται αντικειμενική συνάρτηση του προβλήματος (objective function). Συνήθως ως αντικειμενική συνάρτηση στα προβλήματα βέλτιστου σχεδιασμού των κατασκευών χρησιμοποιείται το ελάχιστο βάρος της κατασκευής. Για το παράδειγμα της κοίλης διατομής το βάρος της δοκού είναι συνάρτηση της εξωτερικής διαμέτρου D και του πάχους t που αποτελούν και τις μεταβλητές σχεδιασμού του προβλήματος:

$$m = \rho.h.A = \rho.h.\pi t(D - t)$$ (h: μήκος της δοκού, ρ: ειδικό βάρος)

Δηλαδή, η αντικειμενική συνάρτηση εξαρτάται από τις μεταβλητές σχεδιασμού του προβλήματος. Για κάθε σχεδιασμό, που προκύπτει θέτοντας αριθμητικές τιμές στις μεταβλητές σχεδιασμού, παίρνουνε μια συγκεκριμένη τιμή της αντικειμενικής συνάρτησης. Έτσι, είναι εύκολο να συγκρίνουμε διάφορους πιθανούς σχεδιασμούς, έχοντας ως κριτήριο την αντικειμενική συνάρτηση: ο βέλτιστος σχεδιασμός είναι εκείνος που ελαχιστοποιεί την τιμή της συνάρτησης.

2.4 Εφαρμογή

Για να γίνουν κατανοητές οι έννοιες που ορίστηκαν ως τώρα, θα θεωρήσουμε το παράδειγμα ενός πακτωμένου υποστυλώματος κοίλης διατομής μιας μεταλλικής κατασκευής, το οποίο έχει ύψος h και καταπονείται από το αξονικό φορτίο P. Το ζητούμενο είναι ο σχεδιασμός (διαστασιολόγηση) ενός στύλου ελαχίστου βάρους χωρίς τον κίνδυνο αστοχίας σε θλίψη ή λυγισμό θεωρούμενου ως δοκού.

α) <u>Μεταβλητές σχεδιασμού:</u> Η επιλογή των μεταβλητών σχεδιασμού δεν είναι μονοσήμαντη, αλλά μπορούμε να επιλέξουμε ως μεταβλητές την εξωτερική διάμετρο D και το πάχος t.

β) <u>Αντικειμενική συνάρτηση</u> : Είναι η συνολική μάζα του στύλου.

$$m = \rho.h.A = \rho.h.\pi t(D - t)$$ (ρ: ειδικό βάρος)

γ) <u>Περιορισμοί</u>

i) <u>έλεγχος θλίψης</u> : Πρέπει $\sigma \leq \sigma_{\varepsilon\pi} \Rightarrow$

$$\frac{P}{A} \leq \sigma_{\varepsilon\pi} \Rightarrow \frac{P}{\pi t(D - t)} \leq \sigma_{\varepsilon\pi}$$

ii) <u>έλεγχος λυγισμού</u> : Πρέπει $P \leq P_{cr} \Rightarrow$

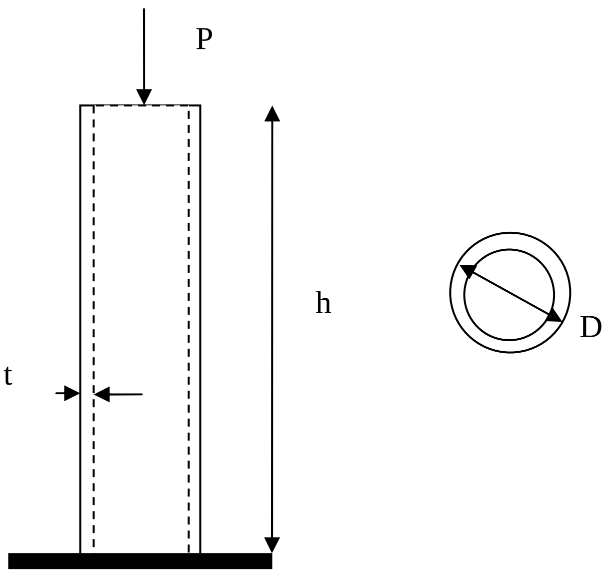
(P_{cr}: κρίσιμο φορτίο λυγισμού)

$$P \leq \frac{\pi^2 EI}{4l^2} \Rightarrow P \leq \frac{\pi^3 E}{256l^2}(D^4 - (D - 2t)^4)$$

όπου, η ροπή αδρανείας τέθηκε $I = \dfrac{\pi}{64}(D^4 - (D - 2t)^4)$

iii) <u>κατασκευαστικές απαιτήσεις</u> : Η διάμετρος και το πάχος θα πρέπει να βρίσκονται μεταξύ συγκεκριμένων ορίων.

$D_{\min} \le D \le D_{\max}$ και $t_{\min} \le t \le t_{\max}$



Πακτωμένος Στύλος κοίλης διατομής

2.5 ΜΑΘΗΜΑΤΙΚΗ ΔΙΑΤΥΠΩΣΗ

Γενικά το πρόβλημα βέλτιστου σχεδιασμού διατυπώνεται μαθηματικά ως εξής:

Να βρεθεί το διάνυσμα των μεταβλητών σχεδιασμού $x^T = (x_1,\ x_2,\ \ldots,\ x_n)$ για το οποίο ελαχιστοποιείται η αντικειμενική συνάρτηση

$$f(\mathbf{x}) = f(x_1,\, x_2,\, \ldots,\, x_n)$$

και ικανοποιούνται κ περιορισμοί ισότητας

$$h_j = h_j(x_1,\, x_2,\, \ldots,\, x_n) = 0, \qquad j = 1, 2, \ldots, \kappa$$

και οι m περιορισμοί ανισότητας

$$g_i(x_1,\, x_2,\, \ldots,\, x_n) \le 0, \qquad i = 1, 2, \ldots, m$$

καθώς και οι περιορισμοί των μεταβλητών σχεδιασμού

$$x_{rl}^{l} \le x_r \le x_r^u, \qquad r = 1, 2, ..., n$$

όπου x_{rl}^{l}, x_r^u το κάτω και άνω όριο των τιμών των μεταβλητών σχεδιασμού αντίστοιχα.

Όλα τα προβλήματα βέλτιστου σχεδιασμού μπορούν να αναχθούν στο παραπάνω μαθηματικό πρόβλημα, και άρα μπορούν να χρησιμοποιηθούν όλες οι μέθοδοι επίλυσης που έχουν αναπτυχθεί στη θεωρία της βελτιστοποίησης γιαυτό.

Ο αριθμός των περιορισμών ισότητας πρέπει να είναι μικρότερος ή το πολύ ίσος με τον αριθμό των μεταβλητών σχεδιασμού, δηλ. $k \le n$. Όταν είναι κ>n, το πρόβλημα έχει πλεονάζοντες περιορισμούς ισότητας και δεν υπάρχει λύση.

Μερικά προβλήματα βέλτιστου σχεδιασμού δεν έχουν καθόλου περιορισμούς, και ονομάζονται προβλήματα βελτιστοποίησης χωρίς περιορισμούς (unconstrained optimization problems), αντίθετα με τα προβλήματα βελτιστοποίησης με περιορισμούς (constrained optimization problems).

Αν οι f(**x**), h_j(**x**), g_i(**x**) είναι γραμμικές συναρτήσεις των μεταβλητών σχεδιασμού **x**, τότε το πρόβλημα ονομάζεται πρόβλημα γραμμικού προγραμματισμού (linear programming problem). Αν κάποια από τις συναρτήσεις είναι μη γραμμική τότε έχουμε ένα πρόβλημα μη γραμμικού προγραμματισμού (nonlinear programming problem).

3

ΓΡΑΜΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ

3.1 ΕΙΣΑΓΩΓΗ

Πολλά προβλήματα βέλτιστου σχεδιασμού κατασκευών ανάγονται στην ελαχιστοποίηση μιας γραμμικής συνάρτησης των μεταβλητών σχεδιασμού που υπόκεινται σε ένα πλήθος γραμμικών περιορισμών ως προς τις ίδιες μεταβλητές. Ένα πρόβλημα βέλτιστου σχεδιασμού του οποίου η αντικειμενική συνάρτηση και οι περιορισμοί είναι γραμμικές συναρτήσεις των μεταβλητών σχεδιασμού, ονομάζεται πρόβλημα *γραμμικού προγραμματισμού* (linear programming problem). Το πρόβλημα του γραμμικού προγραμματισμού κατέχει κεφαλαιώδη θέση στη θεωρία βελτιστοποίησης ενώ βρίσκει πολλές εφαρμογές σε διάφορες επιστημονικές περιοχές. Η ανάλυση του γραμμικού προγραμματισμού είναι σημαντική για δύο επιπλέον λόγους. Ο πρώτος είναι ότι συχνά ένα μη γραμμικό πρόβλημα βελτιστοποίησης μετατρέπεται σε μία ακολουθία προσεγγιστικών γραμμικών προβλημάτων και δεύτερον, οι τεχνικές που αναπτύχθηκαν για την επίλυση του γραμμικού προβλήματος επεκτείνονται και στην επίλυση πιο σύνθετων μη γραμμικών προβλημάτων.

3.2 ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

Η τυπική μορφή του προβλήματος γραμμικού προγραμματισμού ορίζεται ως εξής:

Ελαχιστοποιήστε τη συνάρτηση: $c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$

Που υπόκειται στους περιορισμούς:

$$\begin{aligned}
\alpha_{11} x_1 + \alpha_{12} x_2 + \cdots + \alpha_{1n} x_n &= b_1 \\
\alpha_{21} x_1 + \alpha_{22} x_2 + \cdots + \alpha_{2n} x_n &= b_2 \\
. \qquad\qquad\qquad & \quad . \\
. \qquad\qquad\qquad & \quad . \\
. \qquad\qquad\qquad & \quad . \\
\alpha_{m1} x_1 + \alpha_{m2} x_2 + \cdots + \alpha_{mn} x_n &= b_m
\end{aligned} \tag{3.1}$$

Και $x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0,$

Όπου οι συντελεστές α_{ij}, b_i, c_i είναι δεδομένοι πραγματικοί αριθμοί και ζητούνται οι τιμές των μεταβλητών x_i στο πεδίο των πραγματικών αριθμών που ελαχιστοποιούν την δεδομένη συνάρτηση.

3.2.1 Διατύπωση προβλήματος με μορφή μητρώων

Σε μητρωική μορφή το παραπάνω πρόβλημα γράφεται:

$$\text{Ελαχιστοποιήστε:} \qquad c^T x \tag{3.2}$$

Που υπόκειται στους περιορισμούς: $Ax = b$ και $x \geq 0$,

Όπου x είναι ένα διάνυσμα n - διαστάσεων, c^T μια γραμμή n - διαστάσεων, A ένα ορθογωνικό μητρώο διαστάσεων $m \times n$, και b ένα διάνυσμα m - διαστάσεων.

Στην παραπάνω τυπική μορφή του προβλήματος ανάγονται και άλλες μεγάλες κατηγορίες προβλημάτων με κατάλληλες προσαρμογές.

Ανισοτικοί περιορισμοί

Για το παρακάτω πρόβλημα και ζητείται η αναγωγή του σε τυπικό πρόβλημα γραμμικού προγραμματισμού.

Ελαχιστοποιήστε τη συνάρτηση: $c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$

Που υπόκειται στους περιορισμούς:

$$\begin{array}{ll} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n \leq b_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n \leq b_2 \\ . \qquad\qquad\qquad . \\ . \qquad\qquad\qquad . \\ . \qquad\qquad\qquad . \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n \leq b_m \end{array} \tag{3.3}$$

και $x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0,$

το πρόβλημα με ανισοτικούς περιορισμούς μικρότερους ορισμένων τιμών μπορεί να αναχθεί στο τυπικό πρόβλημα με την εισαγωγή m πρόσθετων μεταβλητών – ελλειμμάτων (slack variables) που καθιστούν τις ανισότητες ισότητες, δηλαδή το πρόβλημα μπορεί να διατυπωθεί ως εξής:

Ελαχιστοποιήστε τη συνάρτηση: $c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$

Που υπόκειται στους περιορισμούς:

$$
\begin{array}{l}
\alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n + y_1 \qquad\qquad = b_1 \\
\alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n \qquad + y_2 \qquad = b_2 \\
. \qquad\qquad\qquad\qquad . \\
. \qquad\qquad\qquad\qquad . \\
. \qquad\qquad\qquad\qquad . \\
\alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n \qquad\qquad + y_m = b_m
\end{array}
\tag{3.4}
$$

και $x_1 \geq 0, x_2 \geq 0, ..., x_n \geq 0,$

και $y_1 \geq 0, y_2 \geq 0, ..., y_m \geq 0,$

Θεωρώντας ότι το πρόβλημα στην παραπάνω μορφή του έχει $n+m$ αγνώστους δηλ. $x_1, x_2, ..., x_n, y_1, y_2, ..., y_m$, διατυπώνεται στην τυπική μορφή ενός προβλήματος γραμμικού προγραμματισμού. Σε μητρωική μορφή το μητρώο A των συντελεστών των περιορισμών, διαστάσεων $m \times (n+m)$ διαχωρίζεται σε $[A, I]$ όπου το A είναι διαστάσεων $m \times n$ και I το μοναδιαίο μητρώο διαστάσεων $m \times m$.

Στη περίπτωση που ορισμένοι ανισοτικοί περιορισμοί είναι της μορφής:

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n \geq b_i \tag{3.5}$$

είναι φανερό ότι μπορούν να διατυπωθούν κατά ισοδύναμο τρόπο ως εξής:

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n - y_i = b_i \tag{3.6}$$

όπου $y_i \geq 0$. Οι πρόσθετες μεταβλητές που ορίζονται σ'αυτή την περίπτωση καλούνται πλεονάσματα (surplus variables) τα οποία αφαιρούμενα αποκαθιστούν την ισότητα στον περιορισμό.

Επίσης πολλαπλασιασμός ενός περιορισμού με πλην ένα (-1) ισοδυναμεί με μεταλλαγή των ελλειμματικών μεταβλητών σε πλεονασματικές μεταβλητές και αντίστροφα, γεγονός που επιτρέπει πάντοτε την αναγωγή του προβλήματος στην τυπική μορφή του προβλήματος γραμμικού προγραμματισμού.

Στην περίπτωση που μια μεταβλητή δεν είναι θετική ή μηδέν (μη-αρνητική) ο πλέον πρόσφορος τρόπος αναγωγής του προβλήματος στην τυπική μορφή είναι η απαλοιφή της μεταβλητής και ενός περιορισμού που την περιλαμβάνει, δηλ. αν για παράδειγμα η μεταβλητή x_1 είναι ελεύθερη τότε χρησιμοποιώντας έναν περιορισμό

$$\alpha_{i1}x_1 + \alpha_{i2}x_2 + \cdots + \alpha_{in}x_n = b_i \tag{3.7}$$

όπου $\alpha_{i1} \neq 0$, η μεταβλητή x_1 μπορεί να εκφραστεί ως γραμμικός συνδυασμός των υπολοίπων συν κάποια σταθερά.. Η έκφραση αυτή αν χρησιμοποιηθεί για την αντικατάσταση της μεταβλητής x_1 σε όλους τους περιορισμούς οδηγεί στη διατύπωση ενός προβλήματος γραμμικού προγραμματισμού ως προς τις $n-1$ μεταβλητές δηλ. $x_2,...,x_n$ και $m-1$ περιορισμούς. Η τιμή της μεταβλητής x_1 προσδιορίζεται από την παραπάνω έκφραση μετά την επίλυση του μειωμένου προβλήματος γραμμικού προγραμματισμού.

Βασικές Λύσεις

Ας θεωρήσουμε το σύστημα των ισοτήτων:

$$[A]\{x\} = \{b\} \tag{3.8}$$

όπου x είναι ένα διάνυσμα n-διαστάσεων, b ένα διάνυσμα m-διαστάσεων και A ένα μητρώο διαστάσεων $m \times n$. Το παραπάνω σύστημα έχει νόημα στην περίπτωση που οι μεταβλητές είναι περισσότερες από τους περιορισμούς δηλ. $n > m$. Τότε υπάρχει η δυνατότητα επιλογής μεταξύ των μεταβλητών, ενώ για την περίπτωση $n = m$ μιλάμε για την επίλυση του συστήματος των n αλγεβρικών εξισώσεων με n αγνώστους, από τη λύση του οποίου προκύπτουν οι τιμές των $x_i, i = 1, n$ οι οποίες αφού αντικατασταθούν στην συνάρτηση δίνουν τη μοναδική τιμή της συναρτήσεως προς ελαχιστοποίηση. Όλα αυτά έχουν βεβαίως νόημα για την περίπτωση όπου οι m εξισώσεις είναι γραμμικά ανεξάρτητες μεταξύ τους.

Αν από τις n στήλες του μητρώου A επιλεγούν m ανεξάρτητες στήλες τότε μορφώνεται ένα τετραγωνικό μητρώο διαστάσεων $m \times m$ το οποίο καλείται βάση $[B]$. Από τη λύση του συστήματος

$$[B]\{x_B\} = \{b\} \tag{3.9}$$

προκύπτει το διάνυσμα x_B διαστάσεων m. Θέτοντας $x = (x_B, 0)$, δηλ. συμπληρώνοντας τις υπόλοιπες μεταβλητές ως μηδενικές ορίζουμε μία λύση του προβλήματος

$$[A]\{x\} = \{b\} \tag{3.10}$$

Μια τέτοια λύση ονομάζεται *βασική λύση* (basic solution) που αντιστοιχεί στο αντιστρέψιμο μητρώο B. Οι μεταβλητές που συμμετέχουν στη βάση B καλούνται βασικές μεταβλητές. Είναι φανερό ότι για ένα πρόβλημα γραμμικού προγραμματισμού με n μεταβλητές και m περιορισμούς είναι δυνατή η μόρφωση ενός πλήθους από βασικές λύσεις που αντιστοιχούν στους συνδυασμούς των m στηλών από n στήλες, δηλ:

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!} \tag{3.11}$$

Όλες οι βασικές μεταβλητές μίας βασικής λύσης δεν είναι κατ'ανάγκη μη-μηδενικές. Αν μία ή περισσότερες βασικές μεταβλητές μιας βασικής λύσης είναι μηδέν τότε η λύση αναφέρεται ως εκφυλισμένη βασική λύση (degenerate basic solution).

Οι λύσεις του τυπικού προβλήματος γραμμικού προγραμματισμού θα πρέπει επίσης να είναι μεγαλύτερες-ίσες του μηδενός. Μία λύση του συστήματος (9) αναφέρεται ως *αποδεκτή λύση* (feasible solution). Αν η λύση είναι επίσης βασική λύση τότε αναφέρεται ως *βασική αποδεκτή λύση* (basic feasible solution)

Εφαρμογή

Ζητείται να βρεθούν οι τιμές των μεταβλητών σχεδιασμού x_1 και x_2 για τις οποίες ελαχιστοποιείται η αντικειμενική συνάρτηση:

$$f = -400x_1 - 600x_2$$

σύμφωνα με τους περιορισμούς: $x_1 + x_2 \le 16$

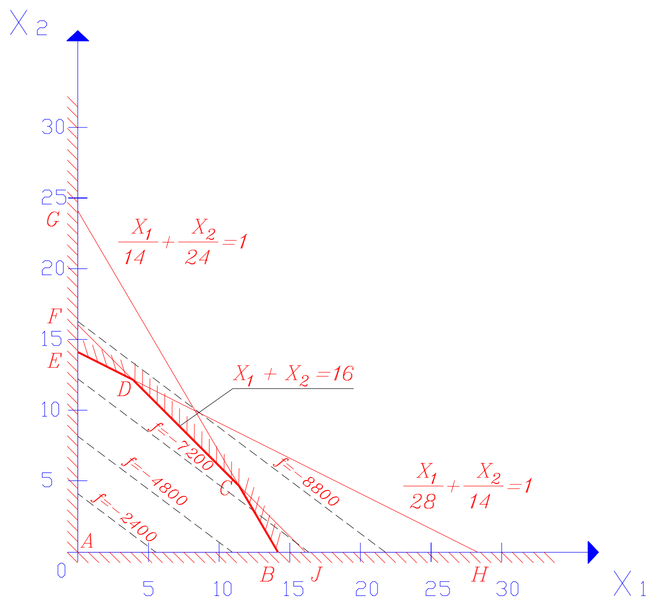
$$\frac{1}{28}x_1 + \frac{1}{14}x_2 \le 1$$

$$\frac{1}{14}x_1 + \frac{1}{24}x_2 \le 1$$

$$x_1, x_2 \ge 0$$

Λύση

Η γραφική λύση του προβλήματος δίνεται στο σχήμα 1, όπου έχουν σχεδιαστεί όλοι οι περιορισμοί και οι ισοϋψείς της αντικειμενικής συνάρτησης. Η επιτρεπτή περιοχή των λύσεων του προβλήματος περικλείεται από το πολύγωνο ABCDE και όπως είναι φανερό η κορυφή D δίνει τη βέλτιστη λύση.



Σχήμα 1. Γραφική επίλυση προβλήματος Γραμμικού Προγραμματισμού

Για να μετατρέψουμε το πρόβλημα στην τυπική μορφή προβλήματος γραμμικού προγραμματισμού εισάγουμε τις βοηθητικές μεταβλητές σχεδιασμού x_3, x_4, x_5. Έτσι, οι περιορισμοί ανισότητας μετασχηματίζονται σε περιορισμούς ισότητας.

$$x_1 + x_2 + x_3 = 16$$

$$\frac{1}{28}x_1 + \frac{1}{14}x_2 + x_4 = 1$$

$$\frac{1}{14}x_1 + \frac{1}{24}x_2 + x_5 = 1$$

$x_i \geq 0$ i=1 έως 5

Παρατηρούμε ότι ο αριθμός των μεταβλητών σχεδιασμού (n=5) είναι μεγαλύτερος από τον αριθμό των περιορισμών (m=3). Έτσι, το σύστημα των περιορισμών έχει άπειρες λύσεις. Αυτό φαίνεται μεταφέροντας τις μεταβλητές x_1, x_2 στα δεύτερα μέλη των περιορισμών:

$$x_3 = 16 - x_1 - x_2$$

$$x_4 = 1 - \frac{1}{28}x_1 - \frac{1}{14}x_2$$

$$x_5 = 1 - \frac{1}{14}x_1 - \frac{1}{24}x_2$$

Στις εξισώσεις αυτές, οι μεταβλητές x_1 και x_2 είναι ανεξάρτητες, και έτσι ορίζουν μια απειρία λύσεων για τις x_3, x_4 και x_5.

Η λύση που έχει ενδιαφέρον για τα προβλήματα γραμμικού προγραμματισμού προκύπτει θέτοντας p μεταβλητές σχεδιασμού ίσες με το μηδέν και υπολογίζοντας τις υπόλοιπες, όπου p είναι η διαφορά του αριθμού των μεταβλητών (n) και του αριθμού των περιορισμών (m), δηλαδή p=n-m (στο συγκεκριμένο παράδειγμα: p=5-3=2). Έτσι, προκύπτει μοναδική λύση για τις υπόλοιπες μεταβλητές. Η λύση αυτή ονομάζεται βασική λύση (basic solution). Για παράδειγμα, αν θέσουμε x_1=0 και x_2=0 θα προκύψουν x_3=16, x_4=1, x_5=1. Οι μεταβλητές που τίθενται ίσες με το μηδέν καλούνται μη βασικές, ενώ οι υπόλοιπες βασικές.

Συνολικά, υπάρχουν 10 βασικές λύσεις (συνδυασμός των 5 μεταβλητών ανά δύο), οι οποίες φαίνονται στον παρακάτω πίνακα.

No.	x_1	x_2	x_3	x_4	x_5	F	
1	0	0	16	1	1	0	(A) Αποδεκτή
2	0	14	2	0	5/12	-8400	(E) Αποδεκτή
3	0	16	0	-2/7	1/3	-	Μη αποδεκτή
4	0	24	-8	-5/7	0	-	Μη αποδεκτή
5	16	0	0	3/7	-2/7	-	Μη αποδεκτή
6	14	0	2	½	0	-5600	(B) Αποδεκτή
7	28	0	-12	0	-1	-	Μη αποδεκτή

No.	x_1	x_2	x_3	x_4	x_5	F	
8	4	12	0	0	3/14	-8800	(D) Αποδεκτή
9	11,2	48	0	1/5	0	-7360	(C) Αποδεκτή
10	130/17	168/17	-26/17	0	0	-	Μη αποδεκτή

Από τις 10 λύσεις του συγκεκριμένου προβλήματος, οι 5 παραβιάζουν τους περιορισμούς $x_i \geq 0$.

Οι υπόλοιπες λύσεις είναι αποδεκτές και όπως παρατηρούμε, αντιστοιχούν στις κορυφές του πολυγώνου που περικλείει την αποδεκτή περιοχή ABCDE. Οι λύσεις αυτές ονομάζονται βασικές αποδεκτές λύσεις (basic feasible solutions). Τέλος, μετακινούμενοι προς τη διεύθυνση μείωσης των ισοϋψών της αντικειμενικής συνάρτησης, φαίνεται ότι η βέλτιστη λύση εμφανίζεται στην κορυφή D του πολυγώνου.

3.3 Θεμελιώδες Θεώρημα του Γραμμικού Προγραμματισμού

Το θεμελιώδες θεώρημα του γραμμικού προγραμματισμού αναδεικνύει την πρωταρχική σημασία των βασικών αποδεκτών λύσεων για την εύρεση της βέλτιστης λύσης του τυπικού προβλήματος γραμμικού προγραμματισμού. Το θεώρημα αποδεικνύει ότι η βέλτιστη λύση είναι πάντοτε μια βασική αποδεκτή λύση και άρα κατά την αναζήτηση της βέλτιστης λύσης απαιτείται μόνο η θεώρηση των βασικών αποδεκτών λύσεων. Το θεώρημα αναδεικνύει επίσης το τυπικό πρόβλημα γραμμικού προγραμματισμού ως πρόβλημα συνδυαστικής βελτιστοποίησης (combinatorial optimization) αφού ανάγει το πρόβλημα της εύρεσης της βέλτιστης λύσης μεταξύ ενός πεπερασμένου συνόλου βασικών αποδεκτών λύσεων. Η προσπάθεια όμως ανάπτυξης ενός αλγορίθμου που διερευνά τις βασικές αποδεκτές λύσεις και προσδιορίζει την βέλτιστη λύση είναι υπολογιστικά δαπανηρή ιδίως για μεγάλα προβλήματα. Επιπλέον η τυχαία εναλλαγή των βάσεων δεν οδηγεί κατ’ανάγκη σε αποδεκτές λύσεις. Απαιτείται λοιπόν ένας αλγόριθμος ο οποίος ξεκινώντας από μία βασική αποδεκτή λύση να μεταβαίνει σε μία άλλη βασική αποδεκτή λύση και ταυτόχρονα να ελαττώνει την τιμή της αντικειμενικής συνάρτησης για να αποφεύγει τις άσκοπες περιπλανήσεις.

Θεμελιώδες Θεώρημα Γραμμικού Προγραμματισμού. *Για το τυπικό πρόβλημα γραμμικού προγραμματισμού ισχύουν τα εξής:*

(i) *αν υπάρχει αποδεκτή λύση, υπάρχει και βασική αποδεκτή λύση*

(ii) *αν υπάρχει μια βέλτιστη αποδεκτή λύση, υπάρχει μια βέλτιστη βασική αποδεκτή λύση.*

Η απόδειξη του θεμελιώδους θεωρήματος γίνεται με αλγεβρικό ή γεωμετρικό τρόπο και δεν αναπτύσσεται στις σημειώσεις αυτές.

3.3.1 Μετάβαση από Βασική-Αποδεκτή λύση σε άλλη

Για να αναδειχθεί ο τρόπος με τον οποίο μεταβαίνουμε από τη μία βασική-αποδεκτή λύση σε άλλη, ας θεωρήσουμε ότι υπάρχει μια βασική-αποδεκτή λύση $x = (x_1, x_2, ..., x_m, 0, 0, ..., 0)$. Η λύση αυτή ικανοποιεί τη σχέση:

$$x_1\{\alpha_1\} + x_2\{\alpha_2\} + \cdots + x_m\{\alpha_m\} = \{b\} \tag{3.12}$$

Θεωρώντας ότι η λύση δεν είναι εκφυλισμένη δηλ. $x_i > 0 \;\; i = 1,2,...,m$ αποφασίζουμε να εισάγουμε στη λύση την στήλη $\{\alpha_q\}$που δεν μετέχει στην αρχική λύση. Εφόσον τα διανύσματα $\{\alpha_i\}, \, i = 1,2,...,m$ αποτελούν βάση στο χώρο των m-διαστάσεων μπορούν να χρησιμοποιηθούν για να εκφράσουν και το διάνυσμα $\{\alpha_q\}$ως εξής:

$$\{\alpha_q\} = y_{1q}\{\alpha_1\} + y_{2q}\{\alpha_2\} + \cdots + y_{mq}\{\alpha_m\} \tag{3.13}$$

Πολλαπλασιάζοντας την σχέση (13) με την ποσότητα $\varepsilon \geq 0$ και αφαιρώντας από την σχέση (12) λαμβάνουμε:

$$(x_1 - \varepsilon\, y_{1q})\{\alpha_1\} + (x_2 - \varepsilon\, y_{2q})\{\alpha_2\} + \cdots + (x_m - \varepsilon\, y_{mq})\{\alpha_m\} + \varepsilon\{\alpha_q\} = \{b\} \tag{3.14}$$

Η παραπάνω σχέση εκφράζει το διάνυσμα $\{b\}$ ως γραμμικό συνδυασμό m+1 διανυσμάτων. Για $\varepsilon = 0$ η σχέση δίνει την αρχική βασική αποδεκτή λύση. Για θετικές τιμές του ε η λύση είναι αποδεκτή αλλά όχι βασική καθώς εμπλέκει m+1 διανύσματα. Αυξάνοντας την τιμή του ε υπάρχει μία τιμή που μηδενίζει έναν τουλάχιστον από τους συντελεστές των βασικών διανυσμάτων της αρχικής λύσης. Η τιμή αυτή είναι:

$$\varepsilon = \min_i \left\{ \frac{x_i}{y_{iq}} : y_{iq} > 0 \right\} \tag{3.15}$$

για την περίπτωση αυτή προκύπτει μία νέα βασική-αποδεκτή λύση στην οποία εισάγεται το διάνυσμα $\{\alpha_q\}$στη βάση και απομακρύνεται το διάνυσμα p που αντιστοιχεί στην παραπάνω τιμή του ε . Αν κανένας από τους συντελεστές y_{iq} δεν είναι θετικός τότε για αυξανόμενο ε όλοι οι συντελεστές στην σχέση (15) αυξάνουν γεγονός που σημαίνει ότι το σύνολο των αποδεκτών λύσεων του προβλήματος δεν είναι φραγμένο.

3.4 Κανονική Μορφή προβλήματος Γραμμικού Προγραμματισμού

Θεωρούμε το σύστημα των αλγεβρικών εξισώσεων

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &= b_1 \\ \alpha_{21}x_1 + \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &= b_2 \\ . \qquad\qquad\qquad & \;\; . \\ . \qquad\qquad\qquad & \;\; . \\ . \qquad\qquad\qquad & \;\; . \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \cdots + \alpha_{mn}x_n &= b_m \end{aligned} \tag{3.16}$$

όπου $m \leq n$, ή σε μητρωική μορφή

$$[A]\{x\} = \{b\} \tag{3.17}$$

Αν οι εξισώσεις είναι γραμμικά ανεξάρτητες τότε γνωρίζουμε και από τα συστήματα n εξισώσεων με n αγνώστους ότι μπορούμε να αντικαταστήσουμε μία εξίσωση με ένα μη-μηδενικό πολλαπλάσιο της συν ένα γραμμικό συνδυασμό από άλλες εξισώσεις του συστήματος. Η ιδιότητα αυτή αποτελεί τη βάση της απαλοιφής κατά Gauss με σκοπό την τριγωνική μορφή του συστήματος. Στο γραμμικό προγραμματισμό ακολουθούμε την ίδια διαδικασία με σκοπό την διαγωνοποίηση του συστήματος ως προς κάποια βασική αποδεκτή λύση ώστε το τυπικό πρόβλημα να περιέλθει στην παρακάτω μορφή που καλείται κανονική μορφή του προβλήματος γραμμικού προγραμματισμού.

$$\begin{aligned} x_1 \qquad\qquad & + y_{1,m+1}x_{m+1} + y_{1,m+2}x_{m+2} + \cdots + y_{1,n}x_n = y_{10} \\ x_2 \qquad & + y_{2,m+1}x_{m+1} + y_{2,m+2}x_{m+2} + \cdots + y_{2,n}x_n = y_{20} \\ & \quad . \qquad\qquad\qquad\qquad\qquad . \\ & \quad . \qquad\qquad\qquad\qquad\qquad . \\ & \quad . \qquad\qquad\qquad\qquad\qquad . \\ x_m & + y_{m,m+1}x_{m+1} + y_{m,m+2}x_{m+2} + \cdots + y_{m,n}x_n = y_{m0} \end{aligned} \tag{3.18}$$

Στην παραπάνω μορφή οι μεταβλητές $x_1, x_2, ..., x_m$ αναφέρονται ως βασικές ενώ οι υπόλοιπες ως μη-βασικές. Η βασική λύση που αντιστοιχεί στην παραπάνω κανονική μορφή είναι:

$$x_1 = y_{10},\; x_2 = y_{20},\; ...,\; x_m = y_{m0},\; x_{m+1} = 0, \ldots ,\; x_n = 0 \tag{3.19}$$

Αν πινακοποιήσουμε τους συντελεστές του συστήματος προκύπτει ο παρακάτω πίνακας:

$$\begin{array}{ccccccc}
1 & 0 \cdots 0 & y_{1,m+1} & y_{1,m+2} & \cdots y_{1,n} & y_{10} \\
0 & 1 \cdots 0 & y_{2,m+1} & y_{2,m+2} & \cdots y_{2,n} & y_{20} \\
. & & . & & & . \\
. & & . & & & . \\
. & & . & & & . \\
0 & 0 \cdots 1 & y_{m,m+1} & y_{m,m+2} & \cdots y_{m,n} & y_{m0}
\end{array} \tag{3.20}$$

Η μορφή αυτή είναι χρήσιμη για την εναλλαγή της βάσης. Ας υποθέσουμε ότι μία μη-βασική μεταβλητή είναι να γίνει βασική και μία βασική να γίνει μη-βασική. Η διαδικασία είναι η εξής. Αν η μη-βασική μεταβλητή αντιστοιχεί στη στήλη q και η βασική στη στήλη p τότε η εναλλαγή μπορεί να γίνει μόνο Αν $y_{pq} \neq 0$. Διαιρούμε την γραμμή p με τον συντελεστή y_{pq} με σκοπό να κάνουμε τον συντελεστή της μη-βασικής μεταβλητής q μονάδα στην p εξίσωση. Στη συνέχεια αφαιρούμε κατάλληλα πολλαπλάσια της γραμμής p από όλες τις άλλες γραμμές ώστε οι συντελεστές στη στήλη q να μηδενίζονται. Οι νέοι συντελεστές του νέου κανονικού προβλήματος αναλυτικά έχουν ως εξής:

$$\begin{cases} y'_{ij} = y_{ij} - \dfrac{y_{pj}}{y_{pq}} y_{iq}, \qquad i \neq p \\ y'_{pj} = \dfrac{y_{pj}}{y_{pq}} \end{cases} \tag{3.21}$$

Οι εξισώσεις αυτές αναφέρονται ως εξισώσεις περιστροφής (pivot equations) το δε στοιχείο y_{pq} στοιχείο περιστροφής (pivot element).

3.5 Μέθοδος Simplex

Οι σχέσεις των προηγουμένων παραγράφων αντιστοιχούν στις εξισώσεις των περιορισμών. Εδώ ενδιαφέρει η ανάπτυξη μιας μεθόδου που προσδιορίζει την βασική-αποδεκτή λύση που καθιστά την αντικειμενική συνάρτηση ελάχιστη. Έτσι στον προηγούμενο πίνακα μπορούμε να προσθέσουμε ακόμη μια σειρά με τους συντελεστές της αντικειμενικής συνάρτησης $c_i \;\; i = 1,2,...,n$ οι οποίοι με δεδομένη βασική-αποδεκτή λύση τροποποιούνται ώστε οι συντελεστές των βασικών μεταβλητών να μηδενιστούν έτσι ο πίνακας γίνεται:

$$
\begin{array}{cccccccc}
\alpha_1 & \alpha_2 \ldots \alpha_m & \alpha_{m+1} & \alpha_{m+2} & \ldots & \alpha_n & & b \\
\hline
1 & 0 \cdots 0 & y_{1,m+1} & y_{1,m+2} & \cdots & y_{1,n} & & y_{10} \\
0 & 1 \cdots 0 & y_{2,m+1} & y_{2,m+2} & \cdots & y_{2,n} & & y_{20} \\
. & & . & & & & & . \\
. & & . & & & & & . \\
. & & . & & & & & . \\
0 & 0 \cdots 1 & y_{m,m+1} & y_{m,m+2} & \cdots & y_{m,n} & & y_{m0} \\
\hline
0 & 0 \cdots 0 & r_{m+1} & r_{m+2} & \cdots & r_n & & -z_0
\end{array}
\tag{3.22}
$$

οι συντελεστές r_j αναφέρονται ως συντελεστές σχετικού κόστους, ενώ η τιμή της αντικειμενικής συνάρτησης που αντιστοιχεί στη συγκεκριμένη βασική-αποδεκτή λύση είναι η z_0. Οι συντελεστές σχετικού κόστους εκφράζουν την ευαισθησία της αντικειμενικής συνάρτησης ως προς την συγκεκριμένη μη-βασική μεταβλητή. Ενδιαφέρουν έτσι οι αρνητικές τιμές των συντελεστών που σημαίνουν ότι εισαγωγή της αντίστοιχης μη-βασικής μεταβλητής στην νέα βασική-αποδεκτή λύση θα έχει μειωτικό αποτέλεσμα στην αντικειμενική συνάρτηση. Από όλες τις αρνητικές τιμές συνήθως επιλέγεται η μικρότερη τιμή που υποδεικνύει και την αντίστοιχη μη-βασική μεταβλητή που θα εισαχθεί στην νέα βασική-αποδεκτή λύση. Το επόμενο βήμα αποσκοπεί στο να καθορίσει ποια βασική μεταβλητή θα αποσυρθεί από την νέα βασική-αποδεκτή λύση. Ο προσδιορισμός του στοιχείου περιστροφής (pivot element) γίνεται υπολογίζοντας τους λόγους $\dfrac{y_{i0}}{y_{iq}}$ για όλους τους θετικούς συντελεστές $y_{iq}, i = 1,2,...,m$ και επιλέγοντας τον μικρότερο έστω στη σειρά p. Εφαρμόζοντας την διαδικασία περιστροφής (pivoting) με βάση το στοιχείο αυτό προκύπτει μια νέα βασική-αποδεκτή λύση. Τροποποιώντας κατά τον ίδιο τρόπο και την τελευταία γραμμή προκύπτει στην κάτω δεξιά γωνία η νέα τιμή της αντικειμενικής συνάρτησης. Συνοψίζοντας τα βήματα της μεθόδου Simplex είναι τα εξής:

Βήμα 0. Από τον πίνακα με την μορφή (3.22) οι συντελεστές σχετικού κόστους προκύπτουν με αναγωγή της σειράς ώστε οι συντελεστές των βασικών μεταβλητών να μηδενίζονται.

Βήμα 1. Αν κάθε συντελεστής $r_j \geq 0$, stop, η λύση είναι η βέλτιστη

Βήμα 2. Επέλεξε την στήλη q που αντιστοιχεί στον ελάχιστο αρνητικό συντελεστή σχετικού κόστους r_q

Βήμα 3. Υπολόγισε τους λόγους y_{i0}/y_{iq} για όλα τα θετικά y_{iq} της στήλης q. Αν κανένα y_{iq} δεν είναι θετικό, stop το πρόβλημα είναι μη-φραγμένο. Αν όχι επέλεξε ως p τον δείκτη ι που αντιστοιχεί στον μικρότερο συντελεστή.

Βήμα 4. Εφάρμοσε την διαδικασία περιστροφής ως προς το στοιχείο y_{pq} τροποποιώντας όλες τις γραμμές συμπεριλαμβανομένης και της τελευταίας.

Με βάση τον παραπάνω αλγόριθμο που βασίζεται στο θεμελιώδες θεώρημα του γραμμικού προγραμματισμού και υποθέτοντας ότι δεν παράγονται εκφυλισμένες (degenerate) λύσεις το πρόβλημα οδηγείται στη βέλτιστη λύση σε πεπερασμένο αριθμό επαναλήψεων ή σύμφωνα με το βήμα 3 στο συμπέρασμα ότι το πρόβλημα είναι μη-φραγμένο.

Εφαρμογή: Να μεγιστοποιηθεί η συνάρτηση $3x_1 + x_2 + x_3$ που υπόκειται στους περιορισμούς:

$$\begin{aligned}
&2x_1 + x_2 + x_3 \le 2 \\
&x_1 + 2x_2 + 3x_3 \le 5 \\
&2x_1 + 2x_2 + x_3 \le 6 \\
&x_1 \ge 0,\ x_2 \ge 0,\ x_3 \ge 0
\end{aligned} \tag{3.23}$$

Για να μετατρέψουμε το παραπάνω πρόβλημα σε τυπικό πρόβλημα γραμμικού προγραμματισμού πολλαπλασιάζουμε την αντικειμενική συνάρτηση με –1 για να μετατρέψουμε το πρόβλημα της μεγιστοποίησης σε πρόβλημα ελαχιστοποίησης. Για να μετατρέψουμε τους ανισοτικούς περιορισμούς σε ισοτικούς εισάγουμε τρεις νέες μη-αρνητικές μεταβλητές x_4, x_5, x_6. Έχουμε έτσι τον αρχικό πίνακα Simplex ως εξής:

	α_1	α_2	α_3	α_4	α_5	α_6	B
	2	1	1	1	0	0	2
	1	2	3	0	1	0	5
	2	2	1	0	0	1	6
r^T	-3	-1	-3	0	0	0	0

Η παραπάνω μορφή είναι κανονική μορφή με βασικές μεταβλητές τις νέο-εισαγμένες μεταβλητές x_4, x_5, x_6. Σε αυτή μπορεί να εφαρμοστούν τα βήματα της μεθόδου Simplex. Επιλέγουμε μία από τις πρώτη ή τρίτη στήλη, έστω την τρίτη στήλη που αντιστοιχεί σε συντελεστή σχετικού κόστους (-3) για να εισαχθεί στη νέα βασική λύση. Στη συνέχεια υπολογίζονται οι λόγοι (2/1, 5/3, 6/1) και επιλέγεται ο 5/3 ως ελάχιστος. Επιλέγεται ως στοιχείο περιστροφής το στοιχείο (2,3) και εφαρμόζονται οι σχέσεις της περιστροφής (pivoting).

	α_1	α_2	α_3	α_4	α_5	α_6	B
	5/3	1/3	0	1	-1/3	0	1/3
	1/3	2/3	1	0	1/3	0	5/3
	5/3	4/3	0	0	-1/3	1	13/3
r^T	-2	1	0	0	1	0	5

Ο παραπάνω πίνακας Simplex αντιστοιχεί στην βασική-αποδεκτή λύση με βάση τις στήλες 4,3,6 και τιμή της αντικειμενικής συνάρτησης –6.

Επειδή υπάρχουν αρνητικοί συντελεστές σχετικού κόστους η διαδικασία συνεχίζεται με την επιλογή της πρώτης στήλης για εισαγωγή στη βάση. Στη συνέχεια υπολογίζονται οι λόγοι (3/15,15/3,13/5) και επιλέγεται η θέση του μικρότερου που στην συγκεκριμένη περίπτωση είναι η πρώτη. Έτσι το στοιχείο στη θέση (1,1) αποτελεί τον πόλο για την δεύτερη περιστροφή που έχει ως εξής:

	α_1	α_2	α_3	α_4	α_5	α_6	B
	1	1/5	0	3/5	-1/5	0	1/5
	0	3/5	1	-1/5	2/5	0	8/5
	0	1	0	-1	0	1	4
r^T	0	7/5	0	6/5	3/5	0	27/5

Η παραπάνω κανονική μορφή είναι και αυτή που αντιστοιχεί στη βέλτιστη βασική αποδεκτή λύση με τιμή της αντικειμενικής συνάρτησης –27/5 καθόσον όλοι οι συντελεστές σχετικού κόστους είναι θετικοί. Οι βασικές μεταβλητές είναι οι $x_1 = 1/5,\ x_3 = 8/5,\ x_6 = 4$ και όλες είναι αποδεκτές δηλ. μεγαλύτερες του μηδενός, ενώ μη βασικές μεταβλητές και άρα μηδενικές είναι οι $x_2 = 0,\ x_4 = 0,\ x_5 = 0$.

3.6 Εκφυλισμένες λύσεις

Κατά την διαδικασία της μεθόδου Simplex ενδέχεται κάποιες ενδιάμεσες λύσεις να είναι εκφυλισμένες λύσεις δηλ. λύσεις με μηδενικό συντελεστή σε μία ή περισσότερες βασικές μεταβλητές. Η μέθοδος Simplex δεν αστοχεί σ'αυτές τις περιπτώσεις ακολουθεί όμως άσκοπους κύκλους για να ξαναγυρίσει στη αρχική εκφυλισμένη λύση. Για την αποφυγή των άσκοπων αυτών κύκλων με εκφυλισμένες λύσεις έχουν αναπτυχθεί διαφορές μέθοδοι. Μεταξύ αυτών η μέθοδος Bland είναι η πιο απλή. Σύμφωνα με τη μέθοδο αυτή:

α) η στήλη που επιλέγεται να εισαχθεί στη βάση κατά την διαδικασία της μεθόδου Simplex είναι η στήλη με τον μικρότερο δείκτη, $j = \min\{j : r_j < 0\}$

β) στη περίπτωση που υπάρχει επιλογή στο καθορισμό της στήλης που αποχωρεί από την βάση τότε επιλέγεται αυτή με τον μικρότερο δείκτη.

Ο απλός αυτός κανόνας αρκεί να αποτρέψει την μέθοδο από κύκλους με εκφυλισμένες λύσεις.

3.7 Μητρωική Μορφή της Μεθόδου Simplex

Η μέθοδο Simplex με την μορφή πίνακα είναι χρήσιμη αλλά αντιμετωπίζει το πρόβλημα στο επίπεδο του στοιχείου του πίνακα χωρίς να προσφέρει την εποπτεία που παρέχει η μητρωική διατύπωση. Επιπλέον στην κανονική μορφή του πίνακα δεν είναι δυνατή η ενσωμάτωση των συμπερασμάτων από την επίλυση γραμμικών συστημάτων αλγεβρικών εξισώσεων. Έτσι χρήσιμη είναι η διατύπωση του προβλήματος σε μητρωική μορφή.

Ας θεωρήσουμε ότι το μητρώο A και τα διανύσματα x και c^T του τυπικού προβλήματος γραμμικού προγραμματισμού διαχωρίζονται ως εξής:

$$\begin{aligned} [A] &= [B, D] \\ \{x\} &= \{x_B, x_D\} \\ \{c\}^T &= [c_B^T, c_D^T] \end{aligned} \tag{3.24}$$

Έτσι το τυπικό πρόβλημα γραμμικού προγραμματισμού διατυπώνεται ως εξής:

Ελαχιστοποίησε την συνάρτηση: $\{c_B\}^T\{x_B\} + \{c_D\}^T\{x_D\}$

Που υπόκειται στους περιορισμούς: $[B]\{x_B\} + [D]\{x_D\} = \{b\}$ και $\{x_B\} \geq 0, \quad \{x_D\} \geq 0$

Μία βασική αποδεκτή λύση προκύπτει ως $\{x_B\} = [B]^{-1}\{b\}, \; \{x_D\} = 0$. Για οποιαδήποτε τιμή των μη-βασικών μεταβλητών οι βασικές μεταβλητές προσδιορίζονται από τη σχέση:

$$\{x_B\} = [B]^{-1}\{b\} - [B]^{-1}[D]\{x_D\} \tag{3.25}$$

Αντικαθιστώντας την παραπάνω γενική τιμή των βασικών μεταβλητών στην αντικειμενική συνάρτηση προκύπτει:

$$\begin{aligned} z &= \{c_B\}^T \{[B]^{-1}\{b\} - [B]^{-1}[D]\{x_D\}\} + \{c_D\}^T \{x_D\} \\ &= \{c_B\}^T [B]^{-1}\{b\} + \left(\{c_D\}^T - \{c_B\}^T [B]^{-1}[D]\right)\{x_D\} \end{aligned} \tag{3.26}$$

Η σχέση αυτή εκφράζει την αντικειμενική συνάρτηση ως προς τις μη-βασικές μεταβλητές. Το μητρώο συντελεστής του διανύσματος των μη-βασικών μεταβλητών εκφράζει τους συντελεστές σχετικού κόστους των μεταβλητών αυτών και ορίζεται ως εξής:

$$\{r_D\}^T = \{c_D\}^T - \{c_B\}^T [B]^{-1}[D] \tag{3.27}$$

Η μικρότερη αρνητική τιμή του διανύσματος κόστους θα προσδιορίσει την στήλη που θα εισαχθεί στη νέα βασική-αποδεκτή λύση.

Η μητρωική μορφή του πίνακα Simplex έχει ως εξής:

$$\begin{bmatrix} [A] & \{b\} \\ \{c\}^T & 0 \end{bmatrix} = \begin{bmatrix} [B] & [D] & \{b\} \\ \{c_B\}^T & \{c_D\}^T & 0 \end{bmatrix} \tag{3.28}$$

Η μορφή αυτή δεν αντιστοιχεί στην κανονική μορφή του προβλήματος γραμμικού προγραμματισμού η οποία προκύπτει πολλαπλασιάζοντας την πρώτη γραμμή με το αντίστροφο του B και απαλείφοντας την γραμμή στη θέση (2,1).

$$T = \begin{bmatrix} I & [B]^{-1}[D] & [B]^{-1}\{b\} \\ 0 & \{c_D\}^T - \{c_B\}^T [B]^{-1}[D] & -\{c_B\}^T [B]^{-1}\{b\} \end{bmatrix} \tag{3.29}$$

Η σχέση αυτή αποτελεί τον πίνακα Simplex σε κανονική μορφή. Η μητρωική διατύπωση της μεθόδου προσφέρεται για την ανάπτυξη μεθόδων που εξοικονομούν αριθμητικές πράξεις και διατυπώνουν αποδοτικές μεθόδους επίλυσης του προβλήματος του γραμμικού προγραμματισμού για την γενική περίπτωση όπως και για περιπτώσεις ειδικής μορφής των μητρώων. Μία τέτοια μέθοδος είναι η τροποποιημένη μέθοδος Simplex με βάση την οποία επιλύονται τα περισσότερα προβλήματα.

3.8 Διερεύνηση λύσεων του προβλήματος

Ένα πρόβλημα γραμμικού προγραμματισμού ενδέχεται να έχει μία ή περισσότερες λύσεις ή ακόμη και καμία λύση όπως προκύπτει από τα παρακάτω παραδείγματα:

3.8.1 Μοναδική λύση

Θεωρούμε το εξής πρόβλημα:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$f = -4x_1 - x_2 + 50$$

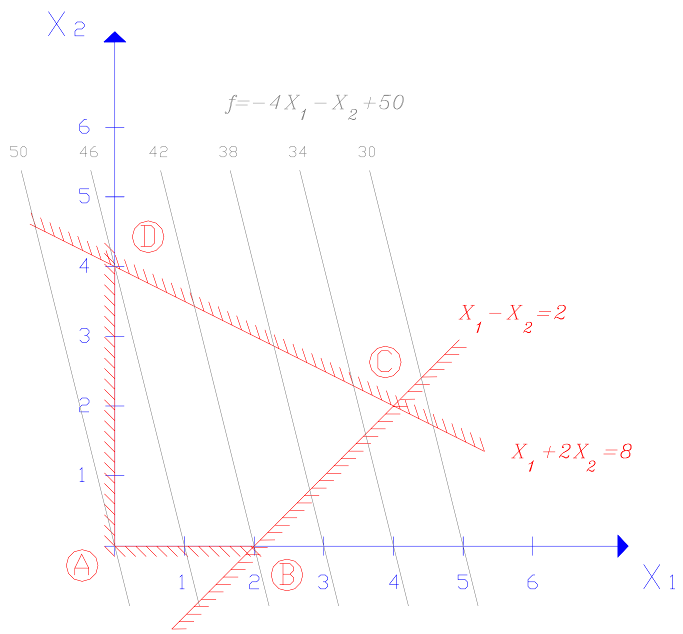
που υπόκειται στους περιορισμούς:

$$x_1 - x_2 \leq 2$$

$$x_1 + 2x_2 \leq 8$$

$$x_1 \geq 0 \quad x_2 \geq 0$$

Ο διδιάστατος χώρος λύσεων γι αυτό το πρόβλημα φαίνεται στο σχήμα (α). Παρατηρούμε ότι οι ισοϋψείς της αντικειμενικής συνάρτησης και τα σύνορα των περιορισμών είναι ευθείες γραμμές. Από το σχήμα είναι φανερό ότι η ελάχιστη τιμή της αντικειμενικής συνάρτησης είναι 32.0 και εμφανίζεται στη θέση x_1=4.0 και x_2=2.0.



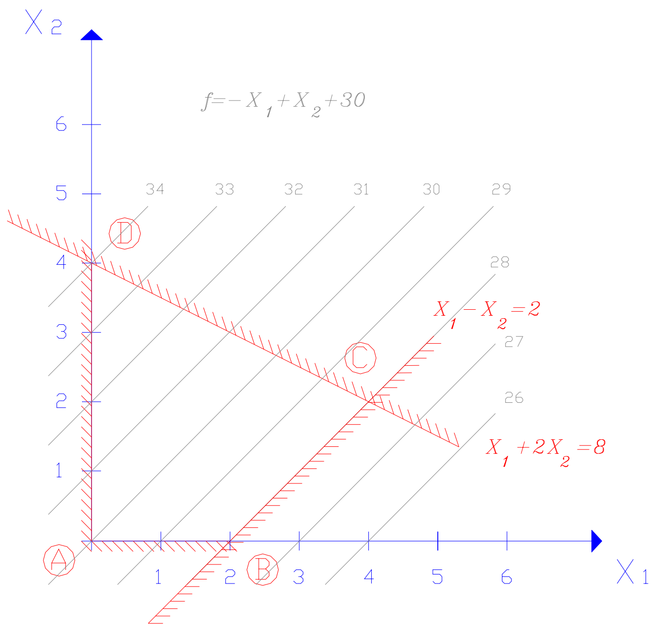
Σχήμα 2. Πρόβλημα Γραμμικού Προγραμματισμού με μοναδική λύση

3.8.2 Μη μοναδική λύση

Αν αντικαταστήσουμε την αντικειμενική συνάρτηση του προβλήματος με την

$$f = -x_1 + x_2 + 30$$

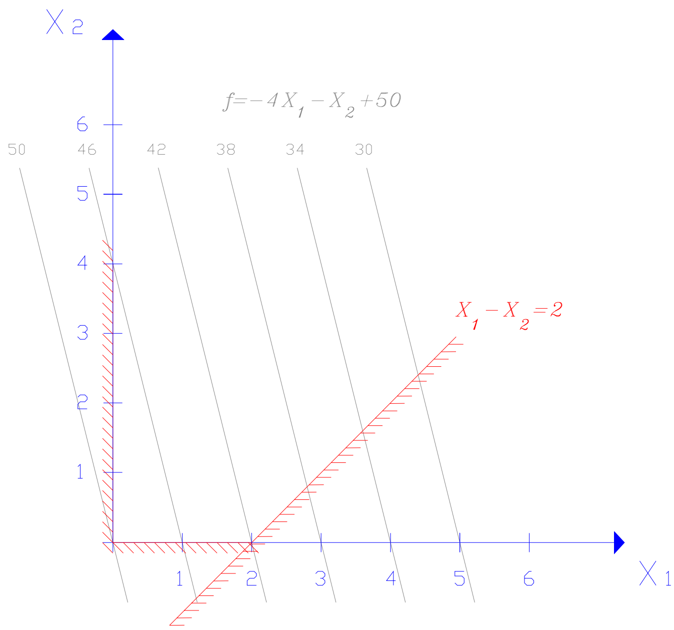
Ο χώρος λύσεων παρουσιάζεται στο σχήμα (b). Παρατηρούμε ότι οι συντελεστές των x_1 και x_2 είναι αντίθετοι των αντιστοίχων συντελεστών του πρώτου περιορισμού, συνεπώς οι ισοϋψείς της αντικειμενικής συνάρτησης είναι παράλληλες προς το σύνορο του πρώτου περιορισμού. Έτσι, οποιοσδήποτε σχεδιασμός (x_1,x_2) επί του συνόρου του περιορισμού αποτελεί λύση του προβλήματος.



Σχήμα 3. Πρόβλημα Γραμμικού Προγραμματισμού με μη-μοναδική λύση

3.8.3 Απεριόριστη λύση

Αν παραλειφθεί ο δεύτερος περιορισμός του προβλήματος, ο χώρος των λύσεων παίρνει την μορφή του σχήματος (c). Όπως παρατηρούμε, η μεταβλητή σχεδιασμού x_1 μπορεί να αυξηθεί απεριόριστα, με αποτέλεσμα η αντικειμενική συνάρτηση να μειώνεται χωρίς περιορισμό μέχρι το άπειρο.



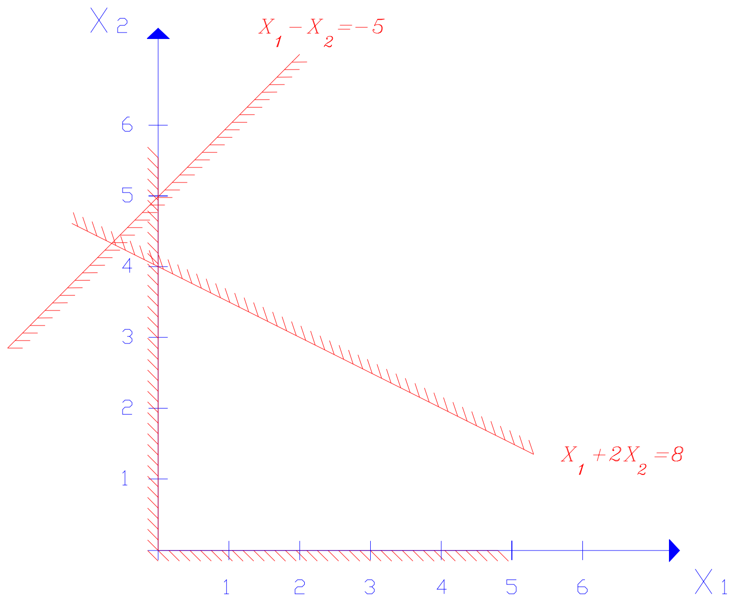
Σχήμα 4. Πρόβλημα Γραμμικού Προγραμματισμού με απεριόριστη λύση

3.8.4 Μη επιτρεπτή λύση

Αντικαθιστώντας τον πρώτο περιορισμό του προβλήματος με τον περιορισμό

$$x_1 - x_2 \leq -5$$

(αλλάζοντας δηλαδή μόνο το δεύτερο μέλος του), διαμορφώνεται ο χώρος των λύσεων του σχήματος (d), Το αποτέλεσμα είναι ότι δεν υπάρχει πλέον σχεδιασμός που να ικανοποιεί όλους τους περιορισμούς, δηλαδή δεν υπάρχει επιτρεπτή λύση.



Σχήμα 5. Πρόβλημα Γραμμικού Προγραμματισμού χωρίς λύση

Στη μέθοδο simplex ξεκινάμε με μια βασική επιτρεπτή λύση, που είναι η κορυφή ενός πολύεδρου (που περικλείει την επιτρεπτή περιοχή), και μετακινούμαστε σε γειτονικές κορυφές με στόχο την μείωση της αντικειμενικής συνάρτησης. Αυτό επιτυγχάνεται αντικαθιστώντας μια βασική μεταβλητή με μια μη βασική. Δεδομένου ότι είναι δυνατόν να υπάρχουν αρκετές γειτονικές κορυφές, είναι φυσικό να αναζητούμε εκείνη που οδηγεί σε μεγαλύτερη μείωση της αντικειμενικής συνάρτησης.

Το ερώτημα που ανακύπτει είναι ποια μη βασική μεταβλητή θα μετατραπεί σε βασική και βέβαια ποια βασική μεταβλητή θα αντικατασταθεί από αυτήν. Με τα παρακάτω παραδείγματα θα φανούν αναλυτικά τα βήματα που ακολουθούνται κατά την επίλυση ενός προβλήματος γραμμικού προγραμματισμού με τη μέθοδο simplex.

Παράδειγμα

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$z = -2x_1 - x_2$$

σύμφωνα με τους περιορισμούς

$$4x_1 + 3x_2 \leq 12$$

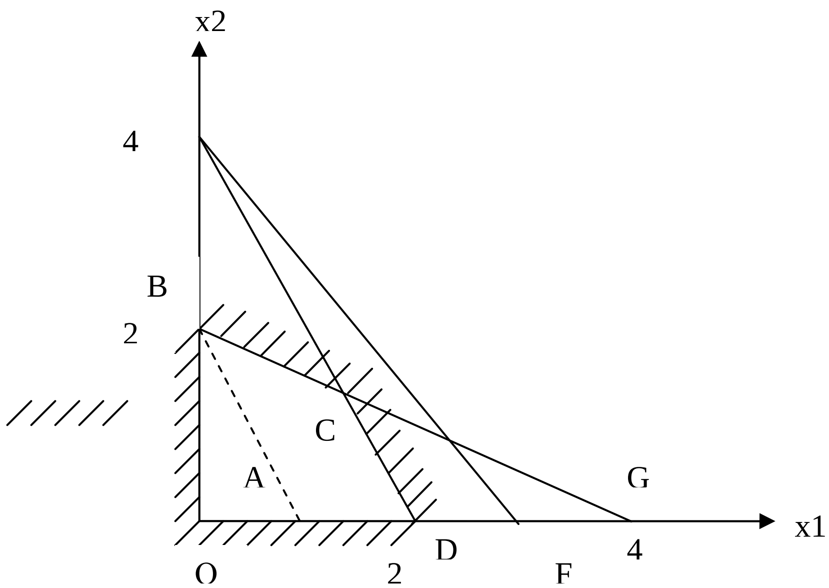
$$2x_1 + x_2 \leq 4$$

$$x_1 + 2x_2 \leq 4$$

$$x_1, x_2 \geq 0$$

<u>Λύση</u>

Η γραφική λύση του προβλήματος απεικονίζεται στο παρακάτω σχήμα. Όπως φαίνεται, το πρόβλημα έχει άπειρες λύσεις (βλ. παράγραφο 2.7.2) κατά μήκος της γραμμής CD (όπου $z^*=4$) διότι η αντικειμενική συνάρτηση είναι παράλληλη προς τον δεύτερο περιορισμό.



Σύμφωνα με την μέθοδο simplex, ακολουθούμε τα παρακάτω βήματα:

1. Γράφουμε το πρόβλημα στην τυπική μορφή εισάγοντας τις βοηθητικές μεταβλητές x_3, x_4 και x_5. Οι περιορισμοί γίνονται:

$$4x_1 + 3x_2 + x_3 = 12$$

$$2x_1 + x_2 + x_4 = 4$$

$$x_1 + 2x_2 + x_5 = 4$$

$$x_i \geq 0 \quad \text{i=1 έως 5}$$

Στην συνέχεια, συντάσσουμε τον πίνακα simplex:

Βασικές μεταβλητές	x_1	x_2	x_3	x_4	x_5	b	λόγος b_i/α_{i1}
X_3	4	3	1	0	0	12	12/4=3
X_4	2	1	0	1	0	4	4/2=2
X_5	1	2	0	0	1	4	4/1=4
	$\boxed{-2}$	-1	0	0	0	f-0	

2. Για να ξεκινήσει η μέθοδος simplex, απαιτείται μια βασική επιτρεπτή λύση. Αυτή είναι ήδη διαθέσιμη, από τον πίνακα simplex και αντιστοιχεί στο σημείο A (0,0) του σχήματος.

Βασικές μεταβλητές: x_3=12, x_4=4, x_5=4 (οι τιμές των βασικών μεταβλητών προκύπτουν απευθείας εξισώνοντας τες με τις αντίστοιχες τιμές της στήλης b).

μη βασικές μεταβλητές: x_1=0, x_2=0 (εξ' ορισμού οι μη βασικές μεταβλητές είναι μηδενικές)

3. Ελέγχουμε την τελευταία γραμμή του πίνακα, η οποία αναφέρεται στην αντικειμενική συνάρτηση. Προφανώς, οι μόνοι μη μηδενικοί όροι βρίσκονται στις στήλες των μη βασικών μεταβλητών σχεδιασμού. Αν όλοι οι μη μηδενικοί όροι είναι μη αρνητικοί, τότε έχουμε αποκτήσει μια βέλτιστη λύση και η μέθοδος simplex μπορεί να τερματιστεί.

4. Στο συγκεκριμένο παράδειγμα, υπάρχουν αρνητικοί όροι στις στήλες των μη βασικών μεταβλητών της τελευταίας γραμμής, συνεπώς η βασική επιτρεπτή λύση που αντιστοιχεί στο σημείο A δεν είναι η βέλτιστη. Θα συνεχίσουμε λοιπόν την αναζήτηση σε μια άλλη βασική επιτρεπτή λύση (κορυφή του πολυέδρου). Για να γίνει αυτό, επιλέγουμε μια μη βασική μεταβλητή για να την μετατρέψουμε σε βασική. Η μη βασική μεταβλητή θα πρέπει να βρίσκεται σε στήλη που αντιστοιχεί σε αρνητικό συντελεστή στην αντικειμενική συνάρτηση. Στο παράδειγμα, και οι δύο μη βασικές μεταβλητές πληρούν αυτήν την προϋπόθεση, εκλέγουμε όμως (εμπειρικά) την μεταβλητή με τον μικρότερο συντελεστή (μεγαλύτερο κατ απόλυτη τιμή) στην αντικειμενική συνάρτηση.

Έτσι, θα μετατρέψουμε την μη βασική μεταβλητή x_1 (συντελεστής στην z: -2) σε βασική. Για το λόγο αυτό, τοποθετούμε το συντελεστή -2 σε ένα πλαίσιο.

5. Για να προσδιορίσουμε ποια βασική μεταβλητή θα γίνει μη βασική, υπολογίζουμε τους λόγους των σταθερών όρων b δια των θετικών συντελεστών της στήλης x_1 όπως φαίνεται στον πίνακα. Η σειρά του πίνακα με τον μικρότερο λόγο είναι η δεύτερη, που αντιστοιχεί στην μεταβλητή x_4. Έτσι, θα αντικατασταθεί η βασική x_1 (1η μη βασική στήλη) και το οδηγό στοιχείο θα είναι το $\alpha_{21}=2$ το οποίο σημειώνεται στον πίνακα με κύκλο. (Η θέση του οδηγού στοιχείου προκύπτει και ως η τομή της γραμμής και της στήλης που θα χρησιμοποιηθούν για την οδήγηση. (απαλοιφή που εδώ είναι η γραμμή 2 και η στήλη 1).

6. Κάνουμε απαλοιφή στη στήλη 1 χρησιμοποιώντας τη γραμμή 2. Διαιρούμε τη 2η γραμμή με 2, την πολλαπλασιάζουμε με 4 και την αφαιρούμε από την 1. Έτσι το στοιχείο $\alpha_{11}=4$ του πίνακα γίνεται $\alpha'_{11}=0$ και το $\alpha_{21}=2$ γίνεται $\alpha'_{21}=1$. Όμοια διαδικασία ακολουθείται για τη γραμμή 3 και τη γραμμή της αντικειμενικής συνάρτησης, οπότε προκύπτει ο νέος πίνακας simplex:

Βασικές μεταβλητές	x_1	x_2	x_3	x_4	x_5	b
X_3	0	1	1	-2	0	4
X_1	1	0.5	0	0.5	0	2
X_5	0	1.5	0	-0.5	1	2
	0	0	0	1	0	f+4

Με τον τρόπο αυτό, έχει ληφθεί μια νέα βασική επιτρεπτή λύση, η οποία είναι:

Βασικές μεταβλητές : $x_3=4$, $x_1=2$, $x_5=2$

μη βασικές μεταβλητές : $x_2=0$ $x_4=0$

που αντιστοιχεί στο σημείο D(2,0) στο οποίο η αντικειμενική συνάρτηση έχει μειωθεί από f=0 σε f=-4. Όπως φαίνεται από τον νέο πίνακα, όλοι οι συντελεστές της τελευταίας σειράς είναι μη αρνητικοί, οπότε δεν είναι δυνατή επιπλέον μείωση της αντικειμενικής συνάρτησης. Συνεπώς, το σημείο D αποτελεί τη βέλτιστη λύση του προβλήματος.

Το γεγονός ότι ο συντελεστής της τελευταίας σειράς που αντιστοιχεί στην μη βασική μεταβλητή x_2 είναι μηδέν, υποδηλώνει ότι το πρόβλημα πιθανόν να έχει πολλαπλές λύσεις (όπως και συμβαίνει). Από πρακτική άποψη, η κατάσταση αυτή είναι επιθυμητή, διότι δίνει εναλλακτικές λύσεις στον μηχανικό.

3.9 Μέθοδοι Simplex δύο φάσεων

Σε πολλά προβλήματα γραμμικού προγραμματισμού, υπάρχουν περιορισμοί ισότητας ή περιορισμοί ανισότητας της «μορφής $\geq$».

Συνήθως, οι ανισοτικοί περιορισμοί μετατρέπονται σε «μορφής $\leq$ «πολλαπλασιάζοντάς τους με μείον ένα, υπάρχει όμως ο κίνδυνος οι σταθεροί όροι b να γίνουν αρνητικοί. Επειδή περιορισμοί τέτοιων μορφών δεν αντιστοιχούν στην τυπική μορφή των προβλημάτων γραμμικού προγραμματισμού, η αρχική βάση επιτρεπτή λύση που απαιτείται δεν είναι άμεσα διαθέσιμη από τον πίνακα simplex. Για να αναχθούν οι περιορισμοί στην τυπική μορφή, προστίθενται σ΄αυτούς οι μη αρνητικές τεχνητές μεταβλητές (artificial variables).

Οι τεχνητές μεταβλητές διευρύνουν το πολύεδρο της επιτρεπτής περιοχής. Έτσι, η αρχική βασική επιτρεπτή λύση που παίρνουμε από τον πίνακα, αντιστοιχεί σε μιά κορυφή του νέου διευρυμένου πολύεδρου. Το πρόβλημα τώρα είναι - διατρέχοντας τις κορυφές του διευρυμένου πολύεδρου - να φτάσουμε όσο το δυνατόν γρηγορότερα στο πραγματικό πολύεδρο και να αναζητήσουμε εκεί τη βέλτιστη τιμή. Όταν φτάσουμε στο πραγματικό πολύεδρο, το διευρυμένο θα απομακρυνθεί, απομακρύνοντας τις τεχνητές μεταβλητές.

Για να εξαλείψουμε τις τεχνητές μεταβλητές , ορίζουμε μια τεχνητή συνάρτηση, γνωστή ως τεχνητή αντικειμενική συνάρτηση (artificial cost function), η οποία είναι το άθροισμα όλων των τεχνητών μεταβλητών. Έτσι ένα πρόβλημα n μεταβλητών και m περιορισμών που υποθέτουμε ότι απαιτούν m τεχνητές μεταβλητές γράφεται:

Να ελαχιστοποιηθεί η τεχνητή αντικειμενική συνάρτηση.

$$w = x_{n+1} + x_{n+2} + \cdots + x_{n+m}$$

σύμφωνα με τους περιορισμούς

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + x_{n+2} = b_2$$

. . . .

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{n+m} = b_m$$

$x_i \geq 0$ i=1έως (n+m)

Όμως γνωρίζουμε ότι η αντικειμενική συνάρτηση ενός προβλήματος γραμμικού προγραμματισμού πρέπει να αποτελείται μόνο από μη βασικές μεταβλητές σχεδιασμού. Έτσι, αντικαθιστούμε τις βασικές μεταβλητές της w με τη βοήθεια των περιορισμών. Για παράδειγμα, η βασική μεταβλητή x_{n+1} είναι, βάσει του πρώτου περιορισμού:

$$x_{n+1} = b_1 - a_{11}x_1 - a_{12}x_2 - \cdots - a_{1n}x_n$$

Η συνάρτηση w γίνεται λοιπόν :

$$w = \sum_{i=1}^{m} b_i - \sum_{j=1}^{n} \sum_{i=1}^{m} \alpha_{ij} x_j$$

εξαρτάται δηλαδή μόνο από τις n μη βασικές μεταβλητές.

Τώρα, το πρόβλημα βρίσκεται στην κανονική μορφή και μπορεί να ξεκινήσει η Φάση Ι κατά την οποία η πραγματική αντικειμενική συνάρτηση αντιμετωπίζεται ως περιορισμός. Στο τέλος της Φάσης Ι, όλες οι τεχνητές μεταβλητές έχουν γίνει μη βασικές, συνεπώς έχουν μηδενική τιμή. Έτσι, η τεχνητή συνάρτηση w γίνεται w=0 , αφού αποτελεί άθροισμα των τεχνητών μεταβλητών. Όταν γίνει w=0, έχουμε μεταπηδήσει σε μιά κορυφή του πραγματικού πολύεδρου, οπότε συνεχίζουμε στη Φάση ΙΙ χωρίς την w, προσπαθώντας να ελαχιστοποιήσουμε την f.

Παράδειγμα

Να βρεθεί η βέλτιστη λύση για το ακόλουθο πρόβλημα γραμμικού προγραμματισμού:

Να μεγιστοποιηθεί η αντικειμενική συνάρτηση

$$z = y_1 + 2y_2$$

σύμφωνα με τους περιορισμούς

$$3y_1 + 2y_2 \le 12$$

$$2y_1 + 3y_2 \ge 6$$

$$y_1 \ge 0 \quad \text{και} \quad y_2 \in R$$

Λύση

Το πρόβλημα μεγιστοποίησης μπορεί να αναχθεί σε πρόβλημα ελαχιστοποίησης πολλαπλασιάζοντας την αντικειμενική συνάρτηση με -1, οπότε γίνεται: $z=-y_1-2y_2$. Η μεταβλητή y μπορεί να πάρει οποιοδήποτε πρόσημο, επειδή όμως οι μεταβλητές σχεδιασμού πρέπει να είναι θετικές, την εκφράζουμε ως διαφορά δύο μη αρνητικών μεταβλητών , δηλαδή $y_2=y_2^+-y_2^-$

Για να γραφεί το πρόβλημα σύμφωνα με την τυπική μορφή, θέτουμε.

$x_1=y_1$, $x_2=y_2^+$ και $x_3=y_2^-$ ενώ προσθέτουμε και τις βοηθητικές μεταβλητές x_4 και x_5. Έτσι, το πρόβλημα γράφεται στη μορφή: *

$$z = -x_1 - 2x_2 + 2x_3$$

$$3x_1 + 2x_2 - 2x_3 + x_4 = 12$$

$$2x_1 + 3x_2 - 3x_3 - x_5 = 6$$

$$x_i \geq 0 \quad \text{i=1 έως 5}$$

Το πρόβλημα όμως εξακολουθεί να μην βρίσκεται στην κανονική μορφή, αφού ο συντελεστής της βασικής μεταβλητής x_5 είναι -1 αντί για +1.

Έτσι, εισάγουμε στο δεύτερο περιορισμό (που είναι τύπου «$\geq$») την τεχνητή μεταβλητή x_6, οπότε γίνεται:

$$2x_1 + 3x_2 - 3x_3 - x_5 + x_6 = 6$$

Παράλληλα, ορίζεται η τεχνητή αντικειμενική συνάρτηση $w=x_6$.

Όμως, η x_6 είναι βασική μεταβλητή, ενώ η αντικειμενική συνάρτηση πρέπει να εξαρτάται μόνο από μη βασικές μεταβλητές. Έτσι, γράφουμε την x_6 συναρτήσει των μη βασικών x_1, x_2, x_3, x_5 (η x_5 θεωρείται μη βασική μετά την ένταξη της x_6): χρησιμοποιώντας το δεύτερο περιορισμό:

$$x_6 = 6 - 2x_1 - 3x_2 + 3x_3 + x_5$$

Συνεπώς :

$$w = 6 - 2x_1 - 3x_2 + 3x_3 + x_5$$

Από τον πρώτο πίνακα simplex που φαίνεται παρακάτω, οι βασικές μεταβλητές είναι $x_4=12$ και $x_6=6$ ενώ οι μη βασικές είναι $x_1=x_2=x_3=x_5=0$. Επίσης είναι w=6 και f=0.

Η κατάσταση αυτή, αντιστοιχεί στο σημείο D(0,0) του σχήματος 5, το οποίο βρίσκεται στην μη επιτρεπτή περιοχή. Κατά τα γνωστά, το στοιχείο οδηγός είναι το α_{22}(2η στήλη : μικρότερος συντελεστής, 2η γραμμή: μικρότερος λόγος) που δείχνει ότι η μεταβλητή x_2 θα γίνει βασική και η x_6 θα γίνει μη βασική. Μετά την οδήγηση, παίρνουμε το δεύτερο πίνακα simplex που αντιστοιχεί στο σημείο A.

Όπως παρατηρούμε, όλοι οι συντελεστές της τελευταίας γραμμής είναι μη αρνητικοί και είναι w=0, συνεπώς έχει βρεθεί μία αρχική βασική επιτρεπτή λύση για το πραγματικό πρόβλημα. Αυτό είναι το τέλος της Φάσης I.

Πλέον, η τεχνητή μεταβλητή x_6 και η τεχνητή αντικειμενική συνάρτηση w μπορούν να παραλειφθούν και αφού στην προτελευταία γραμμή του δεύτερου πίνακα υπάρχει αρνητικός συντελεστής συνεχίζεται κανονικά η Φάση II που δίνει ως ελάχιστο το σημείο B (0,6)

Βασικές μεταβλητές	x_1	x_2	x_3	X_4	x_5	x_6	B
X_4	3	2	-2	1	0	0	12
X_6	2	3	-3	0	-1	1	6
	-1	-2	2	0	0	0	f-0
	$\boxed{-2}$	-3	3	0	1	0	w-6
x_4	5/3	0	0	1	2/3	-2/3	8
x_2	2/3	1	-1	0	-1/3	1/3	2
	1/3	0	0	$\boxed{0}$	-2/3	2/3	f+4
	0	0	0	0	0	1	w-0
x_5	5/2	0	0	3/2	1	-1	12
x_2	3/2	1	-1	½	0	0	6
	2	0	0	1	0	0	f+12

* Αν πολλαπλασιάζαμε το δεύτερο περιορισμό με μείον, θα είχαμε $-2y_1-3y_2<=-6$ οπότε το πρόβλημα δεν θα ήταν στην κανονική μορφή, αφού σ' αυτήν πρέπει όλοι οι συντελεστές b να είναι θετικοί.

3.10 Αλγόριθμοι δύο φάσεων

ΦΑΣΗ Ι.

Βήμα 1: Εισάγουμε τις βοηθητικές μεταβλητές. Ελέγχουμε αν οι σταθεροί όροι b είναι μη αρνητικοί.

Βήμα 2: Εισάγουμε τις τεχνητές μεταβλητές για τους περιορισμούς ισότητας και τους περιορισμούς τύπου «$\geq$». Έτσι, το σύστημα των εξισώσεων βρίσκεται στην κανονική μορφή. Ορίζουμε την τεχνητή αντικειμενική συνάρτηση w ως άθροισμα των τεχνητών μεταβλητών. Απαλείφουμε τις τεχνητές μεταβλητές από την w, ώστε αυτή να εκφραστεί συναρτήσει μόνο των μη βασικών μεταβλητών.

Βήμα 3: Γράφουμε τον πίνακα simplex, στον οποίο στην προτελευταία σειρά βρίσκεται η πραγματική αντικειμενική συνάρτηση και στην τελευταία η τεχνητή.

Βήμα 4: Καθορίζουμε την μη βασική μεταβλητή που θα μετατραπεί σε βασική από τον περισσότερο αρνητικό συντελεστή της τελευταίας γραμμής του πίνακα.

Βήμα 5: Υπολογίζουμε τους λόγους των σταθερών όρων b με τους θετικούς συντελεστές της στήλης που αντιστοιχεί στην παραπάνω μη βασική μεταβλητή.[*] Ο μικρότερος λόγος, μέσω της γραμμής στην οποία βρίσκεται, καθορίζει την μεταβλητή που θα γίνει μη βασική.

Βήμα 6: Η τομή της στήλης και της γραμμής που καθορίστηκαν στο βήμα 5, καταδεικνύουν το στοιχείο οδηγό. Έτσι μπορούμε να εφαρμόσουμε το βήμα οδήγησης,

Βήμα 7: Αν στο νέο πίνακα simplex υπάρχουν αρνητικοί συντελεστές στην τελευταία γραμμή, επιστρέφουμε στο βήμα 4.

Βήμα 8: Αν όλοι οι συντελεστές είναι μη αρνητικοί και είναι w=0, η φάση Ι έχει ολοκληρωθεί και έχει βρεθεί μια αρνητική βασική επιτρεπτή λύση για το πραγματικό πρόβλημα. Αν είναι $w \neq 0$, η λύση του προβλήματος είναι μη επιτρεπτή.

ΦΑΣΗ ΙΙ

Στον τελικό πίνακα της Φάσης Ι, η τελευταία γραμμή παραλείπεται. Συνεχίζουμε ακολουθώντας τα βήματα 4,5,6 της Φάσης Ι για την πραγματική αντικειμενική συνάρτηση.

[*] Αν όλοι οι συντελεστές της στήλης είναι αρνητικοί, τότε το πρόβλημα έχει απεριόριστη λύση.

3.11 Ο ΔΥΪΣΜΟΣ ΣΤΟ ΓΡΑΜΜΙΚΟ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ

Σε κάθε πρόβλημα γραμμικού προγραμματισμού, αντιστοιχεί ένα άλλο πρόβλημα γραμμικού προγραμματισμού το οποίο είναι το <u>δυϊκό</u> του (dual). Το αρχικό πρόβλημα ονομάζεται <u>πρωτογενές</u> (primal) . Αν το πρωτογενές πρόβλημα έχει n μεταβλητές και m περιορισμούς, το δυϊκό του έχει n περιορισμούς και m μεταβλητές. Η λύση καθενός από αυτά τα προβλήματα μπορεί να οδηγήσει με ευκολία στη λύση του άλλου, κάτι που μας προσφέρει μεγάλη ευελιξία, αφού αρκεί να λύσουμε το ευκολότερο από τα δύο.

Αποδεικνύεται ότι κάθε επιτρεπτή λύση του ευθέως προβλήματος φράσσει την αντικειμενική συνάρτηση του δυϊκού προβλήματος καθώς επίσης και το δυικό πρόβλημα του δυϊκού προβλήματος είναι το ευθύ πρόβλημα.

Για να γίνει κατανοητή η συσχέτιση των δύο προβλημάτων ας επιδιώξουμε να βρούμε άνω και κάτω όρια της αντικειμενικής συνάρτησης του πρωτογενούς προβλήματος.

Ας θεωρήσουμε το παρακάτω πρόβλημα:

Να μεγιστοποιηθεί συνάρτηση $z = 4x_1 + x_2 + 3x_3$

που υπόκειται στους περιορισμούς

$$
\begin{aligned}
x_1 + 4x_2 \quad & \leq 1 \\
3x_1 - x_2 + x_3 & \leq 3 \qquad (1.1) \\
x_1, x_2, x_3 & \geq 0
\end{aligned}
$$

Λόγω της μορφής των ανισοτικών περιορισμών παρατηρούμε ότι κάθε επιτρεπτή λύση παρέχει ένα κάτω όριο για την αντικειμενική συνάρτηση z. Για παράδειγμα η λύση $(x_1, x_2, x_3) = (1, 0, 0)$ δηλώνει ότι $z \geq 4$. Επίσης η επιτρεπτή λύση $(x_1, x_2, x_3) = (0, 0, 3)$ δίδει ως κάτω όριο της αντικειμενικής συνάρτησης την τιμή $z \geq 9$. Για να προκύψει το συμπέρασμα ότι η τιμή 9 είναι ένα καλό κάτω όριο χρειάζεται να προσδιορίσουμε και κάποιο άνω όριο. Για να προκύψει κάτι τέτοιο και χωρίς να αλλάξουμε την φορά των ανισοτήτων των περιορισμών επιδιώκουμε να φράξουμε την αντικειμενική συνάρτηση όπως για παράδειγμα πολλαπλασιάζοντας τον πρώτο περιορισμό με 2 και προσθέτοντας τον δεύτερο πολλαπλασιασμένο με τρία λαμβάνουμε την παρακάτω έκφραση:

$$
\begin{array}{l}
\quad 2\,(\ \ x_1 + 4x_2 \qquad) \le 2\ (1) \\
\underline{+\,3\,(3x_1 - \ x_2 + x_3) \le 3\ (3)} \\
\qquad 11x_1 + 5x_2 + 3x_3 \ \le 11
\end{array}
\tag{1.2}
$$

Έτσι επειδή κάθε μεταβλητή είναι μη αρνητική από την σύγκριση προκύπτει ότι:

$$4x_1 + x_2 + 3x_3 \le 11x_1 + 5x_2 + 3x_3 \le 11 \tag{1.3}$$

δηλαδή $z \le 11$ και άρα έχουμε περιορίσει τη εύρεση της βέλτιστης λύσης μεταξύ των τιμών 9 και 11. Για να μειωθεί το χάσμα των παραπάνω ορίων μπορούμε να επαναλάβουμε την ίδια διαδικασία εισάγοντας αντί των συγκεκριμένων τιμών μεταβλητές οι οποίες να παρέχουν την καλύτερη τιμή του άνω ορίου. Έτσι πολλαπλασιάζουμε τους περιορισμούς με τις μη αρνητικές μεταβλητές y_1 και y_2 ώστε να μην μεταβάλλεται η φορά των ανισοτήτων.

$$
\begin{array}{l}
\quad y_1\ (\ \ x_1 + 4x_2 \qquad) \le y_1\ (1) \\
\underline{+\,y_2\ (3x_1 - \ x_2 + x_3) \le y_2\ (3)} \\
(y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + y_2x_3 \le y_1 + 3y_2
\end{array}
\tag{1.4}
$$

Στην επιδίωξή μας να φράξουμε την αντικειμενική συνάρτηση θέτουμε τους συντελεστές των μεταβλητών x_i μεγαλύτερους ίσους με αυτούς της αντικειμενικής συνάρτησης άρα:

$$
\begin{array}{l}
y_1 + 3y_2 \ge 4 \\
4y_1 - y_2 \ge 1 \\
\qquad y_2 \ge 3
\end{array}
\tag{1.5}
$$

Με αυτές τις προϋποθέσεις η αντικειμενική συνάρτηση φράσσεται ως εξής:

$$z = 4x_1 + x_2 + 3x_3 \le (y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + y_2x_3 \le y_1 + 3y_2 \le y_1 + 3y_2$$

και άρα αναζητώντας το ελάχιστο δυνατό άνω όριο διατυπώνουμε το πρόβλημα:

να ελαχιστοποιηθεί η συνάρτηση $y_1 + 3y_2$

που υπόκειται στους περιορισμούς $y_1 + 3y_2 \geq 4$

$$4y_1 - y_2 \geq 1$$

$$y_2 \geq 3$$

$$y_1,\, y_2 \geq 0$$

που αποτελεί το δυικό πρόβλημα του ευθέως προβλήματος σύμφωνα με τη διατύπωση του γενικού δυϊκού προβλήματος που ακολουθεί.

3.11.1 Το Γενικό Δυικό Πρόβλημα

Η κανονική μορφή του πρωτογενούς προβλήματος γραμμικού προγραμματισμού είναι :

Να βρεθούν τα $x_1, x_2, ..., x_n$ για τα οποία <u>μεγιστοποιείται</u> η πρωτογενής αντικειμενική συνάρτηση:

$$z_p = c_1 x_1 + ... + c_n x_n = \sum_{i=1}^{n} c_i x_i = c^{\mathrm{T}} x$$

σύμφωνα με τους περιορισμούς $\sum_{j=1}^{n} a_{ij} x_j \leq b_i\,,\ \ i = 1, 2, \ldots, m$

δηλαδή

$$\alpha_{11} x_1 + ... + \alpha_{1n} x_n \leq b_1$$

$$. \ (Ax \leq b)$$

$$\alpha_{m1} x_1 + ... + \alpha_{mn} x_n \leq b_m$$

και $x_j \geq 0 \qquad j = 1, 2, \ldots, n$

Για τους σταθερούς όρους b_i δεν υπάρχει περιορισμός στο πρόσημο.

Αντίστοιχα η κανονική μορφή του δυϊκού προβλήματος, είναι:

Να βρεθούν οι τιμές των δυϊκών μεταβλητών $y_1, y_2, \ldots, y_m$ για τις οποίες ελαχιστοποιείται η δυϊκή αντικειμενική συνάρτηση:

$$f_d = b_1 y_1 + \ldots + b_m y_m = \sum_{i=1}^{m} b_i y_i = b^{\mathrm{T}} y$$

και υπόκειται στους περιορισμούς $\sum_{j=1}^{m} a_{ij} y_j \leq c_i$, $i = 1, 2, \ldots, n$

δηλαδή

$$\alpha_{11} y_1 + \ldots + \alpha_{m1} y_m \geq c_1$$

$$\ldots\ldots\ldots\ldots\ldots\ldots \quad (A^{\mathrm{T}} y \geq c)$$

$$\alpha_{1n} y_1 + \ldots + \alpha_{mn} y_m \geq c_n$$

και $y_i \geq 0 \qquad i = 1, 2, \ldots, m$

Παρατηρούμε ότι μεταξύ του πρωτογενούς και του δυϊκού προβλήματος υπάρχουν οι παρακάτω συσχετίσεις:

- Ο αριθμός των δυϊκών μεταβλητών είναι ίσος με τον αριθμό των πρωτογενών περιορισμών. Αντίστοιχα, ο αριθμός των πρωτογενών μεταβλητών είναι ίσος με τον αριθμό των δυϊκών περιορισμών.
- Η μεγιστοποίηση της πρωτογενούς αντικειμενικής συνάρτησης αντικαθίσταται με την ελαχιστοποίηση της δυϊκής αντικειμενικής συνάρτησης.
- Οι συντελεστές c_i της πρωτογενούς αντικειμενικής συνάρτησης γίνονται σταθεροί όροι των δυϊκών περιορισμών. Αντίστοιχα, οι σταθεροί όροι b_i των πρωτογενών περιορισμών γίνονται συντελεστές της δυϊκής αντικειμενικής συνάρτησης.
- Το μητρώο των συντελεστών Α των πρωτογενών περιορισμών είναι το ανάστροφο του μητρώου Α των δυϊκών περιορισμών.
- Επίσης στο πρωτογενές πρόβλημα οι περιορισμοί είναι ανισότητες «τύπου $\leq$« ενώ στο δυϊκό είναι «τύπου $\geq$».

Εδώ πρέπει να τονίσουμε ότι η κανονική μορφή του πρωτογενούς προβλήματος είναι διαφορετική από την κανονική μορφή του προβλήματος γραμμικού προγραμματισμού, χωρίς να βλάπτεται η γενικότητα του προβλήματος.

Αν σε ένα πρόβλημα βέλτιστου σχεδιασμού υπάρχει περιορισμός ισότητας, αυτός αντικαθίσταται από ένα ζεύγος ανισοτικών περιορισμών. Για παράδειγμα, ο περιορισμός $2x_1+3x_2=5$ αντικαθίσταται από το ζεύγος $2x_1 + 3x_2 \geq 5$ και $2x_1 + 3x_2 \leq 5$. Αντίστοιχα, πολλαπλασιάζουμε μια από τις δυο ανισότητες με -1 ώστε να έχουμε την κανονική μορφή πρωτογενούς ή δυϊκού προβλήματος.

3.11.2 Προσδιορισμός πρωτογενούς λύσης από δυϊκή λύση

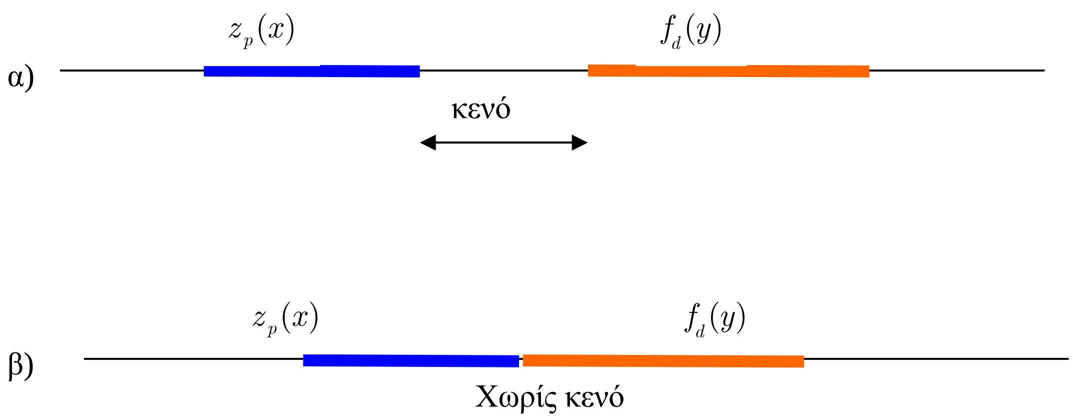
Θεώρημα 1: (Ασθενές Θεώρημα Δυϊσμού). Για κάθε **x** και **y** επιτρεπτές (feasible) λύσεις του πρωτογενούς και του δυϊκού προβλήματος αντίστοιχα ισχύει ότι:

$$z_p(x) \leq f_d(y) \tag{1.7}$$

Η απόδειξη έχει ως εξής:

$$z_p(x) = \sum_{i=1}^{n} c_i x_i \leq \sum_{i=1}^{n}\left(\sum_{j=1}^{m} y_j a_{ij}\right)x_i = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_{ij} x_i\right)y_j \leq \sum_{j=1}^{m} b_j y_j = f_d(y)$$

Γραφικά προκύπτει η εξής εικόνα στον άξονα των πραγματικών αριθμών.



Σχ. Πιθανή διάταξη τιμών αντικειμενικής συνάρτησης πρωτεγενούς και δυικού προβλήματος

Ώστε για κάθε επιτρεπτή λύση να ισχύει η σχέση (1.7), ενώ τα διαστήματα ενδέχεται να εκτείνονται στο άπειρο και προς τα αριστερά ή δεξιά για το πρωτογενές και δυικό πρόβλημα αντίστοιχα.

Από τις δύο παραπάνω περιπτώσεις αποδεικνύεται ότι για την περίπτωση του πρωτογενούς και δυικού προβλήματος δεν υπάρχει κενό μεταξύ των λύσεων και ότι η βέλτιστη (μέγιστο) λύση του πρωτογενούς ισούται με την βέλτιστη (ελάχιστο) λύση του δυικού προβλήματος γεγονός που επιτρέπει και τον έλεγχο δύο λύσεων $\mathbf{x}$ και $\mathbf{y}$ αν είναι βέλτιστες αντίστοιχα για το πρωτογενές και δυικό πρόβλημα σύμφωνα με το ισχυρό θεώρημα του δυισμού.

Θεώρημα 2. (Ισχυρό Θεώρημα Δυισμού). Αν τα διανύσματα **x** και **y** είναι επιτρεπτές λύσεις του πρωτογενούς και του δυϊκού προβλήματος αντίστοιχα και ισχύει η σχέση $f_d(x) = z_p(y)$ τότε τα διανύσματα **x** και **y** αποτελούν τις βέλτιστες λύσεις για το πρωτογενές και δυικό πρόβλημα αντίστοιχα. Η απόδειξη του θεωρήματος δεν παρατίθεται προκύπτει όμως από τα βήματα της μεθόδου Simplex καθόσον κατά την παραγωγή της βέλτιστης λύσης του πρωτογενούς προβλήματος παράγεται και η βέλτιστη λύση του δυικού με χρήση της σχέσης $f_d(x) = z_p(y)$.

Το ισχυρό θεώρημα του δυισμού μας εξασφαλίζει ότι όταν το πρωτογενές πρόβλημα έχει βέλτιστη λύση τότε και το δυικό πρόβλημα έχει βέλτιστη λύση χωρίς την ύπαρξη χάσματος μεταξύ των τιμών της αντικειμενικής τους συνάρτησης. Για την περίπτωση όμως που το πρωτογενές πρόβλημα για παράδειγμα έχει απεριόριστη λύση από το ασθενές θεώρημα προκύπτει ότι το δυικό πρόβλημα δεν έχει λύση και αντίστροφα αν το δυικό πρόβλημα έχει απεριόριστη λύση τότε το πρωτογενές δεν έχει λύση. Δηλαδή στο πεδίο των πραγματικών αριθμών η πρώτη περίπτωση αντιστοιχεί στο σημείο $+\infty$, ενώ η δεύτερη στο σημείο $-\infty$ εξαφανίζοντας το διάστημα των δυικών λύσεων ή των πρωτογενών λύσεων αντίστοιχα.

Τέλος υπάρχει η περίπτωση όπου και το πρωτογενές και το δυικό πρόβλημα δεν έχουν λύση. Για παράδειγμα:

Μεγιστοποιείστε την συνάρτηση: $2x_1 - x_2$

Που υπόκειται στους περιορισμούς: $x_1 - x_2 \leq 1$

$$-x_1 + x_2 \leq -2$$

$$x_1, x_2 \geq 0$$

Στην περίπτωση αυτή θεωρούμε ότι υφίσταται ένα τεράστιο κενό από $-\infty$ έως $+\infty$ που εκτοπίζει τα διαστήματα και των δύο προβλημάτων.

Θεώρημα 3 : Λύση του πρωτογενούς προβλήματος από το δυϊκό

Αν ο δυϊκός περιορισμός i είναι «τύπου >» στο ελάχιστο $(\sum_{i=1}^{m} \alpha_{ij} y_i > c_j)$,

τότε η πρωτογενής μεταβλητή i είναι μη βασική (x_j=0).

Αν η δυϊκή μεταβλητή i είναι βασική (y_i>0), τότε ο πρωτογενής περιορισμός i ικανοποιείται με ισότητα $(\sum_{j=1}^{n} \alpha_{ij} x_j = b_i)$

Η εφαρμογή των παραπάνω θεωρημάτων θα γίνει στο ακόλουθο παράδειγμα.

Παράδειγμα

Να μεγιστοποιηθεί η αντικειμενική συνάρτηση

$$z_p = 5x_1 - 2x_2$$

σύμφωνα με τους περιορισμούς

$$2x_1 + x_2 \le 9$$

$$x_1 - 2x_2 \le 2$$

$$-3x_1 + 2x_2 \le 3$$

$$x_1, x_2 \ge 0$$

Να λυθεί το πρωτογενές και το δυϊκό πρόβλημα.

Λύση

Το πρόβλημα βρίσκεται στην κανονική πρωτογενή μορφή. Θα χρησιμοποιήσουμε την μέθοδο simplex για να ελαχιστοποιήσουμε την

$$f_p = 5x_1 + 2x_2$$

Έτσι, προκύπτουν οι παρακάτω πίνακες simplex.

Βασικές μεταβλητές	x_1	x_2	x_3	x_4	x_5	B
X_3	2	1	1	0	0	9
X_4	1	-2	0	1	0	2
X_5	-3	2	0	0	1	3
	-5	2	0	0	0	$f_ρ$-0
X_3	0	5	1	-2	0	5
X_1	1	-2	0	1	0	2
X_5	0	-4	0	3	1	9
	0	-8	0	5	0	$f_ρ$+10
X_2	0	1	0.2	-0.4	0	1
X_1	1	0	0.4	0.2	0	4
X_5	0	0	0.8	1.4	1	13
	0	0	1.6	1.8	0	$f_ρ$+18

Από τον τελευταίο πίνακα προκύπτει ότι το μέγιστο εμφανίζεται στο σημείο (4,1) και είναι $z_ρ$=18 (ελάχιστο $f_ρ$=-18)

Το δυϊκό πρόβλημα προκύπτει εύκολα από το πρωτογενές. Υπάρχουν τρεις πρωτογενείς περιορισμοί, άρα και τρεις δυϊκές μεταβλητές. Υπάρχουν δυο πρωτογενείς μεταβλητές, άρα και δυο δυϊκοί περιορισμοί. Οι συντελεστές της αντικειμενικής συνάρτησης είναι οι σταθεροί όροι του πρωτογενούς προβλήματος, ενώ οι συντελεστές των περιορισμών είναι οι όροι του ανάστροφου μητρώου των συντελεστών των πρωτογενών περιορισμών. Τέλος, οι σταθεροί όροι είναι οι συντελεστές της αντικειμενικής συνάρτησης του πρωτογενούς προβλήματος. Έτσι, το δυϊκό πρόβλημα είναι:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$f_d = 9y_1 + 2y_2 + 3y_3$$

σύμφωνα με τους περιορισμούς

$$2y_1 + y_2 - 3y_3 \geq 5$$

$$y_1 - 2y_2 + 2y_3 \geq -2$$

Γράφοντας τους περιορισμούς με την τυπική μορφή της μεθόδου simplex, έχουμε:

$$2y_1 + y_2 - 3y_3 - y_4 = 5 \quad \Rightarrow 2y_1 + y_2 - 3y_3 - y_4 + y_6 = 5 \qquad y_i \geq 0 \ ,$$

$$y_1 - 2y_2 + 2y_3 - y_5 = -2 \ \Rightarrow -y_1 + 2y_2 - 2y_3 + y_5 = 2 \qquad \text{i=1έως 6}$$

όπου οι y_4 και y_5 είναι βοηθητικές μεταβλητές, ενώ η y_6 είναι τεχνητή μεταβλητή.* Η τεχνητή αντικειμενική συνάρτηση είναι:

$w = y_6 = 5 - 2y_1 - y_2 + 3y_3 + y_4$ (από τον πρώτο περιορισμό).

Οι πίνακες simplex της μεθόδου δύο φάσεων είναι:

Βασικές μεταβλητές	y_1	y_2	Y_3	Y_4	y_5	y_6	B
Y_6	2	1	-3	-1	0	1	5
Y_5	-1	2	-2	0	1	0	2
	9	2	3	0	0	0	f_d-0
	-2	-1	3	1	0	0	w-5
Y_1	1	0.5	-1.5	-0.5	0	0.5	2.5
Y_5	0	2.5	-3.5	-0.5	1	0.5	4.5
	0	-2.5	16.5	4.5	0	-4.5	f_d-22.5

* Στην πρώτη σχέση προσθέσαμε την τεχνητή μεταβλητή y6 διότι ο συντελεστής της βασικής μεταβλητής y4 είναι -1 (αντί +1). Θα μπορούσαμε να πολλαπλασιάσουμε τα δύο μέλη με -1 (όπως στη δεύτερη εξίσωση) αλλά τότε θα γινόταν αρνητικός ο ελεύθερος όρος 5.

	0	0	0	0	0	1	w-0
Y_1	1	0	-0.8	-0.4	-0.2	0.4	1.6
Y_2	0	1	-1.4	-0.2	0.4	0.2	1.8
	0	0	13.0	4.0	1.0	-4.0	f_d-18

Από τον τελευταίο πίνακα προκύπτει ελάχιστη τιμή f_d=18 στη θέση (1.6,1.8).

Παρατηρούμε ότι στη βέλτιστη λύση και των δύο προβλημάτων είναι $f_d=z_p$, δηλαδή ικανοποιούνται τα θεωρήματα 1 και 2. Επίσης, σύμφωνα με το θεώρημα 3, δεδομένου ότι οι δυϊκές μεταβλητές y_1 και y_2 είναι βασικές, οι αντίστοιχοι περιορισμοί 1 και 2 του πρωτογενούς προβλήματος ικανοποιούνται ως ισότητες:

$$2x_1+x_2=9$$

$$x_1-2x_2=2$$

Από το σύστημα προκύπτουν x_1=4 και x_2=1, δηλαδή το δυϊκό πρόβλημα έχει την ίδια λύση με το πρωτογενές.

3.12 Λύση του πρωτογενούς προβλήματος με τη βοήθεια του δυϊκού πίνακα

Παρατηρώντας τον τελικό πίνακα του δυϊκού προβλήματος στο παραπάνω παράδειγμα, διαπιστώνουμε ότι τα στοιχεία της τελευταίας γραμμής αντιστοιχούν με τα στοιχεία της τελευταίας στήλης του τελικού πίνακα του πρωτογενούς προβλήματος. Δηλαδή η λύση του πρωτογενούς προβλήματος μπορεί να προκύψει απευθείας από τον πίνακα simplex του δυϊκού προβλήματος- και αντίστροφα. Έτσι διατυπώνουμε το θεώρημα:

Θεωρούμε ότι έχει λυθεί το δυϊκό ενός πρωτογενούς προβλήματος με τη μέθοδο simplex. Τότε, η τιμή της i πρωτογενούς μεταβλητής ισούται με το συντελεστή της βοηθητικής μεταβλητής που αντιστοιχεί στον i δυϊκό περιορισμό στην τελευταία γραμμή του (τελικού) πίνακα. Επίσης, αν η δυϊκή μεταβλητή είναι μη βασική, τότε ο συντελεστής που της αντιστοιχεί ισούται με την τιμή της βοηθητικής μεταβλητής του αντίστοιχου πρωτογενούς περιορισμού.

Θα εφαρμόσουμε το θεώρημα στο επόμενο παράδειγμα.

<u>Παράδειγμα</u>

Να λυθεί το παρακάτω πρόβλημα γραμμικού προγραμματισμού και να βρεθεί η δυϊκή λύση του από τον τελικό πρωτογενή πίνακα. Να μεγιστοποιηθεί η αντικειμενική συνάρτηση

$$z_p = x_1 + 4x_2$$

σύμφωνα με τους περιορισμούς

$$x_1 + 2x_2 \leq 5$$

$$2x_1 + x_2 = 4$$

$$x_1 - x_2 \geq 1$$

$$x_1, x_2 \geq 0$$

Λύση

Όπως είδαμε στην παράγραφο 4.5, ο περιορισμός ισότητας αντικαθίσταται από δυο περιορισμούς ανισότητας. Έτσι όμως επιβαρύνουμε το πρόβλημα με μία επιπλέον μεταβλητή. Γι αυτό, θα λύσουμε το πρόβλημα χωρίς να αντικαταστήσουμε τον περιορισμό ισότητας εκμεταλλευόμενοι το γεγονός ότι η κανονική μορφή της μεθόδου simplex απαιτεί μόνο ισότητες.

Εισάγοντας βοηθητικές και τεχνητές μεταβλητές το πρόβλημα γίνεται:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση.

$$f_p=-x_1-4x_2$$

Σύμφωνα με τους περιορισμούς

$$x_1+2x_2+x_3=5$$

$$2x_1+x_2+x_5=4$$ *

$$x_1-x_2-x_4+x_6=1$$

$x_i \geq 0$ i=1 έως 6 (x_3, x_4: βοηθητικές μεταβλητές, x_5, x_6: τεχνητές μεταβλητές)

Η τεχνητή αντικειμενική συνάρτηση είναι $w=x_5+x_6=(4-2x_1-x_2)+(1-x_1+x_2+x_4)=5-3x_1+x_4$

* Όπως αναφέρθηκε στο βήμα 2 του αλγόριθμου (παράγραφος 4.4.7) εισάγονται τεχνητές μεταβλητές και για τους περιορισμούς ισότητας.

Οι πίνακες simplex που προκύπτουν είναι:

Βασικές μεταβλητές	x_1	x_2	x_3	X_4	x_5	x_6	b
X_3	1	2	1	0	0	0	5
X_5	2	1	0	0	1	0	4
X_6	1	-1	0	-1	0	1	1
	-1	-4	0	0	0	0	f_ρ-0
	-3	0	0	1	0	0	w-5
X_3	0	3	1	1	0	-1	4
X_5	0	3	0	2	1	-2	2
X_1	1	-1	0	-1	0	1	1
	0	-5	0	-1	0	1	f_ρ+1
	0	-3	0	-2	0	3	w-2
X_2	0	0	1	-1	-1	1	2
X_2	0	1	0	2/3	1/3	-2/3	2/3
X_1	1	0	0	-1/3	1/3	1/3	5/3
	0	0	0	7/3	5/3	-7/3	f_ρ+13/3
	0	0	0	0	1	1	w-0

Χρησιμοποιώντας το θεώρημα, οι δυϊκές μεταβλητές για τους τρεις περιορισμούς του προβλήματος είναι:

α) Στον πρώτο περιορισμό εμφανίζεται η βοηθητική μεταβλητή x_3. Ο συντελεστής της x_3 στην γραμμή της αντικειμενικής συνάρτησης στον τελευταίο πίνακα, είναι μηδέν. Άρα, η πρώτη δυϊκή μεταβλητή είναι $y_1=0$.

β) στο δεύτερο περιορισμό εμφανίζεται η τεχνητή μεταβλητή x_5, η οποία έχει συντελεστή 5/3, συνεπώς $y_2=5/3$

γ) Στον τρίτο περιορισμό εμφανίζεται η βοηθητική μεταβλητή x_4, που έχει συντελεστή 7/3, οπότε $y_3=7/3$

Δηλαδή, η λύση του δυϊκού προβλήματος που αντιστοιχεί στην πρωτογενή λύση (5/3, 2/3) είναι η (0,5/3).

3.13 Επίλυση Προβλημάτων Γραμμικού Προγραμματισμού στον Υπολογιστή

3.13.1 Χρήση του προγράμματος Excel της Microsoft

Με βάση το πρόγραμμα excel σε αντίστοιχα κελιά εισάγονται οι αρχικές τιμές των μεταβλητών σχεδιασμού, στη συνέχεια σε άλλο κελί εισάγεται η αντικειμενική συνάρτηση με βάση τα κελιά των μεταβλητών σχεδιασμού. Εισάγονται επίσης οι εκφράσεις των αριστερών μελών των περιορισμών. Με βάση τα στοιχεία αυτά καλείται το πρόγραμμα που εκτελεί την γενική βελτιστοποίηση και φέρει το όνομα Solver. Το αντίστοιχο παράθυρο του Solver ζητά τις θέσεις των μεταβλητών σχεδιασμού, της αντικειμενικής συνάρτησης και συμπληρώνονται οι περιορισμοί. Με την ολοκλήρωση της εισαγωγής των στοιχείων ενεργοποιείται το πρόγραμμα και μετά από κάποιο αριθμό επαναλήψεων παρέχει την απάντηση. Σε ειδικά φύλλα παρατίθενται πρόσθετες πληροφορίες για την πορεία επίλυσης του προβλήματος. Η επίλυση με βάση το πρόγραμμα excel βασίζεται σε ένα γενικό πρόγραμμα βελτιστοποίησης αναφορές στο οποίο υπάρχουν στις επεξηγήσεις του προγράμματος. Από την διερεύνηση της λύσης της παραγράφου 3.10 προκύπτει ότι στην περίπτωση μη μοναδικής λύσης ο αλγόριθμος παρέχει μία λύση χωρίς ένδειξη όμως ότι δεν είναι η μοναδική.

Μεταβλητές Σχεδιασμού	X_1 :	4
	X_2 :	2
Αντικειμενική Συνάρτηση	Obj:	32
Περιορισμοί	g_1 :	2
	g_2 :	8
	g_3 :	4
	g_4 :	2

3.13.2 Χρήση προγραμμάτων συμβολικού προγραμματισμού

Τα διάφορα προγράμματα συμβολικού προγραμματισμού διαθέτουν μεθόδους για την επίλυση γραμμικών προβλημάτων. Η σύνταξη των εντολών είναι συνήθως απλή και άμεση. Το πρόγραμμα Maple για παράδειγμα διαθέτει τις εντολές minimize και maximize για την επίλυση γραμμικών προβλημάτων ελαχιστοποίησης ή μεγιστοποίησης αντίστοιχα. Ο τρόπος με τον οποίο καλείται η εντολή έχει ως εξής:

minimize(f, C)

minimize(f , C, vartype)

minimize(f , C, vartype, 'NewC', 'transform')

Παράμετροι:

f - γραμμική έκφραση

C - λίστα γραμμικών περιορισμών

vartype - (προαιρετική) NONNEGATIVE ή UNRESTRICTED

NewC - (προαιρετική) ένα όνομα

transform - (προαιρετική) ένα όνομα

Η εντολή επιστρέφει την λύση ή είναι κενή στη περίπτωση που δεν ορίζεται αποδεκτή λύση, ή NULL για την περίπτωση που το πρόβλημα είναι μη-φραγμένο.

Ως εφαρμογή επιλύουμε το παράδειγμα: Να μεγιστοποιηθεί η συνάρτηση $3x_1 + x_2 + x_3 \leq 0$

που υπόκειται στους περιορισμούς:

$$\begin{aligned} &2x_1 + x_2 + x_3 \leq 2 \\ &x_1 + 2x_2 + 3x_3 \leq 5 \\ &2x_1 + 2x_2 + x_3 \leq 6 \\ &x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned} \tag{24}$$

```
> with(simplex):
> maximize( 3*x1+x2+3*x3, {2*x1+x2+x3+x4 = 2, x1+2*x2+3*x3+x5 = 5,
2*x1+2*x2+x3+x6 =6}, NONNEGATIVE );

         {x2 = 0, x3 = 8/5, x1 = 1/5, x6 = 4, x4 = 0, x5 = 0}
```

4

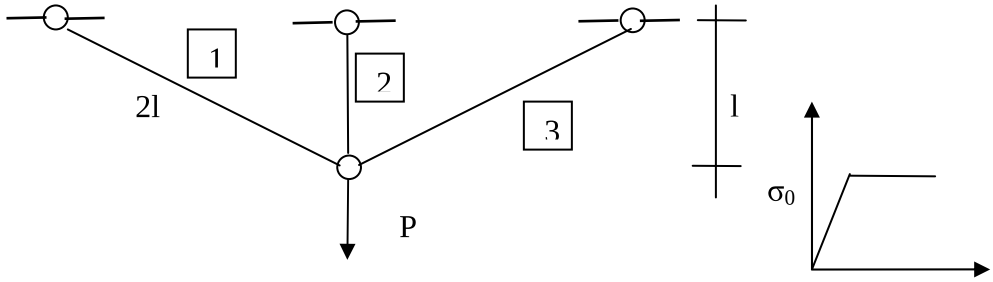
ΕΦΑΡΜΟΓΕΣ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

4.1..Γενικά

Το πρόβλημα του γραμμικού προγραμματισμού κατέχει σημαντική θέση στη θεωρία βελτιστοποίησης και βρίσκει εφαρμογή σε μεγάλο εύρος. Στα προβλήματα βέλτιστου σχεδιασμού των κατασκευών παρέχει την βάση για την επίλυση προβλημάτων πλαστικής ανάλυσης που αποτελούν μαθήματα του προπτυχιακού κύκλου προγραμμάτων σπουδών Πολιτικών Μηχανικών και για το λόγο αυτό δεν θα αναφερθούμε εδώ στη κατηγορία των προβλημάτων πλαστικής ανάλυσης και εύρεσης του μηχανισμού κατάρρευσης ραβδωτών φορέων.

4.2 Κάτω Όριο Οριακού Φορτίου Επιπέδου Δικτυώματος

Θεωρούμε το δικτύωμα του σχήματος (α) τα μέλη του οποίου συμπεριφέρονται ελαστικά – απολύτων πλαστικά ακολουθώντας τον νόμο του σχήματος (β).



(β)

(α)

Οι παραμορφώσεις των μελών για την συγκεκριμένη γεωμετρία είναι:

$$\varepsilon_2 = \frac{v}{l}, \qquad \varepsilon_1 = \varepsilon_3 = \frac{v}{2 \cdot 2l} = \frac{v}{4l}$$

Οι δε δυνάμεις στην ελαστική περιοχή δίδονται από τις σχέσεις:

$$N_2 = \frac{EA}{l} v, \qquad N_1 = N_3 = \frac{EA}{4l} v = 0.25 N_2$$

Από την ισορροπία του ελεύθερου κόμβου κατά y προκύπτει:

$$P = N_2 + \frac{1}{2}\left(N_1 + N_3\right) = 1.25 N_2$$

από όπου προκύπτει ότι:

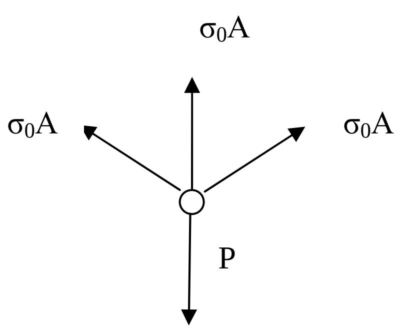
$$N_2 = 0.8P$$
$$N_1 = N_3 = 0.2P$$

Η κατάσταση αυτή ισχύει μέχρι την πρώτη διαρροή που θα συμβεί στο πλέον εντεινόμενο μέλος 2 οπότε:

$$N_2 = \sigma_0 A$$

Το δε φορτίο θα έχει την τιμή: $P = 1.25 \sigma_0 A$

Τέλος με σταδιακή αύξηση του φορτίου θα προκύψει η οριακή κατάσταση κατά την οποία όλες οι δυνάμεις των ράβδων θα έχουν φτάσει στη διαρροή και θα ισχύει:



οπότε το οριακό φορτίο προκύπτει από την εξίσωση ισορροπίας ως $P_u = 2\sigma_0 A$

Γενικότερα αν θεωρήσουμε επιπλέον ένα οριζόντιο φορτίο P στον ελεύθερο κόμβο από τις εξισώσεις ισορροπίας θα έχουμε:

$$N_2 + \frac{1}{2}\left(N_1 + N_3\right) - P = 0$$
$$\frac{\sqrt{3}}{2}\left(N_1 - N_3\right) - P = 0$$

ενώ θα πρέπει να ισχύουν οι παρακάτω περιορισμοί:

$$-\sigma_0 A \leq N_i \leq \sigma_0 A, \qquad i = 1, 2, 3$$

Το πρόβλημα της εύρεσης του κάτω ορίου του οριακού φορτίου

Μπορεί να διατυπωθεί ως πρόβλημα γραμμικού προγραμματισμού ως εξής:

Μεγιστοποιείστε το φορτίο P

Που υπόκειται στους περιορισμούς:

$$N_2 + \frac{1}{2}\left(N_1 + N_3\right) - P = 0$$
$$\frac{\sqrt{3}}{2}\left(N_1 - N_3\right) - P = 0$$

και

$$-\sigma_0 A \leq N_i \leq \sigma_0 A, \qquad i = 1, 2, 3$$

που κατ' ουσίαν εκφράζει το θεώρημα του κάτω ορίου της πλαστικής ανάλυσης.

Το ίδιο πρόβλημα μπορεί να διατυπωθεί για ένα οποιοδήποτε επίπεδο δικτύωμα με m ελεύθερους κόμβους $i = 1, 2, \ldots, m$ και δεδομένες διατομές A_j, $j = 1, 2, \ldots, r$ όπου r το πλήθος των μελών του δικτυώματος και δεδομένα φορτία P_i στους m ελεύθερους αν θεωρήσουμε ως μεταβλητές σχεδιασμού τον πολλαπλασιαστικό συντελεστή της φόρτισης λ και τις τιμές των δυνάμεων των μελών του δικτυώματος N_j, $j = 1, 2, \ldots, r$ οπότε διατυπώνεται το παρακάτω πρόβλημα γραμμικού προγραμματισμού:

Μεγιστοποιείστε τον συντελεστή λ

Που υπόκειται στους περιορισμούς ισότητας :

$$\sum_{j=1}^{r} e_{ij} N_j = \lambda P_i^{x,y}, \quad i = 1, 2, \ldots, m \tag{4.1}$$

που εκφράζουν τις εξισώσεις ισορροπίας στις δύο διευθύνσεις x και y, ενώ e_{ij} είναι τα συνημίτονα κατευθύνσεως των συντρεχόντων μελών.

και του ανισοτικούς περιορισμούς:

$$A_j \sigma_{cj} \le N_j \le A_j \sigma_{yj}, \qquad j = 1, 2, \ldots, r \tag{4.2}$$

Μπορεί επίσης να διατυπωθεί ως πρόβλημα γραμμικού προγραμματισμού και το πρόβλημα ελαχίστου βάρους του δικτυώματος αν θεωρηθούν ως μεταβλητές σχεδιασμού αφενός οι αξονικές δυνάμεις των μελών, αλλά και οι διατομές τους οπότε η αντικειμενική συνάρτηση ελαχίστου βάρους είναι:

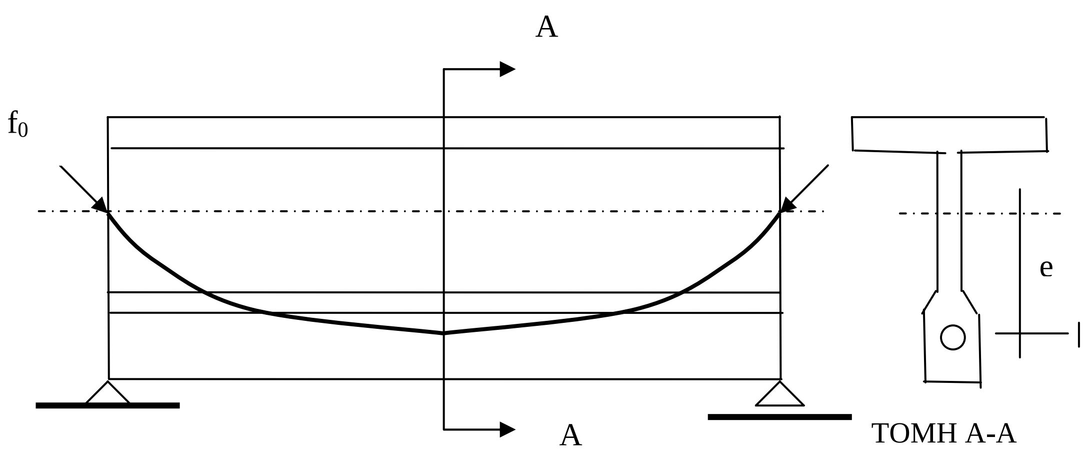
Ελαχιστοποιείστε το βάρος του δικτυώματος:

$$\min \sum_{j=1}^{r} \rho_j A_j l_j$$

που υπόκειται στους περιορισμούς (4.1) και (4.2) για σταθερή τιμή του πολλαπλασιαστή λ.

4.3 Βέλτιστη χάραξη τένοντα προεντεταμένης δοκού

Ένα άλλο κλασικό πρόβλημα βέλτιστου σχεδιασμού αφορά την βέλτιστη χάραξη παραβολικού τένοντα της αμφιέρειστης δοκού του σχήματος.



Ως μεταβλητές σχεδιασμού θεωρούνται η εκκεντρότητα e και η δύναμη προεντάσεως f_0.

Οι μεταβλητές αυτές όμως διατυπώνουν τους περιορισμούς τάσεων ως μη γραμμικές εκφράσεις καθόσον συστηματικά εμφανίζεται η ροπή ef_0. Η μη γραμμικότητα αυτή αίρεται εύκολα αν θεωρήσουμε ως μεταβλητή σχεδιασμού την ποσότητα $m = ef_0$ οπότε το πρόβλημα ανάγεται σε πρόβλημα γραμμικού προγραμματισμού ως εξής:

Ελαχιστοποιείστε την δύναμη προέντασης

$$\min f(f_0, m) = f_0$$

που υπόκειται στους περιορισμούς:

$$\sigma^{li} \leq -\frac{af_0}{A} \pm \frac{m_{ei} - am}{z} \leq \sigma^{ui} \text{ και}$$

$$\delta^{li} \leq \delta_{ei} + A \cdot k \cdot m \leq \delta^{ui}$$

$$m^l \leq m \leq m^u$$

$$f_0 \geq 0, \qquad i = 1, 2, \ldots, nl \text{ πλήθος φορτίσεων¨}$$

όπου Α η ενεργή διατομή α ο συντελεστής απομείωσης λόγω ερπυσμού, κ σταθερά της διατομής – υλικού και m_{ei}, δ_{ei} η μέγιστη εξωτερική ροπή και βύθιση αντίστοιχα.

5

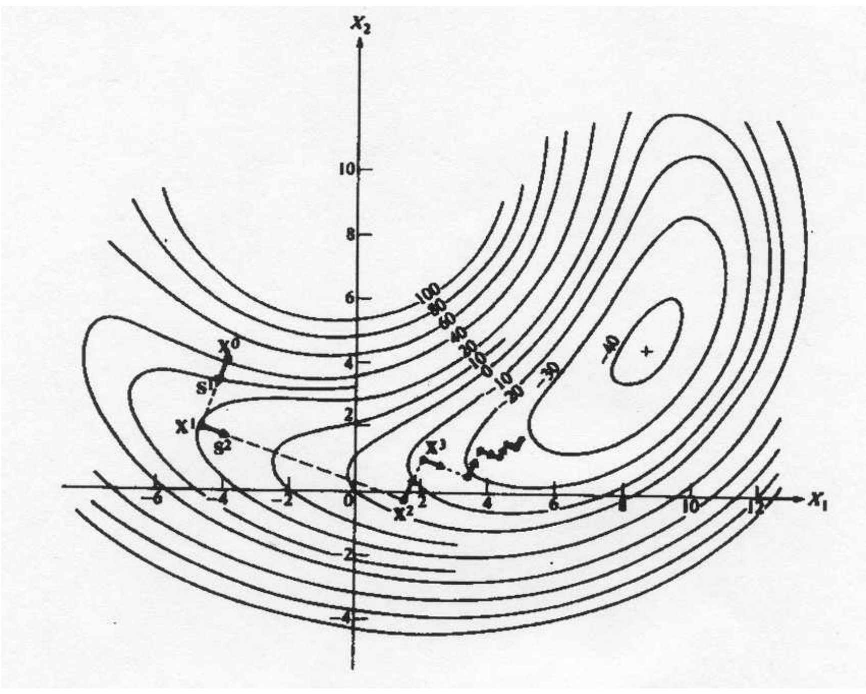
ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΜΕΣΩ ΕΠΑΝΑΛΗΠΤΙΚΗΣ ΔΙΑΔΙΚΑΣΙΑΣ

Πολλές μέθοδοι βελτιστοποίησης βασίζονται στην έννοια της επαναληπτικής διαδικασίας. Σύμφωνα μ' αυτήν, απαιτείται η επιλογή μιας αρχικής λύσης (σχεδιασμού) X^0 το προβλήματος. Ξεκινώντας από την λύση αυτήν, ο σχεδιασμός βελτιώνεται μέσω επαναληπτικής διαδικασίας. Η πιο συνηθισμένη μορφή αυτής της διαδικασίας εκφράζεται από τη σχέση

$$x^q = x^{q-1} + aS^q \tag{5.1}$$

όπου q είναι ο αριθμός της επανάληψης και **S** είναι το διάνυσμα της διεύθυνσης αναζήτησης στο χώρο των λύσεων. Ο συντελεστής α^* ορίζει την απόσταση κίνησης επί της διεύθυνσης **S**.

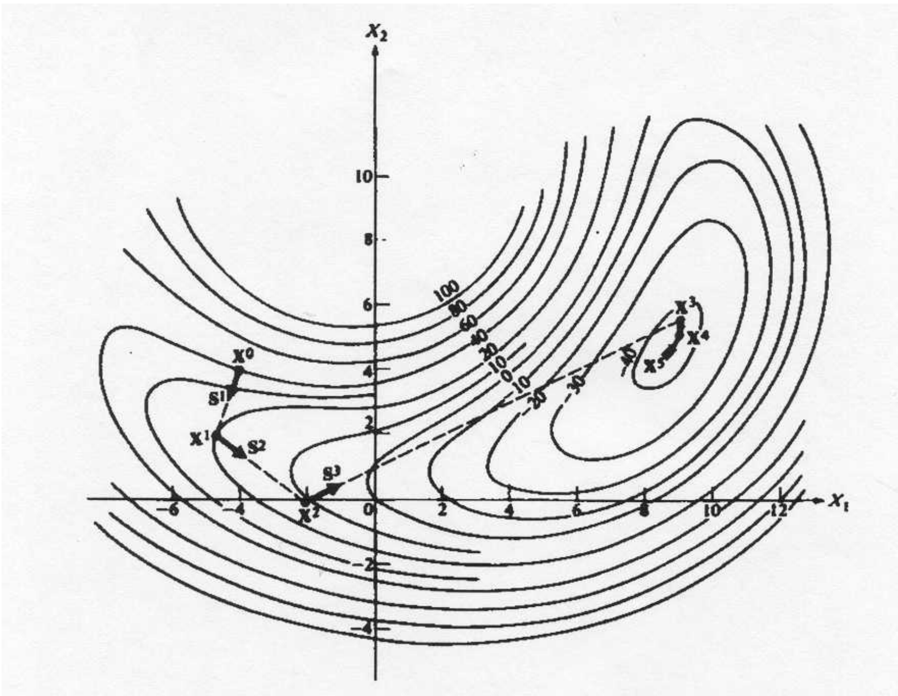
Για να δούμε πώς εφαρμόζεται η επαναληπτική διαδικασία που ορίζει η παραπάνω σχέση, ας θεωρήσουμε το πρόβλημα δυο διαστάσεων του σχήματος.



Ξεκινάμε από το σημείο $\mathbf{X}^0$=(2.0,1.0) και θα προσπαθήσουμε να μειώσουμε την τιμή της αντικειμενικής συνάρτησης. Η αρχική διεύθυνση αναζήτησης είναι $\mathbf{S}^1$=(-1.0,-0.5). Η επιλογή της διεύθυνσης **S** είναι τέτοια ώστε να μειώνεται η τιμή της αντικειμενικής συνάρτησης χωρίς να παραβιάζονται οι περιορισμοί. Στο συγκεκριμένο παράδειγμα το διάνυσμα $\mathbf{S}^1$ αποτελεί το αντίθετο της κλίσης (gradient) της αντικειμενικής συνάρτησης. Υπενθυμίζεται ότι η κλίση μιας συνάρτησης n μεταβλητών f(x) σ' ένα σημείο $\mathbf{x}^*$, είναι το παρακάτω διάνυσμα:

$$\nabla f(x^*) = \begin{bmatrix} \dfrac{\partial f(x^*)}{\partial x_1} \\ \dfrac{\partial f(x^*)}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(x^*)}{\partial x_n} \end{bmatrix} \tag{5.2}$$

Γεωμετρικά, το διάνυσμα της κλίσης μιας συνάρτησης είναι κάθετο στο εφαπτομενικό επίπεδο επί του σημείου $\mathbf{x}^*$, όπως φαίνεται στο σχήμα για μια συνάρτηση τριών μεταβλητών. Το σημαντικό χαρακτηριστικό του διανύσματος κλίσης είναι ότι ορίζει τη διεύθυνση μέγιστης αύξησης της τιμής της συνάρτησης.



Έτσι, επιστρέφοντας στο παράδειγμα, το S^1 ως το αντίθετο της κλίσης της αντικειμενικής συνάρτησης, ορίζει τη διεύθυνση μέγιστης μείωσης της αντικειμενικής συνάρτησης (steepest descent direction). Τώρα, απομένει να προσδιορίσουμε το συντελεστή $α^*$ της εξίσωσης (2.1) ώστε η συνάρτηση κόστους να μειώνεται χωρίς να παραβιάζονται οι περιορισμοί.

Για α=0, βρισκόμαστε στο αρχικό σημείο X^0. Από το σχήμα παίρνουμε F(α)=10.0 και g(α)=-1.0.

Τώρα, θα θέσουμε αυθαίρετες δοκιμαστικές τιμές για το α

1) <u>α=1.0</u> $X^1=X^0+αS^1=(2.0,1.0)+1.0(-1.0,-0.5)=(1.0,0.5)$

F(α)=8.4 g(α)=-0.2

Η αντικειμενική συνάρτηση έχει περιθώρια μείωσης

2) <u>α=1.5</u> $X^1=(2.0,1.0)+1.5(-1.0,-0.5)=(0.50, 0.25)$

F(α)=7.6 g(α)=0.2

Η αντικειμενική συνάρτηση μειώθηκε, αλλά παραβιάζεται ο περιορισμός

3) <u>α*=1.25</u> $X^1=(2.0,1.0)+1.25(-1.0,-0.5)=(0.750, 0.375)$

F(α*)=8.0 g(α*)=0.0

Όπως είδαμε στην παράγραφο 2.6.1, ο περιορισμός είναι ενεργός, συνεπώς η μείωση της αντικειμενικής συνάρτησης είναι η μεγαλύτερη δυνατή κατά μήκος της διεύθυνσης S.

Παρατηρούμε ότι αναζητώντας τη βέλτιστη λύση κατά μήκος μιας καθορισμένης διεύθυνσης, ουσιαστικά έχουμε μετατρέψει το πρόβλημα από πρόβλημα n μεταβλητών X σε πρόβλημα μιας μεταβλητής α. Τα προβλήματα αυτά ονομάζονται προβλήματα μονοδιάστατης αναζήτησης και εξετάζονται στο επόμενο κεφάλαιο.

Έχοντας ορίσει το σημείο X^1, αναζητούμε μια νέα διεύθυνση αναζήτησης τέτοια ώστε να μειώνεται ακόμα περισσότερο η τιμή της αντικειμενικής συνάρτησης χωρίς να παραβιάζονται οι περιορισμοί. Με αυτό τον τρόπο, η σχέση (5.1) χρησιμοποιείται επαναληπτικά μέχρι την εύρεση του βέλτιστου σχεδιασμού.

5.1 ΥΠΑΡΞΗ ΚΑΙ ΜΟΝΑΔΙΚΟΤΗΤΑ ΤΗΣ ΒΕΛΤΙΣΤΗΣ ΛΥΣΗΣ

Στην εφαρμογή των μεθόδων βελτιστοποίησης σε προβλήματα πρακτικού ενδιαφέροντος, σπάνια είναι δυνατό να εξασφαλίσουμε την εύρεση της απολύτως βέλτιστης λύσης (καθολικό ελάχιστο). Αυτό οφείλεται στην ύπαρξη πολλαπλών λύσεων του προβλήματος ή ενδεχομένως στην εξαιρετικά αργή σύγκλιση των αλγορίθμων βελτιστοποίησης κατά την εφαρμογή τους σε ένα πρόβλημα. Από πρακτική σκοπιά, η καλύτερη προσέγγιση ενός προβλήματος είναι να ξεκινήσουμε τη διαδικασία βελτιστοποίησης από διαφορετικές αρχικές διευθύνσεις αναζήτησης και αν καταλήξουμε τελικά στην ίδια λύση, θα είμαστε πιο σίγουροι ότι αυτή αποτελεί την πραγματική βέλτιστη λύση. Ωστόσο, είναι δυνατόν να ελέγξουμε μαθηματικά αν υπάρχει τουλάχιστον σχετικό ελάχιστο. Με άλλα λόγια, μπορούμε να ορίσουμε αναγκαίες συνθήκες για την ύπαρξη βέλτιστης λύσης και κάτω από συγκεκριμένες συνθήκες να αποδείξουμε ότι αυτές είναι και ικανές ώστε να εξασφαλίσουμε ότι η λύση αποτελεί καθολικό ελάχιστο. Οι συνθήκες αυτές εξηγούνται παρακάτω.

5.2 Προβλήματα χωρίς περιορισμούς

Ας θεωρήσουμε ένα πρόβλημα βελτιστοποίησης χωρίς περιορισμούς. Είναι γνωστό ότι η αντικειμενική συνάρτηση F(X) ελαχιστοποιείται όταν η πρώτη παράγωγός της ως προς την μεταβλητή X μηδενίζεται. Αν η αντικειμενική συνάρτηση εξαρτάται από περισσότερες μεταβλητές σχεδιασμού, τότε αντίστοιχα η κλίση της συνάρτησης πρέπει να μηδενίζεται, δηλαδή :

$$\nabla F(x) = 0$$

Ωστόσο, αυτή η συνθήκη είναι μόνο αναγκαία και έτσι δεν εξασφαλίζει την υπαρξη ελαχίστου. Αυτό φαίνεται καθαρά στο παρακάτω σχήμα όπου απεικονίζεται μία συνάρτηση μιας μεταβλητής. Η παράγωγος (κλίση) της συνάρτησης στα A,B και C μηδενίζεται, αλλά αυτό δεν σημαίνει ότι και τα τρία σημεία είναι ελάχιστα.

Γνωρίζουμε όμως ότι για να έχει μία συνάρτηση μιας μεταβλητής ελάχιστο, θα πρέπει η δεύτερη παράγωγός της ως προς την ανεξάρτητη μεταβλητή να είναι θετική, και αυτό ισχύει μόνο για το σημείο A του σχ. 10.

Γενικότερα, για μια συνάρτηση n μεταβλητών, η ικανή συνθήκη αφορά το μητρώο των δευτέρων μερικών παραγώγων της συνάρτησης που ονομάζεται μητρώο Hessian(Hessian matrix) και ορίζεται ως εξής:

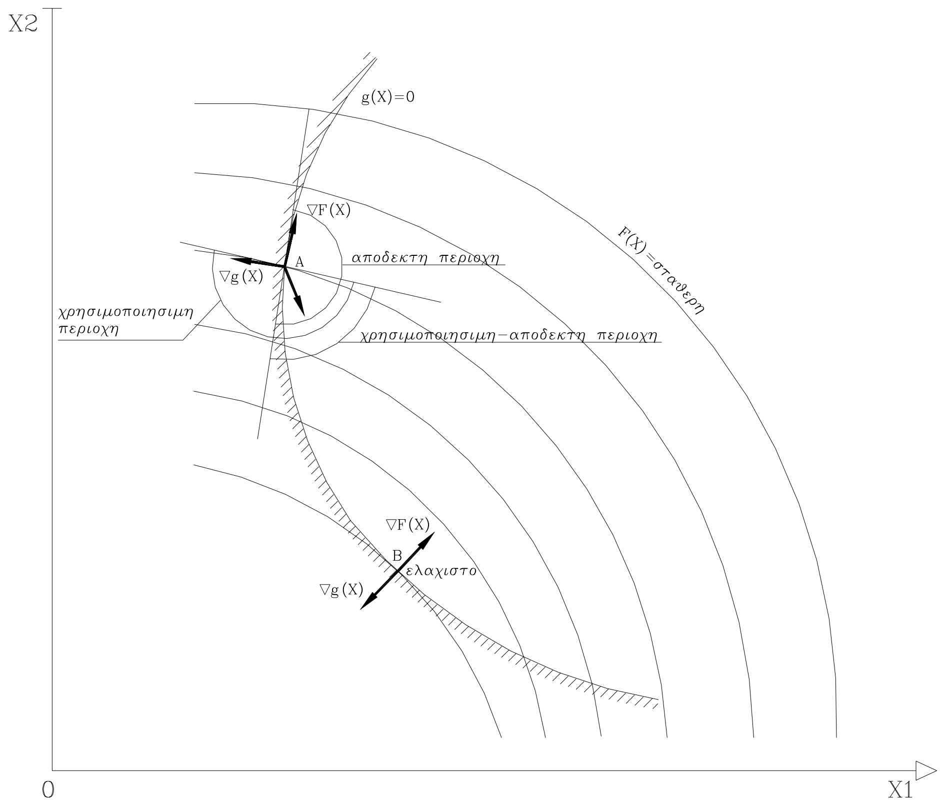
$$H = \begin{bmatrix} \frac{\partial^2 F(x)}{\partial x_1^2} \frac{\partial^2 F(x)}{\partial x_1 \partial x_2} \frac{\partial^2 F(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 F(x)}{\partial x_2 \partial x_1} \frac{\partial^2 F(x)}{\partial x_2^2} \cdots \frac{\partial^2 F(x)}{\partial x_2 \partial x_n} \\ \vdots\vdots\vdots \\ \frac{\partial^2 F(x)}{\partial x_n \partial x_1} \frac{\partial^2 F(x)}{\partial x_n \partial x_2} \frac{\partial^2 F(x)}{\partial x_n^2} \end{bmatrix} \qquad (5.3)$$

Αν το μητρώο Hessian είναι θετικά ορισμένο (δηλ. όλες οι κύριες ελάσσονες ορίζουσές του είναι θετικές) για όλες τις πιθανές τιμές του διανύσματος των μεταβλητών σχεδιασμού X, τότε εξασφαλίζεται ότι ο σχεδιασμός αντιστοιχεί σε σημείο που είναι καθολικό ελάχιστο.

5.3 Προβλήματα με περιορισμούς

Τώρα, ας θεωρήσουμε ένα πρόβλημα με περιορισμούς. Σύμφωνα με όσα αναφέρθηκαν παραπάνω, θα ήταν λογικό να συμπεράνουμε ότι ο προσδιορισμός του ελαχίστου εξαρτάται αποκλειστικά από την αντικειμενική συνάρτηση F(X). Αυτό όμως δεν είναι αλήθεια. Οι συναρτήσεις των περιορισμών παίζουν σημαντικό ρόλο στον προσδιορισμό της βέλτιστης λύσης. Θα προσπαθήσουμε λοιπόν να ορίσουμε τις αναγκαίες συνθήκες για την εύρεση του ελαχίστου ενός προβλήματος με περιορισμούς, με τη βοήθεια του σχήματος.

Ας θεωρήσουμε το σχεδιασμό που καθορίζεται από το σημείο A επί του ενεργού περιορισμού. Προκειμένου να βελτιώσουμε αυτόν το σχεδιασμό, θα πρέπει να ορίσουμε μια διεύθυνση αναζήτησης S η οποία θα μειώνει την τιμής της αντικειμενικής συνάρτησης ενώ παράλληλα δεν θα παραβιάζει τον ενεργό περιορισμό. Ένα διάνυσμα που απλά μειώνει την τιμή της αντικειμενικής συνάρτησης ορίζει την περιοχή όλων των χρησιμοποιήσιμων διευθύνσεων η οποία ονομάζεται χρησιμοποιήσιμη περιοχή (usable sector) (βλ. Σχήμα).



Παρατηρούμε ότι το εσωτερικό γινόμενο κάθε διανύσματος S που βρίσκεται στη χρησιμοποιήσιμη περιοχή με την κλίση της αντικειμενικής συνάρτησης θα είναι αρνητικό ή μηδέν (δηλαδή $\nabla F(X) \cdot S \leq 0$. Με άλλα λόγια, η γωνία μεταξύ των δύο διανυσμάτων πρέπει να κείται μεταξύ 90^o και 270^o.

Παρομοίως, η εφαπτομένη στο Α επί της καμπύλης του περιορισμού ορίζει, κατά τα γνωστά, την επιτρεπτή περιοχή, όπου μία διεύθυνση S ειναι επιτρεπτή, όταν μία μικρή μετακίνηση επί αυτής δεν παραβιάζει τον περιορισμό. Σ΄αυτήν την περίπτωση, το εσωτερικό γινόμενο του S με την κλίση του περιορισμού είναι αρνητικό ή μηδέν (S.∇g(X)≤0), οπότε η γωνία μεταξύ των δύο διανυσμάτων θα κείται επίσης μεταξύ 90^o και 270^o.

Έτσι, προκειμένου ένα διάνυσμα S να αποφέρει έναν βελτιωμένο σχεδιασμό χωρίς να παραβιάζει τον περιορισμό, θα πρέπει να βρίσκεται ταυτόχρονα και στην επιτρεπτή και στη χρησιμοποιήσιμη περιοχή. Αν το διάνυσμα βρίσκεται κοντά στην εφαπτομένη επί του περιορισμού, μία μικρή μετακίνηση, κατά μήκος του οδηγεί σε μεγάλη μείωση της τιμής της αντικειμενικής συνάρτησης, παράλληλα όμως οδηγεί και σε παραβίαση του περιορισμού. Από την άλλη μεριά, αν το διάνυσμα S βρίσκεται κοντά στην εφαπτομένη επί της ισοϋψούς της συνάρτησης κόστους, μια μικρή μετακίνηση δεν θα παραβιάζει τον περιορισμό, αλλά ταυτόχρονα δεν θα μειώνει γρήγορα την τιμή της αντικειμενικής συνάρτησης.

Η απαίτηση από μία διεύθυνση S να είναι ταυτόχρονα χρησιμοποιήσιμη και επιτρεπτή, εκφράζεται μαθηματικά ως εξής:

Χρησιμοποιήσιμη διεύθυνση:

$$\nabla F(X) \cdot S \leq 0 \tag{5.4}$$

$$\nabla g(X) \cdot S \leq 0 \tag{5.5}$$

(για τους j ενεργούς περιορισμούς $g_i(X) = 0$

Τώρα ας θεωρήσουμε το σημείο Β του σχήματος το οποίο αντιστοιχεί στο βέλτιστο σχεδιασμό του παραδείγματος. Στο Β, η κλίση της αντικειμενικής συνάρτησης και η κλίση του περιορισμού βρίσκονται σε αντίθετες διευθύνσεις. Συνεπώς, το μόνο πιθανό διάνυσμα S που είναι παράλληλα επιτρεπτό και χρησιμοποιήσιμο, θα είναι ακριβώς εφαπτόμενο τόσο στο σύνορο του περιορισμού όσο και στην ισοϋψή της αντικειμενικής συνάρτησης, δηλαδή θα σχηματίζει γωνία 90^o και με τις δύο κλίσεις. Αυτή η συνθήκη για τη βέλτιστη λύση εκφράζεται μαθηματικά ως εξής:

$$\nabla F(X) + \lambda \nabla g_j(X) = 0 \text{ και } \lambda \geq 0 \tag{5.6}$$

και αν υπάρχουν παραπάνω από ένας ανισοτικοί περιορισμοί είναι περιορισμοί ισότητας h_k είναι

$$\nabla F(x) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x) + \sum_{k=1}^{l} \lambda_{k+m} \nabla h_k(x) = 0 \qquad (5.7)$$

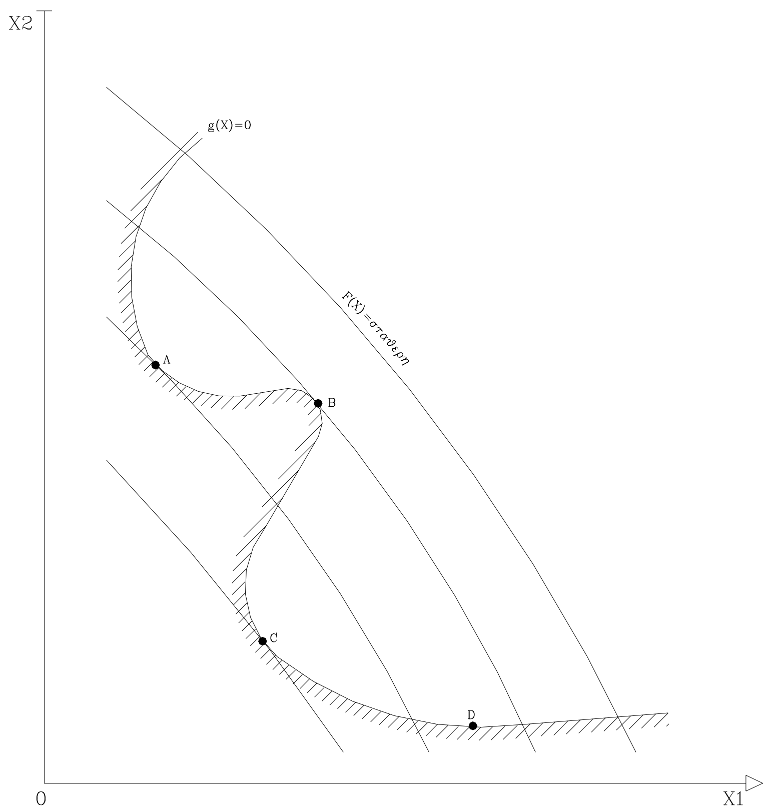
με $\lambda_j \geq 0$ και λ_{m+k} ελεύθερο προσήμου

(οι συντελεστές λ αποτελούν τους πολλαπλασιαστές Lagrance)

Παρατηρούμε ότι αν δεν υπάρχουν περιορισμοί, η εξίσωση (5.7) οδηγεί στην απαίτηση μηδενισμού της κλίσης.

Ωστόσο, αν και η εξίσωση (5.7) αποτελεί μία αναγκαία συνθήκη εύρεσης τοπικού ελαχίστου, με κανένα τρόπο δεν εξασφαλίζει ότι αυτό είναι το καθολικό ελάχιστο, δηλαδή η (5.7) δεν είναι και ικανή συνθήκη.

Αυτό φαίνεται καθαρά στο επόμενο σχήμα, όπου και τα τρία σημεία A,B,C ικανοποιούν την εξίσωση (5.7), αλλά μόνο το C είναι το πραγματική καθολικό ελάχιστο για μειούμενη προς τα κάτω αντικειμενική συνάρτηση.



Τοπικά και καθολικά ακρότατα

5.4 Συνθήκες Kuhn-Tucker

Η εξίσωση (5.7) είναι η τρίτη από το σύνολο των αναγκαίων συνθηκών που πρέπει να ικανοποιούνται σε ένα σημείο προκειμένου αυτό να αποτελεί τοπικά ελάχιστο. Οι τρεις αυτές σχέσεις διατυπώνονται παρακάτω και αποτελούν τις αναγκαίες συνθήκες Kuhn-Tucker. Η έννοιά τους είναι ότι αν ένα σημείο X* αποτελεί τη βέλτιστη λύση ενός προβλήματος , θα πρέπει να ικανοποιεί αυτές τις τρείς σχέσεις:

1. Ο σχεδιασμός που αντιστοιχεί στο X^* είναι επιτρεπτός (5.8)

2. $\lambda_j g_j\left(X^*\right) = 0, \qquad j = 1, 2, \ldots, m$ (5.9)

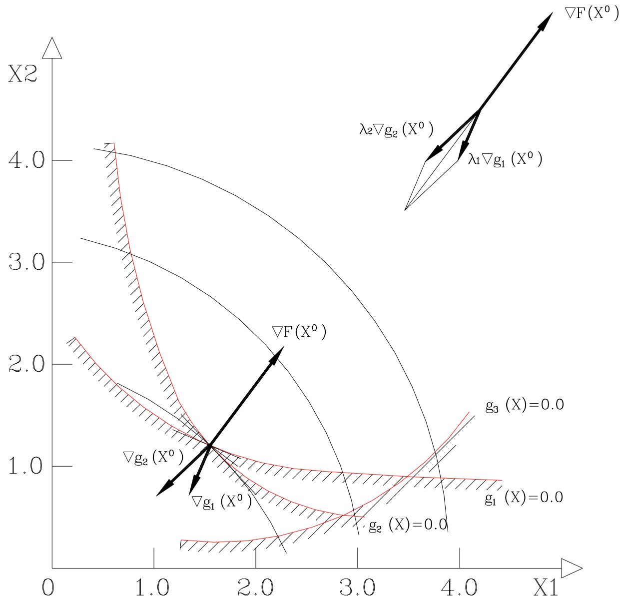
3. $\nabla F(x^*) + \sum_{j=1}^{m} \lambda_j \nabla g_j(x^*) + \sum_{k=1}^{l} \lambda_{k+m} \nabla h_k(x^*) = 0$ (5.10)

με $\lambda_j \geq 0$ και λ_{m+k} ελεύθερο προσήμου

Η σχέση (5.8) διατυπώνει την προφανή απαίτηση ο βέλτιστος σχεδιασμός να ικανοποιεί όλους του περιορισμούς. Η εξίσωση (5.9) σημαίνει ότι αν ένας περιορισμός δεν είναι ενεργός, δηλαδή $g_j(X) < 0$, τότε ο αντίστοιχος πολλαπλασιαστής Lagrange μηδενίζεται, δηλαδή $\lambda_j = 0$.

Η γεωμετρική ερμηνεία των συνθηκών Kuhn-Tucker θα γίνει με βάση το παρακάτω σχήμα, το οποίο απεικονίζει ένα πρόβλημα βελτιστοποίησης με δύο μεταβλητές σχεδιασμού και τρεις ανισοτικούς περιορισμούς.

Όπως φαίνεται, στη βέλτιστη λύση (σημείο X*) ο περιορισμός g_3 δεν είναι ενεργός, συνεπώς βάσει της (5.9) έχουμε λ_3=0.



Απεικόνιση Συνθηκών Kuhn-Tucker για δύο ενεργούς περιορισμούς.

Η σχέση (5.10) εκφράζει ουσιαστικά την απαίτηση το διάνυσμα κλίσης της αντικειμενικής συνάρτησης στη βέλτιστη λύση Χ* να έχει αντίθετη φορά από το άθροισμα των διανυσμάτων κλίσης των ενεργών περιορισμών. Ανατρέχοντας στο σχήμα αυτό είναι φανερό και διατυπώνεται μαθηματικά από τη σχέση:

$$\nabla F\left(X^*\right) + \lambda_1 g_1\left(X^*\right) + \lambda_2 g_2\left(X^*\right) = 0, \qquad \lambda_1, \lambda_2 > 0$$

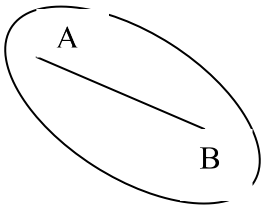
Τέλος, δεδομένου ότι οι περιορισμοί, g_1 και g_2 είναι ενεργοί οπότε

g_1 (Χ*)=0 και g_2(Χ*)=0, η δεύτερη σχέση των συνθηκών Kuhn-Tucker ικανοποιείται, ενώ προφανώς ικανοποιείται και η πρώτη σχέση αφού το Χ* βρίσκεται στην επιτρεπτή περιοχή.

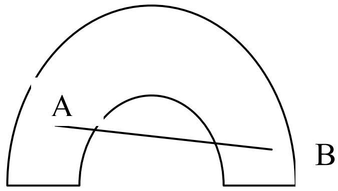
Θα πρέπει επίσης να σημειώσουμε ότι οι πολλαπλασιαστές Lagrange είναι μονοσήμαντα ορισμένοι βάσει των κλίσεων της αντικειμενικής συνάρτησης και των περιορισμών, έτσι ώστε να ικανοποιείται η (5.10).

Το ερώτημα που τίθεται είναι πότε οι συνθήκες Kuhn - Tucker είναι ταυτόχρονα αναγκαίες και ικανές για να οδηγήσουν στην εύρεση του καθολικού ελαχίστου. Για να απαντήσουμε στο ερώτημα αυτό, θα πρέπει πρώτα να αναφερθούμε στην έννοια της κυρτότητας.

Ας θεωρήσουμε ότι στο σχήμα α) ενώνουμε τα τυχαία σημεία A και B με μία ευθεία. Είναι φανερό, ότι οποιοδήποτε σημείο επί του ευθύγραμμου τμήματος AB ανήκει στην επιτρεπτή περιοχή. Από την άλλη μεριά, αν ενώσουμε τα σημεία A και B του σχήματος β), παρατηρούμε ότι κάποια σημεία του AB βρίσκονται εντός της επιτρεπτής περιοχής.



α)



β)

Κυρτό και μη κυρτό χωρίο

6

ΠΡΟΣΕΓΓΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

6.1. ΕΙΣΑΓΩΓΗ

Τα περισσότερα προβλήματα βέλτιστου σχεδιασμού κατασκευών είναι μη γραμμικά. Για την επίλυση τους, έχουν αναπτυχθεί δύο κύριες κατηγορίες μεθόδων. Η πρώτη περιλαμβάνει μεθόδους που μετατρέπουν το πρόβλημα βέλτιστου σχεδιασμού σε μία αλληλουχία απλούστερων προβλημάτων χωρίς περιορισμούς, οι λύσεις των οποίων συγκλίνουν στη λύση του αρχικού προβλήματος. Η κατηγορία αυτή γενικά απαιτεί πολλές επιλύσεις και δεν ενδείκνυται για προβλήματα βέλτιστου σχεδιασμού.

Στο κεφάλαιο αυτό θα ασχοληθούμε με τη δεύτερη κατηγορία, που περιλαμβάνει τις λεγόμενες προσεγγιστικές μεθόδους (direct methods) και ιδιαίτερα τις μεθόδους διαδοχικού γραμμικού ή δευτεροβάθμιου προγραμματισμού (sequential linear or quadratic programming methods). Σύμφωνα με τις μεθόδους αυτές το μη-γραμμικό πρόβλημα διαδοχικά προσεγγίζεται από ένα γραμμικό πρόβλημα ή ένα δευτεροβάθμιο πρόβλημα τα οποία επιλύονται εύκολα.

Το μη-γραμμικό πρόβλημα βέλτιστού σχεδιασμού γενικά διατυπώνεται ως εξής:

Να βρεθεί ένα διάνυσμα $x = (x_1, x_2, ..., x_n)$ που αντιπροσωπεύει ένα σχεδιασμό - για το οποίο ελαχιστοποιείται η αντικειμενική συνάρτηση:

$$f = f(x)$$

που υπόκειται στους περιορισμούς ισότητας

$$h_i(x) = 0 \qquad i = 1,2,..., p$$

και τους περιορισμούς ανισότητας

$$g_i(x) \le 0 \qquad i = 1,2,..., m$$

ενώ κάθε μεταβλητή σχεδιασμού βρίσκεται στο διάστημα

$$x_{il} \le x_i \le x_{iu} \qquad i = 1,2,...,n$$

όπου x_{il} και x_{iu} είναι αντίστοιχα η ελάχιστη και η μέγιστη επιτρεπόμενη τιμή για την μεταβλητή x_i .

6.2 ΓΕΝΙΚΕΣ ΕΝΝΟΙΕΣ

Στις παρακάτω παραγράφους θα αναπτυχθούν ορισμένες βασικές έννοιες που απαιτούνται για την κατανόηση της μεθόδου διαδοχικού γραμμικού προγραμματισμού SLP (Sequential Linear Programming). Η ανάπτυξη των εννοιών θα είναι αναλυτική, διότι πολλές από αυτές χρησιμοποιούνται και σε άλλες μεθόδους βελτιστοποίησης.

6.2.1 Ανάπτυξη σε σειρές Taylor

Όπως αναφέρθηκε, η μέθοδος SLP βασίζεται στην προσέγγιση των μη γραμμικών συναρτήσεων του προβλήματος με γραμμικές. Η μετατροπή γίνεται αναπτύσσοντας τις συναρτήσεις σε σειρές Taylor.

Γενικά, μία συνάρτηση μπορεί να προσεγγισθεί από μία σειρά όρων γύρω από ένα σημείο στη γειτονιά του σημείου, με όρους την τιμή της συνάρτησης και των παραγώγων της στο συγκεκριμένο σημείο, αναπτύσσοντας τη συνάρτηση σε σειρά Taylor.

Ας θεωρήσουμε τη συνάρτηση μίας μεταβλητής $f(x)$. Αναπτύσσοντας σε σειρά Taylor στο σημείο x^* έχουμε:

$$f(x) = f(x^*) + \frac{df(x^*)}{dx}(x - x^*) + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2}(x - x^*)^2 + R \tag{6.1}$$

όπου R είναι οι όροι ανωτέρας τάξεως, που παραλείπονται. Αν θέσουμε $x\text{-}x^*=d$, η παραπάνω σχέση γίνεται:

$$f(x^* + d) = f(x^*) + \frac{df(x^*)}{dx} d + \frac{1}{2}\frac{d^2 f(x^*)}{dx^2} d^2 + R \tag{6.2}$$

Για μια συνάρτηση δύο μεταβλητών $f(x_1,x_2)$, η ανάπτυξη κατά Taylor είναι:

$$f(x_1, x_2) = f(x_1^*, x_2^*) + \frac{\partial f}{\partial x_1}(x_1 - x_1^*) + \frac{\partial f}{\partial x_2}(x_2 - x_2^*) +$$

$$+\frac{1}{2}[\frac{\partial^2 f}{\partial x_1^2}(x_1-x_1^*)^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(x_1-x_1^*)(x_2-x_2^*)+\frac{\partial^2 f}{\partial x_2^2}(x_2-x_2^*)^2]+R \quad (6.3)$$

όπου οι μερικές παράγωγοι υπολογίζονται στη θέση (x_1^*,x_2^*). Η παραπάνω έκφραση μπορεί να γραφεί με μορφή αθροισμάτων:

$$f(x_1,x_2)=f(x_1^*,x_2^*)+\sum_{i=1}^{2}\frac{\partial f}{\partial x_1}(x_1-x_1^*)+$$

$$+\frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\partial^2 f}{\partial x_i\partial x_j}(x_i-x_i^*)(x_j-x_j^*)+R \quad (6.4)$$

Με τη βοήθεια της κλίσης ∇f (gradient) και του μητρώου Hessian H, μπορούμε να αναπτύξουμε τη διανυσματική συνάρτηση f σε σειρά Taylor με τη μορφή μητρώων:

$$f(x)=f(x^*)+\nabla f^{\mathrm{T}}(x-x^*)+\frac{1}{2}(x-x^*)^{\mathrm{T}}H(x-x^*)+R \quad (6.5)$$

όπου στη γενικότερη περίπτωση τα x, x* είναι διανύσματα n διαστάσεων και τα μητρώα ∇f και H είναι αντίστοιχα $n\times 1$ και $n\times m$ διαστάσεων. Ορίζοντας $x\text{-}x^*=d$, η προηγούμενη σχέση γίνεται:

$$f(x^*+d)=f(x^*)+\nabla f^{\mathrm{T}}(d)+\frac{1}{2}(d)^{\mathrm{T}}H(d)+R \quad (6.6)$$

Αν από την εξίσωση (6.8) κρατήσουμε μόνο τους γραμμικούς όρους, θα έχουμε μια γραμμική έκφραση της μη γραμμικής συνάρτησης f .

6.2.2 Αλγόριθμος άμεσων μεθόδων βελτιστοποίησης (direct methods)

Στη μέθοδο SLP, και γενικότερα σε όλες τις άμεσες μεθόδους βελτιστοποίησης (direct methods), εκλέγουμε μια αρχική λύση του προβλήματος και με την βοήθεια επαναληπτικής διαδικασίας προσεγγίζουμε την πραγματική λύση* . Η μαθηματική έκφραση της παραπάνω διαδικασίας είναι:

$$x^{(k+1)}=x^{(k)}+\Delta x^{(k)} \qquad k=0,1,2,\ldots \quad (6.7)$$

όπου ο εκθέτης k εκφράζει τον αριθμό της επανάληψης και το διάνυσμα $\Delta x^{(\kappa)}$ εκφράζει μια μικρή μεταβολή της τρέχουσας λύσης. Η μεταβολή της λύσης αναλύεται ως εξής:

* Η επαναληπτική διαδικασία για την εύρεση της λύσης έχει διατυπώθει στην παράγραφο 2.8

$$\Delta x^{(k)} = \alpha_k d^{(k)} \tag{6.8}$$

όπου $d^{(k)}$ είναι η διεύθυνση αναζήτησης της νέας λύσης και α_κ είναι ένας συντελεστής που ονομάζεται βήμα, και καθορίζει την απόσταση κίνησης πάνω στην διεύθυνση αναζήτησης. Όπως φαίνεται στο σχήμα 1, το σημείο B είναι η τρέχουσα λύση της επανάληψης k, το $d^{(k)}$ είναι η διεύθυνση αναζήτησης της νέας λύσης και το α είναι το βήμα, δηλαδή η απόσταση της νέας λύσης C από την προηγούμενη B (επί της διεύθυνσης αναζήτησης). Άρα, η βελτίωση της λύσης γίνεται προσδιορίζοντας σε κάθε επανάληψη τη διεύθυνση αναζήτησης και το βήμα.

Σχήμα 1

Τα τέσσερα βασικά βήματα του αλγορίθμου των άμεσων μεθόδων είναι:

Γραμμικοποίηση της αντικειμενικής συνάρτησης και των περιορισμών στην εκάστοτε εκτίμηση της λύσης (κατά Taylor).

Ορισμός του προβλήματος καθορισμού της διεύθυνσης αναζήτησης χρησιμοποιώντας τις γραμμικοποιημένες εκφράσεις του βήματος 1.

Επίλυση του προβλήματος του βήματος 2 και συνεπώς καθορισμός της διεύθυνσης αναζήτησης

Υπολογισμός του βήματος α_κ

6.2.3 Γραμμικοποίηση μη γραμμικών προβλημάτων βελτιστοποίησης

Όπως είδαμε στην παράγραφο 5.2.2 σε κάθε επανάληψη μιας άμεσης μεθόδου επίλυσης, ορίζουμε ένα υποπρόβλημα η λύση του οποίου δίνει τη μεταβολή της λύσης Δx. Το υποπρόβλημα προκύπτει αναπτύσσοντας τις μη γραμμικές συναρτήσεις του προβλήματος σε σειρές Taylor.

Κάθε μέθοδος ξεκινά με μια εκτίμηση της λύσης, η οποία προσεγγίζεται καλύτερα μετά από ορισμένες επαναλήψεις. Έστω $x^{(\kappa)}$ η εκτίμηση της λύσης στην k επανάληψη και $\Delta x^{(\kappa)}$ η μεταβολή της λύσης. Αναπτύσσουμε την αντικειμενική συνάρτηση και τις συναρτήσεις των περιορισμών κατά Taylor στο σημείο $x^{(\kappa)}$ και έτσι αποκτάμε το παρακάτω υποπρόβλημα:

Να ελαχιστοποιηθεί η γραμμικοποιημένη αντικειμενική συνάρτηση:

$$f(x^{(k)} + \Delta x^{(k)}) \cong f(x^{(k)}) + \nabla f^{\mathrm{T}}(x^{(k)})\Delta x^{(k)} \tag{6.9}$$

σύμφωνα με τους γραμμικοποιημένους περιορισμούς ισότητας

$$h_j(x^{(k)} + \Delta x^{(k)}) \cong h_j(x^{(k)}) + \nabla h_j^{\mathrm{T}}(x^{(k)})\Delta x^{(k)} = 0 \quad \text{j=11 έως ρ} \tag{6.10}$$

και τους γραμμικοποιημένους περιορισμούς ανισότητας

$$g_j(x^{(k)} + \Delta x^{(k)}) \cong g_j(x^{(k)}) + \nabla g_j^{\mathrm{T}}(x^{(k)})\Delta x^{(k)} \leq 0 \quad \text{j=11 έως m} \tag{6.11}$$

όπου ∇f , ∇h_j και ∇g_j είναι οι κλίσεις της αντικειμενικής συνάρτησης, του j περιορισμού ισότητας και του j περιορισμού ανισότητας αντίστοιχα, ενώ τα σύμβολα «$\cong$» δείχνουν ότι έχουν παραληφθεί οι μη γραμμικοί όροι της ανάπτυξης κατά Taylor.

Στην συνέχεια θα ορίσουμε κάποιες απλοποιητικές εκφράσεις που θα χρησιμοποιηθούν αργότερα.

$f_k = f\left(x^{(k)}\right)$: τιμή της αντικειμενικής συνάρτησης στην τρέχουσα τιμή της λύσης $x^{(k)}$

$e_j = -h_j\left(x^{(k)}\right)$: τιμή του j περιορισμού ισότητας στην τρέχουσα τιμή της λύσης $x^{(k)}$, με αρνητικό πρόσημο

$b_j = -g\left(x^{(k)}\right)$: τιμή του j ανισοτικού περιορισμού στην τρέχουσα τιμή της λύσης $x^{(k)}$, με αρνητικό πρόσημο

$c_i = \dfrac{\partial f(x^{(k)})}{\partial x_i}$: ο i όρος του διανύσματος της κλίσης της αντικειμενικής συνάρτησης στην τρέχουσα τιμή της λύσης $x^{(k)}$.

$n_{ij} = \dfrac{\partial h_j(x^{(k)})}{\partial x_i}$: ο i όρος του διανύσματος της κλίσης του περιορισμού ισότητας στην τρέχουσα τιμή της λύσης $x^{(k)}$.

$\alpha_{ij} = \dfrac{\partial g_j(x^{(k)})}{\partial x_i}$: ο i όρος του διανύσματος της κλίσης του περιορισμού ανισότητας στην τρέχουσα τιμή της λύσης $x^{(k)}$.

$d_i = \Delta x_i^{(k)}$: Το i στοιχείο του διανύσματος n διαστάσεων d της μεταβολής της λύσης (σχεδιασμού)

Με τη βοήθεια των παραπάνω απλοποιητικών εκφράσεων, το υποπρόβλημα προσδιορισμού της μεταβολής της λύσης Δx=d γίνεται:

Να ελαχιστοποιηθεί η γραμμικοποιημένη αντικειμενική συνάρτηση

$$\overline{f} = \sum_{i=1}^{n} c_i d_i \qquad \underline{(\overline{f} = c^{\mathrm{T}} d)} \tag{6.12}$$

σύμφωνα με τους γραμμικοποιημένους περιορισμούς ισότητας

$$\sum_{i=1}^{n} n_{ij} d_i = e_i \qquad \text{j=1 έως ρ} \quad \left(N^T d = e \right) \tag{6.13}$$

και τους γραμμικοποιημένους ανισοτικούς περιορισμούς

$$\sum_{i=1}^{n} \alpha_{ij} d_i \le b_j \qquad \text{j=1 έως m} \quad \underline{(A^T d \le b)} \tag{6.14}$$

όπου οι στήλες του μητρώου N(nXp) είναι οι κλίσεις των περιορισμών ισότητας και οι στήλες του μητρώου A(nXm) είναι οι κλίσεις των ανισοτικών περιορισμών. Δηλαδή είναι αντίστοιχα

$$n^{(j)} = \left(\frac{\partial h_j}{\partial x_1}, \frac{\partial h_j}{\partial x_2}, \ldots, \frac{\partial h_j}{\partial x_n}\right)$$

$$a^{(j)} = \left(\frac{\partial g_j}{\partial x_1}, \frac{\partial g_j}{\partial x_2}, \ldots, \frac{\partial g_j}{\partial x_n}\right) \tag{6.15}$$

οπότε τα μητρώα N και A γράφονται:

$$\mathrm{N} = [\mathrm{n}^{(j)}]_{(\mathrm{nXp})} \qquad \mathrm{A} = [\mathrm{a}^{(j)}]_{\mathrm{nXm}} \tag{6.16}$$

<u>Παράδειγμα:</u> Διατύπωση ενός γραμμικοποιημένου υποπροβλήματος.

Ας θεωρήσουμε το παρακάτω πρόβλημα βέλτιστου σχεδιασμού:

Να ελαχιστοποιηθεί η συνάρτηση κόστους

$$f(x) = x_1^2 + x_2^2 - 3x_1 x_2$$

σύμφωνα με τους περιορισμούς

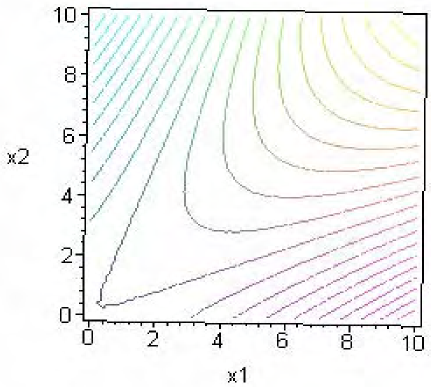
$$g_1(x) = \frac{1}{6} x_1^2 + \frac{1}{6} x_2^2 - 1.0 \le 0$$

$$g_2(x) = -x_1 \le 0$$

$$g_3(x) = -x_2 \leq 0$$

Ζητείται να γραμμικοποιηθεί η συνάρτηση κόστους και οι συναρτήσεις των περιορισμών στο σημείο $x^{(0)}=(1,1)$ και να γραφεί το υποπρόβλημα με τη μορφή των σχέσεων (6.14) έως (6.16).

Λύση Η γραφική επίλυση του προβλήματος φαίνεται στο παρακάτω σχήμα 2. Όπως βλέπουμε, η βέλτιστη λύση βρίσκεται στο σημείο $(\sqrt{3}, \sqrt{3})$ όπου η αντικειμενική συνάρτηση έχει τιμή $f=-3$.



Οι κλίσεις της αντικειμενικής συνάρτησης και των περιορισμών, είναι:

$$\nabla f^T = (2x_1 - 3x_2, 2x_2 - 3x_1)$$

$$\nabla g_1{}^T = (\frac{2}{6}x_1, \frac{2}{6}x_2)$$

$$\nabla g^T{}_2 = (-1, 0)$$

$$\nabla g^T{}_3 = (0, -1)$$

Υπολογίζουμε τις τιμές της αντικειμενικής συνάρτησης και των περιορισμών, καθώς και των κλίσεων τους στο σημείο (1,1)

$f(x^{(0)})=-1.0$

$b_1=-g_1(x^{(0)})=2/3$

$b_2=-g_2(x^{(0)})=1$

$b_3=-g_3(x^{(0)})=1$

$$\nabla f(x^{(0)}) = c = (-1,-1)$$

$$\nabla g_1(x^{(0)}) = c = (\frac{1}{3}, \frac{1}{3})$$

Το μητρώο A και το διάνυσμα b της εξίσωσης (6.16) είναι:

$$A = \begin{bmatrix} \frac{1}{3} & -1 & 0 \\ \frac{1}{3} & 0 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix}$$

Έτσι, το γραμμικοποιημένο υποπρόβλημα υπολογισμού του Δx, γράφεται:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = [-1 -1]\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \Rightarrow \overline{f} = -d_1 - d_2$$

σύμφωνα με τους περιορισμούς

$$\begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \le \begin{bmatrix} \frac{2}{3} \\ 1 \\ 1 \end{bmatrix} \Rightarrow \begin{matrix} \frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3} \\ -d_1 \le 1 \\ -d_2 \le 1 \end{matrix}$$

Μπορούμε επίσης να γράψουμε το πρόβλημα σε όρους των μεταβλητών σχεδιασμού x_1 και x_2. Για να γίνει αυτό, αντικαθιστούμε τις συναρτήσεις του προβλήματος με τα αναπτύγματά τους σε σειρές Taylor στο σημείο $x^{(0)}=(1,1)$.

$$\overline{f}(x_1, x_2) = f(x^{(0)}) + \nabla f(x - x^{(0)}) =$$

$$= -1 + [-1 \quad -1]\begin{bmatrix} (x_1 - 1) \\ (x_2 - 1) \end{bmatrix} = -x_1 - x_2 + 1$$

$$\overline{g}_1(x_1,x_2) = g_1(x^{(0)}) + \nabla g_1(x - x^{(0)}) =$$

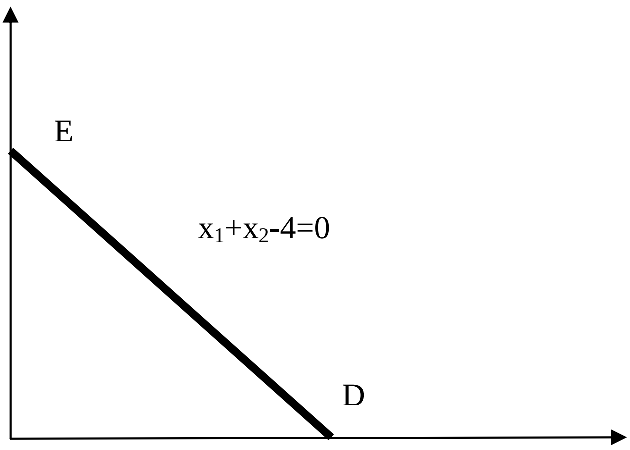
$$= -\frac{2}{3} + \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} (x_1 - 1) \\ (x_2 - 1) \end{bmatrix} = \frac{1}{3}(x_1 + x_2 - 4) \le 0$$

$$\overline{g_2} = -x_1 \le 0$$

$$\overline{g_3} = -x_2 \le 0$$

Όπως φαίνεται, οι συναρτήσεις δευτέρου βαθμού f και g_1 έχουν προσεγγισθεί με γραμμικές συναρτήσεις των x_1 και x_2.

Η νέα επιτρεπτή περιοχή που ορίζεται από το γραμμικοποιημένο στο σημείο (1,1) πρόβλημα φαίνεται στο σχήμα 3. Επειδή η γραμμικοποιημένη συνάρτηση κόστους είναι παράλληλη με τον πρώτο (γραμμικοποιημένο) περιορισμό* , η βέλτιστη λύση είναι οποιοδήποτε σημείο της ευθείας D-E



Εδώ πρέπει να σημειώσουμε ότι η γραμμική προσέγγιση των συναρτήσεων με τη βοήθεια των σειρών Taylor εξαρτάται από το σημείο στο οποίο θα γίνει η ανάπτυξη. Έτσι, στο παράδειγμα, η επιτρεπτή περιοχή θα αλλάξει αν αναπτύξουμε τις συναρτήσεις σε άλλο σημείο.

* Ο πολλαπλασιασμός μιας συνάρτησης με μια θετική σταθερά δεν επηρεάζει τη γραφική παράστασή της (βλ. παρατήρηση παρ. 5.4.3.2)

6.3 ΜΕΘΟΔΟΣ ΔΙΑΔΟΧΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (SLP)

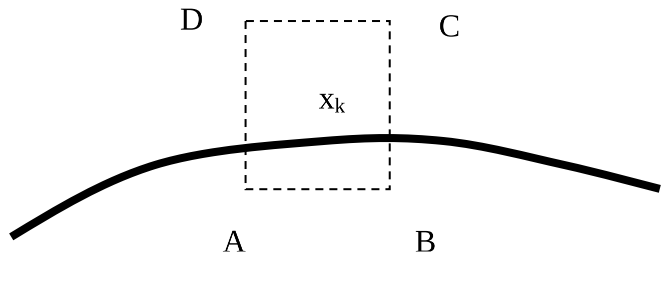
6.3.1 Βασική ιδέα της μεθόδου

Οι εξισώσεις (6.14) έως (6.16) του υποπροβλήματος προσδιορισμού της λύσης d είναι γραμμικές ως προς τις μεταβλητές d_i, οπότε μπορεί να χρησιμοποιηθεί η μέθοδος γραμμικού προγραμματισμού για την επίλυση ως προς d_i. Οι μέθοδοι που χρησιμοποιούν το γραμμικό προγραμματισμό σε κάθε επανάληψη για τον υπολογισμό της μεταβλητής της λύσης **d**, ονομάζονται μέθοδοι διαδοχικού γραμμικού προγραμματισμού (sequential linear programming - SLP).

Για να λύσουμε το υποπρόβλημα προσδιορισμού της μεταβολής **d** της λύσης με την μέθοδο simplex, θα πρέπει οι σταθεροί όροι e_j και b_j των εξισώσεων (6.15) και (6.16) να είναι μη αρνητικοί. <u>Επίσης, θα πρέπει να τεθούν κάποια όρια στην μεταβολή της λύσης **d**, διότι αν αυτή γίνει αρκετά μεγάλη, τότε παύουν να έχουν νόημα οι γραμμικές προσεγγίσεις που έγιναν.</u> (Υπενθυμίζεται ότι $\mathbf{d}=\mathbf{x}-\mathbf{x}^*$ και ότι η ανάπτυξη κατά Taylor γίνεται στη γειτονιά του $\mathbf{x}^*$). Τα όρια που τίθενται είναι η μέγιστη επιτρεπόμενη μείωση $\Delta_{il}^{(k)}$ και η μέγιστη επιτρεπόμενη αύξηση $\Delta_{iu}^{(k)}$ στην i μεταβλητή σχεδιασμού κατά την κ επανάληψη. Ο περιορισμός που προκύπτει, είναι:

$$-\Delta_{il}^{(k)} \le d_i \le \Delta_{iu}^{(k)} \qquad \text{i=1 έως n}$$

Στο σχήμα 4 φαίνεται καθαρά, ότι η αναζήτηση της νέας λύσης πρέπει να γίνει σε σημείο εντός του ορθογωνίου ABCD.



Η επιλογή των ορίων Δ είναι καθοριστικής σημασίας, αφού επηρεάζει σημαντικά την επιτυχία της μεθόδου. Όμως, ο καθορισμός των ορίων απαιτεί κάποια εξοικείωση με την μέθοδο, γι' αυτό δεν πρέπει να διστάσουμε να δοκιμάσουμε διαφορετικές τιμές των Δ αν οδηγηθούμε σε αποτυχία ή ακατάλληλο σχεδιασμό. <u>Συνήθως οι τιμές των Δ_{il} και Δ_{iu} εκλέγονται ως κάποιο κλάσμα των τιμών των μεταβλητών σχεδιασμού.</u>

Το υποπρόβλημα προσδιορισμού του **d** είναι πρακτικά ένα πρόβλημα γραμμικού προγραμματισμού με μεταβλητές d_i. Επειδή οι τιμές των μεταβλητών σχεδιασμού του αρχικού προβλήματος μπορεί να αυξάνονται ή να μειώνονται, οι μεταβλητές d_i μπορεί να είναι θετικές ή αρνητικές. Έτσι, κάθε μεταβλητή d_i μπορεί να αντικατασταθεί από τη διαφορά $d_i=d_i^+-d_i^-$, όπως είδαμε στο παράδειγμα της παραγράφου 4.4.6. Επομένως το υποπρόβλημα των εξισώσεων (6.14) έως (6.16) έχει μετατραπεί στην τυπική μορφή της μεθόδου Simplex.

Η μέθοδος SLP ολοκληρώνεται όταν:

1. Ολοι οι περιορισμοί ικανοποιούνται. Αυτό μπορεί να εκφραστεί ως:

$$g_i \le \varepsilon_1, \quad i=1 \text{ έως } m$$

$$\text{και } |h_i| \le \varepsilon_1, \quad i=1 \text{ έως } p$$

όπου $\varepsilon_1>0$ είναι ένας προκαθορισμένος μικρός αριθμός που ορίζει την ανοχή για την παραβίαση του περιορισμού.

2. Η μεταβολή της λύσης είναι σχεδόν μηδενική. Δηλαδή

$\|d\| \le \varepsilon_2$, όπου $\varepsilon_2>0$ είναι ένας προκαθορισμένος μικρός αριθμός.

6.3.2 Ο αλγόριθμος της μεθόδου SLP

<u>Βήμα 1:</u> Αρχική εκτίμηση της λύσης, $\mathbf{x}^{(0)}$. Θέτουμε κ=0 και καθορίζουμε δύο μικρούς θετικούς αριθμούς ε_1 και ε_2

<u>Βήμα 2:</u> Υπολογίζουμε τις ποσότητες f_κ, b_j (με j=1 έως m και e_j (με j=1 έως p)

<u>Βήμα 3:</u> Υπολογίζουμε τις κλίσεις της αντικειμενικής συνάρτησης και των συναρτήσεων των περιορισμών στην τρέχουσα λύση $\mathbf{x}^{(\kappa)}$, δηλαδή τις ποσότητες

c_i με i=1 έως n

n_{ij} με j=1 έως p και i=1 έως n

α_{ij} με j=1 έως m και i=1 έως n

<u>Βήμα 4:</u> Επιλέγουμε τα κατάλληλα όρια $\Delta_{il}^{(k)}$ και $\Delta_{iu}^{(k)}$ ως κλάσματα των μεταβλητών d_i

<u>Βήμα 5:</u> Ορίζουμε το υποπρόβλημα γραμμικού προγραμματισμού των εξισώσεων (6.14) έως (6.16).

Βήμα 6: Αν απαιτείται, μετασχηματίζουμε το υποπρόβλημα στην κανονική μορφή της μεθόδου simplex, και τελικά υπολογίζουμε το $\mathbf{d}^{(k)}$.

Βήμα 7: Ελέγχουμε τη σύγκλιση της μεθόδου όπως είδαμε παραπάνω.

Βήμα 8: Προχωρούμε στην επόμενη επανάληψη

$$\mathbf{x}^{(k+1)}=\mathbf{x}^{(k)}+\mathbf{d}^{(k)}$$

και συνεχίζουμε στο βήμα 2.

Παράδειγμα 1

Ας θεωρήσουμε το ίδιο παράδειγμα. Θα ορίσουμε το γραμμικοποιημένο υποπρόβλημα στο σημείο (3,3) και θα το λύσουμε θέτοντας τα κατάλληλα όρια Δ.

Λύση

Για να ορίσουμε το γραμμικοποιημένο υποπρόβλημα, υπολογίζουμε τις ακόλουθες ποσότητες στο σημείο (3,3):

$$f(3,3)=3^2+3^2-3.3.3=-9$$

$$g_1(3,3)=1/6\ 3^2+1/6\ 3^2-1=2>0 \text{ (παραβίαση)}$$

$$g_2(3,3)=-3<0 \text{ (ανενεργός περιορισμός)}$$

$$g_3(3,3)=-3<0 \text{ (ανενεργός)}$$

$$c=\nabla f=(2x_1-3x_2,\, 2x_2-3x_1)=(-3,-3)$$

$$\nabla g_1=(\frac{2}{6}x_1,\, \frac{2}{6}x_2)=(1,1)$$

$$\nabla g_2=(-1,\, 0)$$

$$\nabla g_3=(0,\,-1)$$

Το σημείο (3,3) βρίσκεται στην μη επιτρεπτή περιοχή αφού ο πρώτος περιορισμός παραβιάζεται. Το υποπρόβλημα που ορίζεται από τις εξισώσεις (6.14) και (6.16) είναι:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = [-3 \; -3]\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \Rightarrow \overline{f} = -3d_1 - 3d_2$$

σύμφωνα με τους περιορισμούς

$$\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \le \begin{bmatrix} -2 \\ 3 \\ 3 \end{bmatrix} \Rightarrow \begin{matrix} d_1 + d_2 \le -2 \\ -d_1 \le 3 \\ -d_2 \le 3 \end{matrix}$$

Το υποπρόβλημα έχει μόνο δύο μεταβλητές, συνεπώς μπορεί να λυθεί γραφικά, όπως φαίνεται στο σχήμα 5. Η επιτρεπτή περιοχή βρίσκεται εντός του τριγώνου ABC. Η συνάρτηση κόστους είναι παράλληλη στην ευθεία B-C, συνεπώς κάθε σημείο αυτής ελαχιστοποιεί τη συνάρτηση. Μπορούμε να επιλέξουμε d_1=-1 και d_2=-1 ως λύση που ικανοποιεί όλους τους γραμμικοποιημένους περιορισμούς (και δίνει $\overline{f} = 6$).

Αν επιλέξουμε τα όρια Δ ως κλάσμα 100% των μεταβλητών σχεδιασμού x_1=3 και x_2=3, έχουμε

$$-3 \le d_1 \le 3 \quad \text{και} \quad -3 \le d_2 \le 3$$

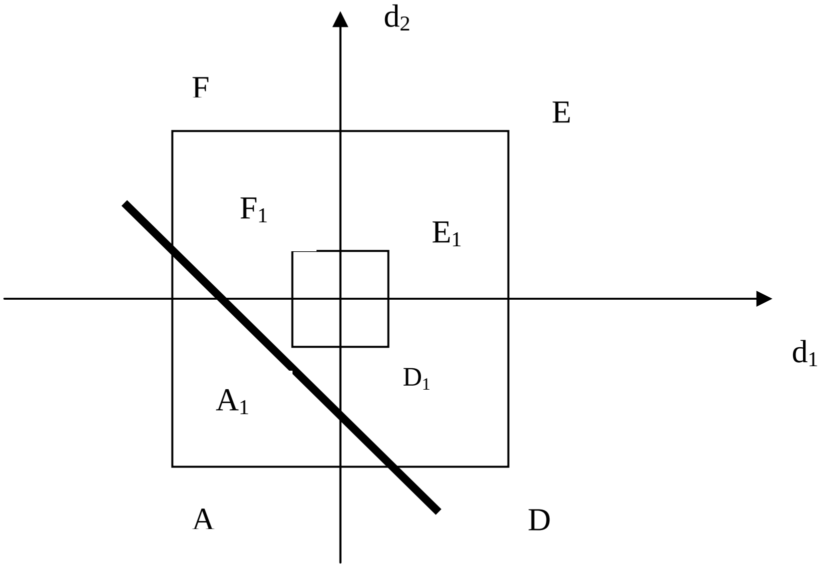
και η λύση πρέπει να είναι στην περιοχή ADEF (σχ. 5).

Αν θέσουμε τα όρια Δ ως κλάσμα 20% των μεταβλητών σχεδιασμού, έχουμε:

$$-0.6 \le d_1 \le 0.6 \quad \text{και} \quad -0.6 \le d_2 \le 0.6$$

Σ’ αυτήν την περίπτωση, η λύση πρέπει να κείται στην περιοχή $A_1D_1E_1F_1$. Από το σχήμα 5 φαίνεται ότι δεν υπάρχει επιτρεπτή λύση στο γραμμικοποιημένο υποπρόβλημα, αφού η περιοχή $A_1D_1E_1F_1$δεν τέμνει την ευθεία B-C. Έτσι, πρέπει να μεγαλώσουμε την περιοχή $A_1D_1E_1F_1$ αυξάνοντας τα όρια Δ. Συμπεραίνουμε δηλαδή, ότι αν τα όρια Δ είναι πολύ περιοριστικά, ενδέχεται το υποπρόβλημα να μην έχει λύση.

Αν επιλέξουμε d_1=-1 και d_2=-1, τότε η νέα προσέγγιση της λύσης είναι: $\mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\mathbf{d}^{(0)}$ = (3,3)+(-1,-1) = (2,2) που εξακολουθεί να βρίσκεται στη μη επιτρεπτή περιοχή. Δηλαδή αν και οι γραμμικοποιημένοι περιορισμοί ικανοποιούνται, ο πραγματικός περιορισμός g_1 παραβιάζεται.



Παράδειγμα 2

Ας θεωρήσουμε το παράδειγμα της παραγράφου 5.2.4. Θα εφαρμόσουμε τη μέθοδο SLP για μία επανάληψη θεωρώντας $\varepsilon_1=\varepsilon_2=0.001$ και εκλέγοντας όρια Δ που επιτρέπουν μεταβολή της λύσης κατά 15%. Επίσης θεωρούμε ως αρχική εκτίμηση τη λύση $\mathbf{x}^{(0)}=(1,1)$.

Λύση

Η αρχική εκτίμηση αποτελεί μία επιτρεπτή λύση για το πρόβλημα, όπως φαίνεται στο σχήμα 2. Το γραμμικοποιημένο υποπρόβλημα με τα κατάλληλα όρια Δ στις μεταβολές d_1 και d_2, στο σημείο $\mathbf{x}^{(0)}$ είναι:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = -d_1 - d_2$$

σύμφωνα με τους περιορισμούς

$$\frac{1}{3}d_1 + \frac{1}{3}d_2 \le \frac{2}{3}$$

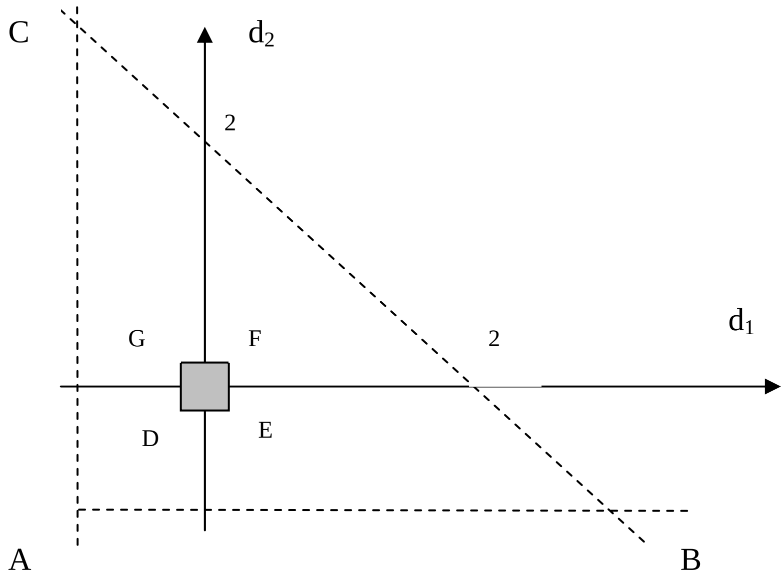
$$-(1+d_1) \le 0$$

$$-(1+d_2) \le 0$$

$$-0.15 \le d_1 \le 0.15$$

$$-0.15 \le d_2 \le 0.15$$

Η γραφική λύση του γραμμικοποιημένου υποπροβλήματος φαίνεται στο σχήμα 6. Τα όρια Δ της μεταβολής της λύσης ορίζουν ως περιοχή λύσεων το τετράγωνο DEFG. Η βέλτιστη λύση του προβλήματος είναι το σημείο E όπου d_1=0.15 και d_2=0.15. Όπως είναι φανερό, στο συγκεκριμένο πρόβλημα μπορούν να τεθούν και πολύ μεγαλύτερα όρια Δ.



Η ίδια λύση του γραμμικοποιημένου υποπροβλήματος προκύπτει και με τη χρήση της μεθόδου simplex. Πρέπει όμως να σημειώσουμε ότι στο γραμμικοποιημένο υποπρόβλημα οι μεταβολές d_1 και d_2 είναι ελεύθερες σε πρόσημο. Έτσι, για να λύσουμε το πρόβλημα με την μέθοδο simplex πρέπει να ορίσουμε νέες μεταβλητές A, B, C καιD τέτοιες ώστε:

d_1=A-B , d_2=C-D με A, B, C και D ≥ 0

Εισάγοντας τις νέες μεταβλητές στο υποπρόβλημα, έχουμε:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = -A + B - C + D$$

σύμφωνα με τους περιορισμούς

$$\frac{1}{3}(A - B + C - D) \le \frac{2}{3}$$

$$-A + B \leq 1.0$$

$$-C + D \leq 1.0$$

$$A - B \leq 0.15$$

$$B - A \leq 0.15$$

$$C - D \leq 0.15$$

$$D - C \leq 0.15$$

$$A, B, C, D \geq 0$$

Η λύση του προβλήματος γραμμικού προγραμματισμού που προέκυψε με τη βοήθεια της μεθόδου Simplex, είναι A=0.15, B=0, C=0.15 και D=0

Συνεπώς:

d_1=A-B=0.15

d_2=C-D=0.15

Η νέα προσέγγιση της λύσης είναι:

$\mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\mathbf{d}^{(0)} = (1.15,1.15)$

Στη νέα προσέγγιση αντιστοιχούν:

$f(\mathbf{x}^{(1)})=-1.3225$ (έναντι $f(\mathbf{x}^{(0)})=-0.3$)

$g_1(\mathbf{x}^{(1)})=-0.6166$

Παρατηρούμε ότι η συνάρτηση κόστους έχει μειωθεί για την νέα εκτίμηση της λύσης χωρίς να παραβιάζεται ο περιορισμός. Αυτό δείχνει ότι η λύση $\mathbf{x}^{(1)}$ είναι βελτιωμένη συγκριτικά με την αρχική $\mathbf{x}^{(0)}$. Ομως, δεδομένου ότι η νόρμα $\|d\| = 0.212$ είναι μεγαλύτερη από την επιτρεπτή ανοχή ε_2=0.001, απαιτούνται περισσότερες επαναλήψεις για την ολοκλήρωση της μεθόδου SLP.

Τέλος, σημειώνουμε ότι το γραμμικοποιημένο υποπρόβλημα στο σημείο (1,1) μπορεί να γραφεί και συναρτήσει των πραγματικών μεταβλητών σχεδιασμού. Αυτό έγινε στο παράδειγμα της παρ. 5.2.4 και προέκυψε:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} - x_1 - x_2 + 1$$

σύμφωνα με τους περιορισμούς

$$\overline{g_1} = \frac{1}{3}(x_1 + x_2 - 4) \leq 0$$

$$\overline{g_2} = -x_1 \leq 0$$

$$\overline{g_3} = -x_2 \leq 0$$

Επίσης, βάσει της σχέσης $-\Delta_{il} \leq x_i - x_i^{(0)} \leq \Delta_{iu}$ έχουμε:

$$-0.15 \leq (x_1 - 1) \leq 0.15 \quad \text{ή} \quad 0.85 \leq x_1 \leq 1.15$$

$$-0.15 \leq (x_2 - 1) \leq 0.15 \quad \text{ή} \quad 0.85 \leq x_2 \leq 1.15$$

Λύνοντας αυτό το υποπρόβλημα, μπορεί να προκύψει απευθείας η νέα λύση, που και εδώ είναι η (1.15,1.15).

6.4 ΜΕΘΟΔΟΣ ΔΙΑΔΟΧΙΚΟΥ ΔΕΥΤΕΡΟΒΑΘΜΙΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

Οπως φάνηκε στην προηγούμενη παράγραφο, η μέθοδος SLP είναι μιά απλή επέκταση της μεθόδου γραμμικού προγραμματισμού για την επίλυση μη-γραμμικών προβλημάτων βελτιστοποίησης. Ωστόσο, η μέθοδος έχει κάποια μειονεκτήματα, τα οποία αντιμετωπίζονται με την επίλυση ενός υποπροβλήματος δευτεροβάθμιου προγραμματισμού (quadratic programming subproblem - QP).

Ενα μειονέκτημα της μεθόδου SLP είναι ότι εξαρτάται από την επιλογή των κατάλληλων ορίων Δ για τις μεταβολές των λύσεων **d**. Επίσης, δεν μπορεί να αποδειχθεί η σύγκλιση της μεθόδου σε ένα σχετικό ελάχιστο ξεκινώντας από μία αυθαίρετα εκλεγμένη λύση. Τέλος, τα όρια Δ πρέπει να εκλέγονται εκ νέου σε κάθε επανάληψη της μεθόδου.

Για να αντιμετωπιστούν αυτές οι δυσκολίες, έχουν αναπτυχθεί άλλες μέθοδοι για την επίλυση του υποπροβλήματος προσδιορισμού των μεταβολών των λύσεων. Έτσι, τα όρια Δ της σχέσης $-\Delta_{il}^{(k)} \le d_i \le \Delta_{iu}^{(k)}$ εγκαταλείπονται και εισάγεται ένας περιορισμός της μορφής

$$\|d\| \le \xi \tag{6.17}$$

όπου $\|d\|$ είναι το μήκος του διανύσματος αναζήτησης **d** και ξ είναι ένας μικρός θετικός αριθμός. Δηλαδή, η ανισότητα (6.17) θέτει έναν περιορισμό στο μήκος του διανύσματος μεταβολής της λύσης **d**. Χρησιμοποιώντας τον ορισμό του μήκους διανύσματος και υψώνοντας τα δύο μέλη στο τετράγωνο, η σχέση (6.17) γίνεται:

$$0.5\sum_{i=1}^{n} (d_i)^2 \le \xi^2 \quad (0.5 d^{\mathrm{T}} d \le \xi^2) \tag{6.18}$$

Ο συντελεστής 0.5 εισάγεται για την απαλοιφή του συντελεστή 2 κατά τη διάρκεια υπολογισμών που θα γίνουν αργότερα. Δεν επηρεάζει τους υπολογισμούς για την εύρεση του διανύσματος αναζήτησης, αλλά επηρεάζει το βήμα ξ.

Το σχήμα 7 δείχνει το δευτεροβάθμιο περιορισμό του **d**, ο οποίος είναι πολύ διαφορετικός από το γραμμικό περιορισμό στο σχήμα 4. Όπως φαίνεται, η νέα λύση απαιτείται να βρίσκεται εντός κύκλου διαμέτρου $\sqrt{2}\xi$ με κέντρο την τρέχουσα λύση $\mathbf{x}^{(k)}$. Έτσι το υποπρόβλημα που πρέπει να λυθεί ορίζεται ως εξής:

Να υπολογιστούν οι μεταβολές d_1, d_2, ..., d_n που ελαχιστοποιούν τη γραμμικοποιημένη συνάρτηση κόστους

$$\overline{f} = \sum_{i=1}^{n} c_i d_i \quad (\overline{f} = c^{\mathrm{T}} d) \tag{6.19}$$

σύμφωνα με τους γραμμικοποιημένους περιορισμούς ισότητας

$$\sum_{i-1}^{n} n_{ij} d_i = e_j \quad \text{j=1 έως p} \qquad (N^{\mathrm{T}} d = e) \tag{6.20}$$

τους γραμμικοποιημένους ανισοτικούς περιορισμούς

$$\sum_{i-1}^{n} \alpha_{ij} d_i \le b_j \quad \text{=1 έως m} \qquad (A^{\mathrm{T}} d \le b) \tag{6.21}$$

και το δευτεροβάθμιο περιορισμό βήματος

$$0.5 \sum_{i=1}^{n} (d_i^2) \le \xi^2 (0.5 d^{\mathrm{T}} d \le \xi^2) \tag{6.22}$$

Το υποπρόβλημα είναι μη γραμμικό, λόγω της μη γραμμικότητας του περιορισμού του βήματος. Έτσι, για την επίλυσή του δεν μπορεί να εφαρμοστεί η μέθοδος simplex.

Σχήμα 7

6.4.1 Υποπρόβλημα δευτεροβάθμιου προγραμματισμού (QP)

Ένα πρόβλημα δευτεροβάθμιου προγραμματισμού έχει δευτεροβάθμια συνάρτηση κόστους και γραμμικούς περιορισμούς. Το μη γραμμικό υποπρόβλημα προσδιορισμού της διεύθυνσης **d**, που ορίζεται από τις εξισώσεις (6.19) έως (6.22) δεν είναι αυτής της μορφής, αλλά μπορεί να μετασχηματιστεί. Το μειονέκτημα του υποπροβλήματος που ορίζεται από αυτές τις σχέσεις είναι ότι μπορεί να μη δίνει λύση αν το ξ είναι πολύ μικρό και η τρέχουσα λύση είναι μη επιτρεπτή. Αυτό φαίνεται από το σχήμα 7. Αν ο κύκλος δεν τέμνει την επιτρεπτή περιοχή (π.χ. η ακτίνα του είναι πολύ μικρή), τότε το υποπρόβλημα δεν έχει επιτρεπτή λύση. Μετασχηματίζοντας όμως το υποπρόβλημα στη μορφή QP (που φαίνεται παρακάτω), εξασφαλίζεται η ύπαρξη λύσης η οποία είναι μοναδική αν το πρόβλημα είναι κυρτό (βλ. παρ. 2.9.3). Επίσης, υπάρχουν πολλές αριθμητικές μέθοδοι επίλυσης του προβλήματος QP. Στη συνέχεια θα αναφερθούμε στη τροποποιημένη μέθοδο simplex και για απλούστερα προβλήματα στη χρήση των συνθηκών Kuhn-Tucker (για τις συνθήκες Kuhn-Tucker βλ. παρ. 2.9.2).

Είναι δυνατόν να ορίσουμε το παρακάτω ισοδύναμο υποπρόβλημα QP η λύση του οποίου είναι ίδια με του υποπροβλήματος των σχέσεων (6.19) έως (6.20):

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = c^{\mathrm{T}} d + 0.5 d^{\mathrm{T}} d \tag{6.23}$$

σύμφωνα με τους περιορισμούς

$$N^{\mathrm{T}} d = e \tag{6.24}$$

$$A^{\mathrm{T}} d \leq b \tag{6.25}$$

<u>Παράδειγμα:</u> Διατύπωση ενός προβλήματος QP

Ας θεωρήσουμε το παρακάτω πρόβλημα βελτιστοποίησης:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$f(x)=2x_1^3+15x_2^2-8x_1x_2-4x_1$$

σύμφωνα με τους περιορισμούς

$$h(x) = x_1^2 + x_1 x_2 + 1.0 = 0$$

$$g(x) = x_1 - \frac{1}{4} x_2^2 - 1.0 \leq 0$$

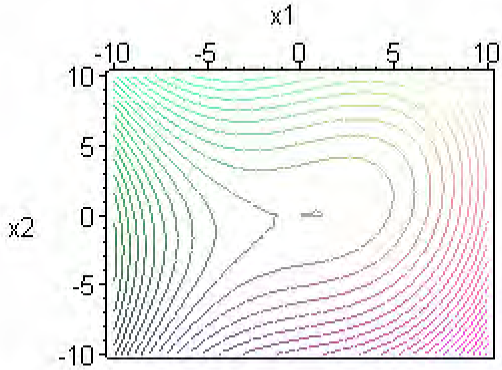
Ζητείται να γραμμικοποιηθούν η αντικειμενική συνάρτηση και οι συναρτήσεις των περιορισμών στο σημείο (1,1) και να διατυπωθεί το υποπρόβλημα QP.

<u>Λύση</u>

Το παρακάτω σχήμα δίνει τη γραφική λύση του προβλήματος. Προφανώς ο περιορισμός ισότητας h πρέπει να ικανοποιείται, οπότε η βέλτιστη λύση βρίσκεται επί των καμπυλών h=0. Όπως φαίνεται από το σχήμα, υπάρχουν δύο βέλτιστες λύσεις:

Σημείο Α: $\mathbf{x}^*=(1,-2)$, $f(\mathbf{x}^*) = 74$

Σημείο Β: $\mathbf{x}^*=(-1,2)$, $f(\mathbf{x}^*) = 78$



Οι κλίσεις της συνάρτησης κόστους και των συναρτήσεων των περιορισμών είναι:

$$\nabla f = (6x_1^2 - 8x_2 - 4, 30x_2 - 8x_1)$$

$$\nabla h = (2x_1 + x_2, x_1)$$

$$\nabla g = (1, -x_2 / 2)$$

Οι τιμές των συναρτήσεων κόστους και των συναρτήσεων των περιορισμών καθώς και οι τιμές των κλίσεων τους στο σημείο (1,1) είναι:

f(1,1)=5

$h(1,1) = 3 \neq 0$ (παραβίαση)

$g(1,1) = -0.25 < 0$ (ανενεργός περιορισμός)

$c = \nabla f = (-6, 22)$

$\nabla h = (3, 1)$

$\nabla g = (1, -0.5)$

Χρησιμοποιώντας όρια Δ της τάξης 50%, το υποπρόβλημα γραμμικού προγραμματισμού (LP) των σχέσεων (6.14) έως (6.16) γίνεται:

Να ελαχιστοποιηθεί η γραμμικοποιημένη συνάρτηση κόστους

$$\overline{f} = -6d_1 + 22d_2$$

σύμφωνα με τους γραμμικοποιημένους περιορισμούς

$$3d_1 + d_2 = -3$$

$$d_1 - 0.5d_2 \leq 0.25$$

$$-0.5 \leq d_1 \leq 0.5$$

$$-0.5 \leq d_2 \leq 0.5$$

Το υποπρόβλημα δευτεροβάθμιου προγραμματισμού (QP) των σχέσεων (6.23) έως (6.25) είναι:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$\overline{f} = (-6d_1 + 22d_2) + 0.5(d_1^2 + d_2^2)$$

σύμφωνα με τους περιορισμούς

$$3d_1 + d_2 = -3$$

$$d_1 - 0.5d_2 \leq 0.25$$

Οι λύσεις των παραπάνω υποπροβλημάτων LP και QP, φαίνονται στα σχήματα 8 και 9 αντίστοιχα. Η λύση πρέπει να ικανοποιεί τον περιορισμό ισότητας και επίσης να βρίσκεται στην επιτρεπτή περιοχή που ορίζει ο περιορισμός ανισότητας. Έτσι, η λύση θα κείται στην ευθεία G-C. Από το σχήμα 8 φαίνεται ότι με όρια 50%, το γραμμικοποιημένο υποπρόβλημα έχει μη επιτρεπτή λύση, διότι το τετράγωνο HIJK που αντιστοιχεί στα όρια Δ δεν τέμνει την ευθεία G-C. Θέτουμε λοιπόν τα όρια Δ ως κλάσμα 100% των μεταβλητών σχεδιασμού (x_1=1, x_2=1), οπότε το σημείο L δίνει τη βέλτιστη λύση ως:

$$d_1 = -\frac{2}{3}, \qquad d_2 = -1.0, \qquad \overline{f} = -18$$

Σχήμα 9

Στο πρόβλημα QP οι περιορισμοί είναι οι ίδιοι, αλλά όπως φαίνεται από τις ισοϋψείς της συνάρτησης κόστους, η συνάρτηση κόστους είναι δευτεροβάθμιοι. Επίσης, στο υποπρόβλημα QP δεν αντιμετωπίζεται η χρονοβόρα διαδικασία επιλογής των ορίων Δ. Η βέλτιστη λύση βρίσκεται στο σημείο:

$$d_1 = -0.5,\ d_2 = -1.5,\ \overline{f} = -28.75$$

<u>Η διεύθυνση αναζήτησης που προέκυψε από το υποπρόβλημα QP είναι μοναδική, σε αντίθεση με το LP όπου εξαρτάται από τον καθορισμό των ορίων Δ.</u> Έτσι, οι διευθύνσεις που προσδιορίζονται από τα δύο υποπροβλήματα είναι γενικά διαφορετικές.

6.4.3 Λύση προβλημάτων δευτεροβάθμιου προγραμματισμού (QP)

Η μέθοδος επίλυσης των προβλημάτων δευτεροβάθμιου προγραμματισμού (QP) αποτελεί μία απλή επέκταση της μεθόδου simplex. Καταρχήν, θα ορίσουμε το πρόβλημα QP, ως εξής[*] :

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$q(x) = c^T x + 0.5 x^T H x \tag{6.26}$$

σύμφωνα με τους περιορισμούς ισότητας και ανισότητας

$$A^{\mathrm{T}} x \le b \tag{6.27}$$

$$N^{\mathrm{T}} x = e \tag{6.28}$$

$$\text{και} \qquad x \ge 0 \tag{6.29}$$

Στη διατύπωση του προβλήματος, οι γραμμικοί περιορισμοί ανισότητας εκφράζονται στην «μορφή $\le$», επειδή αυτή τη μορφή έχουν οι αντίστοιχοι περιορισμοί στη διατύπωση των συνθηκών Kuhn-Tucker. Επίσης, γίνεται η παραδοχή ότι το Hessian μητρώο **H** είναι τουλάχιστον θετικά ημιορισμένο, ώστε το πρόβλημα να είναι κυρτό και έτσι οποιαδήποτε λύση του να αντιπροσωπεύει ένα καθολικό ελάχιστο.

Διατύπωση συνθηκών Kuhn-Tucker για το πρόβλημα QP

Ενας τρόπος επίλυσης του προβλήματος QP είναι να διατυπώσουμε πρώτα γι' αυτό τις συνθήκες Kuhn-Tucker και στη συνέχεια να τις μετασχηματίσουμε έτσι ώστε να μπορούν να χρησιμοποιηθούν στη Φάση Ι της μεθόδου simplex. Για να γράψουμε τις αναγκαίες συνθήκες

[*] Σημειώνεται ότι τα c, b, e, A, N είναι διανύσματα ή μητρώα των οποίων τα στοιχεία είναι κάποιοι σταθεροί όροι και δεν έχουν καμμία σχέση με τα αντίστοιχα μητρώα της παρ. 5.2.4. Τα μητρώα αυτά χρησιμοποιούνται απλώς για να δώσουν μητρωική έκφραση στην αντικειμενική συνάρτηση και τους περιορισμούς.

Kuhn-Tucker, εισάγουμε τις βοηθητικές μεταβλητές **s** για τους περιορισμούς ανισότητας της σχέσης (6.27) και έτσι τους μετατρέπουμε σε περιορισμούς ισότητας:

$$\mathbf{A}^{T}\mathbf{x}+\mathbf{s}=\mathbf{b} \quad \text{με} \quad s \geq 0$$

ή λύνοντας ως προς s για τον j περιορισμό ανισότητας

$$s_j = b_j - \sum_{i=1}^{n} \alpha_{ij} x_i \qquad (\mathbf{s}=\mathbf{b}-\mathbf{A}^{T}\mathbf{x})$$

Τώρα, θα ορίσουμε τη συνάρτηση Lagrange για το πρόβλημα QP

$$L=\mathbf{c}^{T}\mathbf{x}+0.5\mathbf{x}^{T}\mathbf{H}\mathbf{x}+\mathbf{u}^{T}(\mathbf{A}^{T}\mathbf{x}+\mathbf{s}-\mathbf{b})-\zeta^{T}\mathbf{x}+\mathbf{v}^{T}(\mathbf{N}^{T}\mathbf{x}-\mathbf{e})$$

όπου **u,v** και ζ είναι οι πολλαπλασιαστές Lagrange για τους περιορισμούς ανισότητας, τους περιορισμούς ισότητας και τον περιορισμό $-x \leq 0$ αντίστοιχα.

Οι αναγκαίες συνθήκες Kuhn-Tucker δίνουν:

$$\frac{\partial L}{\partial x} = c + Hx + Au - \zeta + Nv = 0 \tag{6.30}$$

$$\mathbf{A}^{T}\mathbf{x}+\mathbf{s}-\mathbf{b}=\mathbf{0} \tag{6.31}$$

$$\mathbf{N}^{T}\mathbf{x}-\mathbf{e}=\mathbf{0} \tag{6.32}$$

$$u_i s_i=0, \quad i=1 \text{ έως } m \tag{6.33}$$

$$\zeta_i x_i=0, \quad i=1 \text{ έως } n \tag{6.34}$$

$$s_i, u_i \geq 0, \quad i=1 \text{ έως } m \tag{6.35}$$

$$\zeta_i \geq 0 \quad i=1 \text{ έως } n \tag{6.36}$$

Δεδομένου ότι οι πολλαπλασιαστές Lagrange **v** για τους περιορισμούς ισότητας είναι ελεύθεροι σε πρόσημο, μπορούν να αναλυθούν ως εξής:

$$\mathbf{v}=\mathbf{y}-\mathbf{z} \quad \text{με} \quad y, z \geq 0$$

Γράφοντας τις εξισώσεις (6.30) έως (6.32) σε μορφή μητρώου, παίρνουμε:

$$\begin{bmatrix} H & A & -I_{(n)} & O_{(nXm)} & N & -N \\ A^{\mathrm{T}} & O_{(mXm)} & O_{(mXn)} & I_m & O_{(mXp)} & O_{(mXp)} \\ N^{\mathrm{T}} & O_{(pXm)} & O_{(pXn)} & O_{(pXm)} & O_{(pXp)} & O_{(pXp)} \end{bmatrix} \begin{bmatrix} x \\ u \\ \zeta \\ s \\ y \\ z \end{bmatrix} = \begin{bmatrix} -c \\ b \\ e \end{bmatrix} \quad \text{ή}$$

$$\mathbf{Bx} = \mathbf{D} \tag{6.37}$$

Οι συνθήκες Kuhn-Tucker έχουν αναχθεί πλέον σε ένα πρόβλημα εύρεσης της λύσης **x** του συστήματος (6.33), σύμφωνα με τους περιορισμούς (6.33) έως (6.36). Για τις νέες μεταβλητές x_i οι συμπληρωματικές συνθήκες των εξισώσεων (6.33) και (6.34) γίνονται:

$$X_i \; X_{n+m+i}=0 \qquad \text{i=1 έως (n+m)} \tag{6.38}$$

$$x_i \geq 0 \qquad \text{i=1 έως (2n+2m+2p)} \tag{6.39}$$

Η μέθοδος simplex για την επίλυση του προβλήματος QP

Η λύση του γραμμικού συστήματος (6.37) που ικανοποιεί τις συνθήκες (6.38) και (6.39), είναι η λύση του προβλήματος QP. Η συνθήκη (6.38) είναι μη γραμμική ως προς τις μεταβλητές x_i οπότε η μέθοδος Simplex δεν μπορεί να χρησιμοποιηθεί για την επίλυση του συστήματος. Όμως, έχει αναπτυχθεί μία τροποποιημένη μέθοδος simplex σύμφωνα με την οποία το σύστημα συγκλίνει σε μία λύση αν το μητρώο **H** είναι θετικά ορισμένο ή αν το **H** είναι θετικά ημιορισμένο και το διάνυσμα **c** της σχέσης (6.26) είναι μηδενικό.

Η μέθοδος βασίζεται στη Φάση Ι της μεθόδου Simplex όπου ορίστηκε μια τεχνητή μεταβλητή για κάθε περιορισμό ισότητας, και μία τεχνητή αντικειμενική συνάρτηση. Ακολουθώντας αυτή τη διαδικασία εισάγουμε μία τεχνητή μεταβλητή Y_i για κάθε εξίσωση του συστήματος (6.37):

$$\mathbf{BX}+\mathbf{Y}=\mathbf{D} \tag{6.40}$$

όπου το **Y** είναι ένα διάνυσμα διάστασης (n+m+p). Με τον τρόπο αυτό έχουμε ορίσει τις μεταβλητές X_i ως μη βασικές, και τις Y_i ως βασικές. Σημειώνεται ότι όλα τα στοιχεία του **D** πρέπει να είναι μη αρνητικά ώστε η αρχική βασική λύση να είναι επιτρεπτή, ειδάλλως πολλαπλασιάζουμε με -1 την αντίστοιχη εξίσωση.

Η τεχνητή αντικειμενική συνάρτηση ορίζεται ως εξής:

$$w = \sum_{i=1}^{n+m+p} Y_i \tag{6.41}$$

Σύμφωνα με τη μέθοδο simplex, εκφράζουμε την τεχνητή συνάρτηση κόστους συναρτήσει μόνο των μη βασικών μεταβλητών απαλείφοντας τις βασικές μεταβλητές Y_i μέσω της (6.40)

$$w = \sum_{i=1}^{n+m+p} D_i - \sum_{j=1}^{2(n+m+p)} \sum_{i=1}^{n+m+p} B_{ij} x_j \; = w_0 + \sum_{j=1}^{2(n+m+p)} c_j x_j \tag{6.42}$$

όπου w_0 είναι μία αρχική τιμή της τεχνητής συνάρτησης κόστους και c_j είναι το άθροισμα των στοιχείων της j στήλης του μητρώου **B** με αντίθετο πρόσημο.

Πριν προχωρήσουμε στη Φάση Ι της μεθόδου simplex εξασφαλίζουμε την ικανοποίηση της συνθήκης (6.38) ελέγχοντας αν οι X_i και X_{n+m+i} είναι ταυτόχρονα μη βασικές. Αν είναι ταυτόχρονα βασικές, τότε μία από αυτές έχει μηδενική τιμή.

Η διαδικασία επίλυσης των προβλημάτων QP συνοψίζεται στα παρακάτω βήματα:

1. Με βάση το δεδομένο πρόβλημα QP των σχέσεων (6.26) έως (6.29) ορίζουμε το μητρώο **B** της σχέσης (6.37).

2. Ορίζουμε το διάνυσμα **D** της σχέσης (6.37), εξασφαλίζοντας ότι όλα τα στοιχεία του είναι μη αρνητικά, δηλαδή $D_i \geq 0$.

3. Υπολογίζουμε τις τρέχουσες τιμές της τεχνητής αντικειμενικής συνάρτησης και των συντελεστών c_j και w_0 της εξίσωσης (6.42).

4. Ολοκληρώνουμε τη Φάση Ι της μεθόδου simplex. Αν η τιμή της τεχνητής αντικειμενικής συνάρτησης είναι διάφορη του μηδενός και οι συντελεστές c_j και w_0 είναι μη αρνητικοί ή αν δεν μπορεί να προσδιοριστεί το στοιχείο οδηγός, τότε το πρόβλημα QP δεν έχει επιτρεπτή λύση.

5. Αν βρεθεί μία επιτρεπτή λύση -από τις εξισώσεις (6.37) έως (6.39)- τότε, χρησιμοποιώντας τον ορισμό του διανύσματος **X** της εξίσωσης (6.37), υπολογίζουμε τις τιμές των μεταβλητών σχεδιασμού **X**, των πολλαπλασιαστών Lagrange **u**, **v**, **ζ** και των βοηθητικών μεταβλητών **s**.

Η παραπάνω μεθοδολογία θα επεξηγηθεί στο παρακάτω παράδειγμα:

<u>Παράδειγμα</u>

Να λυθεί το ακόλουθο πρόβλημα QP:

Να ελαχιστοποιηθεί η αντικειμενική συνάρτηση

$$f(x)=(x_1-3)^2+(x_2-3)^2$$

σύμφωνα με τους περιορισμούς

$$x_1 + x_2 \le 4$$

$$x_1 - 3x_2 = 1$$

$$x_1, x_2 \ge 0$$

<u>Λύση</u>

Το πρόβλημα έχει δύο μεταβλητές σχεδιασμού συνεπώς είναι αρκετά απλό ώστε να μπορεί να επιλυθεί με την βοήθεια των συνθηκών Kuhn-Tucker. Έτσι, θα λύσουμε αρχικά το πρόβλημα χρησιμοποιώντας τη μέθοδο simplex και στη συνέχεια θα επαληθεύσουμε τις συνθήκες Kuhn-Tucker.

Η συνάρτηση κόστους αναπτύσσεται ως εξής:

$$f(\mathbf{x})=x_1^2-6x_1+x_2^2-6x_2+18$$

Καταρχήν θα αγνοήσουμε τον σταθερό όρο 18 της συνάρτησης κόστους* και θα ελαχιστοποιήσουμε την ακόλουθη δευτεροβάθμια συνάρτηση που έχει τη μορφή της εξίσωσης (6.26)

$$f(x) = [-6\ -6]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + 0.5[x_1\ x_2]\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

όπου $c = \begin{bmatrix} -6 \\ -6 \end{bmatrix}$ $\qquad H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\qquad$ b=[4]

* βλ. παρατήρηση στο τέλος της παραγράφου.

$$N=\begin{bmatrix}1\\-3\end{bmatrix} \qquad \text{e}=[1]$$

Χρησιμοποιώντας αυτές τις ποσότητες, το μητρώο **B** και τα διανύσματα **D** και **x** της εξίσωσης (6.37) είναι:

$$\mathbf{B}=\left|\begin{array}{cc|c|cc|c|c|c} 2 & 0 & 1 & -1 & 0 & 0 & 1 & -1 \\ 0 & 2 & 1 & 0 & -1 & 0 & -3 & 3 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 1 & -3 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right|$$

$$\mathbf{D}=[6 \quad 6 \ |4| \ 1]^T$$

$$\mathbf{x}=[x_1 \quad x_2 \ |u_1| \ \zeta_1 \quad \zeta_2 \ |s_1| \ y_1 \quad z_1]^T$$

Παρατηρούμε ότι όλοι οι όροι του διανύσματος **D** είναι μη αρνητικοί, συνεπώς δεν απαιτείται ο πολλαπλασιασμός κάποιας γραμμής του συστήματος με -1.

Ο πίνακας της σελίδας 89 δείχνει τις τέσσερις επαναλήψεις της μεθόδου simplex που οδηγούν στη βέλτιστη λύση. Η τεχνητή αντικειμενική συνάρτηση προκύπτει από τη σχέση (6.42) όπου υπενθυμίζουμε ότι η w_0 είναι το άθροισμα των όρων του διανύσματος **D** και c_j είναι το αντίθετο του αθροίσματος των όρων κάθε στήλης του μητρώου **B**, καθεμιά από τις οποίες αντιστοιχεί σε κάθε μεταβλητή σχεδιασμού. Έτσι, η τεχνητή αντικειμενική συνάρτηση προκύπτει απ’ ευθείας:

$$w=17-4x_1-2x_3+x_4+x_5-x_6+2x_7-2x_8$$

Επίσης, οι συνθήκες συμπληρωματικότητας απαιτούν:

$$X_1X_4=0$$

$$X_2X_5=0$$

$$X_3X_6=0$$

Οι σχέσεις αυτές ικανοποιούνται μόνο αν οι X_1 και X_4, X_2 και X_5, X_3 και X_6 δεν είναι ταυτόχρονα βασικές μεταβλητές* .

Στο τέλος των τεσσάρων επαναλήψεων, όλες οι τεχνητές μεταβλητές είναι μη βασικές και η τεχνητή συνάρτηση κόστους είναι μηδενική. Συνεπώς, προκύπτει η λύση:

$$X_1 = \frac{13}{4},\ X_2 = \frac{3}{4},\ X_3 = \frac{3}{4},\ X_8 = \frac{5}{4}$$

$$X_4=0 \quad , \quad X_5=0 \quad , \quad X_6=0 \quad , \quad X_7=0$$

Χρησιμοποιώντας τον ορισμό του διανύσματος **X** από την εξίσωση (6.37) προκύπτει η λύση του δοσμένου προβλήματος QP.

$$x_1 = \frac{13}{4},\ x_2 = \frac{3}{4},\ u_1 = \frac{3}{4},\ \zeta_1 = 0,\ \zeta_2 = 0$$

$$s_1 = 0,\ y_1 = 0,\ z_1 = \frac{5}{4},\ v_1 = y_1 - z_1 = -\frac{5}{4}$$

$$f\left(\frac{13}{4}, \frac{3}{4}\right) = \frac{41}{8}$$

Τώρα, θα επαληθεύσουμε τις συνθήκες Kuhn-Tucker. Η συνάρτηση Lagrange του προβλήματος είναι -κατά τα γνωστά:

$$L=(x_1-3)^2+(x_2-3)^2+u_1(x_1+x_2-4)+v_1(x_1-3x_2-1)-\zeta_1 x_1-\zeta_2 x_2$$

Δεδομένου ότι η βοηθητική μεταβλητή $s_1=x_6$ του περιορισμού ανισότητας είναι μηδενική, ο περιορισμός πρέπει να ικανοποιείται ως ισότητα. Έτσι, λύνοντας το σύστημα των δύο περιορισμών παίρνουμε $x_1=13/4$ και $x_2=3/4$. Επίσης, αφού $x_1>0$ και $x_2>0$ δεν απαιτούνται οι πολλαπλασιαστές ζ_1 και ζ_2 δηλαδή $\zeta_1=\zeta_2=0$. Παραγωγίζοντας τη συνάρτηση Lagrange παίρνουμε:

$$\frac{\partial L}{\partial x_1} = 2(x_1 - 3) + u_1 + v_1 = 0$$

$$\frac{\partial L}{\partial x_2} = 2(x_2 - 3) + u_1 - 3v_1 = 0$$

* Πρέπει δηλαδή μια από αυτές τις μεταβλητές σε κάθε σχέση να είναι μη βασική ώστε αυτόματα να είναι

Αντικαθιστώντας τις x_1=13/4 , x_2=3/4 , u_1=3/4 και v_1=-5/4 στις προηγούμενες σχέσεις, αυτές ικανοποιούνται. Δηλαδή η λύση που προέκυψε από τη μέθοδο simplex ικανοποιεί τις συνθήκες Kuhn-Tucker. Επίσης, δεδομένου ότι η αντικειμενική συνάρτηση είναι απολύτως κυρτή, η λύση που προκύπτει είναι μοναδική.

Παρατήρηση: Αν προσθέσουμε μία σταθερή ποσότητα στην αντικειμενική συνάρτηση f(**x**), η λύση $\mathbf{x}^*$ του προβλήματος δεν αλλοιώνεται, αν και η τιμή της f(**x**) μεταβάλλεται. Γραφικά, η μορφή της καμπύλης της συνάρτησης δεν αλλάζει, αλλάζει όμως η κλίμακα των αξόνων συντεταγμένων. Στο παρακάτω σχήμα, φαίνεται η επίδραση σε συνάρτηση μίας μεταβλητής. Επίσης, αν πολλαπλασιάσουμε την f(x) με μια θετική σταθερά, η λύση x* επίσης δεν αλλοιώνεται, αλλάζει όμως η τιμή της αντικειμενικής συνάρτησης.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y_1	y_2	y_3	y_4	D
Y_1	2	0	1	-1	0	0	1	-1	1	0	0	0	6
Y_2	0	2	1	0	-1	0	-3	3	0	1	0	0	6
Y_3	1	1	0	0	0	1	0	0	0	0	1	0	4
Y_4	**1**	-3	0	0	0	0	0	0	0	0	0	1	1
	-4	0	-2	1	1	-1	2	-2	0	0	0	0	w-17
Y_1	0	6	1	-1	0	0	1	-1	1	0	0	-2	4
Y_2	0	2	1	0	-1	0	-3	3	0	1	0	0	6
Y_3	0	4	0	0	0	1	0	0	0	0	1	-1	3
X_1	1	-3	0	0	0	0	0	0	0	0	0	1	1
	0	**-12**	-2	1	1	-1	2	-2	0	0	0	4	w-13
X_2	0	1	1/6	-1/6	0	0	1/6	-1/6	1/6	0	0	-1/3	2/3
Y_2	0	0	2/3	1/3	-1	0	-10/3	10/3	-1/3	1	0	2/3	14/3

μηδενική.

Y_3	0	0	-2/3	2/3	0	1	-2/3	**2/3**	-2/3	0	1	1/3	1/3
X_1	1	0	1/2	-1/2	0	0	1/2	-1/2	1/2	0	0	0	3
	0	0	0	-1	1	-1	4	**-4**	2	0	0	0	w-5
X_2	0	1	0	0	0	1/4	0	0	0	0	1/4	-1/4	3/4
Y_2	0	0	**4**	-3	-1	-5	0	0	3	1	-5	-1	3
X_8	0	0	-1	1	0	3/2	-1	1	-1	0	3/2	1/2	1/2
X_1	1	0	0	0	0	3/4	0	0	0	0	3/4	4/4	1 3/4
	0	0	**-4**	3	1	5	0	0	-2	0	6	2	w-3
X_2	0	1	0	0	0	1/4	0	0	0	0	1/4	-1/4	3/4
X_3	0	0	1	-3/4	-1/4	-5/4	0	0	3/4	1/4	-5/4	-1/4	3/4
X_8	0	0	0	1/4	-1/4	1/4	-1	1	-1/4	1/4	1/4	1/4	5/4
X_1	1	0	0	0	0	3/4	0	0	0	0	3/4	1/4	13/4
	0	0	0	0	0	0	0	0	1	1	1	1	w-0

6.5 ΜΕΘΟΔΟΣ ΕΠΙΤΡΕΠΤΩΝ ΔΙΕΥΘΥΝΣΕΩΝ (MFD)

6.5.1 Γενική διατύπωση

Η μέθοδος των επιτρεπτών διευθύνσεων (Method of Feasible Directions - MFD) εφαρμόζεται συνήθως για προβλήματα με ανισοτικούς περιορισμούς, π.χ. να βρεθεί το διάνυσμα των μεταβλητών σχεδιασμού **X** έτσι ώστε να ελαχιστοποιείται η αντικειμενική συνάρτηση

$$Z = F(X)$$

σύμφωνα με τους ανισοτικούς περιορισμούς

$$g_j(X) \leq 0 \qquad j = 1, ..., m$$

Η βασική ιδεά της μεθόδου είναι η κίνηση από μια επιτρεπτή λύση σε μια βελτιωμένη επιτρεπτή λύση. Έτσι, γνωρίζοντας μια επιτρεπτή λύση $\mathbf{X}_q$, προσδιορίζουμε μια επιτρεπτή διεύθυνση $\mathbf{S}_q$

τέτοια ώστε για ένα μικρό βήμα $\alpha > 0$ προς αυτήν τη διεύθυνση, να ικανοποιούνται οι παρακάτω απαιτήσεις:

(i) η νέα λύση $\mathbf{X}_{q+1} = \mathbf{X}_q + \alpha\mathbf{S}_q$ να είναι επιτρεπτή, και

(ii) η νέα τιμή της αντικειμενικής συνάρτησης να είναι μικρότερη από την προηγούμενη, δηλαδή

$$F(\mathbf{X}_{q+1}) < F\,(\mathbf{X}_q)$$

Έχοντας προσδιορίσει την επιτρεπτή διεύθυνση (που είναι ουσιαστικά μια διεύθυνση αναζήτησης λύσης, όπως ορίστηκε στην παράγραφο 5.2.2) εφαρμόζουμε την μέθοδο πολυωνυμικής προσέγγισης (κεφάλαιο 3) για τον υπολογισμό του βήματος α, δηλαδή του μήκους κίνησης επί της διεύθυνσης S_q. Έτσι, παίρνουμε μια νέα επιτρεπτή λύση X_{q+1} και η διαδικασία συνεχίζεται.

Για τον προσδιορισμό του διανύσματος S, πρέπει να ικανοποιούνται οι ακόλουθες συνθήκες:

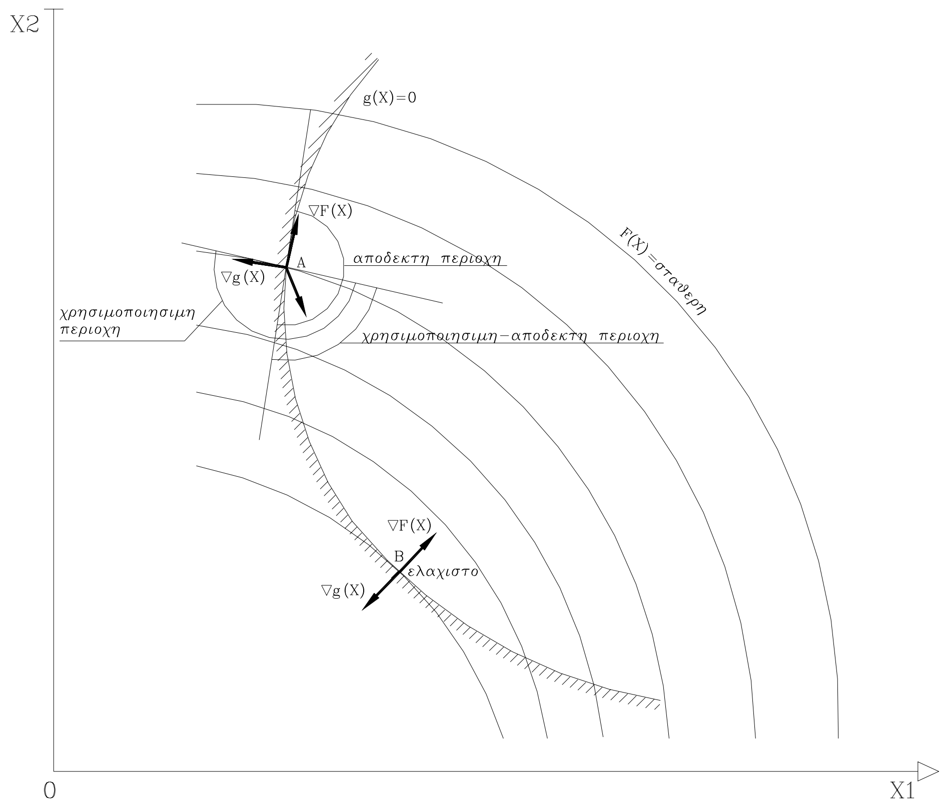
(α) Η διεύθυνση **S** πρέπει να είναι επιτρεπτή (feasible), ώστε για μια μικρή μετακίνηση από την λύση $\mathbf{X}_q$ κατά μήκος της **S** να μην βρισκόμαστε εκτός της επιτρεπτής περιοχής. Αυτή η απαίτηση ικανοποιείται αν ισχύει:

$$\mathbf{S}^T \cdot \nabla g_j < 0 \tag{6.43}$$

Αν ένας περιορισμός είναι γραμμικός ή δεν είναι κυρτός (π.χ. ο περιορισμός g_3 του σχήματος 1) τότε απαιτείται:

$$\mathbf{S}^T \cdot \nabla g_j \leq 0 \tag{6.44}$$

Η ερμηνεία αυτής της συνθήκης είναι ότι το διάνυσμα **S** πρέπει να σχηματίζει αμβλεία γωνία με την κάθετο ∇g_j σε κάθε σημείο του περιορισμού (εκτός από τις περιπτώσεις g_2, g_3, του Σχ. 1 όπου αυτή η γωνία μπορεί να είναι ακριβώς 90°). Όπως φαίνεται και στο σχήμα 1, κάθε διάνυσμα που ικανοποιεί αυτήν τη συνθήκη κείται στην επιτρεπτή περιοχή.



β) Η διεύθυνση S πρέπει να είναι χρησιμοποιήσιμη (usable), ώστε να οδηγεί σε βελτίωση (μείωση) της τιμής της αντικειμενικής συνάρτησης. Αυτή η απαίτηση ικανοποιείται αν:

$$\mathbf{S}^{T} \cdot \nabla F < 0 \qquad (6.45)$$

Ένα διάνυσμα S που ικανοποιεί την παραπάνω ανισότητα είναι εγγυημένο ότι οδηγεί —με κάποιο βήμα α > 0— σε μια νέα λύση $\mathbf{X}_{q+1}$ η οποία μειώνει την τιμή της αντικειμενικής συνάρτησης (βλ. Σχ. 2).

Και εδώ, η ερμηνεία της συνθήκης είναι ότι το διάνυσμα **S** πρέπει να σχηματίζει αμβλεία γωνία με την κάθετο σε κάθε σημείο μιας ισοϋψούς (∇F).

Ένα διάνυσμα **S** που ικανοποιεί και τις δύο συνθήκες (α) και (β), ορίζει μια επιτρεπτή - χρησιμοποιήσιμη διεύθυνση (feasible - usable direction) που ξεκινά από το σημείο $\mathbf{X}_q$. Η μέθοδος MFD οδηγεί σε μια διαδοχή επιτρεπτών σχεδιασμών που βελτιώνουν την τιμή της συνάρτησης κόστους και ορίζονται από την κίνηση επί των διευθύνσεων **S**.

6.5.2 Επιλογή της διεύθυνσης αναζήτησης.

Στην περίπτωση που δεν υπάρχουν ενεργοί περιορισμοί στο σημείο $\mathbf{X}_q$, μπορούμε να επιλέξουμε ως διεύθυνση αναζήτησης της διεύθυνση της μέγιστης μείωσης της συνάρτησης κόστους (steepest descent direction) που είναι το αντίθετο της κλίσης της F, δηλαδή **S** = -∇F. Στη γενικότερη περίπτωση όμως, όπου οι περιορισμοί είναι ενεργοί, οι συνθήκες των σχέσεων (6.43) και (6.45) πρέπει να ικανοποιούνται.

Από τις δυο αυτές σχέσεις παρατηρούμε ότι η μεγαλύτερη μείωση της F επιτυγχάνεται βρίσκοντας το διάνυσμα **S** που ελαχιστοποιεί την ποσότητα της (6.45)* ενώ η (6.44) ικανοποιείται ακριβώς με ισότητα. Σ΄ αυτήν την περίπτωση, η διεύθυνση αναζήτησης είναι ακριβώς εφαπτόμενη στο σύνορο του περιορισμού. Υποθέτοντας ότι ο περιορισμός είναι μη γραμμικός και κυρτός, μια μικρή κίνηση κατά τη διεύθυνση αναζήτησης θα παραβίαζε τον περιορισμό (βλ. Σχ. 3). Αυτό δεν είναι επιθυμητό, γι΄ αυτό πρέπει να βρούμε έναν τρόπο ώστε να "σπρώξουμε" τη διεύθυνση S μακριά από το σύνορο του περιορισμού. Έτσι, εισάγουμε ένα συντελεστή θ (push-off factor) στην (6.44):

$$\mathbf{S}^T \cdot \nabla g_j + \theta \leq 0 \tag{6.46}$$

Ο συντελεστής θ δεν είναι μια γεωμετρική γωνία, αλλά απλά μια μη αρνητική σταθερά. Η ανωτέρω σχέση εξασφαλίζει ότι το γινόμενο $\mathbf{S}^T \cdot \nabla g_j$ είναι αρνητικό, δηλαδή ισχύει η (6.43), οπότε η γωνία του διανύσματος **S** με την κάθετο στον περιορισμό είναι αμβλεία.

Στην πράξη, επιδιώκουμε η σταθερά θ να εξαρτάται από τη διεύθυνση μέγιστης κλίσης της F (steepest descent). Έτσι, έχοντας υπόψιν ότι το γινόμενο $\mathbf{S}^T \cdot \nabla F$ είναι αρνητικό, μπορούμε να τροποποιήσουμε την (6.46) ως εξής:

$$\mathbf{S}^T \cdot \nabla g_j \; - \; [\mathbf{S}^T \cdot \nabla F]\, \theta \leq 0 \tag{6.47}$$

Όπως προαναφέρθηκε, η ποσότητα της (6.45) πρέπει να ελαχιστοποιηθεί (ώστε να πλησιάσουμε τη διεύθυνση μέγιστης μείωσης της F, δηλαδή την - ∇F). Αυτό μπορεί να γίνει ισοδύναμα μεγιστοποιώντας το συντελεστή β της ακόλουθης ανισότητας:

$$\mathbf{S}^T \cdot \nabla F + \beta \leq 0 \tag{6.48}$$

Είναι φανερό, ότι για να γίνει το β μέγιστο, η (6.48) πρέπει να ικανοποιείται με ισότητα, δηλαδή $\beta = -\mathbf{S}^T \cdot \nabla F$. Έτσι, η σχέση (6.47) γίνεται:

$$\mathbf{S}^T \cdot \nabla g_j + \theta\, \beta \leq 0 \tag{6.49}$$

Το σχήμα 3 δείχνει την επιρροή του συντελεστή θ στον προσδιορισμό του διανύσματος **S**. Για θ = 0, το διάνυσμα **S** είναι εφαπτόμενο στο σύνορο του περιορισμού, ενώ όσο το θ μεγαλώνει, το διάνυσμα τείνει να γίνει εφαπτόμενο στην ισοϋψή της αντικειμενικής συνάρτησης. Γενικά,

* Δηλαδή η γωνία ∇F και S να είναι όσο το δυνατόν πιο αμβλεία.

χρησιμοποιούμε την τιμή θ=1.0, για την οποία διχοτομείται (περίπου) η γωνία μεταξύ της ισοϋψούς και του συνόρου του περιορισμού.

Είναι φανερό, ότι αν έχουμε ένα διάνυσμα **S** τέτοιο ώστε ο συντελεστής β να μεγιστοποιείται στη σχέση (6.48) και ο πρώτος όρος της (6.49) να είναι αρνητικός, τότε ικανοποιούνται οι συνθήκες (α) και (β) της παραγράφου 5.5.1, έχουμε δηλαδή μια επιτρεπτή - χρησιμοποιήσιμη διεύθυνση (feasible - usable direction).

Ωστόσο, αν πολλαπλασιάσουμε το διάνυσμα **S** με ένα θετικό αριθμό, οι συνθήκες θα εξακολουθούν να ικανοποιούνται. Συνεπώς, αν θέλουμε να μεγιστοποιήσουμε τη σταθερά β, η λύση θα είναι απεριόριστη αφού το μέγεθος του **S** θα αυξάνει συνεχώς. Για να αποφύγουμε αυτή την κατάσταση, πρέπει να θέσουμε όρια στο διάνυσμα **S**.

Τώρα, μπορούμε να διατυπώσουμε το πρόβλημα εύρεσης διεύθυνσης ως εξής:

Να μεγιστοποιηθεί η σταθερά β

σύμφωνα με τους περιορισμούς

$$\mathbf{S}^{T} \cdot \nabla g_j + \theta_j \, \beta \leq 0, \quad j = 1, ..., J \tag{6.50}$$

$$\mathbf{S}^{T} \cdot \nabla F + \beta \leq 0 \tag{6.51}$$

$$-1 \leq S_i \leq 1, \quad i = 1, ..., n \tag{6.52}$$

όπου J είναι το σύνολο των ενεργών περιορισμών.

Εδώ πρέπει να τονιστεί ξανά ότι όσο μεγαλύτερο γίνεται το β τόσο μικρότερο (περισσότερο αρνητικό) γίνεται το γινόμενο $\mathbf{S}^{T} \cdot \nabla F$ οπότε το διάνυσμα **S** πλησιάζει περισσότερο στη διεύθυνση $-\nabla F$. Δηλαδή, σκοπός της μεγιστοποιήσης του β είναι η εύρεση μιας διεύθυνσης που βρίσκεται κοντά στη διεύθυνση της μέγιστης κλίσης της F^{*}.

Το πρόβλημα εύρεσης διεύθυνσης αποτελεί ένα πρόβλημα γραμμικού προγραμματισμού (LP). Ορίζοντας τις νέες μεταβλητές **S´** ως

$$S'_i = S_i + 1$$

και εισάγοντας βοηθητικές μεταβλητές (slack variables), το πρόβλημα διατυπώνεται στην κανονική μορφή και έτσι μπορεί να εφαρμοστεί η μέθοδος simplex.

6.5.3 Επιλογή του βήματος α

Υποθέτοντας ότι μια επιτρεπτή - χρησιμοποιήσιμη διεύθυνση **S** έχει βρεθεί, το πρόβλημα είναι τώρα η εκλογή του βήματος α. Στόχος είναι ο υπολογισμός του α έτσι ώστε η νέα λύση $\mathbf{X}_{q+1}$ να είναι επιτρεπτή και η συνάρτηση κόστους F να ελαχιστοποιηθεί, χωρίς να απαιτείται μεγάλος αριθμός επαναλήψεων της διαδικασίας. Οι δυο περιπτώσεις που θα εξετάσουμε φαίνονται στο σχήμα 4 και είναι:

(α) Το σημείο $\mathbf{X}_{q+1}$ βρίσκεται στο σύνορο της επιτρεπτής περιοχής.

Για να βρούμε το $\mathbf{X}_{q+1}$ στο σύνορο, εκλέγουμε βήμα όσο το δυνατόν μεγαλύτερο χωρίς να παραβιάζονται οι περιορισμοί. Γενικά, θεωρούμε ένα δοκιμαστικό βήμα α και ελέγχουμε τους περιορισμούς. Αν αυτοί παραβιάζονται, μειώνουμε το α και ελέγχουμε ξανά. Αν το νέο σημείο βρίσκεται εντός της επιτρεπτής περιοχής, αυξάνουμε το α. Τέλος, αν το σημείο είναι επί του συνόρου, δηλαδή τουλάχιστον ένας από τους περιορισμούς είναι ενεργός, αναζητούμε μια νέα διεύθυνση **S**.

Πρέπει να σημειωθεί, ωστόσο, ότι ακόμα και αν το ελάχιστο της F κατά μήκος του **S** βρίσκεται στο σύνορο, είναι καλύτερο να βρίσκουμε ένα σημείο μέσα στην επιτρεπτή περιοχή και μετά να κάνουμε ένα βήμα επί της διεύθυνσης $-\nabla F$.

Ένα πρόβλημα που τίθεται είναι το πώς θα προσδιορίσουμε αν ένας περιορισμός g_j είναι ενεργός ή όχι. Αν απαιτούνται πολύ μικρές τιμές του g_j, είναι πιθανό να χρειάζονται αρκετές επαναλήψεις της μεθόδου. Έτσι, ορίζουμε κάποια περιθώρια για να κάνουμε την αναζήτηση πιο αποτελεσματική. Για παράδειγμα, μπορούμε να ορίσουμε ότι ο περιορισμός g_j είναι ενεργός αν

$$-\varepsilon_g \le g_j \le 0, \qquad (6.53)$$

ή

$$|g_j| \le \varepsilon_g \qquad (6.54)$$

όπου ε_g είναι η ανοχή του περιορισμού. Η σχέση (6.53) απαιτεί το σημείο να βρίσκεται απαραίτητα στην επιτρεπτή περιοχή, ενώ η (6.54) επιτρέπει μια μικρή παραβίαση του περιορισμού. Τα κριτήρια αυτά δίνουν κάποια ευρύτητα στα όρια του περιορισμού και έτσι διευκολύνουν την εύρεση ενός σημείου που ικανοποιεί τον περιορισμό.

(β) Το σημείο $\mathbf{X}_{q+1}$ βρίσκεται εντός της επιτρεπτής περιοχής.

[*] Με αδρές γραμμές, μπορούμε να πούμε ότι η (5.50) εξασφαλίζει ότι η S βρίσκεται επί της επιτρεπτής περιοχής ενώ η (5.51) εξασφαλίζει ότι η S μειώνει την τιμή της αντικειμενικής συνάρτησης όσο το δυνατόν περισσότερο.

Η δεύτερη πιθανή περίπτωση του σχήματος 4 είναι το ελάχιστο της F να βρίσκεται κατά μήκος του **S** μέσα στην επιτρεπτή περιοχή. Στην περίπτωση αυτή, η τιμή του βήματος α* προσδιορίζεται χρησιμοποιώντας την μέθοδο πολυωνυμικής προσέγγισης (που αποτελεί την μέθοδο μονοδιάστατης αναζήτησης που αναπτύχθηκε στο κεφάλαιο 3).

Παράδειγμα

Στο παράδειγμα αυτό θα εφαρμόσουμε την μέθοδο επιτρεπτών διευθύνσεων (MFD) για να ελαχιστοποιήσουμε τη συνάρτηση κόστους

$$Z = X_1^2 + X_2^2$$

σύμφωνα με τους περιορισμούς

$$g_1 = X_1^2 / 20 - X_2 + 1 \leq 0$$

$$g_2 = X_2^2 / 20 - X_1 + 1 \leq 0$$

Λύση

Υποθέτουμε ότι η αρχική επιτρεπτή λύση είναι $X_1^T = (6, 3)$. Δεδομένου ότι σ΄ αυτήν τη θέση δεν υπάρχει ενεργός περιορισμός (βλ. Σχ. 6), επιλέγουμε ως διεύθυνση αναζήτησης τη διεύθυνση μέγιστης κλίσης της F, δηλαδή:

$$S_1 = -\nabla F = -\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = -\begin{bmatrix} 12 \\ 6 \end{bmatrix}$$

Η νέα λύση είναι:

$$x_2 = x_1 + \alpha S_1 = \begin{bmatrix} 6 \\ 3 \end{bmatrix} - \alpha \begin{bmatrix} 12 \\ 6 \end{bmatrix}$$

Ελαχιστοποιώντας την F(α) κατά μήκος της ευθείας που ορίζει η παραπάνω σχέση, βρίσκουμε το σημείο $\mathbf{X_2}^T = (2.764, 1.382)$ επί του συνόρου. Για να προσδιορίσουμε τη διεύθυνση $\mathbf{S_2}$, λύνουμε το ακόλουθο πρόβλημα γραμμικού προγραμματισμού, που διατυπώνεται στο σημείο $\mathbf{X_2}$:

Να μεγιστοποιηθεί η ποσότητα β

σύμφωνα με τους περιορισμούς:

$$[S_1, S_2]\begin{bmatrix}0.276\\-1\end{bmatrix} + \beta \leq 0, \quad \text{(ενεργός περιορισμός: } g_1\text{)}$$

$$[S_1, S_2]\begin{bmatrix}5.528\\2.764\end{bmatrix} + \beta \leq 0$$

$$-\begin{bmatrix}1\\1\end{bmatrix} \leq \begin{bmatrix}S_1\\S_2\end{bmatrix} \leq \begin{bmatrix}1\\1\end{bmatrix}$$

Η λύση του προβλήματος είναι $\mathbf{S_2}^T = [S_1, S_2] = [-1.0, 1.0]$ και β = 1.276. Το σημείο $\mathbf{X_3}$ προκύπτει ομοίως, ως εξής:

$$x_3 = x_2 + \alpha S_2 = \begin{bmatrix}2.764\\1.382\end{bmatrix} + \alpha\begin{bmatrix}-1.0\\1.0\end{bmatrix}$$

Η τιμή του βήματος α* υπολογίζεται ελαχιστοποιώντας την αντικειμενική συνάρτηση κατά μήκος αυτής της ευθείας. Το αποτέλεσμα είναι $\mathbf{X_3}^T = [2.073, 2.073]$. Από το σχήμα παρατηρούμε ότι στο $\mathbf{X_3}$ δεν υπάρχει ενεργός περιορισμός, συνεπώς η νέα διεύθυνση αναζήτησης είναι:

$$S_3 = -\nabla F = -\begin{bmatrix}4.146\\4.146\end{bmatrix}$$

και η νέα λύση είναι:

$$x_4 = x_3 + \alpha S_3 = \begin{bmatrix}2.073\\2.073\end{bmatrix} - \alpha\begin{bmatrix}4.146\\4.146\end{bmatrix}$$

Ομοίως, ελαχιστοποιώντας την αντικειμενική συνάρτηση κατά μήκος του S_3, παίρνουμε:

$$x_4 = \begin{bmatrix}1.056\\1.056\end{bmatrix}$$

που αποτελεί τη βέλτιστη λύση.

6.6 Προσέγγιση του προβλήματος κατά Zoutendijk

Η διατύπωση του προβλήματος με τις εξισώσεις (6.50) έως (6.52) έχει ένα σημαντικό μειονέκτημα. Το μειονέκτημα έγκειται στα όρια που τίθενται για τη διεύθυνση **S**, τα οποία βάσει της σχέσης (6.52) ορίζουν το τετράγωνο του σχήματος 7α. Από τις σχέσεις (6.50) και (6.51) φαίνεται ότι μεγιστοποιώντας το συντελεστή β, οι πρώτοι όροι γίνονται συνεχώς μεγαλύτεροι και πιο αρνητικοί. Ένας τρόπος να γίνει αυτό είναι αυξάνοντας το μέγεθος του διανύσματος **S** όσο το δυνατόν περισσότερο. Όμως, από το σχήμα 7α φαίνεται ότι το μεγαλύτερο μέγεθος του **S** προκύπτει όταν το **S** στρέφεται προς κάποια από τις γωνίες του τετραγώνου. Δηλαδή, η επιλογή των ορίων βάσει της (6.52) επηρεάζει τον καθορισμό της διεύθυνσης αναζήτησης. Αυτό όμως δεν είναι επιθυμητό αφού οι συντελεστές θ_j έχουν εισαχθεί ώστε να ελέγχουν τη διεύθυνση αναζήτησης. Ένας τρόπος να αποφύγουμε την εξάρτηση της διεύθυνσης αναζήτησης από τα όρια που τίθενται, φαίνεται στο σχήμα 6 με τον κύκλο που προκύπτει φράσσοντας την **S** με την Ευκλείδια νόρμα:

$$\mathbf{S} \cdot \mathbf{S} \leq 1 \tag{6.55}$$

Δυστυχώς, χρησιμοποιώντας την (6.55) με τις (6.50) και (6.51) προκύπτει ένα πρόβλημα που είναι γραμμικό εκτός από το δευτεροβάθμιο περιορισμό (6.55), συνεπώς η μέθοδος γραμμικού προγραμματισμού δεν μπορεί να εφαρμοστεί άμεσα. Το πρόβλημα αντιμετωπίστηκε από τον Zoutendijk με τη βοήθεια των συνθηκών Kuhn - Tucker.

Καταρχήν, γράφουμε τις εξισώσεις (6.50), (5,51) και (6.55) σε μητρωική μορφή. Ορίζουμε τα μητρώα **p**, **y** και **A** ως εξής:

$$y = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \\ \beta \end{bmatrix} \qquad p = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \qquad A = \begin{bmatrix} \nabla^T g_1(x)\theta_1 \\ \nabla^T g_2(x)\theta_2 \\ \vdots \\ \nabla^T g_j(x)\theta_j \\ \nabla^T F(x) 1 \end{bmatrix}$$

οπότε, το πρόβλημα εύρεσης της διεύθυνσης αναζήτησης γίνεται:

Να μεγιστοποιηθεί το γινόμενο $\mathbf{p} \cdot \mathbf{y}$

σύμφωνα με τους περιορισμούς

$$\mathbf{A} \cdot \mathbf{y} \leq 0 \tag{6.56}$$

$$\mathbf{y} \cdot \mathbf{y} \leq 1 \tag{6.57}$$

Η σχέση (6.57) είναι ισοδύναμη με την (6.55) εκτός από έναν συντελεστή:

$$\mathbf{y} \cdot \mathbf{y} = \mathbf{S} \cdot \mathbf{S} + \beta^2 \leq 1$$

Έτσι, όταν επιλυθεί το πρόβλημα, το διάνυσμα **S** θα κανονικοποιηθεί σε μια τιμή διαφορετική της μονάδας (δηλαδή $\mathbf{S} \cdot \mathbf{S} \leq 1 - \beta^2$), αλλά η μορφή της κανονικοποίησης είναι η ίδια.

Οι συνθήκες Kuhn - Tucker γράφονται:

$$\mathbf{p} - \mathbf{A}^T\mathbf{u} - \mathbf{\Psi y} = 0 \tag{6.58}$$

$$\mathbf{u}^T \mathbf{A y} = 0 \tag{6.59}$$

$$\mathbf{u} \geq 0 \tag{6.60}$$

$$\mathbf{\Psi} \geq 0 \tag{6.61}$$

όπου **u** και **Ψ** είναι οι πολλαπλασιαστές Lagrange.

Θέτουμε $\mathbf{z} = -\mathbf{Ay}$

οπότε, από την (6.56) είναι $\mathbf{z} \geq 0$ και οι παραπάνω σχέσεις γίνονται:

$$\mathbf{p} = \mathbf{A}^T\mathbf{u} + \mathbf{\Psi y} \quad \mathbf{u} \geq 0 \quad \mathbf{\Psi} \geq 0 \quad \mathbf{u} \cdot \mathbf{z} = 0$$

Πολλαπλασιάζοντας με **- A** παίρνουμε:

$$-\mathbf{Ap} = -\mathbf{A}\,\mathbf{A}^T\mathbf{u} - \mathbf{\Psi A y}$$

Επίσης θέτουμε:

$$\mathbf{V} = -\mathbf{\Psi A y} = \mathbf{\Psi z} \text{ και } \mathbf{c} = -\mathbf{Ap}$$

οπότε τελικά οι σχέσεις (6.58) έως (6.61) γίνονται:

$$-\mathbf{A}\,\mathbf{A}^T\mathbf{u} + \mathbf{V} = \mathbf{c} \tag{6.62}$$

$$\mathbf{u} \geq 0 \quad \mathbf{V} \geq 0 \quad \mathbf{u} \cdot \mathbf{V} = 0 \tag{6.63}$$

Αν θέσουμε $\mathbf{B} = -\mathbf{A}\,\mathbf{A}^T$, η σχέση (6.62) γίνεται

$$[BI]\begin{Bmatrix} u \\ v \end{Bmatrix} = c \tag{6.64}$$

Το πρόβλημα μπορεί να λυθεί πλέον με τη μέθοδο γραμμικού προγραμματισμού, οπότε προκύπτει τελικά το διάνυσμα **u**, το οποίο χρησιμοποιείται στη σχέση (6.58) για την εύρεση του y:

$$\mathbf{\Psi y} = \mathbf{p} - \mathbf{A}^T\mathbf{u} \qquad (6.65)$$

Η τιμή του **Ψ** δεν είναι σημαντική, συνεπώς λαμβάνεται ως μονάδα. Τα n πρώτα στοιχεία του **y** αποτελούν το ζητούμενο διάνυσμα της διεύθυνσης αναζήτησης **S**.

6.7 Παρατηρήσεις επί της μεθόδου μέγιστης κλίσεως

6.7.1 Εισαγωγή

Η μέθοδος μέγιστης κλίσης που εφαρμόστηκε στην μέθοδο MFD, ανήκει στην κατηγορία των μεθόδων πρώτης τάξεως (first-order methods) οι οποίες χρησιμοποιούν τις ευαισθησίες ή κλίσεις των συναρτήσεων (gradient information) για την εύρεση της διεύθυνσης αναζήτησης της βέλτιστης λύσης. Όπως θα δούμε στο Κεφάλαιο 7, ο υπολογισμός των κλίσεων δεν είναι πάντοτε εύκολη υπόθεση, μιας και απαιτούν μια σύνθετη αναλυτική διαδικασία. Επιπλέον οι μέθοδοι πρώτης τάξεως δεν αποδίδουν καλά για συναρτήσεις με μη συνεχείς πρώτες παραγώγους. Παρ' όλα αυτά οι μέθοδοι πρώτης τάξεως αποδίδουν καλύτερα από τις πιθανοτικές μεθόδους μηδενικής τάξεως και τις περίπλοκες μεθόδους δευτέρας τάξεως.

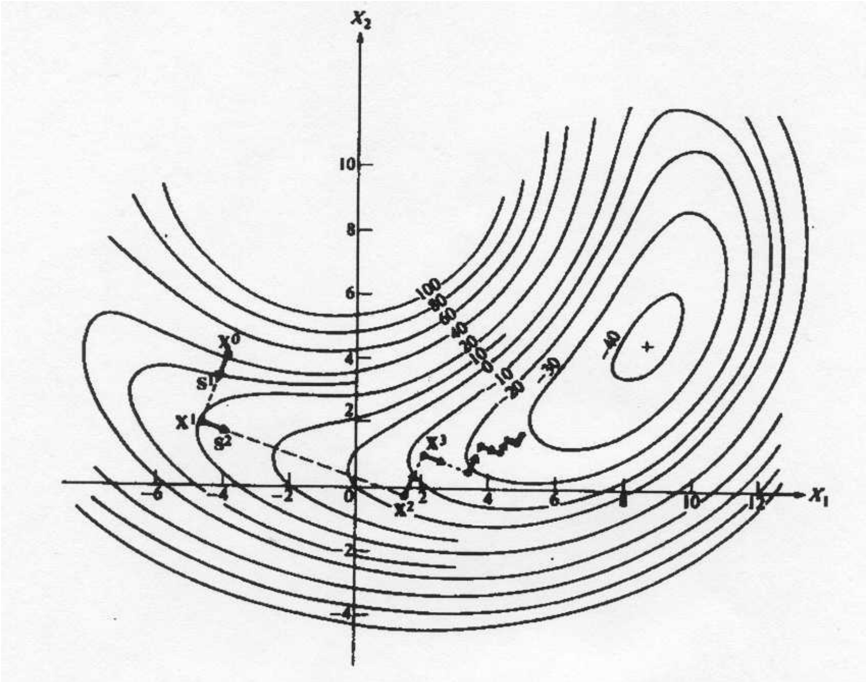
6.7.2 Η μέθοδος μέγιστης κλίσεως

Η μέθοδος μέγιστης κλίσης είναι η πιο γνωστή και παράλληλα η λιγότερο αποδοτική από τις μεθόδους πρώτης τάξεως. Ωστόσο είναι σημαντική διότι συνήθως χρησιμοποιήται για να δώσει την αρχική διεύθυνση αναζήτησης βάσει της οποίας εφαρμόζονται άλλες μέθοδοι.

Ως γνωστόν, στην μέθοδο μέγιστης κλίσης η διεύθυνση αναζήτησης είναι η αντίθετη της κλίσης της αντικειμενικής συνάρτησης. Δηλαδή, για την επανάληψη q:

$$S^q = -\nabla F(X^q)$$

Το διάνυσμα **S** χρησιμοποιείται στη γνωστή σχέση (2.1) με την οποία πραγματοποιείται η μονοδιάστατη αναζήτηση. Στο παρακάτω σχήμα φαίνεται η εφαρμογή της μεθόδου μέγιστης κλίσης σ' ένα πρόβλημα.



Όπως φαίνεται, ο ρυθμός σύγκλισης της μεθόδου είναι μικρός. Αυτό οφείλεται στο ότι η μέθοδος μέγιστης κλίσης δεν χρησιμοποιεί πληροφορίες από προηγούμενες επαναλήψεις ώστε να επιταχυνθεί η σύγκλιση. Έτσι, η μέθοδος μέγιστης κλίσης δεν συνιστάται γενικά, αν και χρησιμοποιείται συχνά για την εύρεση της αρχικής διεύθυνσης αναζήτησης πριν την εφαρμογή πιο ισχυρών αλγορίθμων.

6.7.3 Η μέθοδος συζυγών διευθύνσεων

Η μέθοδος συζυγών διευθύνσεων (conjugate direction method) των Fletcher - Reeves, απαιτεί μόνο μια μικρή μετατροπή της μεθόδου μέγιστης κλίσεως η οποία όμως βελτιώνει δραστικά το ρυθμό σύγκλισής της. Η βασική θεώρηση της μεθόδου είναι ότι επιλέγουμε διευθύνσεις αναζήτησης οι οποίες είναι συζυγείς* . Αυτό επιτυγχάνεται ορίζοντας ως αρχική διεύθυνση αναζήτησης το διάνυσμα μέγιστης κλίσης και για τις επόμενες επαναλήψεις μια συζυγή διεύθυνση που είναι:

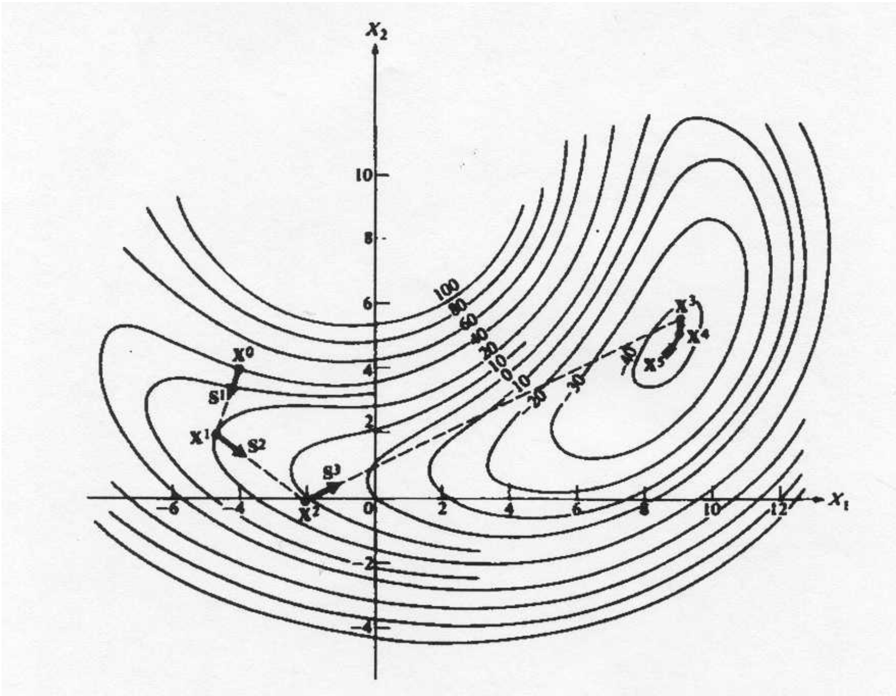
$$S^q = -\nabla F(X^q) + \beta_q S^{q-1}$$

όπου η παράμετρος β_q ορίζεται ως:

* Δυο διευθύνσεις S^i και S^j είναι συζυγείς όταν: $(S^i)^T H S^j = 0$. Μια μαθηματική ανάλυση του θέματος δεν είναι επί του παρόντος, αυτό που πρέπει να ξέρουμε όμως είναι ότι μια δευτεροβάθμια συνάρτηση μπορεί ελαχιστοποιηθεί με n ή λιγότερες συζηγείς διευθύνσεις όπου n είναι ο αριθμός των μεταβλητών σχεδιασμού.

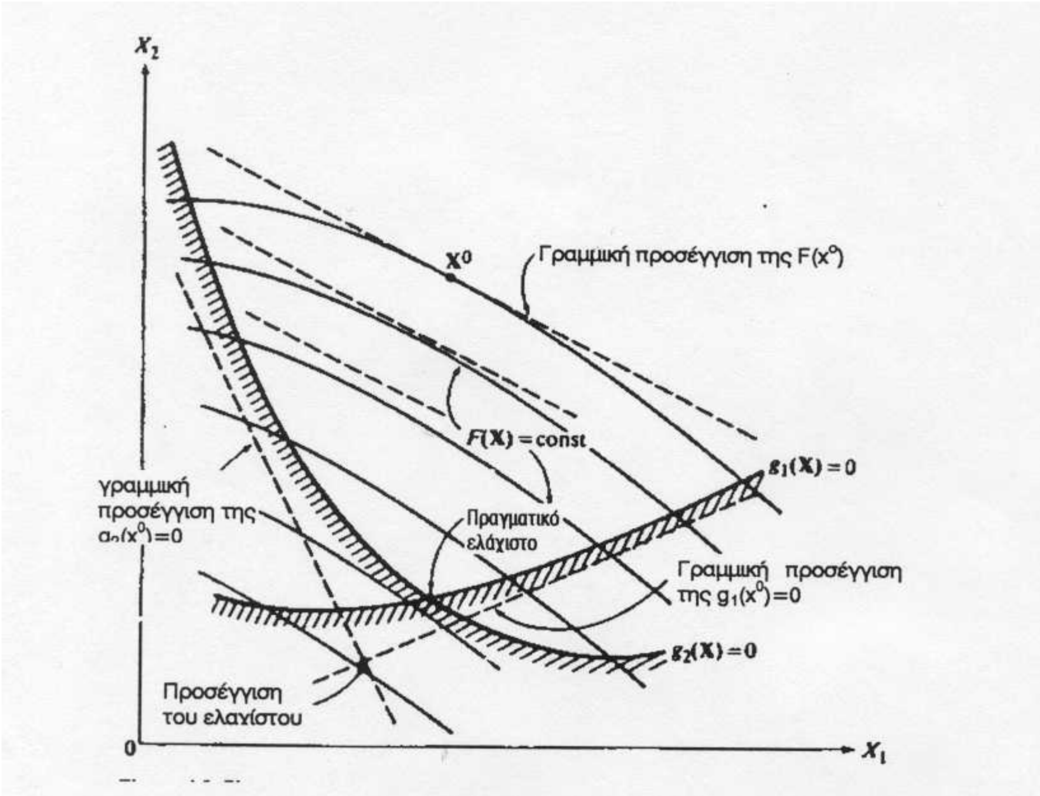
$$\beta^{q} = \frac{\left|\nabla F(X^{q})\right|^{2}}{\left|\nabla F(X^{q-1})\right|^{2}}$$

Στο σχήμα φαίνεται η εφαρμογή της μεθόδου, όπου είναι προφανής η βελτίωση του ρυθμού σύγκλισης.



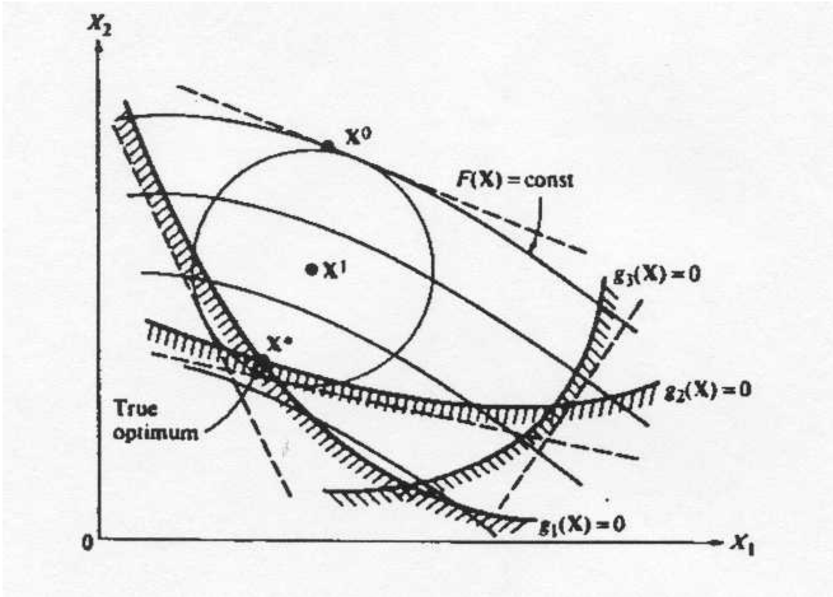
6.8 Η ΜΕΘΟΔΟΣ ΤΩΝ ΚΕΝΤΡΩΝ

Στην παράγραφο 5.3 είδαμε ότι η μέθοδος SLP (Sequential Linear Programming) οδηγεί σε μια διαδοχή βελτιωμένων αλλά μη επιτρεπτών σχεδιασμών, στη συνήθη περίπτωση κυρτού χώρου λύσεων. Αυτό φαίνεται καθαρά στο σχήμα 7β παρατηρώντας ότι η τομή των γραμμικών προσεγγίσεων των περιορισμών βρίσκεται πάντοτε στην εξωτερική πλευρά των αρχικών, μη γραμμικών (και κυρτών) περιορισμών.

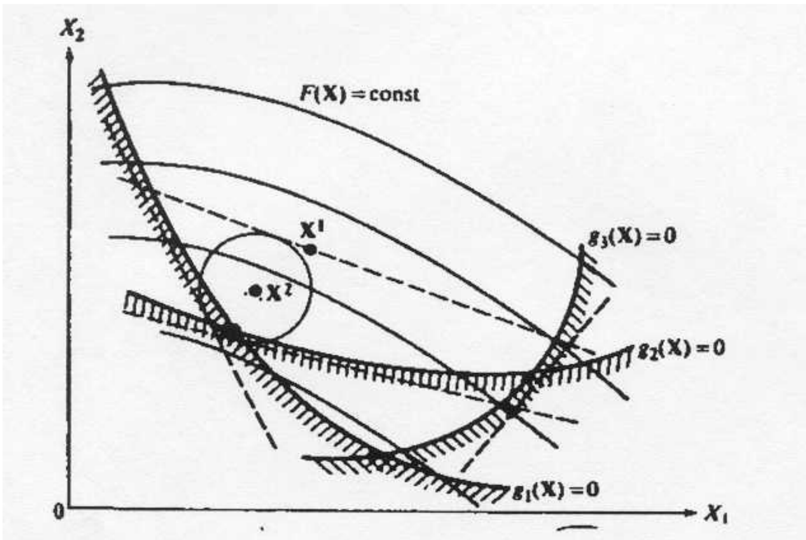


Αλλά, από την άποψη του μηχανικού, είναι συχνά επιθυμητό να προσεγγίζουμε το βέλτιστο σχεδιασμό μέσω μιας διαδοχής επιτρεπτών σχεδιασμών. Επίσης, είναι επιθυμητό αυτή η διαδοχή των λύσεων να ακολουθεί μια πορεία προς το "κέντρο" του χώρου λύσεων. Μ΄ αυτόν τον τρόπο, αν η διαδικασία βελτιστοποίησης τερματιστεί προσωρινά, ή για κάποιο λόγο ο τελικός σχεδιασμός κριθεί μη αποδεκτός, θα έχει ήδη καταγραφεί μια σειρά επιτρεπτών λύσεων από τις οποίες μπορεί να επιλεγεί η κατάλληλη. Τις δυνατότητες αυτές προσφέρει η μέθοδος των κέντρων (method of centers) η οποία είναι ουσιαστικά μια μέθοδος SLP.

Ας θεωρήσουμε το πρόβλημα του σχήματος 8, το οποίο απεικονίζει ένα μη γραμμικό πρόβλημα με δυο μεταβλητές σχεδιασμού και τρεις περιορισμούς. Οι διακεκομμένες γραμμές είναι οι γραμμικές προσεγγίσεις του προβλήματος στο σημείο **X** °. Η βασική έννοια της μεθόδου είναι ότι αναζητούμε τον μεγαλύτερο κύκλο που προσαρμόζεται στο γραμμικοποιημένο χώρο και στη συνέχεια προχωρούμε στο κέντρο αυτού του κύκλου. Αυτό φαίνεται στο σχήμα 8 όπου ο κύκλος εφάπτεται στη γραμμή της αντικειμενικής συνάρτησης και στα σύνορα των περιορισμών g_1 (**X**) και g_2 (**X**).

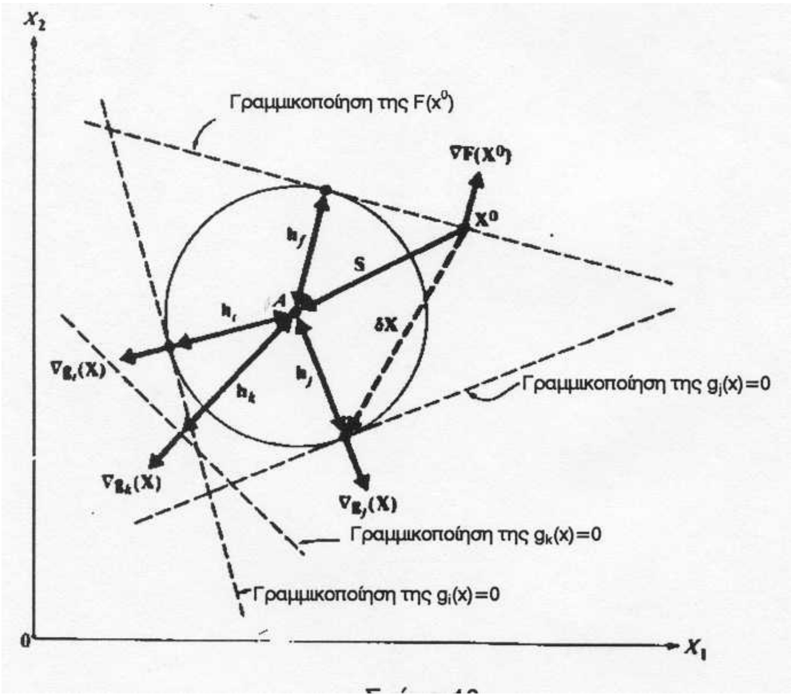


Τώρα, κινούμαστε προς τη λύση $\mathbf{X}^1$ στο κέντρο του κύκλου, και το πρόβλημα γραμμικοποιείται εκ νέου. Η γραμμικοποιημένη ισοϋψής της αντικειμενικής συνάρτησης στο $\mathbf{X}^1$ αντιμετωπίζεται τώρα ως σύνορο, όπως φαίνεται στο σχήμα 9. Το κέντρο του νέου κύκλου αποτελεί τη νέα λύση $\mathbf{X}^2$. Η ίδια διαδικασία συνεχίζεται μέχρι να επιτευχθεί σύγκλιση της λύσης σε ένα καθορισμένο διάστημα.



Ο αλγόριθμος της μεθόδου τροποποιείται ελαφρώς όταν η αρχική λύση είναι μη επιτρεπτή. Στην περίπτωση αυτή, η γραμμικοποιημένη ισοϋψής της αντικειμενικής συνάρτησης δεν αντιμετωπίζεται ως σύνορο, αλλά αγνοείται και αναζητείται ο μεγαλύτερος κύκλος εντός της επιτρεπτής περιοχής. Κατά τα γνωστά, η νέα λύση είναι το κέντρο αυτού του κύκλου. Επίσης, προκειμένου να αποφευχθεί η εύρεση απεριόριστης λύσης, τίθενται όρια, όπως στην μέθοδο SLP.

Το ερώτημα που τίθεται τώρα είναι πώς προσδιορίζεται ο μεγαλύτερος κύκλος που προσαρμόζεται στο χώρο που ορίζουν οι γραμμικοποιημένοι περιορισμοί. Η απάντηση σ΄ αυτό το ερώτημα θα γίνει με τη βοήθεια του σχήματος 10.



Ας θεωρήσουμε ένα σημείο Α στην επιτρεπτή περιοχή. Η απόσταση h_j είναι η απόσταση από τον περιορισμό g_j (**X**) και η απόσταση h_f είναι απόσταση από τη γραμμικοποιημένη ισοϋψή της αντικειμενικής συνάρτησης. Το διάνυσμα της διεύθυνσης αναζήτησης **S** από το **X** ° στο σημείο Α είναι το ζητούμενο μέγεθος —αν Α είναι το κέντρο του μεγαλύτερου κύκλου που προσαρμόζεται στο χώρο των λύσεων. Από τη γεωμετρία του σχήματος 10 προκύπτουν:

$$h_f = -\frac{\nabla F(x^0)\cdot S}{\left|\nabla F(x^0)\right|} \tag{6.66}$$

και $$h_j = \frac{\nabla g_j(x)\cdot\ \ x}{\left|\nabla g_j(x)\right|} - \frac{\nabla g_j(x)\cdot S}{\left|\nabla g_j(x)\right|} \tag{6.67}$$

Αλλά, παρατηρώντας ότι **δX** είναι το διάνυσμα κίνησης από τη λύση **X**° στο σύνορο του περιορισμού g_j (**X**) = 0, προκύπτει ότι:

g_j (**X**) ≈ g_j (**X** °) + $\nabla \mathbf{g}_j$ (**X** °) · **δX** = 0 ⇒

$\nabla \mathbf{g}_j$ (**X** °) · **δX** = **-** g_j (**X** °)

Έτσι, η σχέση (6.67) γίνεται:

$$h_j = \frac{g_j(x^0) + g_j(x^0) \cdot S}{\left|\nabla g_j(x^0)\right|} \tag{6.68}$$

Τώρα, ας θεωρήσουμε r την ακτίνα του μεγαλύτερου κύκλου που προσαρμόζεται στο χώρο λύσεων χωρίς να παραβιάζονται οι περιορισμοί ή να αυξάνεται η τιμή της αντικειμενικής συνάρτησης. Προφανώς, θα πρέπει:

$$h_j - r \geq 0, \quad j = 1, m \tag{6.69}$$

$$h_f - r \geq 0 \tag{6.70}$$

Χρησιμοποιώντας τις σχέσεις (6.66), (6.68), (6.69) και (6.70) και γνωρίζοντας ότι επιθυμούμε να μεγιστοποιήσουμε την ακτίνα r, μπορούμε να ορίσουμε το ακόλουθο γραμμικό πρόβλημα:

Να μεγιστοποιηθεί η ακτίνα r

σύμφωνα με τους περιορισμούς

$$\nabla\mathbf{F}\,(\mathbf{X}^{\circ}) \cdot \mathbf{S} + \left|\nabla\mathbf{F}\,(\mathbf{X}^{\circ})\right|\, r \leq 0$$

$$\nabla\mathbf{g}_j\,(\mathbf{X}^{\circ}) \cdot \mathbf{S} + \left|\nabla\mathbf{g}_j\,(\mathbf{X}^{\circ})\right|\, r \leq -\,g_j\,(\mathbf{X}^{\circ}) \quad j = 1, m$$

Το πρόβλημα αυτό αποτελεί ένα πρόβλημα γραμμικού προγραμματισμού στο οποίο η τιμή r και οι όροι του **S** αποτελούν τις μεταβλητές σχεδιασμού.

Επίσης, αν η αντικειμενική συνάρτηση εξασφαλίζει τις απαιτήσεις για την ύπαρξη καθολικού ελαχίστου που αφορά προβλήματα χωρίς περιορισμούς, δηλαδή αν το μητρώο Hessian της F είναι θετικά ορισμένο για όλους τους πιθανούς σχεδιασμούς, τότε λέμε ότι η αντικειμενική συνάρτηση είναι κυρτή.

Αν λοιπόν η αντικειμενική συνάρτηση και οι περιορισμοί είναι κυρτοί, τότε οι αναγκαίες συνθήκες Kuhn-Tucker είναι και ικανές. Στην περίπτωση αυτή, το ελάχιστο ενός προβλήματος είναι ταυτόχρονα και καθολικό ελάχιστο.

6.9 ΧΡΗΣΗ ΥΠΟΡΟΥΤΙΝΩΝ IMSL στην γλώσσα Fortran

Στους παρακάτω πίνακες παρουσιάζονται συνοπτικά οι υπορουτίνες που διατίθενται από την βιβλιοθήκη IMSL (International Mathematical and Statistical Library) για την επίλυση προβλημάτων βελτιστοποίησης.

Ελαχιστοποίηση χωρίς Περιορισμούς (Unconstrained Minimization)

Συναρτήσεις μίας μεταβλητής (Univariate Function)

Χρησιμοποιώντας μόνο τιμές συναρτήσεων	UVMIF
Χρησιμοποιώντας τιμές συναρτήσεων και πρώτων παραγώγων	UVMID
Μη-λείες συναρτήσεις	UVMGS

Συναρτήσεις πολλών μεταβλητών Multivariate Function

Χρησιμοποιώντας κλίσεις με πεπερασμένες διαφορές	UMINF
Χρησιμοποιώντας αναλυτικές κλίσεις	UMING
Χρησιμοποιώντας προσέγγιση πεπερασμένων διαφορών για την Hessian	UMIDH
Χρησιμοποιώντας αναλυτικό Hessian	UMIAH
Χρησιμοποιώντας συζυγείς κατευθύνσεις με κλίσεις εκφρασμένες με πεπερασμένες διαφορές	UMCGF
Χρησιμοποιώντας συζυγείς κατευθύνσεις με αναλυτικές κλίσεις	UMCGG
Μη λείες συναρτήσεις	UMPOL

Μη γραμμικά ελάχιστα τετράγωνα (Nonlinear Least Squares)

Χρησιμοποιώντας Ιακωβιανή με πεπερασμένες διαφορές	UNLSF
Χρησιμοποιώντας αναλυτική Ιακωβιανή	UNLSJ

Ελαχιστοποίηση με απλά όρια (Minimization with Simple Bounds)

Χρησιμοποιώντας πεπερασμένες διαφορές για τις κλίσεις	BCONF
Χρησιμοποιώντας αναλυτικές κλίσεις	BCONG
Χρησιμοποιώντας πεπερασμένες διαφορές για την Hessian	BCODH
Χρησιμοποιώντας αναλυτική έκφραση για την Hessian	BCOAH
Μη λείες συναρτήσεις	BCPOL
Μη γραμμικά ελάχιστα τετράγωνα χρησιμοποιώντας πεπερασμένες διαφορές για την Ιακωβιανή	BCLSF
Μη γραμμικά ελάχιστα τετράγωνα χρησιμοποιώντας αναλυτική Ιακωβιανή	BCLSJ

Γραμμικά Περιορισμένη Ελαχιστοποίηση

Linearly Constrained Minimization

Πυκνός Γραμμικός Προγραμματισμός	DLPRS
Δευτεροβάθμιος Προγραμματισμός	QPROG
Γενική αντικειμενική συνάρτηση με πεπερασμένες διαφορές για τις κλίσεις	LCONF
Γενική αντικειμενική συνάρτηση με αναλυτικές για τις κλίσεις	LCONG

Μη γραμμικά Περιορισμένη Ελαχιστοποίηση

Nonlinearly Constrained Minimization

Χρησιμοποιώντας πεπερασμένες διαφορές για τις κλίσεις	NCONF
Χρησιμοποιώντας αναλυτικές κλίσεις	NCONG

Βοηθητικές Ρουτίνες - Service Routines

Κεντρικές διαφορές για τις κλίσεις	CDGRD
Έμπροσθιες διαφορές για τις κλίσεις	FDGRD
Εμπρόσθιες διαφορές για την Hessian	FDHES
Εμπρόσθιες διαφορές για την Hessian χρησιμοποιώντας αναλυτικές κλίσεις	GDHES
Εμπρόσθιες διαφορές για την Ιακωβιανή	FDJAC
Check user-supplied gradient	CHGRD
Check user-supplied Hessian	CHHES
Check user-supplied Jacobian	CHJAC
Generate starting points	GGUES

6.9.1 Επίλυση Προβλήματος Γραμμικού Προγραμματισμού με βάση την αναθεωρημένη μέθοδο Simplex.

Τρόπος Χρήσης

CALL DLPRS (M, NVAR, A, LDA, BL, BU, C, IRTYPE, XLB, XUB, OBJ, XSOL, DSOL)

Παράμετροι

M — Number of constraints. (Input)

NVAR — Number of variables. (Input)

A — `M` by `NVAR` matrix containing the coefficients of the `M` constraints. (Input)

LDA — Leading dimension of `A` exactly as specified in the dimension statement of the calling program. (Input) `LDA` must be at least `M`.

BL — Vector of length `M` containing the lower limit of the general constraints; if there is no lower limit on the `I`-th constraint, then `BL(I)` is not referenced. (Input)

BU — Vector of length `M` containing the upper limit of the general constraints; if there is no upper limit on the `I`-th constraint, then `BU(I)` is not referenced; if there are no range constraints, `BL` and `BU` can share the same storage locations. (Input)

C — Vector of length `NVAR` containing the coefficients of the objective function. (Input)

IRTYPE — Vector of length `M` indicating the types of general constraints in the matrix `A`. (Input)

Let `R(I)` = `A(I, 1) * XSOL(1)` + ... + `A(I, NVAR) * XSOL(NVAR)`. Then, the value of `IRTYPE(I)` signifies the following:

IRTYPE(I)	I-th Constraint
0	BL(I).EQ.R(I).EQ.BU(I)
1	R(I).LE.BU(I)
2	R(I).GE.BL(I)
3	BL(I).LE.R(I).LE.BU(I)

XLB — Vector of length `NVAR` containing the lower bound on the variables; if there is no lower bound on a variable, then 1.0E30 should be set as the lower bound. (Input)

XUB — Vector of length `NVAR` containing the upper bound on the variables; if there is no upper bound on a variable, then –1.0E30 should be set as the upper bound. (Input)

OBJ — Value of the objective function. (Output)

XSOL — Vector of length `NVAR` containing the primal solution. (Output)

DSOL — Vector of length `M` containing the dual solution. (Output)

Σχόλια

1. Automatic workspace usage is

DLPRS   M * M + 57 * M + 3 * NVAR units, or
DDLPRS   2 * M * M + 85 * M + 3 * NVAR units.

Workspace may be explicitly provided, if desired, by use of D2PRS/DD2PRS. The reference is

```
CALL D2PRS (M, NVAR, A, LDA, BL, BU, C, IRTYPE, XLB,
            XUB, OBJ, XSOL, DSOL, AWK, LDAWK, WK,
            IWK)
```

The additional arguments are as follows:

AWK — Real work array of dimension 1 by 1. (AWK is not used in the new implementation of the revised simplex algorithm. It is retained merely for calling sequence consistency.)

LDAWK — Leading dimension of AWK exactly as specified in the dimension statement of the calling program. LDAWK should be 1. (LDAWK is not used in the new implementation of the revised simplex algorithm. It is retained merely for calling sequence consistency.)

WK — Real work vector of length M * (M + 28).

IWK — Integer work vector of length 29 * M + 3 * NVAR.

2. Informational errors

Type	Code	
3	1	The problem is unbounded.
4	2	Maximum number of iterations exceeded.
3	3	The problem is infeasible.
4	4	Numerical difficulty occurred; using double precision may help.
4	5	The bounds are inconsistent.

Algorithm

The routine DLPRS uses a revised simplex method to solve linear programming problems, i.e., problems of the form

$$\min_{x \in R^n} c^T x$$

subject to $b_l \le Ax \le b_u$

$x_l \le x \le x_u$

where c is the objective coefficient vector, A is the coefficient matrix, and the vectors b_l, b_u, x_l and x_u are the lower and upper bounds on the constraints and the variables, respectively.

For a complete description of the revised simplex method, see Murtagh (1981) or Murty (1983).

Example

A linear programming problem is solved.

```
      INTEGER    LDA, M, NVAR
      PARAMETER  (M=2, NVAR=2, LDA=M)
C                                  M = number of constraints
```

```
C                                 NVAR = number of variables
C
      INTEGER    I, IRTYPE(M), NOUT
      REAL       A(LDA,NVAR), B(M), C(NVAR), DSOL(M), OBJ, XLB(NVAR),
     &           XSOL(NVAR), XUB(NVAR)
      EXTERNAL   DLPRS, SSCAL, UMACH
C
C                                 Set values for the following problem
C
C                                 Max 1.0*XSOL(1) + 3.0*XSOL(2)
C
C                                 XSOL(1) + XSOL(2) .LE. 1.5
C                                 XSOL(1) + XSOL(2) .GE. 0.5
C
C                                 0 .LE. XSOL(1) .LE. 1
C                                 0 .LE. XSOL(2) .LE. 1
C
      DATA XLB/2*0.0/, XUB/2*1.0/
      DATA A/4*1.0/, B/1.5, .5/, C/1.0, 3.0/
      DATA IRTYPE/1, 2/
C                                 To maximize, C must be multiplied by
C                                 -1.
      CALL SSCAL (NVAR, -1.0E0, C, 1)
C                                 Solve the LP problem.  Since there is
C                                 no range constraint, only B is
C                                 needed.
      CALL DLPRS (M, NVAR, A, LDA, B, B, C, IRTYPE, XLB, XUB, OBJ,
     &            XSOL, DSOL)
C                                 OBJ must be multiplied by -1 to get
C                                 the true maximum.
      OBJ = -OBJ
C                                 DSOL must be multiplied by -1 for
C                                 maximization.
      CALL SSCAL (M, -1.0E0, DSOL, 1)
C                                 Print results
      CALL UMACH (2, NOUT)
      WRITE (NOUT,99999) OBJ, (XSOL(I),I=1,NVAR), (DSOL(I),I=1,M)
C
99999 FORMAT (//, '   Objective       = ', F9.4, //, '   Primal ',
     &       'Solution =', 2F9.4, //, '   Dual solution   =', 2F9.4)
C
      END
```

Output

```
Objective       =    3.5000

Primal Solution =   0.5000   1.0000

Dual solution   =   1.0000   0.0000
```

6.9.2 Επίλυση Δευτοροβάθμιου προβλήματος με γραμμικούς περιορισμούς

Χρήση - Usage

CALL QPROG (NVAR, NCON, NEQ, A, LDA, B, G, H, LDH, DIAG,
SOL, NACT, IACT, ALAMDA)

Arguments

NVAR — The number of variables. (Input)

NCON — The number of linear constraints. (Input)

NEQ — The number of linear equality constraints. (Input)

A — NCON by NVAR matrix. (Input)
The matrix contains the equality constraints in the first NEQ rows followed by the inequality constraints.

LDA — Leading dimension of A exactly as specified in the dimension statement of the calling program. (Input)

B — Vector of length NCON containing right-hand sides of the linear constraints. (Input)

G — Vector of length NVAR containing the coefficients of the linear term of the objective function. (Input)

H — NVAR by NVAR matrix containing the Hessian matrix of the objective function. (Input)
H should be symmetric positive definite; if H is not positive definite, the algorithm attempts to solve the QP problem with H replaced by a H + DIAG * I such that H + DIAG * I is positive definite. See Comment 3.

LDH — Leading dimension of H exactly as specified in the dimension statement of the calling program. (Input)

DIAG — Scalar equal to the multiple of the identity matrix added to H to give a positive definite matrix. (Output)

SOL — Vector of length NVAR containing solution. (Output)

NACT — Final number of active constraints. (Output)

IACT — Vector of length NVAR containing the indices of the final active constraints in the first NACT positions. (Output)

ALAMDA — Vector of length NVAR containing the Lagrange multiplier estimates of the final active constraints in the first NACT positions. (Output)

Σχόλια

1. Automatic workspace usage is

QPROG (3 * NVAR**2 + 11 * NVAR)/2 + NCON units, or
DQPROG (3 * NVAR**2 + 11 * NVAR) + 2 * NCON units.

Workspace may be explicitly provided, if desired, by use of Q2ROG/DQ2ROG. The reference is

```
CALL Q2ROG (NVAR, NCON, NEQ, A, LDA, B, G, H, LDH,
            DIAG, SOL, NACT, IACT, ALAMDA, WK)
```

The additional argument is

WK — Work vector of length (3 * NVAR**2 + 11 * NVAR)/2 + NCON.

2. Informational errors

Type Code

3 1 Due to the effect of computer rounding error, a change in the variables fail to improve the objective function value; usually the solution is close to optimum.

4 2 The system of equations is inconsistent. There is no solution.

3. If a perturbation of H, H + DIAG * I, was used in the QP problem, then H + DIAG * I should also be used in the definition of the Lagrange multipliers.

Algorithm

The routine QPROG is based on M.J.D. Powell's implementation of the Goldfarb and Idnani (1983) dual quadratic programming (QP) algorithm for convex QP problems subject to general linear equality/inequality constraints, i.e., problems of the form

$$\min_{x \in R^n} g^T x + \frac{1}{2} x^T H x$$

subject to $A1x = b1$

$A2x \geq b2$

given the vectors $b1$, $b2$, and g and the matrices H, $A1$, and $A2$. H is required to be positive definite. In this case, a unique x solves the problem or the constraints are inconsistent. If H is not positive definite, a positive definite perturbation of H is used in place of H. For more details, see Powell (1983, 1985).

Παράδειγμα - Example

A quadratic programming problem is solved.

```
C                                  Declare variables
      INTEGER    LDA, LDH, NCON, NEQ, NVAR
      PARAMETER  (NCON=2, NEQ=2, NVAR=5, LDA=NCON, LDH=NVAR)
C
      INTEGER    IACT(NVAR), K, NACT, NOUT
      REAL       A(LDA,NVAR), ALAMDA(NVAR), B(NCON), DIAG, G(NVAR),
     &           H(LDH,LDH), SOL(NVAR)
      EXTERNAL   QPROG, UMACH
C
C                                  Set values of A, B, G and H.
C                                  A = ( 1.0  1.0  1.0  1.0  1.0)
C                                      ( 0.0  0.0  1.0 -2.0 -2.0)
C
C                                  B = ( 5.0 -3.0)
C
C                                  G = (-2.0  0.0  0.0  0.0  0.0)
C
C                                  H = ( 2.0  0.0  0.0  0.0  0.0)
C                                      ( 0.0  2.0 -2.0  0.0  0.0)
C                                      ( 0.0 -2.0  2.0  0.0  0.0)
C                                      ( 0.0  0.0  0.0  2.0 -2.0)
C                                      ( 0.0  0.0  0.0 -2.0  2.0)
C
      DATA A/1.0, 0.0, 1.0, 0.0, 1.0, 1.0, 1.0, -2.0, 1.0, -2.0/
      DATA B/5.0, -3.0/
```

```
      DATA G/-2.0, 4*0.0/
      DATA H/2.0, 5*0.0, 2.0, -2.0, 3*0.0, -2.0, 2.0, 5*0.0, 2.0,
     &     -2.0, 3*0.0, -2.0, 2.0/
C
      CALL QPROG (NVAR, NCON, NEQ, A, LDA, B, G, H, LDH, DIAG, SOL,
     &            NACT, IACT, ALAMDA)
C
      CALL UMACH (2, NOUT)
      WRITE (NOUT,99999) (SOL(K),K=1,NVAR)
99999 FORMAT ('  The solution vector is', /, '  SOL = (', 5F6.1,
     &       '  )')
C
      END
```

Output

```
The solution vector is
SOL = (   1.0   1.0   1.0   1.0   1.0  )
```

6.9.3 Ελαχιστοποίηση οποιασδήποτε αντικειμενικής συνάρτησης που υπόκειται σε γραμμικούς περιορισμούς

Minimize a general objective function subject to linear equality/inequality constraints.

Χρήση

CALL LCONG (FCN, GRAD, NVAR, NCON, NEQ, A, LDA, B, XLB,
XUB, XGUESS, ACC, MAXFCN, SOL, OBJ, NACT, IACT,
ALAMDA)

Arguments

FCN — User-supplied `SUBROUTINE` to evaluate the function to be minimized. The usage is `CALL FCN (N, X, F)`, where

`N` — Value of `NVAR`. (Input)

`X` — Vector of length `N` at which point the function is evaluated. (Input)
`X` should not be changed by `FCN`.

`F` — The computed function value at the point `X`. (Output)

`FCN` must be declared `EXTERNAL` in the calling program.

GRAD — User-supplied `SUBROUTINE` to compute the gradient at the point `X`. The usage is `CALL GRAD (N, X, G)`, where

`N` — Value of `NVAR`. (Input)

`X` — Vector of length `N` at which point the function is evaluated. (Input)
`X` should not be changed by `GRAD`.

`G` — Vector of length `N` containing the values of the gradient of the objective function evaluated at the point `X`. (Output)

`GRAD` must be declared `EXTERNAL` in the calling program.

NVAR — The number of variables. (Input)

NCON — The number of linear constraints (excluding simple bounds). (Input)

NEQ — The number of linear equality constraints. (Input)

A — NCON by NVAR matrix. (Input)
The matrix contains the equality constraint gradients in the first NEQ rows, followed by the inequality constraint gradients.

LDA — Leading dimension of A exactly as specified in the dimension statement of the calling program. (Input)

B — Vector of length NCON containing right-hand sides of the linear constraints. (Input)
Specifically, the constraints on the variables X(I), I = 1, . . . , NVAR are A(K, 1) * X(1) + . . . + A(K, NVAR) * X(NVAR).EQ.B(K), K = 1, . . . , NEQ.A(K, 1) * X(1) + . . . + A(K, NVAR) * X(NVAR).LE.B(K), K = NEQ + 1, . . . , NCON. Note that the data that define the equality constraints come before the data of the inequalities.

XLB — Vector of length NVAR containing the lower bounds on the variables; choose a very large negative value if a component should be unbounded below or set XLB(I) = XUB(I) to freeze the I-th variable. (Input)
Specifically, these simple bounds are XLB(I).LE.X(I), I = 1, . . . , NVAR.

XUB — Vector of length NVAR containing the upper bounds on the variables; choose a very large positive value if a component should be unbounded above. (Input)
Specifically, these simple bounds are X(I).LE. XUB(I), I = 1, . . . , NVAR.

XGUESS — Vector of length NVAR containing the initial guess of the minimum. (Input)

ACC — The nonnegative tolerance on the first order conditions at the calculated solution. (Input)

MAXFCN — On input, maximum number of function evaluations allowed.(Input/ Output)
On output, actual number of function evaluations needed.

SOL — Vector of length NVAR containing solution. (Output)

OBJ — Value of the objective function. (Output)

NACT — Final number of active constraints. (Output)

IACT — Vector containing the indices of the final active constraints in the first NACT positions. (Output)
Its length must be at least NCON + 2 * NVAR.

ALAMDA — Vector of length NVAR containing the Lagrange multiplier estimates of the final active constraints in the first NACT positions. (Output)

Σχόλια

1. Automatic workspace usage is

LCONG NVAR**2 + 11 * NVAR + NCON units, or
DLCONG 2 * (NVAR**2 + 11 * NVAR + NCON) units.

Workspace may be explicitly provided, if desired, by use of L2ONG/DL2ONG. The reference is

```
CALL L2ONG (FCN, GRAD, NVAR, NCON, NEQ, A, LDA, B,
            XLB, XUB, XGUESS, ACC, MAXFCN, SOL, OBJ,
            NACT, IACT, ALAMDA, IPRINT, INFO, WK)
```

The additional arguments are as follows:

IPRINT — Print option (see Comment 3). (Input)

INFO — Informational flag (see Comment 3). (Output)

WK — Real work vector of length NVAR**2 + 11 * NVAR + NCON.

2. Informational errors

Type	Code	
4	4	The equality constraints are inconsistent.
4	5	The equality constraints and the bounds on the variables are found to be inconsistent.
4	6	No vector X satisfies all of the constraints. In particular, the current active constraints prevent any change in X that reduces the sum of constraint violations.
4	7	Maximum number of function evaluations exceeded.
4	9	The variables are determined by the equality constraints.

3. The following are descriptions of the arguments IPRINT and INFO:

IPRINT — This argument must be set by the user to specify the frequency of printing during the execution of the routine LCONG. There is no printed output if IPRINT = 0. Otherwise, after ensuring feasibility, information is given every IABS(IPRINT) iterations and whenever a parameter called TOL is reduced. The printing provides the values of X(.), F(.) and G(.) = GRAD(F) if IPRINT is positive. If IPRINT is negative, this information is augmented by the current values of IACT(K) K = 1, . . . , NACT, PAR(K) K = 1, . . . , NACT and RESKT(I) I = 1, . . . , N. The reason for returning to the calling program is also displayed when IPRINT is nonzero.

INFO — On exit from L2ONG, INFO will have one of the following integer values to indicate the reason for leaving the routine:

INFO = 1 SOL is feasible and the condition that depends on ACC is satisfied.

INFO = 2 SOL is feasible and rounding errors are preventing further progress.

INFO = 3 SOL is feasible but the objective function fails to decrease although a decrease is predicted by the current gradient vector.

INFO = 4 In this case, the calculation cannot begin because LDA is less than NCON or because the lower bound on a variable is greater than the upper bound.

INFO = 5 This value indicates that the equality constraints are inconsistent. These constraints include any components of X(.) that are frozen by setting XL(I) = XU(I).

INFO = 6 In this case, there is an error return because the equality constraints and the bounds on the variables are found to be inconsistent.

INFO = 7 This value indicates that there is no vector of variables that satisfies all of the constraints. Specifically, when this return or an INFO = 6 return occurs, the current active constraints (whose indices are IACT(K), K = 1, . . . , NACT) prevent any change in X(.) that reduces the sum of constraint violations, where only bounds are included in this sum if INFO = 6.

INFO = 8 Maximum number of function evaluations exceeded.

INFO = 9 The variables are determined by the equality constraints.

Algorithm

The routine `LCONG` is based on M.J.D. Powell's `TOLMIN`, which solves linearly constrained optimization problems, i.e., problems of the form

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

subject to $A_1 x = b_1$

$A_2 x \leq b_2$

$x_l \leq x \leq x_u$

given the vectors b_1, b_2, x_l and x_u and the matrices A_1, and A_2.

The algorithm starts by checking the equality constraints for inconsistency and redundancy. If the equality constraints are consistent, the method will revise x_0, the initial guess provided by the user, to satisfy

$A_1 x = b_1$

Next, x_0 is adjusted to satisfy the simple bounds and inequality constraints. This is done by solving a sequence of quadratic programming subproblems to minimize the sum of the constraint or bound violations.

Now, for each iteration with a feasible x_k, let J_k be the set of indices of inequality constraints that have small residuals. Here, the simple bounds are treated as inequality constraints. Let I_k be the set of indices of active constraints. The following quadratic programming problem

$$\min f\left(x^k\right) + d^T \nabla f\left(x^k\right) + \frac{1}{2} d^T B^k d$$

subject to $a_j d = 0 \quad j \in I_k$

$a_j d \leq 0 \quad j \in J_k$

is solved to get (d_k, λ_k) where a_j is a row vector representing either a constraint in A_1or A_2 or a bound constraint on x. In the latter case, the $a_j = e_i$ for the bound constraint $x_i \leq (x_u)_i$ and $a_j = -e_i$ for the constraint $-x_i \leq (-x_l)_i$. Here, e_i is a vector with a 1 as the i-th component, and zeroes elsewhere. λ_k are the Lagrange multipliers, and B_k is a positive definite approximation to the second derivative $\nabla^2 f(x_k)$.

After the search direction d_k is obtained, a line search is performed to locate a better point. The new point $x_{k+1}= x_k + \alpha_k d_k$ has to satisfy the conditions

$$f\left(x^k + \alpha^k d^k\right) \leq f\left(x^k\right) + 0.1\alpha^k \left(d^k\right)^T \nabla f\left(x^k\right)$$

and

$$\left(d^k\right)^T \nabla f\left(x^k + \alpha^k d^k\right) \geq 0.7\left(d^k\right)^T \nabla f\left(x^k\right)$$

The main idea in forming the set J_k is that, if any of the inequality constraints restricts the step-length α_k, then its index is not in J_k. Therefore, small steps are likely to be avoided.

Finally, the second derivative approximation, B_k, is updated by the BFGS formula, if the condition

$$\left(d^k\right)^T \nabla f\left(x^k + \alpha^k d^k\right) - \nabla f\left(x^k\right) > 0$$

holds. Let $x^k \leftarrow x^{k+1}$, and start another iteration.

The iteration repeats until the stopping criterion

$$\left\| \nabla f\left(x^k\right) - A^k \lambda^k \right\|_2 \le \tau$$

is satisfied; here, τ is a user-supplied tolerance. For more details, see Powell (1988, 1989).

Παράδειγμα - Example

The problem from Schittkowski (1987)

min $f(x) = -x_1 x_2 x_3$

subject to $-x_1 - 2x_2 - 2x_3 \le 0$

$x_1 + 2x_2 + 2x_3 \le 72$

$0 \le x_1 \le 20$

$0 \le x_2 \le 11$

$0 \le x_3 \le 42$

is solved with an initial guess $x_1 = 10$, $x_2 = 10$ and $x_3 = 10$.

```
C                                 Declaration of variables
      INTEGER    LDA, NCON, NEQ, NVAR
      PARAMETER  (NCON=2, NEQ=0, NVAR=3, LDA=NCON)
C
      INTEGER    IACT(8), MAXFCN, NACT, NOUT
      REAL       A(NCON,NVAR), ACC, ALAMDA(NVAR), B(NCON), OBJ,
     &           SOL(NVAR), XGUESS(NVAR), XLB(NVAR), XUB(NVAR)
      EXTERNAL   FCN, GRAD, LCONG, UMACH
C
C                                 Set values for the following
problem.
C
C                                 Min  -X(1)*X(2)*X(3)
C
C                                 -X(1) - 2*X(2) - 2*X(3)  .LE.   0
C                                  X(1) + 2*X(2) + 2*X(3)  .LE.  72
C
C                                 0  .LE.  X(1)  .LE.  20
C                                 0  .LE.  X(2)  .LE.  11
C                                 0  .LE.  X(3)  .LE.  42
C
      DATA A/-1.0, 1.0, -2.0, 2.0, -2.0, 2.0/, B/0.0, 72.0/
      DATA XLB/3*0.0/, XUB/20.0, 11.0, 42.0/, XGUESS/3*10.0/
      DATA ACC/0.0/, MAXFCN/400/
C
      CALL UMACH (2, NOUT)
C
      CALL LCONG (FCN, GRAD, NVAR, NCON, NEQ, A, LDA, B, XLB, XUB,
     &            XGUESS, ACC, MAXFCN, SOL, OBJ, NACT, IACT, ALAMDA)
C
      WRITE (NOUT,99998) 'Solution:'
      WRITE (NOUT,99999) SOL
      WRITE (NOUT,99998) 'Function value at solution:'
      WRITE (NOUT,99999) OBJ
      WRITE (NOUT,99998) 'Number of function evaluations:', MAXFCN
      STOP
99998 FORMAT (//, ' ', A, I4)
```

```
99999 FORMAT (1X, 5F16.6)
      END
C
      SUBROUTINE FCN (N, X, F)
      INTEGER    N
      REAL       X(*), F
C
      F = -X(1)*X(2)*X(3)
      RETURN
      END
C
      SUBROUTINE GRAD (N, X, G)
      INTEGER    N
      REAL       X(*), G(*)
C
      G(1) = -X(2)*X(3)
      G(2) = -X(1)*X(3)
      G(3) = -X(1)*X(2)
      RETURN
      END
```

Output

```
Solution:
20.000000       11.000000       15.000000

Function value at solution:
-3300.000000

Number of function evaluations:   5
```

6.9.4 Επίλυση γενικού μη γραμμικού προβλήματος (SQP) αριθμητικές ευαισθησίες

Solve a general nonlinear programming problem using the successive quadratic programming algorithm and a finite difference gradient.

χρήση

CALL NCONF (FCN, M, ME, N, XGUESS, IBTYPE, XLB, XUB,
XSCALE, IPRINT, MAXITN, X, FVALUE)

Arguments

FCN — User-supplied `SUBROUTINE` to evaluate the functions at a given point. The usage is `CALL FCN (M, ME, N, X, ACTIVE, F, G)`, where

`M` — Total number of constraints. (Input)

`ME` — Number of equality constraints. (Input)

`N` — Number of variables. (Input)

`X` — The point at which the functions are evaluated. (Input)
`X` should not be changed by `FCN`.

`ACTIVE` — Logical vector of length `MMAX` indicating the active constraints. (Input)
`MMAX` = MAX(1, `M`)

`F` — The computed function value at the point `X`. (Output)

`G` — Vector of length `MMAX` containing the values of constraints at point `X`. (Output)

`FCN` must be declared `EXTERNAL` in the calling program.

M — Total number of constraints. (Input)

ME — Number of equality constraints. (Input)

N — Number of variables. (Input)

XGUESS — Vector of length `N` containing an initial guess of the computed solution. (Input)

IBTYPE — Scalar indicating the types of bounds on variables. (Input)

`IBTYPE` Action

0 User will supply all the bounds.

1 All variables are nonnegative.

2 All variables are nonpositive.

3 User supplies only the bounds on 1st variable; all other variables will have the same bounds.

XLB — Vector of length `N` containing the lower bounds on variables. (Input, if `IBTYPE` = 0; output, if `IBTYPE` = 1 or 2; input/output, if `IBTYPE` = 3)
If there is no lower bound for a variable, then the corresponding `XLB` value should be set to −1.0E6.

XUB — Vector of length `N` containing the upper bounds on variables. (Input, if `IBTYPE` = 0; output, if `IBTYPE` = 1 or 2; input/output, if `IBTYPE` = 3)
If there is no upper bound for a variable, then the corresponding `XLB` value should be set to 1.0E6.

XSCALE — Vector of length `N` containing the diagonal scaling matrix for the variables. (Input)
All values of `XSCALE` must be greater than zero. In the absence of other information, set all entries to 1.0.

IPRINT — Parameter indicating the desired output level. (Input)

`IPRINT` Action

0 No output printed.

1 Only a final convergence analysis is given.

2 One line of intermediate results are printed in each iteration.

3 Detailed information is printed in each iteration.

MAXITN — Maximum number of iterations allowed. (Input)

X — Vector of length `N` containing the computed solution. (Output)

FVALUE — Scalar containing the value of the objective function at the computed solution. (Output)

Σχόλια

1. Automatic workspace usage is

`NCONF` `N * (3 * N + 38 + MMAX) + 7 * MMAX + 6 * M + MAX(N, M) + 91` units, or

`DNCONF` `2 * N * (3 * N + 38 + MMAX) + 14 * MMAX + 12 * M + MAX(N, M) + 163` units.
`MMAX = MAX(1, M)`

Workspace may be explicitly provided, if desired, by use of N2ONF/DN2ONF. The reference is

```
CALL N2ONF (FCN, M, ME, N, XGUESS, IBTYPE, XLB, XUB,
            XSCALE, IPRINT, MAXITN, X, FVALUE, WK,
            LWK, IWK, LIWK, CONWK)
```

The additional arguments are as follows:

WK — Work vector of length N * (3 * N + 38 + MMAX) + 6 * MMAX + 6 * M + 72

LWK — Length of WK.

IWK — Work vector of length 19 + MAX(M, N).

LIWK — Length of LIWK.

CONWK — Work vector of length MMAX.

2. Informational errors

Type	Code	
4	1	Search direction uphill.
4	2	Line search took more than 5 function calls.
4	3	Maximum number of iterations exceeded.
4	4	Search direction is close to zero.
4	5	The constraints for the QP subproblem are inconsistent.

3. If reverse communication is desired, then N0ONF/DN0ONF can be used rather than using an external subroutine to evaluate the function. The reference is

```
CALL N0ONF (IDO, M, ME, N, IBTYPE, XLB, XUB, IPRINT,
            MAXITN, X, FVALUE, G, DF, DG, LDDG, U,
            C, LDC, D, ACC, SCBOU, MAXFUN, ACTIVE,
            MODE, WK, IWK, CONWK)
```

The additional arguments are as follows:

IDO — Reverse communication parameter indicating task to be done. (Input/Output)

On the initial call, IDO must be set to 0; and initial values must be passed into N0ONF/DN0ONF for X, FVALUE, G, DF, and DG. If the routine returns with IDO = 1, then the user has to compute FVALUE and G with respect to X. If the routine returns with IDO = 2, then the user has to compute DF and DG with respect to X. The user has to call the routine repeatedly until IDO does not equal to 1 or 2.

X — Vector of length N containing the initial guesses to the solution on input and the solution on output. (Input/Output)

FVALUE — Scalar containing the objective function value evaluated at the current X. (Input)

G — Vector of length MMAX containing constraint values at the current X. (Input) MMAX is MAX(1, M).

DF — Vector of length N containing the gradient of the objective function evaluated at the current X. (Input)

DG — Array of dimension MMAX by N containing the gradient of the constraints evaluated at the current X. (Input)

LDDG — Leading dimension of `DG` exactly as specified in the dimension statement of the calling program. (Input)

U — Vector of length `M` + `N` + `N` + 2 containing the multipliers of the nonlinear constraints and of the bounds. (Output)
The first `M` locations contain the multipliers for the nonlinear constraints. The second `N` locations contain the multipliers for the lower bounds. The third `N` locations contain the multipliers for the upper bounds.

C — Array of dimension `N` + 1 by `N` + 1 containing the final approximation to the Hessian. (Output)

LDC — Leading dimension of `C` exactly as specified in the dimension statement of the calling program. (Input)

D — Vector of length `N` + 1 containing the diagonal elements of the Hessian. (Output)

ACC — Final accuracy. (Input)

SCBOU — Scalar containing the scaling variable for the problem function. (Input)
In the absence of further information, `SCBOU` may be set to `1.0E3`.

MAXFUN — Scalar containing the maximum allowable function calls during the line search. (Input)

ACTIVE — Logical vector of length 2 * `MMAX` + 13. (Input/Output)
The first `MMAX` locations are used to determine which gradient constraints are active and must be set to `.TRUE.` on input. If `ACTIVE(I)` is `.TRUE.`, then `DG(I, K)` is evaluated for `K` = 1, `N`. The last `MMAX` + 13 locations are used for workspace.

MODE — Desired solving version for the algorithm. (Input)
If `MODE` = 2; then reverse communication is used. If `MODE` = 3; then reverse communication is used and initial guesses for the multipliers and Hessian matrix of the Lagrange function are provided on input.

WK — Work vector of length 2 * `N` * (`N` + 16) + 4 * `MMAX` + 5 * `M` + 68.

IWK — Work vector of length 19 + `MAX(M, N)`.

CONWK — Work vector of length `M`.

Algorithm

The routine `NCONF` is based on subroutine `NLPQL`, a FORTRAN code developed by Schittkowski (1986). It uses a successive quadratic programming method to solve the general nonlinear programming problem. The problem is stated as follows:

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

$$\text{subject to} \qquad g_j(x) = 0, \text{ for } \quad j = 1, \ldots, m_e$$

$$g_j(x) \geq 0, \text{ for } \quad j = m_e + 1, \ldots, m$$

$$x_l \leq x \leq x_u$$

where all problem functions are assumed to be continuously differentiable. The method, based on the iterative formulation and solution of quadratic programming subproblems, obtains subproblems by using a quadratic approximation of the Lagrangian and by linearizing the constraints. That is,

$$\min_{d \in R^n} \frac{1}{2} d^T B_k d + \nabla f(x_k)^T d$$

$$\text{subject to} \quad \nabla g_j(x_k)^T d + g_j(x_k) = 0, \quad j = 1, \ldots, m_e$$

$$\nabla g_j(x_k)^T d + g_j(x_k) \geq 0, \quad j = m_e + 1, \ldots, m$$

$$x_l - x_k \leq d \leq x_u - x_k$$

where B_k is a positive definite approximation of the Hessian and x_k is the current iterate. Let d_k be the solution of the subproblem. A line search is used to find a new point x_{k+1},

$x_{k+1} = x_k + \lambda d_k, \quad \lambda \in (0, 1]$

such that a "merit function" will have a lower function value at the new point. Here, the augmented Lagrange function (Schittkowski 1986) is used as the merit function.

When optimality is not achieved, B_k is updated according to the modified BFGS formula (Powell 1978). Note that this algorithm may generate infeasible points during the solution process. Therefore, if feasibility must be maintained for intermediate points, then this routine may not be suitable. For more theoretical and practical details, see Stoer (1985), Schittkowski (1983, 1986) and Gill et al. (1985).

Since a finite-difference method is used to estimate the gradient for some single precision calculations, an inaccurate estimate of the gradient may cause the algorithm to terminate at a noncritical point. In such cases, high precision arithmetic is recommended. Also, whenever the exact gradient can be easily provided, routine NCONG should be used instead.

Within `NCONF`, there is a user-callable subroutine `N0ONF` that gives the user the option to use "reverse communication." This option allows the user to evaluate the functions and gradients in the main program. This option is useful when it is difficult to do the function evaluation in the fixed form required by `NCONF`.

Παράδειγμα 1 - Example 1

The problem

$$\min F(x) = (x_1 - 2)^2 + (x_2 - 1)^2$$

$$\text{subject to} \quad g_1(x) = x_1 - 2x_2 + 1 = 0$$

$$g_2(x) = -(x_1^2)/4 - x_2^2 + 1 \geq 0$$

is solved with an initial guess (2.0, 2.0).

```
      INTEGER    IBTYPE, IPRINT, M, MAXITN, ME, N
      PARAMETER  (IBTYPE=0, IPRINT=0, M=2, MAXITN=100, ME=1, N=2)
C
      REAL       FVALUE, X(N), XGUESS(N), XLB(N), XSCALE(N), XUB(N)
      EXTERNAL   FCN, NCONF, WRRRN
C
      DATA XGUESS/2.0E0, 2.0E0/, XSCALE/2*1.0E0/
      DATA XLB/-1.0E6, -1.0E6/, XUB/1.0E6, 1.0E6/
C
      CALL NCONF (FCN, M, ME, N, XGUESS, IBTYPE, XLB, XUB, XSCALE,
     &            IPRINT, MAXITN, X, FVALUE)
C
      CALL WRRRN ('The solution is', N, 1, X, N, 0)
      END
C
      SUBROUTINE FCN (M, ME, N, X, ACTIVE, F, G)
      INTEGER    M, ME, N
      REAL       X(*), F, G(*)
      LOGICAL    ACTIVE(*)
C                                 Himmelblau problem 1
      F = (X(1)-2.0E0)**2 + (X(2)-1.0E0)**2
C
      IF (ACTIVE(1)) G(1) = X(1) - 2.0E0*X(2) + 1.0E0
      IF (ACTIVE(2)) G(2) = -(X(1)**2)/4.0E0 - X(2)**2 + 1.0E0
      RETURN
      END
```

Output

```
The solution is
 1   0.8229
 2   0.9114
```

Παράδειγμα 2 - Example 2

This example uses the reverse communication option to solve the same problem as Example 1.

```
      INTEGER    LDC, LDDG, LWK, M, ME, N
      PARAMETER  (M=2, ME=1, N=2, LDC=N+1, LDDG=M,
     &           LWK=2*N*(N+16)+9*M+68)
C
      INTEGER    IBTYPE, IDO, IPRINT, IWK(19+M), MAXFUN, MAXITN,
     &           MODE
      REAL       ACC, AMACH, C(LDC,N+1), CONWK(M), D(N+1), DF(N),
     &           DG(LDDG,N), FVALUE, G(M), SCBOU, SQRT,
     &           U(M+N+N+2), WK(LWK), X(N), XLB(N), XUB(N)
      LOGICAL    ACTIVE(2*M+13)
      INTRINSIC  SQRT
```

```
      EXTERNAL   AMACH, N0ONF
C
      DATA IBTYPE/3/, MAXITN/100/, MODE/2/, MAXFUN/10/, IPRINT/0/
      DATA X/2.0E0, 2.0E0/, XLB(1)/-1.0E6/, XUB(1)/1.0E6/,
SCBOU/1.0E3/
C                                Set final accuracy (ACC)
      ACC = SQRT(AMACH(4))
C
      ACTIVE(1) = .TRUE.
      ACTIVE(2) = .TRUE.
      IDO       = 0
   10 IF (IDO.EQ.0 .OR. IDO.EQ.1) THEN
C                                Evaluate the function at X.
         FVALUE = (X(1)-2.0E0)**2 + (X(2)-1.0E0)**2
C                                Evaluate the constraints at X.
         G(1) = X(1) - 2.0E0*X(2) + 1.0E0
         G(2) = -(X(1)**2)/4.0E0 - X(2)**2 + 1.0E0
      END IF
C
      IF (IDO.EQ.0 .OR. IDO.EQ.2) THEN
C                                Evaluate the function gradient at
X.
         DF(1) = 2.0E0*(X(1)-2.0E0)
         DF(2) = 2.0E0*(X(2)-1.0E0)
C                                If active evaluate the constraint
C                                gradient at X.
         IF (ACTIVE(1)) THEN
            DG(1,1) = 1.0E0
            DG(1,2) = -2.0E0
         END IF
C
         IF (ACTIVE(2)) THEN
            DG(2,1) = -0.5E0*X(1)
            DG(2,2) = -2.0E0*X(2)
         END IF
      END IF
C                               Call N0ONF for the next update.
C
      CALL N0ONF (IDO, M, ME, N, IBTYPE, XLB, XUB, IPRINT, MAXITN, X,
     &            FVALUE, G, DF, DG, LDDG, U, C, LDC, D, ACC, SCBOU,
     &            MAXFUN, ACTIVE, MODE, WK, IWK, CONWK)
C                               If IDO does not equal 1 or 2, exit.
      IF (IDO.EQ.1 .OR. IDO.EQ.2) GO TO 10
C                                Print the solution
      CALL WRRRN ('The solution is', N, 1, X, N, 0)
C
      END
```

Output

```
The solution is
 1   0.8229
 2   0.9114
```

6.9.5 Επίλυση Γενικού μη γραμμικού προβλήματος (SQP) αναλυτικές ευαισθησίες

Solve a general nonlinear programming problem using the successive quadratic programming algorithm and a user-supplied gradient.

Usage

CALL NCONG (FCN, GRAD, M, ME, N, XGUESS, IBTYPE, XLB, XUB, IPRINT, MAXITN, X, FVALUE)

Arguments

FCN — User-supplied `SUBROUTINE` to evaluate the functions at a given point. The usage is `CALL FCN (M, ME, N, X, ACTIVE, F, G)`, where

`M` — Total number of constraints. (Input)

`ME` — Number of equality constraints. (Input)

`N` — Number of variables. (Input)

`X` — The point at which the functions are evaluated. (Input)
`X` should not be changed by `FCN`.

`ACTIVE` — Logical vector of length `MMAX` indicating the active constraints. (Input)

`F` — The computed function value at the point `X`. (Output)

`G` — Vector of length `MMAX` containing the values of constraints at point `X`. (Output)

`FCN` must be declared `EXTERNAL` in the calling program.

GRAD — User-supplied `SUBROUTINE` to evaluate the gradients at a given point. The usage is `CALL GRAD (M, ME, MMAX, N, X, ACTIVE, F, G, DF, DG)`, where

`M` — Total number of constraints. (Input)

`ME` — Number of equality constraints. (Input)

`MMAX` — Maximum of (1, `M`). (Input)

`N` — Number of variables. (Input)

`X` — Vector of length `N` at which point the function is evaluated. (Input)
`X` should not be changed by `FCN`.

`ACTIVE` — Logical vector of length `MMAX` indicating the active constraints. (Input)

`F` — The computed function value at the point `X`. (Input)

`G` — Vector of length `MMAX` containing the values of the constraints at point `X`. (Input)

`DF` — Vector of length `N` containing the values of the gradient of the objective function. (Output)

`DG` — `MMAX` by `N` array containing the values of the gradients for the active constraints. (Output)

`GRAD` must be declared `EXTERNAL` in the calling program.

M — Total number of constraints. (Input)

ME — Number of equality constraints. (Input)

N — Number of variables. (Input)

XGUESS — Vector of length N containing an initial guess of the computed solution. (Input)

IBTYPE — Scalar indicating the types of bounds on variables. (Input)

IBTYPE	Action
0	User will supply all the bounds.
1	All variables are nonnegative.
2	All variables are nonpositive.
3	User supplies only the bounds on 1st variable, all other variables will have the same bounds.

XLB — Vector of length N containing the lower bounds on the variables. (Input, if IBTYPE = 0; output, if IBTYPE = 1 or 2; input/output, if IBTYPE = 3) If there is no lower bound on a variable, then the corresponding XLB value should be set to –1.0E6.

XUB — Vector of length N containing the upper bounds on the variables. (Input, if IBTYPE = 0; output, if IBTYPE = 1 or 2; input/output, if IBTYPE = 3) If there is no upper bound on a variable, then the corresponding XUB value should be set to 1.0E6.

IPRINT — Parameter indicating the desired output level. (Input)

IPRINT	Action
0	No output printed.
1	Only a final convergence analysis is given.
2	One line of intermediate results is printed for each iteration.
3	Detailed information is printed for each iteration.

MAXITN — Maximum number of iterations allowed. (Input)

X — Vector of length N containing the computed solution. (Output)

FVALUE — Scalar containing the value of the objective function at the computed solution. (Output)

Σχόλια

1. Automatic workspace usage is

NCONG N * (3 * N + 38 + MMAX) + 6 * (M + MMAX) + MAX(N, M) + 91 units, or

DNCONG 2 * N * (3 * N + 38 + MMAX) + 12 * (M + MMAX) + MAX(N, M) + 163 units.
MMAX = MAX(1, M)

Workspace may be explicitly provided, if desired, by use of N2ONG/DN2ONG. The reference is

```
CALL N2ONG (FCN, GRAD, M, ME, N, XGUESS, IBTYPE,
            XLB, XUB, IPRINT, MAXITN, X, FVALUE, WK,
            LWK, IWK, LIWK)
```

The additional arguments are as follows:

WK — Work vector of length N * (3 * N + 38 + MMAX) + 6 * (M + MMAX) + 72.

LWK — Scalar containing the value for the length of WK.

IWK — Work vector of length 19 + MAX(M, N).

LIWK — Scalar containing the value for the length of `IWK`.

2. Informational errors

Type	Code	
4	1	Search direction uphill.
4	2	Line search took more than 5 function calls.
4	3	Maximum number of iterations exceeded.
4	4	Search direction is close to zero.
4	5	The constraints for the `QP` subproblem are inconsistent.

Algorithm

The routine `NCONG` is based on subroutine `NLPQL`, a FORTRAN code developed by Schittkowski (1986). It uses a successive quadratic programming method to solve the general nonlinear programming problem. The problem is stated as follows:

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

$$\begin{aligned}
\text{subject to} \quad & g_j(x) = 0, \text{ for} \quad j = 1, \ldots, m_e \\
& g_j(x) \ge 0, \text{ for} \quad j = m_e + 1, \ldots, m \\
& x_l \le x \le x_u
\end{aligned}$$

where all problem functions are assumed to be continuously differentiable. The method, based on the iterative formulation and solution of quadratic programming subproblems, obtains subproblems by using a quadratic approximation of the Lagrangian and by linearizing the constraints. That is,

$$\min_{d \in \mathbf{R}^n} \frac{1}{2} d^T B_k d + \nabla f(x_k)^T d$$

$$\begin{aligned}
\text{subject to} \quad & \nabla g_j(x_k)^T d + g_j(x_k) = 0, \quad j = 1, \ldots, m_e \\
& \nabla g_j(x_k)^T d + g_j(x_k) \ge 0, \quad j = m_e + 1, \ldots, m \\
& x_l - x_k \le d \le x_u - x_k
\end{aligned}$$

where *B*k is a positive definite approximation of the Hessian and *x*k is the current iterate. Let *d*k be the solution of the subproblem. A line search is used to find a new point *x*k+1,

*x*k+1 = *x*k + λ*d*k,   λ∈ (0, 1]

such that a "merit function" will have a lower function value at the new point. Here, the augmented Lagrange function (Schittkowski 1986) is used as the merit function.

When optimality is not achieved, *B*k is updated according to the modified BFGS formula (Powell 1978). Note that this algorithm may generate infeasible points during the solution process. Therefore, if feasibility must be maintained for intermediate points, then this routine may not be suitable. For more theoretical and practical details, see Stoer (1985), Schittkowski (1983, 1986) and Gill et al. (1985).

Παράδειγμα - Example

The problem

$$\min F(x) = (x_1 - 2)^2 + (x_2 - 1)^2$$

$$\text{subject to} \quad g_1(x) = x_1 - 2x_2 + 1 = 0$$

$$g_2(x) = -x_1^2/4 - x_2^2 + 1 \geq 0$$

is solved with an initial guess (2.0, 2.0).

```
      INTEGER    IBTYPE, IPRINT, M, MAXITN, ME, N
      PARAMETER  (IBTYPE=0, IPRINT=0, M=2, MAXITN=100, ME=1, N=2)
C
      REAL       FVALUE, X(N), XGUESS(N), XLB(N), XUB(N)
      EXTERNAL   FCN, GRAD, NCONG, WRRRN
C
      DATA XGUESS/2.0E0, 2.0E0/
      DATA XLB/-1.0E6, -1.0E6/, XUB/1.0E6, 1.0E6/
C
      CALL NCONG (FCN, GRAD, M, ME, N, XGUESS, IBTYPE, XLB, XUB,
     &             IPRINT, MAXITN, X, FVALUE)
C
      CALL WRRRN ('The solution is', N, 1, X, N, 0)
      END
C
      SUBROUTINE FCN (M, ME, N, X, ACTIVE, F, G)
      INTEGER    M, ME, N
      REAL       X(*), F, G(*)
      LOGICAL    ACTIVE(*)
C                                  Himmelblau problem 1
      F = (X(1)-2.0E0)**2 + (X(2)-1.0E0)**2
C
      IF (ACTIVE(1)) G(1) = X(1) - 2.0E0*X(2) + 1.0E0
      IF (ACTIVE(2)) G(2) = -(X(1)**2)/4.0E0 - X(2)**2 + 1.0E0
      RETURN
      END
C
      SUBROUTINE GRAD (M, ME, MMAX, N, X, ACTIVE, F, G, DF, DG)
      INTEGER    M, ME, MMAX, N
      REAL       X(*), F, G(*), DF(*), DG(MMAX,*)
      LOGICAL    ACTIVE(*)
C
      DF(1) = 2.0E0*(X(1)-2.0E0)
      DF(2) = 2.0E0*(X(2)-1.0E0)
C
      IF (ACTIVE(1)) THEN
         DG(1,1) = 1.0E0
         DG(1,2) = -2.0E0
      END IF
C
      IF (ACTIVE(2)) THEN
         DG(2,1) = -0.5E0*X(1)
         DG(2,2) = -2.0E0*X(2)
      END IF
      RETURN
      END
```

Output

```
The solution is
1   0.8229
2   0.9114
```

7

ΑΝΑΛΥΣΗ ΕΥΑΙΣΘΗΣΙΑΣ

Η πλέον διαδεδομένη μέθοδος ανάλυσης των κατασκευών σήμερα είναι η μέθοδος των πεπερασμένων στοιχείων και ειδικότερα αυτή που βασίζεται στη μέθοδο των μετακινήσεων. Η ανάπτυξη της μεθόδου σε προγράμματα, η πρόσβαση στα οποία δεν είναι ευχερής, θέτει ένα θέμα σχετικά με την χρησιμοποίηση της μεθόδου στον βέλτιστο σχεδιασμό των κατασκευών. Συνήθως επιλέγεται ο διαχωρισμός της ανάλυσης από την βελτιστοποίηση έτσι ώστε να είναι δυνατή η χρήση των έτοιμων προγραμμάτων πεπερασμένων στοιχείων.

7.1 ΕΥΑΙΣΘΗΣΙΑ

Οι μέθοδοι βελτιστοποίησης που βασίζονται στον μαθηματικό προγραμματισμό συνήθως ξεκινούν από κάποια υποψήφια λύση την οποία ακολουθώντας μία επαναληπτική διαδικασία βελτιώνουν μέχρι την επίτευξη ικανοποιητικής λύσης. Κατά την επιλογή της νέας κατεύθυνσης στο χώρο των λύσεων βασίζονται στον υπολογισμό των κλίσεων (gradients) ή ευαισθησιών (sensitivities) των συναρτήσεων της αντικειμενικής συνάρτησης και των περιορισμών.

Οι ευαισθησίες αυτές μπορούν να υπολογιστούν χρησιμοποιώντας αναλυτικές μεθόδους, ημιαναλυτικές και αριθμητικές μεθόδους. Η ακρίβεια των ευαισθησιών είναι ιδιαίτερης σημασίας ιδίως σε περιοχές όπου αλλάζουν πρόσημο. Οι αριθμητικές μέθοδοι είναι οι απλούστερες και στηρίζονται σε κάποιο αριθμητικό ορισμό της παραγώγου που ορίζει την ευαισθησία. Έτσι έχουμε εκφράσεις της ευαισθησίας που βασίζονται σε κεντρικές διαφορές αρχικές και τελικές διαφορές για την έκφραση των ευαισθησιών.

$$\frac{dU_i}{dx_j} \approx \frac{U_i\left(x_j + h\right) - U_i\left(x_j - h\right)}{2h} \tag{7.1}$$

Συχνά όμως ο αριθμητικός υπολογισμός των ευαισθησιών δεν είναι ακριβής ενώ πολλές φορές οδηγεί σε αποτελέσματα ακόμη και αντιθέτου προσήμου. Οι ανακρίβειες αυτές έχουν ως αποτέλεσμα τον αποπροσανατολισμό του αλγορίθμου βελτιστοποίησης από την σωστή πορεία σύγκλισης και την καθυστέρηση ή και αδυναμία εύρεσης λύσης.

Ο αναλυτικός προσδιορισμός των ευαισθησιών από την άλλη πλευρά συχνά οδηγεί σε περίπλοκες εκφράσεις λόγω ης πεπλεγμένης μορφής των εκφράσεων των περιορισμών. Η παραγώγιση με χρήση συμβολικών προγραμμάτων όπως η Mathematica το πρόγραμμα Maple ή άλλα αντίστοιχα είναι πάντοτε επιτυχής και επιπλέον άμεσα μετατρέψιμη σε κώδικα σε μία από τις ευρύτερα διαδεδομένες γλώσσες προγραμματισμού όπως Fortran, C κλπ.

Ιδιαίτερη θέση στην βιβλιογραφία της ανάλυσης ευαισθησίας κατέχουν οι ημι-αναλυτικές μέθοδοι οι οποίες παρέχουν τις ζητούμενες ευαισθησίες με ακρίβεια και οικονομία στους υπολογισμούς, ιδίως για τις συνήθεις περιπτώσεις που οι συναρτήσεις των περιορισμών είναι πεπλεγμένες ως προς τις μεταβλητές σχεδιασμού.

Ανεξάρτητα από την μαθηματική σημασία της ευαισθησίας της αντικειμενικής συνάρτησης και των περιορισμών ως προς τις μεταβλητές σχεδιασμού, ιδιαίτερη σημασία έχει να εκτιμηθεί η πληροφορία που περικλείουν οι ευαισθησίες κατά τον βέλτιστο σχεδιασμό. Ένας απλός τρόπος για να κατανοηθεί η σημασία τους είναι να τις ορίσουμε ως τον λόγο του ποσοστού μεταβολής της αντικειμενικής συνάρτησης ή των περιορισμών για κάποια ποσοστιαία μεταβολή της μεταβλητής σχεδιασμού. Έτσι κατ’αρχήν το πρόσημο της ευαισθησίας φανερώνει την αύξηση ή ελάττωση της αντικειμενικής συνάρτησης ή του περιορισμού για αύξηση της μεταβλητής σχεδιασμού ενώ το μέτρο της ευαισθησίας εκφράζει το μέγεθος της επίδρασης της μεταβολής της μεταβλητής σχεδιασμού στην αντικειμενική συνάρτηση ή τον περιορισμό.

7.2 Ανάλυση Ευαισθησίας

Ας θεωρήσουμε τον περιορισμό $g_i(\mathbf{x},U)$ ως συνάρτηση των μεταβλητών σχεδιασμού και του διανύσματος των μετακινήσεων σε ένα ενδιάμεσο σημείο της διαδικασίας βελτιστοποίησης. Χρησιμοποιώντας τον κανόνα της αλυσιδωτής παραγώγισης, η παράγωγος της g_i ως προς την j μεταβλητή σχεδιασμού είναι:

$$\frac{dg_i}{dx_j} = \frac{\partial g_i}{\partial x_j} + \frac{\partial g_i^T}{\partial U}\frac{dU}{dx_j} \tag{7.2}$$

όπου

$$\frac{dg_i}{dU} = \left[\frac{\partial g_i}{\partial U_1} \; \frac{\partial g_i}{\partial U_2} \; \cdots \; \frac{\partial g_i}{\partial U_l} \right]^T \tag{7.3}$$

και

$$\frac{dU}{dx_j} = \left[\frac{\partial U_1}{\partial x_j} \ \frac{\partial U_2}{\partial x_j} \ \cdots \ \frac{\partial U_l}{\partial x_j}\right]^T \tag{7.4}$$

για να υπολογιστεί η ευαισθησία ενός περιορισμού ως προς μία μεταβλητή σχεδιασμού, απαιτείται να υπολογιστούν οι μερικές παράγωγοι $\partial g_i / \partial x_j$ και $\partial g_i / \partial U$ καθώς και η παράγωγος του διανύσματος των μετακινήσεων dU/dx_j. Οι μερικές παράγωγοι $\partial g_i / \partial x_j$ και $\partial g_i / \partial U$ είναι σχετικά εύκολο να υπολογιστούν αφού η g_i είναι μη πεπλεγμένη συνάρτηση των **x** και **U**. Για να υπολογιστεί η ευαισθησία του διανύσματος των μετακινήσεων ως προς μία μεταβλητή σχεδιασμού, dU/dx_j, παραγωγίζουμε την εξίσωση ισορροπίας ως προς την μεταβλητή σχεδιασμού x_j:

$$\left[\frac{\partial K(x)}{\partial x_j}\right]\{U\} + [K(x)]\left\{\frac{dU}{dx_j}\right\} = \left\{\frac{\partial F}{\partial x_j}\right\} \Rightarrow$$

$$[K(x)]\left\{\frac{dU}{dx_j}\right\} = \left\{\frac{\partial F}{\partial x_j}\right\} - \left[\frac{\partial K(x)}{\partial x_j}\right]\{U\} \Rightarrow \tag{7.5}$$

$$[K(x)]\left\{\frac{dU}{dx_j}\right\} = \{P\} \ , \text{ όπου } \{P\} = \left\{\frac{\partial F}{\partial x_j}\right\} - \left[\frac{\partial K(x)}{\partial x_j}\right]\{U\}$$

Η παραπάνω εξίσωση χρησιμοποιείται για τον υπολογισμό της παραγώγου του διανύσματος των μετακινήσεων dU/dx_j. Η μορφή της εξίσωσης είναι ίδια με την εξίσωση ισορροπίας ενώ στη θέση του φορτίου υπάρχει το λεγόμενο *ψευδοφορτίο* που αποτελείται από δύο όρους, ο πρώτος προκύπτει παραγωγίζοντας το διάνυσμα του φορτίου ως προς την μεταβλητή σχεδιασμού και ο δεύτερος από το γινόμενο του τετραγωνικού μητρώου $\left[\frac{\partial K(x)}{\partial x_j}\right]$ που ονομάζεται ευαισθησία του μητρώου ακαμψίας ως προς την μεταβλητή σχεδιασμού j, επί το διάνυσμα των μετακινήσεων που προκύπτει από την επίλυση των εξισώσεων ισορροπίας. Η παράγωγος του μητρώου ακαμψίας $\partial K(x)/\partial x_j$ ως προς την μεταβλητή σχεδιασμού x_j υπολογίζεται εύκολα εάν είναι γνωστή η μη-πεπλεγμένη σχέση μεταξύ **K** και **x**. Η παραπάνω μορφή διατηρεί το μητρώο ακαμψίας του φορέα ως μητρώο συντελεστών του γραμμικού συστήματος, γεγονός που ανάγει τον υπολογισμό των ευαισθησιών του διανύσματος των μετακινήσεων σε επιλύσεις των εξισώσεων ισορροπίας για μία σειρά από ψευδοφορτία. Στη συνέχεια θα αναφερθούν οι διάφοροι τρόποι μόρφωσης της ευαισθησίας του μητρώου ακαμψίας ενός φορέα κατά τρόπο που να προσφέρεται για προγραμματισμό. Παρατηρούμε επίσης ότι η σχέση (7.5) πρέπει να λυθεί για κάθε μεταβλητή σχεδιασμού, που σημαίνει ότι για κάθε μεταβλητή σχεδιασμού απαιτείται ο υπολογισμός της ευαισθησίας του μητρώου ακαμψίας και με βάση την παραπάνω σχέση ο υπολογισμός ενός ψευδοφορτίου για κάθε μεταβλητή σχεδιασμού και όλα αυτά για κάθε βήμα της επαναληπτικής μεθόδου βελτιστοποίησης. Έτσι, αφού υπολογιστούν οι παράγωγοι dU/dx_j, μπορεί να προσδιοριστεί η κλίση (ευαισθησία) του περιορισμού στη σχέση (7.2). Η παραπάνω διαδικασία ακολουθείται στο παρακάτω παράδειγμα.

Εφαρμογή 1: Υπολογισμός ευαισθησιών πλαισίου με δύο μέλη.

Να υπολογιστεί η κλίση του περιορισμού τάσης $g_1(\mathbf{x},\mathbf{U})$ για το πλαίσιο δυο μελών του παραδείγματος της παραγράφου 7.1.2, στο σημείο (6.35, 6.35, 0.25).

<u>Λύση</u>

Οι μερικές παράγωγοι του περιορισμού της σχέσης (7.2) ως προς **x** και **U** είναι:

$$\frac{\partial g_1}{\partial x} = \frac{1}{\sigma_{\varepsilon\pi}^2}[2\sigma_1 \frac{\partial \sigma_1}{\partial x} + 6\tau \frac{\partial \tau}{\partial x}] \tag{7.4}$$

$$\frac{\partial g_1}{\partial U} = \frac{1}{\sigma_{\varepsilon\pi}^2}[2\sigma_1 \frac{\partial \sigma_1}{\partial U} + 6\tau \frac{\partial \tau}{\partial U}] \tag{7.5}$$

Χρησιμοποιώντας τις σχέσεις ακαμψίας, προκύπτουν οι μερικές παράγωγοι των σ_1 και τ ως προς τις μεταβλητές σχεδιασμού **x** και **U**.

<u>Μερικές παράγωγοι διατμητικής τάσης τ</u> : Παραγωγίζοντας την έκφραση της διατμητικής τάσης ως προς τις μεταβλητές σχεδιασμού **x**, παίρνουμε:

$$\frac{\partial \tau}{\partial x} = \frac{1}{2At}\frac{\partial T}{\partial x} - \frac{T}{2A^2 t}\frac{\partial A}{\partial x} - \frac{T}{2A^2 t}\frac{\partial t}{\partial x} \tag{7.6}$$

όπου οι μερικές παράγωγοι της στρεπτικής ροπής Τ ως προς τις μεταβλητές σχεδιασμού **x** είναι:

$$\frac{\partial T}{\partial x} = -\frac{GU_3}{L}\frac{\partial J}{\partial x} \tag{7.7}$$

όπου

$$\frac{\partial J}{\partial d} = \frac{4t(d-t)(h-t)^2(d+h-2t) - 2t(d-t)^2(h-t)^2}{(d+h-2t)^2} = 14.158$$

$$\frac{\partial J}{\partial h} = \frac{4t(d-t)^2(h-t)(d+h-2t) - 2t(d-t)^2(h-t)^2}{(d+h-2t)^2} = 14.158$$

$$\frac{\partial J}{\partial t} = \frac{2(d-t)^2(h-t)^2 - 4t(d-t)(h-t)^2 - 4t(d-t)^2(h-t)}{(d+h-2t)} -$$

$$-\frac{2t(d-t)^2(h-t)^2(-2)}{(d+h-2t)^2} = 198.218$$

Συνεπώς, η παράγωγος $\partial J \,/\, \partial x$ γράφεται με τη μορφή διανύσματος

$$\frac{\partial J}{\partial x} = \begin{bmatrix} 14.158 \\ 14.158 \\ 198.218 \end{bmatrix}$$

οπότε προκύπτει η σχέση

$$\frac{\partial T}{\partial x} = -\frac{80 \cdot 10^6}{2.5}(0.57275)\begin{bmatrix} 14.158 \\ 14.158 \\ 198.218 \end{bmatrix} = -(18.328)10^6 \begin{bmatrix} 14.158 \\ 14.158 \\ 198.218 \end{bmatrix}$$

Επίσης, πρέπει να υπολογιστούν οι ποσότητες

$\partial A / \partial x$ και $\partial t / \partial x$:

$$\frac{\partial A}{\partial x} = \begin{bmatrix} (h-t) \\ (d-t) \\ -(h-t)-(d-t) \end{bmatrix} = \begin{bmatrix} 6.1 \\ 6.1 \\ -11.7 \end{bmatrix} cm$$

$$\frac{\partial t}{\partial x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Έτσι, η παρακάτω εξίσωση δίνει την μερική παράγωγο της τ ως προς τις **x**

$$\frac{\partial \tau}{\partial x} = \frac{1}{2At}[\frac{\partial T}{\partial x} - \frac{T}{A}\frac{\partial A}{\partial x} - \frac{T}{t}\frac{\partial t}{\partial x}] = \begin{bmatrix} -4.59 \\ -4.59 \\ 9.60 \end{bmatrix} \cdot 10^6$$

Παραγωγίζοντας τώρα τη διατμητική τάση ως προς τις γενικευμένες μετακινήσεις **U,** παίρνουμε:

$$\frac{\partial \tau}{\partial U} = \frac{1}{2At}\frac{\partial T}{\partial U} \tag{7.8}$$

όπου

$$\frac{\partial T}{\partial U} = \begin{bmatrix} 0 \\ 0 \\ -GJ/L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -18.412 \end{bmatrix}$$

Συνεπώς, προκύπτει:

$$\frac{\partial \tau}{\partial U} = \begin{bmatrix} 0 \\ 0 \\ -(0.975) \cdot 10^6 \end{bmatrix}$$

<u>Μερικές παράγωγοι ορθών (καμπτικών) τάσεων:</u> Παραγωγίζοντας την τάση σ_1 της σχέσης ακαμψίας ως προς τις μεταβλητές σχεδιασμού **x**, παίρνουμε:

$$\frac{\partial \sigma_1}{\partial x} = \frac{h}{2I}\frac{\partial M_1}{\partial x} + \frac{M_1}{2I}\frac{\partial h}{\partial x} - \frac{M_1 h}{2I^2}\frac{\partial I}{\partial x} \tag{7.9}$$

όπου οι $\partial M_1 / \partial x, \ \ \partial I / \partial x, \ \ \partial h / \partial x$ είναι:

$$\frac{\partial M_1}{\partial x} = \frac{2E}{L^2}(-3U_1 + U_2 L)\frac{\partial I}{\partial x}$$

$$\frac{\partial I}{\partial d} = \frac{1}{12}[h^3 - (h-2t)^3] = 4.722$$

$$\frac{\partial I}{\partial h} = \frac{1}{4}[dh^2 - (d-2t)(h-2t)^2] = 14.167$$

$$\frac{\partial I}{\partial t} = \frac{(h-2t)^3}{6} + \frac{(h-2t)^2(d-t)}{2} = 132.921$$

$$\frac{\partial h}{\partial x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Αντικαθιστώντας τις παραπάνω ποσότητες στη σχέση (7.20) παίρνουμε

$$\frac{\partial \sigma_1}{\partial x} = \begin{bmatrix} 0 \\ (61.25)10^6 \\ 0 \end{bmatrix}$$

Παραγωγίζοντας την έκφραση της σ_1 της σχέσης ακαμψίας ως προς τις γενικευμένες μετακινήσεις **U**, παίρνουμε

$$\frac{\partial \sigma_1}{\partial U} = \frac{h}{2I}\frac{\partial M_1}{\partial U}$$

όπου η $\frac{\partial M_1}{\partial U}$ προκύπτει από την εξίσωση ακαμψίας ως εξής:

$$\frac{\partial M_1}{\partial U} = \frac{2EI}{L^2}\begin{bmatrix} -3 \\ L \\ 0 \end{bmatrix}$$

Συνεπώς, η $\frac{\partial \sigma_1}{\partial U}$ γίνεται:

$$\frac{\partial \sigma_1}{\partial U} = \frac{Eh}{L^2}\begin{bmatrix} -3 \\ L \\ 0 \end{bmatrix} = \begin{bmatrix} -6.1 \\ 5.1 \\ 0 \end{bmatrix} 10^6$$

Αντικαθιστώντας τα παραπάνω μεγέθη παίρνουμε τελικά τις μερικές παραγώγους των περιορισμών:

$$\frac{\partial g_1}{\partial x} = \begin{bmatrix} 206.77 \\ 6449.77 \\ -432.47 \end{bmatrix}$$

$$\frac{\partial g_2}{\partial U} = \begin{bmatrix} -621.75 \\ 519.83 \\ 43.92 \end{bmatrix}$$

Παράγωγοι των μετακινήσεων: Για τον υπολογισμό των παραγώγων των μετακινήσεων, χρησιμοποιούμε τη σχέση του ψευδοφορτίου (7.5). Στη σχέση αυτή, οι παράγωγοι $\partial F / \partial x_j$ είναι μηδενικές, αφού το διάνυσμα των φορτίων δεν εξαρτάται από τις μεταβλητές σχεδιασμού. Για τον υπολογισμό του άλλου όρου του δεξιού μέρους της (7.5), παραγωγίζουμε το μητρώο ακαμψίας ως προς κάθε μεταβλητή σχεδιασμού.

$$\left[\frac{\partial K(x)}{\partial x_j}\right]\{U\} = 10^3 \begin{bmatrix} -3.1214 & -9.3642 & -87.861 \\ -28.585 & 28.454 & 3.300 \\ 28.585 & -28.454 & -3.300 \end{bmatrix}, \qquad j = 1,2,3$$

όπου κάθε στήλη του μητρώου αντιστοιχεί σε κάθε μεταβλητή σχεδιασμού. Αφού το μητρώο ακαμψίας είναι γνωστό, από το σύστημα της σχέσης (7.5) προκύπτει η ευαισθησία του διανύσματος των μετακινήσεων dU/dx για τις τρεις μεταβλητές σχεδιασμού είναι:

$$\left[\frac{dU}{dx_j}\right] = \begin{bmatrix} 16.909 & 37.645 & 383.42 \\ 0.2443 & 0.4711 & 5.0247 \\ -0.2443 & -0.4711 & -5.0247 \end{bmatrix}, \qquad j = 1,2,3$$

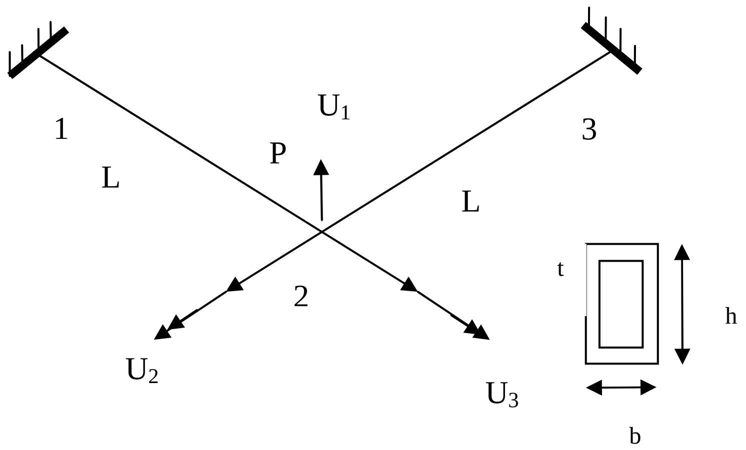
Τέλος, αντικαθιστώντας όλες τις ποσότητες στην εξίσωση (7.13), παίρνουμε τη ζητούμενη ευαισθησία του περιορισμού g_1

$$\left\{\frac{dg_1}{dx}\right\} = \begin{bmatrix} -141.55 \\ -202.87 \\ -3577.30 \end{bmatrix}$$

Παρατηρείστε ότι ο δοσμένος σχεδιασμός **x**=(6.35, 6.35, 0.25) παραβιάζει σοβαρά τον περιορισμό τάσης. Τα πρόσημα των όρων του διανύσματος $dg_1/d\mathbf{x}$ δείχνουν ότι όλες οι μεταβλητές σχεδιασμού πρέπει να αυξηθούν ώστε να μην παραβιάζεται ο περιορισμός. Επιπλέον δείχνουν ότι η αύξηση της τρίτης μεταβλητής σχεδιασμού έχει πολλαπλάσια επίδραση σε σχέση με τις άλλες δύο στην ικανοποίηση του περιορισμού.

Εφαρμογή 2: Σχεδιασμός εσχάρας δυο μελών

Ας θεωρήσουμε το πρόβλημα σχεδιασμού μιας εσχάρας δύο μελών το οποίο υφίσταται φορτία εκτός του επιπέδου της. Ζητείται να διατυπωθεί το πρόβλημα ελαχιστοποίησης του όγκου του πλαισίου σύμφωνα με κατασκευαστικούς περιορισμούς και περιορισμούς τάσης.



Λύση

Τα δύο μέλη του πλαισίου είναι ίδια, με κοίλη ορθογωνική διατομή. Το πρόβλημα έχει τρεις μεταβλητές σχεδιασμού: το πλάτος d, το ύψος h και το πάχος t (σχήμα 1). Έτσι, το διάνυσμα των μεταβλητών σχεδιασμού είναι **x**=(d,h,t).

Η αντικειμενική συνάρτηση του προβλήματος είναι ο όγκος της κατασκευής που είναι μη πεπλεγμένη συνάρτηση ως προς τις μεταβλητές σχεδιασμού:

$$f(x) = 2L(2dt + 2ht - 4t^2)$$

Τα μέλη υφίστανται συγχρόνως καμπτικές και στρεπτικές τάσεις, των οποίων οι μέγιστες τιμές είναι σ και τ αντίστοιχα. Το κριτήριο διαρροής του Mises δίνει: $\sqrt{\sigma^2 + 3\tau^2} \le \sigma_{\varepsilon\pi}$ όπου $\sigma_{\varepsilon\pi}$ είναι η επιτρεπόμενη τάση σχεδιασμού. Κανονικοποιώντας την παραπάνω ανισότητα, προκύπτει ο περιορισμός τάσης:

$$\frac{1}{\sigma_{\varepsilon\pi}^2}(\sigma^2 + 3\tau^2) - 1.0 \le 0 \tag{7.3}$$

Οι τάσεις σ και τ υπολογίζονται από τις καμπτικές και στρεπτικές ροπές στα άκρα των μελών, οι οποίες προκύπτουν χρησιμοποιώντας την μέθοδο πεπερασμένων στοιχείων. Οι τρεις γενικευμένες επικόμβιες μετακινήσεις για το μοντέλο πεπερασμένων στοιχείων του σχήματος είναι η κατακόρυφη μετακίνηση U_1 του κόμβου 2, η στροφή U_2 περί τον άξονα 3-2 και η στροφή U_3 περί τον άξονα 1-2. Η εφαρμογή της εξίσωσης ισορροπίας για το προσομοίωμα πεπερασμένων στοιχείων, δίνει τελικά:

$$\frac{EI}{L^3}\begin{bmatrix} 24 & -6L & 6L \\ -6L & (4L^2+\frac{GJ}{EI}L^2) & 0 \\ 6L & 0 & (4L^2+\frac{GJ}{EI}L^2) \end{bmatrix}\begin{Bmatrix} U_1 \\ U_2 \\ U_3 \end{Bmatrix}=\begin{Bmatrix} P \\ 0 \\ 0 \end{Bmatrix} \tag{7.4}$$

όπου E=200GPα , L=2.5m , G=80GPα , P=-45KN

$$I=\frac{1}{12}[dh^3-(d-2t)(h-2t)^3]$$

$$J=\frac{2t(d-t)^2(h-t)^2}{(d+h-2t)} \qquad \text{(στροφική ροπή αδρανείας)}$$

$$A=(d-t)(h-t)$$

Από την εξίσωση (7.5), το μητρώο ακαμψίας **K**(**x**) και το μητρώο φόρτισης μπορούν να υπολογιστούν. Έτσι, προσδιορίζονται οι μετακινήσεις U_1, U_2, U_3 ως πεπλεγμένες συναρτήσεις των μεταβλητών σχεδιασμού **x**. Η στρεπτική ροπή T και οι καμπτικές ροπές M_1 και M_2 στα άκρα 1 και 2 του μέλους 1-2, μπορούν πλέον να υπολογιστούν:

$$T=-\frac{GJ}{L}U_3 \tag{7.5}$$

$$M_1=\frac{2EI}{L^2}(-3U_1+U_2L) \tag{7.6}$$

$$M_2=\frac{2EI}{L^2}(-3U_1+2U_2L) \tag{7.7}$$

Τώρα, μπορεί να υπολογιστεί η στρεπτική τάση τ και οι καμπτικές τάσεις σ_1 και σ_2:

$$\tau=\frac{T}{2\eta t} \tag{7.8}$$

$$\sigma_1 = \frac{1}{2I} M_1 h \tag{7.9}$$

$$\sigma_2 = \frac{1}{2I} M_2 h \tag{7.10}$$

Έτσι, οι περιορισμοί τάσης στα σημεία 1 και 2 της σχέσης (7.3), γίνονται:

$$g_1(x,U) = \frac{1}{\sigma_{\varepsilon\pi}^2}(\sigma_1^2 + 3\tau^2) - 1.0 \leq 0 \tag{7.11}$$

$$g_2(x,U) = \frac{1}{\sigma_{\varepsilon\pi}^2}(\sigma_2^2 + 3\tau^2) - 1.0 \leq 0 \tag{7.12}$$

Παρατηρούμε ότι αφού οι ροπές T, M_1 και M_2 είναι πεπλεγμένες συναρτήσεις των μεταβλητών σχεδιασμού, και οι τάσεις θα είναι επίσης πεπλεγμένες συναρτήσεις. Οι τάσεις όμως είναι και μη πεπλεγμένες συναρτήσεις των μεταβλητών σχεδιασμού, όπως φαίνεται από τις παραπάνω σχέσεις (7.8) έως (7.10). Συνεπώς, οι περιορισμοί τάσης (7.11) και (7.12) είναι τόσο μη πεπλεγμένες όσο και πεπλεγμένες συναρτήσεις των μεταβλητών σχεδιασμού.

Τέλος, εκτός από τους περιορισμούς τάσης, τίθενται και οι παρακάτω περιορισμοί λόγω κατασκευαστικών απαιτήσεων για τις μεταβλητές σχεδιασμού:

$$2.5 \leq d \leq 10.0$$

$$2.5 \leq h \leq 10.0$$

$$0.1 \leq t \leq 1.0$$

Για να έχουμε μια αριθμητική εφαρμογή των παραπάνω, θα θεωρήσουμε ότι το διάνυσμα των μεταβλητών σχεδιασμού (μια λύση) είναι το σημείο x=(d,h,t)=(6.35, 6.35, 0.25) [cm]. Καταρχήν υπολογίζουμε τις παρακάτω ποσότητες:

$$I = \frac{1}{12}[6.35^4 - 5.84^4] = 38.43cm^4$$

$$J = \frac{1}{12.19}[2(0.25)(6.1)^2(6.1)^2] = 57.54cm^4$$

$A=(6.1)(6.1)=37.16cm^2$

$GJ=(80)10^6(57.54)10^{-8}=46.03KNm^2$

$EI=(200)10^6(38.43)10^{-8}=76.86KNm^2$

$$4L^2+\frac{GJ}{EI}L^2=(4+\frac{46.03}{76.86})(2.5)^2=28.74m^2$$

Χρησιμοποιώντας αυτά τα στοιχεία, η εξίσωση (7.4) γίνεται:

$$4.92\begin{bmatrix}24 & -15 & 15\\ -15 & 28.74 & 0\\ 15 & 0 & 28.74\end{bmatrix}\begin{Bmatrix}U_1\\ U_2\\ U_3\end{Bmatrix}=\begin{Bmatrix}-45\\ 0\\ 0\end{Bmatrix}$$

Λύνοντας το σύστημα, προκύπτουν οι τρεις γενικευμένες μετακινήσεις του κόμβου 2:

U_1=-1,1095m

U_2=-0.57275

U_3=0.57275

Οι σχέσεις (7.5) έως (7.7), αντικαθιστώντας τις U_1, U_2, U_3, γίνονται:

$$T=-\frac{46.03}{2.5}0.57275=-10.55\ kNm$$

$$M_1=\frac{2(76.86)}{2.5^2}(-3(-1.1095)+(-0.57275)2.5)=46.65kNm$$

$$M_2=\frac{2(76.86)}{2.5^2}(-3(-1.1095)+(-0.57275)5)=11.43kNm$$

Από τις σχέσεις (7.9) και (7.10), φαίνεται ότι η τάση σ_1 είναι μεγαλύτερη της σ_2, συνεπώς αρκεί να ελέγξουμε μόνο τον περιορισμό της σχέσης (7.11). Οι τάσεις σ_1 και τ είναι - από τις (7.8) και (7.9):

$$\tau=\frac{-10.55}{2(37.16)(0.25)10^{-6}}=-567.8MPa$$

$$\sigma_1=\frac{(46.65)(6.35)10^{-2}}{2(38.43)10^{-8}}=3854.1MPa$$

Θεωρώντας ότι η επιτρεπόμενη τάση $\sigma_{\varepsilon\pi}$ είναι 275MPa , ο περιορισμός τάσης της σχέσης (7.11) γίνεται:

$$g_1 = \frac{1}{275^2}[3854.1)^2 + 3(-567.8)^2] - 1.0 = 208.2 > 0$$

Συνεπώς, ο περιορισμός παραβιάζεται σοβαρά για τον σχεδιασμό **x**=(6.35,6.35,0.25)

Εφαρμογή 3:

Ας θεωρήσουμε το δικτύωμα τριών μελών του παρακάτω σχήματος (3-bar truss). Τα φορτία F_1 και F_2 αποτελούν δυο ανεξάρτητες φορτίσεις ίσες προς 40 kN. Επίσης δίδονται E=207 GPa, A_1=A_3=2.5 cm^2 και A_2=1.0 cm^2 ως το τρέχον σημείο του διανύσματος των μεταβλητών σχεδιασμού, ενώ οι επιτρεπόμενες τάσεις ορίζονται ως $\sigma^- = -10000N/cm^2$, $\sigma^+ = 14000N/cm^2$.

Αγνοώντας τους δεσμευμένους βαθμούς ελευθερίας που αντιστοιχούν στους κόμβους 2, 3 και 4, τα καθολικά μητρώα ακαμψίας των μελών προκύπτουν συναρτήσει των ελεύθερων βαθμών ελευθερίας του κόμβου 1 ως εξής:

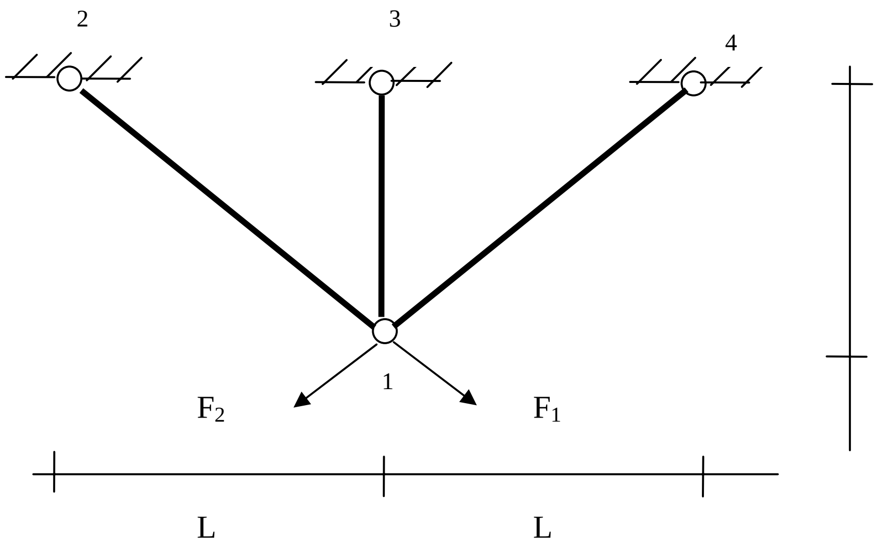
$$k_1 = \begin{bmatrix} \cos\frac{\pi}{4} \\ -\sin\frac{\pi}{4} \end{bmatrix} \frac{A_1 E}{L_1} [\cos\frac{\pi}{4} - \sin\frac{\pi}{4}] = \begin{bmatrix} 18300 & -18300 \\ -18300 & 18300 \end{bmatrix}$$

$$k_2 = \begin{bmatrix} 0 & 0 \\ 0 & 20700 \end{bmatrix}$$

$$k_3 = \begin{bmatrix} 18300 & 18300 \\ 18300 & 18300 \end{bmatrix}$$

Έτσι, το καθολικό μητρώο ακαμψίας του φορέα είναι:

$$k = k_1 + k_2 + k_3 = \begin{bmatrix} 36600 & 0 \\ 0 & 57300 \end{bmatrix}$$



Τα φορτία F_1 και F_2 αναλύονται στις συνιστώσες κατά τους άξονες του καθολικού συστήματος συντεταγμένων, και ορίζουν το μητρώο φόρτισης (το οποίο έχει δυο στήλες —μια για κάθε περίπτωση φόρτισης):

$$\begin{array}{cc} F_1 & F_2 \end{array}$$

$$[P] = \begin{bmatrix} -28300 & 28300 \\ -28300 & -28300 \end{bmatrix}$$

Τώρα, λύνοντας την εξίσωση ισορροπίας ως προς τις μετακινήσεις, παίρνουμε:

$$\begin{array}{cc} U_1 & U_2 \end{array}$$

$$[U] = [K]^{-1}[P] = \begin{bmatrix} \frac{1}{36600} & 0 \\ 0 & \frac{1}{57300} \end{bmatrix} \begin{bmatrix} -28300 & 28300 \\ -28300 & -28300 \end{bmatrix} = \begin{bmatrix} -0.7732 & 0.7732 \\ -0.4939 & -0.4939 \end{bmatrix}$$

Στη συνέχεια υπολογίσουμε τις μετακινήσεις στα τοπικά συστήματα συντεταγμένων των μελών. Θεωρώντας μόνο το μέλος 1, οι μετακινήσεις στον κόμβο 1 του τοπικού συστήματος, είναι:

$$\begin{array}{cc} LC1 & LC2 \end{array}$$

$$u_1^0 = \Lambda u_1 = [\cos\frac{\pi}{4} - \sin\frac{\pi}{4}] \begin{bmatrix} -0.7732 & 0.7732 \\ -0.4939 & -0.4939 \end{bmatrix} = [-0.1975 \quad 0.8960]$$

όπου LC: η περίπτωση φόρτισης (loading condition).

Έτσι, οι δυνάμεις που ασκούνται στο μέλος 1 είναι:

$$f_1 = \frac{A_1 E}{L_1} u_1^0 = 36600[-0.1975 \quad 0.8960] = [-7230 \quad 32790]$$

(LC1, LC2)

Συνεπώς, οι τάσεις του μέλους 1 (πρώτος δείκτης) από τις περιπτώσεις φορτίσεων 1 και 2 (δεύτερος δείκτης) είναι:

$\sigma_{11} = f_{11} / A_1 = -2890 \text{ N / cm}^2$

$\sigma_{12} = f_{12} / A_1 = 13120 \text{ N / cm}^2$

Γνωρίζοντας επίσης τις επιτρεπόμενες τάσεις (βλ. Σχ. 2), μπορούμε πλέον να εκφράσουμε τους (κανονικοποιημένους) περιορισμούς:

$$g_1 = \frac{\sigma_{11}}{\bar{\sigma}^-} - 1 = \frac{-2890}{-10000} - 1 = -0.711$$

$$g_2 = \frac{\sigma_{11}}{\bar{\sigma}^+} - 1 = \frac{-2890}{-14000} - 1 = -1.206$$

$$g_3 = \frac{\sigma_{12}}{\bar{\sigma}^-} - 1 = \frac{13120}{-10000} - 1 = -2.312$$

$$g_4 = \frac{\sigma_{12}}{\bar{\sigma}^+} - 1 = \frac{13120}{-14000} - 1 = -0.063$$

Με τον ίδιο τρόπο υπολογίζονται οι δυνάμεις, οι τάσεις και οι τιμές των περιορισμών για τα μέλη 2 και 3, δίνοντας ένα σύνολο 12 περιορισμών τάσεων για το πρόβλημα βελτιστοποίησης.

Όπως αναφέρθηκε και στην παράγραφο 7.1, η συνήθης αντικειμενική συνάρτηση που επιλέγεται είναι το βάρος του φορέα που υπολογίζεται ως το άθροισμα του βάρους των μελών του:

$$W = \sum_{i=1}^{3} \rho_i A_i L_i = 0.03[2.5(2)(\sqrt{2})(100) + 1.0(100)] = 24.2N$$

Στη συνέχεια θα υπολογίσουμε την κλίση (gradient) του περιορισμού g_4 (εφελκυστική τάση στο μέλος 1 υπό την κατάσταση φόρτισης 2) ως προς την μεταβλητή σχεδιασμού X_1 που είναι το εμβαδόν της διατομής του μέλους 1 (και του μέλους 3, λόγω συμμετρίας). Είναι λοιπόν:

$$\frac{\partial k_1}{\partial x_1} = \begin{bmatrix} 7320 & -7320 \\ -7320 & 7320 \end{bmatrix}$$

$$\frac{\partial k_3}{\partial x_1} = \begin{bmatrix} 7320 & 7320 \\ 7320 & 7320 \end{bmatrix}$$

Συνεπώς

$$\frac{\partial k}{\partial x_1} = \frac{\partial k_1}{\partial x_1} + \frac{\partial k_3}{\partial x_1} = \begin{bmatrix} 14640 & 0 \\ 0 & 14640 \end{bmatrix}$$

Δεδομένου ότι τα φορτία F δεν εξαρτώνται από τη μεταβλητή σχεδιασμού X_1, είναι ∂ **P** / ∂ $X_1 = 0$. Επίσης, αφού το καθολικό μητρώο ακαμψίας του φορέα **K** είναι διαγώνιο, το αντίστροφό του υπολογίζεται εύκολα αναλυτικά και είναι:

$$K^{-1} = \begin{bmatrix} \frac{1}{36600} & 0 \\ 0 & \frac{1}{57300} \end{bmatrix}$$

Από τη σχέση (7.14) παίρνουμε πλέον:

$$\frac{\partial U}{\partial x_1} = -\begin{bmatrix} \frac{1}{36600} & 0 \\ 0 & \frac{1}{57300} \end{bmatrix} \begin{bmatrix} 14620 & 0 \\ 0 & 14620 \end{bmatrix} \begin{bmatrix} 0.7732 \\ -0.4939 \end{bmatrix} = \begin{bmatrix} -0.3093 \\ 0.1262 \end{bmatrix}$$

Παραγωγίζοντας την εξίσωση $\mathbf{u}_1° = \mathbf{\Lambda}\, \mathbf{u}_1$ και ως προς X_1 παίρνουμε:

$$\frac{\partial u_1^0}{\partial x_1} = \Lambda \frac{\partial u_1}{\partial x_1} = -0.3079$$

Στη συνέχεια, παραγωγίζοντας την $f_{12} = (A_1\, E / L_1) \cdot \mathbf{u}_1°$ προκύπτει:

$$\frac{\partial f_{12}}{\partial x_1} = \frac{f_{12}}{A_1} + \frac{A_1 E}{L_1}\frac{\partial u_1^0}{\partial x_1} = \frac{32790}{2.5} + 36600(-0.3079) = 1847$$

Ομοίως, με παραγώγιση ως προς X_1 της $\sigma_{12} = f_{12} / A_1$ έχουμε:

$$\frac{\partial \sigma_{12}}{\partial x_1} = \frac{\partial}{\partial x_1}\frac{f_{12}}{A_1} = \frac{1}{A_1}\frac{\partial f_{12}}{\partial x_1} - \frac{f_{12}}{A_1^2} = \frac{1847}{2.5} - \frac{32790}{6.25} = -4508$$

και τελικά, η κλίση του περιορισμού g_4 ως προς την μεταβλητή σχεδιασμού X_1 είναι:

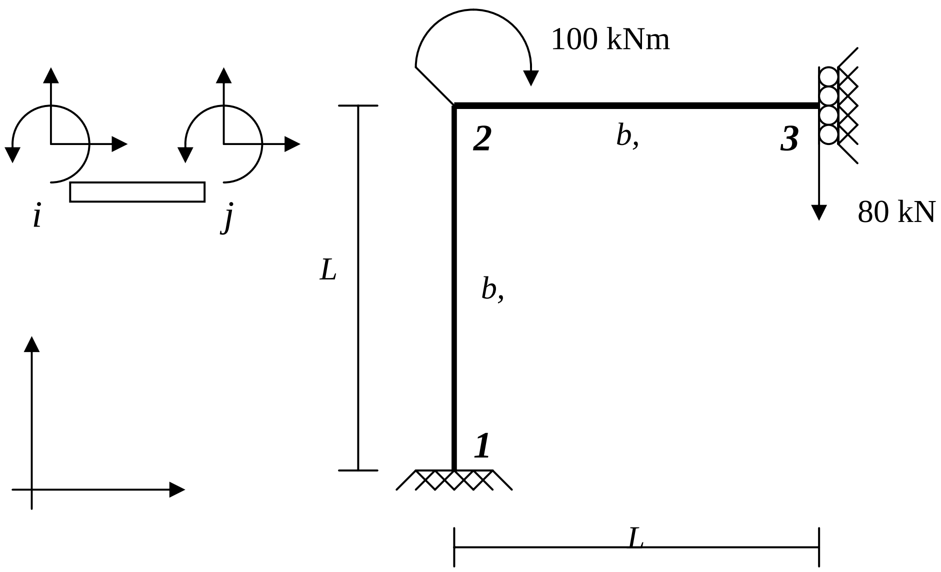
$$\frac{\partial g_4}{\partial x_1} = -\frac{4508}{14000} = -0.322$$

Άσκηση:

Το επίπεδο πλαίσιο του σχήματος με στύλο ορθογωνικής διατομής (b, h_1) και ζύγωμα (b, h_2) φορτίζεται με τη συγκεντρωμένη ροπή 100 kNm στο κόμβο 2 και το συγκεντρωμένο φορτίο P=80 kN στο κόμβο 3, ζητούνται:

Να υπολογιστεί η ευαισθησία του περιορισμού $g = \delta_{y3} - 0.07 \cdot h_2 < 0$ ως προς το ύψος της διατομής h_2, ακολουθώντας την άμεση μέθοδο υπολογισμού ευαισθησιών αριθμητικά στην ενδιάμεση θέση της διαδικασίας βελτιστοποίησης με στοιχεία:

$$b = 0.20m,\ h_1 = 0.4m,\ h_2 = 0.6m,\ kai \quad E = 2.0E8kN/m^2,\ L = 10m.$$



Δίδεται: Μητρώο ακαμψίας αμφίπακτου στοιχείου αμελώντας τις αξονικές παραμορφώσεις:

$$K^e = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}$$

7.3 ΣΥΖΥΓΗΣ ΜΕΘΟΔΟΣ ΥΠΟΛΟΓΙΣΜΟΥ ΕΥΑΙΣΘΗΣΙΩΝ

Στην παράγραφο αυτήν θα αναπτύξουμε τις δυο μεθόδους υπολογισμού των ευαισθησιών (sensitivities) ή κλίσεων των περιορισμών: την <u>άμεση μέθοδο</u> (direct metdod) που αναλύθηκε στην παράγραφο 7.2, και την συζυγή <u>μέθοδο </u>(dummy load method ή adjoint method). Οι δυο μέθοδοι θα εφαρμοστούν στην συνέχεια σε ένα παράδειγμα.

Οι εξισώσεις ισορροπίας ως προς τις επικόμβιες μετακινήσεις διατυπώνονται με τη σχέση (7.1), όπως είδαμε. Επίσης, ένας τυπικός περιορισμός εκφράζεται από τη σχέση (7.2). Χρησιμοποιώντας τον κανόνα της αλυσιδωτής παραγώγισης, παίρνουμε τη σχέση (7.13) την οποία ξαναγράφουμε:

$$\left\{\frac{dg}{dx}\right\} = \left\{\frac{\partial g}{\partial x}\right\} + \{z\}^T\left\{\frac{dU}{dx}\right\} \tag{7.21}$$

όπου $\{z\}$ είναι ένα διάνυσμα με στοιχεία

$$z_i = \frac{\partial g}{\partial U_i}$$

Ο πρώτος όρος του δεξιού μέλους της (7.21) είναι εύκολο να υπολογιστεί, συνεπώς μας ενδιαφέρει μόνο ο δεύτερος όρος. Κατά τα γνωστά, παραγωγίζοντας την εξίσωση (7.1) ως προς x, παίρνουμε:

$$[K(x)]\left\{\frac{dU}{dx_j}\right\} = \left\{\frac{\partial F}{\partial x_j}\right\} - \left[\frac{\partial K(x)}{\partial x_j}\right]\{U\} \Rightarrow \tag{7.22}$$

Πολλαπλασιάζοντας την ανωτέρω σχέση με την ποσότητα $\mathbf{z}^T\mathbf{K}^{-1}$, έχουμε:

$$\{z\}^T\left\{\frac{dU}{dx}\right\} = \{z\}^T[K]^{-1}\left(\left\{\frac{\partial F}{\partial x}\right\} - \left[\frac{\partial K}{\partial x}\right]\{U\}\right) \tag{7.23}$$

Ο υπολογισμός της ποσότητας $\mathbf{z}^T d\mathbf{U}/dx$ μπορεί να γίνει με δυο διαφορετικούς τρόπους. Ο πρώτος βασίζεται στην άμεση μέθοδο, κατά την οποία λύνεται η εξίσωση (7.22) ως προς d**U**/dx και στη συνέχεια πολλαπλασιάζεται με το διάνυσμα **z** (εσωτερικό γινόμενο). Ο δεύτερος τρόπος βασίζεται στην μέθοδο του συζυγούς φορτίου, κατά την οποία ορίζεται το διάνυσμα (adjoint vector) **λ** ως η λύση του συστήματος.

$$[K]\{\lambda\} = \{z\} \qquad \{\lambda\} = [K]^{-1}\{z\} \tag{7.24}$$

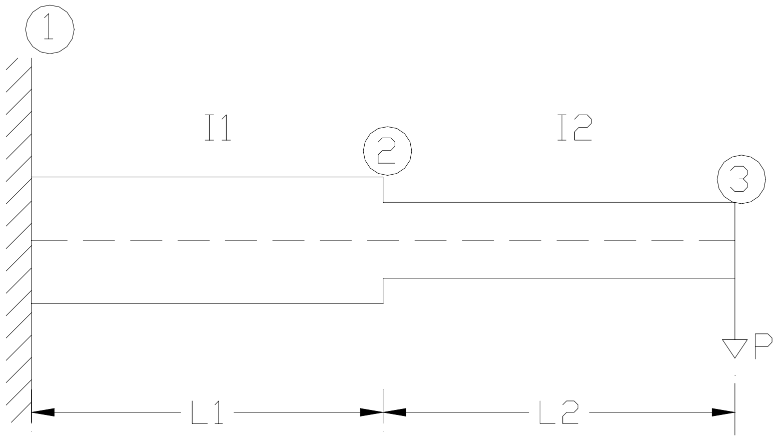
Αντικαθιστώντας τις (7.23) και (7.24) στην (7.21), έχουμε:

$$\frac{dg}{dx} = \frac{\partial g}{\partial x} + \lambda^T \left(\frac{\partial f}{\partial x} - \frac{dK}{dx} U\right) \tag{7.25}$$

Η επίλυση της (7.24) ως προς **λ** είναι αντίστοιχη με την επίλυση ως προς τις μετακινήσεις για διάνυσμα φόρτισης **z**, γι αυτό και η ονομασία της μεθόδου είναι dummy load method (μέθοδος συζυγούς φορτίου)- όπου το φορτίο είναι το διάνυσμα **z**.

Εφαρμογή 4:

Στο παράδειγμα αυτό, θα υπολογίσουμε την κλίση ή ευαισθησία (sensitivity) ενός περιορισμού, που τίθεται στην μετακίνηση του άκρου ενός προβόλου μεταβλητής διατομής, ως προς τη ροπή αδρανείας I_1 και το μήκος l_1 Εννοείται ότι τα μεγέθη I_1 και l_1 αποτελούν μεταβλητές σχεδιασμού του προβλήματος βελτιστοποίησης.



Ο περιορισμός της μετακίνησης στο άκρο είναι:

$$g = w_{tip} - c \le 0$$

Η υποχώρηση του άκρου υπολογίζεται αναλυτικά με τη βοήθεια της θεωρίας δοκού, και είναι:

$$w_{tip} = \frac{P}{3EI_1}(l_1^3 + 3l_1^2 l_2 + 3l_1 l_2^2) + \frac{Pl_2^3}{3EI_2}$$

Έτσι λαμβάνουμε εύκολα:

$$\frac{\partial g}{\partial I_1} = -\frac{P}{3EI_1^2}(l_1^3 + 3l_1^2 l_2 + 3l_1 l_2^2) \tag{7.26}$$

$$\frac{\partial g}{\partial l_1} = \frac{P}{3EI_1}(3l_1^2 + 6l_1 l_2 + 3l_2^2) = \frac{P}{EI_1}(l_1 + l_2)^2 \tag{7.27}$$

Το παραπάνω αποτέλεσμα που προέκυψε αναλυτικά με τη θεωρία δοκού, θα το επαληθεύσουμε τώρα χρησιμοποιώντας τις δυο μεθόδους υπολογισμού ευαισθησίας που αναπτύχθηκαν παραπάνω.

Άμεση Μέθοδος

Προηγείται η ανάλυση της κατασκευής με τη μέθοδο άμεσης ακαμψίας της Μητρωικής Στατικής ή γενικότερα της μεθόδου των πεπερασμένων στοιχείων. Η δοκός θεωρείται ότι αποτελείται από τρεις κόμβους και δύο μέλη, ενώ κάθε κόμβος διαθέτει έναν μεταφορικό w_i (μετατόπιση) και έναν στροφικό βαθμό ελευθερίας θ_i (στροφή). Τα μητρώα ακαμψίας είναι τα ακριβή μητρώα ακαμψίας της μητρωικής στατικής και έχουν ως εξής:

$$K^e = \frac{EI}{l^3}\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}$$

οπότε το καθολικό μητρώο ακαμψίας της δοκού είναι:

$$K = E\begin{bmatrix} 12(I_1/l_1^3 + I_2/l_2^3) & -6(I_1/l_1^2 - I_2/l_2^2) & -12I_2/l_2^3 & 6I_2/l_2^2 \\ & 4(I_1/l_1 + I_2/l_2) & -6I_2/l_2^2 & 2I_2/l_2 \\ & & 12I_2/l_2^3 & -6I_2/l_2^2 \\ & & & 4I_2/l_2 \end{bmatrix}$$

Το διάνυσμα φόρτισης είναι $\mathbf{F}=[0 \quad 0 \quad P \quad 0]^T$, οπότε η λύση του συστήματος (7.1) ως προς τις μετακινήσεις είναι:

$$\begin{Bmatrix} w_2 \\ \theta_2 \\ w_3 \\ \theta_3 \end{Bmatrix} = \frac{P}{E} \begin{Bmatrix} l_1^3 / 3I_1 + l_1^2 l_2 / 2I_1 \\ l_1^2 / 2I_1 + l_1 l_2 / I_1 \\ (l_1^3 + 3l_1^2 l_2 + 3l_1 l_2^2) / 3I_1 + l_2^3 / 3I_2 \\ l_1^2 / 2I_1 + l_1 l_2 / I_1 + l_2^2 / I_2 \end{Bmatrix}$$

Στη συνέχεια, χρησιμοποιώντας αναλυτικές μεθόδους, υπολογίζουμε τις παραγώγους $(\partial K / \partial I_1)U$ και $(\partial K / \partial l_1)U$ ως εξής:

$$\left[\frac{\partial K}{\partial I_1}\right]\{U\} = \frac{E}{l_1^3}\begin{bmatrix} 12 & -6l_1 & 0 & 0 \\ -6l_1 & 4l_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{Bmatrix} w_2 \\ \theta_2 \\ w_3 \\ \theta_3 \end{Bmatrix} = \frac{E}{l_1^3}\begin{Bmatrix} 12w_2 - 6l_1\theta_2 \\ -6l_1 w_2 + 4l_1^2\theta_2 \\ 0 \\ 0 \end{Bmatrix} = \frac{P}{I_1}\begin{Bmatrix} 1 \\ l_2 \\ 0 \\ 0 \end{Bmatrix}$$

όπου έχουν χρησιμοποιηθεί οι εκφράσεις των λύσεων w_2 και θ_2. Ομοίως, είναι:

$$\frac{\partial K}{\partial l_1}U = \frac{EI_1}{l_1^4}\begin{bmatrix} -36 & 12l_1 & 0 & 0 \\ 12l_1 & -4l_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} w_2 \\ \theta_2 \\ w_3 \\ \theta_3 \end{bmatrix} = 4\frac{EI_1}{l_1^4}\begin{bmatrix} -9w_2 + 3l_1\theta_2 \\ 3l_1 w_2 - l_1^2\theta_2 \\ 0 \\ 0 \end{bmatrix} = \frac{P}{l_1}\begin{bmatrix} -6(1 + l_2 / l_1) \\ 2(l_1 + l_2) \\ 0 \\ 0 \end{bmatrix}$$

Χρησιμοποιώντας την άμεση μέθοδο (direct method), παίρνουμε από τη σχέση (7.22) για κάθε μεταβλητή σχεδιασμού:

$$\frac{\partial U}{\partial I_1} = K^{-1}\left(\frac{\partial F}{\partial I_1} - \frac{\partial K}{\partial I_1}U\right) \Rightarrow$$

$$\frac{\partial}{\partial I_1}\begin{bmatrix} w_2 \\ \theta_2 \\ w_3 \\ \theta_3 \end{bmatrix} = K^{-1}\begin{bmatrix} P / I_1 \\ Pl_2 / I_1 \\ 0 \\ 0 \end{bmatrix} = -\frac{P}{EI_1^2}\begin{bmatrix} l_1^2 l_2 / 2 + l_1^3 / 3 \\ l_1 l_2 + l_1^2 / 2 \\ l_1^2 l_2 + l_1 l_2^2 + l_1^3 / 3 \\ l_1 l_2 + l_1^2 / 2 \end{bmatrix}$$

$$\frac{\partial U}{\partial l_1} = K^{-1}\left(\frac{\partial F}{\partial l_1} - \frac{\partial K}{\partial l_1}U\right) \Rightarrow$$

$$\frac{\partial}{\partial l_1}\begin{bmatrix} w_2 \\ \theta_2 \\ w_3 \\ \theta_3 \end{bmatrix} = K^{-1}\begin{bmatrix} -(6P/l_1)(1+l_2/l_1) \\ (2Pl_1)(l_1+l_2) \\ 0 \\ 0 \end{bmatrix} = \frac{P}{EI_1}\begin{bmatrix} l_1^2 + l_1 l_2 \\ l_1 + l_2 \\ (l_1+l_2)^2 \\ l_1 + l_2 \end{bmatrix}$$

Συνεπώς η ευαισθησία του περιορισμού ως προς τις δύο μεταβλητές σχεδιασμού είναι:

$$\frac{\partial g}{\partial I_1} = \frac{\partial w_{tip}}{\partial I_1} = \frac{\partial w_3}{\partial I_1} = -\frac{P}{EI_1^2}(l_1^2 l_2 + l_1 l_2^2 + l_1^3/3)$$

$$\frac{\partial g}{\partial l_1} = \frac{\partial w_{tip}}{\partial l_1} = \frac{\partial w_3}{\partial l_1} = -\frac{P}{EI_1}(l_1 + l_2)^2$$

Τα αποτελέσματα συμφωνούν με τις σχέσεις (7.26) και (7.27) που προέκυψαν αναλυτικά με τη θεωρία δοκού.

Συζυγής Μέθοδος (adjoint Method)

Το διάνυσμα z ορίζεται ως εξής:

$$z^T = \partial g / \partial U = \partial w_{tip} / \partial U = [0 \quad 0 \quad -1 \quad 0] \quad ,$$

οπότε μπορούμε να υπολογίσουμε το διάνυσμα **λ**:

$$\{\lambda\} = [K]^{-1}\{z\} = [K]^{-1}\begin{Bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{Bmatrix} = -\frac{1}{E}\begin{Bmatrix} l_1^3/3I_1 + l_1^2 l_2/2I_1 \\ l_1^2/2I_1 + l_1 l_2/I_1 \\ (l_1^3 + 3l_1^2 l_2 + 3l_1 l_2^2)/3I_1 + l_2^3/3I_2 \\ l_1^2/2I_1 + l_1 l_2/I_1 + l_2^2/I_2 \end{Bmatrix}$$

Έτσι, η εξίσωση (7.25) γράφεται:

$$\frac{\partial g}{\partial I_1} = -\{\lambda\}^T\left[\frac{\partial K}{\partial I_1}\right]\{U\} = -\frac{P}{EI_1}\left(\frac{l_1^3}{3I_1} + \frac{l_1^2 l_2}{2I_1} + \frac{l_1^2 l_2}{2I_1} + \frac{l_1 l_2^2}{I_1}\right) = -\frac{P}{EI_1^2}\left(l_1^2 l_2 + l_1 l_2^2 + \frac{l_1^3}{3}\right)$$

$$\frac{\partial g}{\partial l_1} = -\{\lambda\}^T\left[\frac{\partial K}{\partial I_1}\right]\{U\} = \frac{P}{El_1}(l_1 + l_2)\left(-\frac{2l_1^2}{I_1} - \frac{3l_1}{I_1} + \frac{l_1^2}{I_1} + \frac{2l_1 l_2}{I_1}\right) = \frac{P}{EI_1}(l_1 + l_2)^2$$

Η διαφορά στον υπολογιστικό φόρτο που απαιτείται στις δυο μεθόδους εξαρτάται από τον αριθμό των περιορισμών και των μεταβλητών σχεδιασμού του προβλήματος. Η άμεση μέθοδος απαιτεί την επίλυση της εξίσωσης (7.22) για κάθε μεταβλητή σχεδιασμού, όταν η συζυγής μέθοδος απαιτεί την επίλυση της (7.24) για κάθε περιορισμό. Επομένως, η άμεση μέθοδος είναι πιο αποδοτική όταν ο αριθμός των μεταβλητών σχεδιασμού είναι μικρότερος από τον αριθμό των περιορισμών. Αντίθετα, η συζυγής μέθοδος είναι αποδοτικότερη όταν ο αριθμός των περιορισμών είναι μικρότερος από τον αριθμό των μεταβλητών σχεδιασμού.

Στις συνήθεις εφαρμογές, σημαντικό ρόλο παίζει επίσης και ο αριθμός των ανεξάρτητων φορτίσεων που ασκούνται στην κατασκευή για κάθε μία από τις οποίες ορίζεται ένα διαφορετικό διάνυσμα μετακινήσεων. Ο υπολογιστικός φόρτος για τον υπολογισμό των ευαισθησιών σύμφωνα με την άμεση μέθοδος είναι ανάλογος με τον αριθμό των φορτίσεων, καθόσον το διάνυσμα μετακινήσεων συμμετέχει στον ορισμό του ψευδοφορτίου. Για την συζυγή μέθοδο, όπως προκύπτει από τις παραπάνω σχέσεις το διάνυσμα των μετακινήσεων παρεμβαίνει στον τελικό πολλαπλασιασμό. Επομένως, για φορείς που υπόκεινται σε πολλές ανεξάρτητες φορτίσεις προσφέρεται η χρήση της συζυγούς μεθόδου (adjoint method). Στις περιπτώσεις που οι περιορισμοί αναφέρονται σε συνδυασμούς φορτίσεων όπως αυτοί ορίζονται στους διάφορους κανονισμούς, η όλη ανάλυση ευαισθησίας πραγματοποιείται με βάση τις επιμέρους ανεξάρτητες φορτίσεις και η τελική έκφραση του περιορισμού ορίζεται ως γραμμικός συνδυασμός των επιμέρους εκφράσεων.

8

ΜΕΘΟΔΟΙ ΕΣΩΤΕΡΙΚΟΥ ΣΗΜΕΙΟΥ - Η ΚΕΝΤΡΙΚΗ ΔΙΑΔΡΟΜΗ

Στο κεφάλαιο αυτό θα παρουσιαστεί μία διαφορετική μέθοδος επίλυσης του προβλήματος του γραμμικού προγραμματισμού που ακολουθεί κάποια διαδρομή (path following method) που ανήκει στην κατηγορία των μεθόδων εσωτερικού σημείου (interior point methods). Οι μέθοδοι αυτές κινούνται σε μία ακολουθία εσωτερικών σημείων για να προσδιορίσουν την βέλτιστη λύση του προβλήματος γραμμικού προγραμματισμού. Για την παρουσίαση της μεθόδου απαιτείται η διερεύνηση της κεντρικής διαδρομής (central path) η οποία απαιτεί τη διερεύνηση του προβλήματος με εμπόδια (barrier problem).

Το πρωτογενές πρόβλημα γραμμικού προγραμματισμού ορίζεται ως εξής:

Μεγιστοποιείστε τη συνάρτηση: $c^T x$

που υπόκειται στους περιορισμούς: $Ax \leq b$

$$x \geq 0$$

Το παραπάνω πρόβλημα μπορεί να διατυπωθεί με την μορφή ισοτικών περιορισμών με την εισαγωγή πρόσθετων ελλειμματικών μεταβλητών w ως εξής:

$$c^T x$$

$$Ax + w = b$$

$$x, w \geq 0$$

Αντίστοιχα το δυικό πρόβλημα ορίζεται ως εξής:

Ελαχιστοποιείστε τη συνάρτηση $b^T y$

που υπόκειται στους περιορισμούς: $A^T y \geq c$

και $y \geq 0$

Το παραπάνω πρόβλημα μετά την εισαγωγή των πλεονασματικών μεταβλητών μετατρέπεται σε πρόβλημα με περιορισμούς ισότητας ως εξής:

Ελαχιστοποιείστε τη συνάρτηση: $b^T y$

που υπόκειται στους περιορισμούς: $A^T y - z = c$

και $y, z \geq 0$

Για την διατύπωση του προβλήματος χωρίς τους περιορισμούς μη αρνητικών μεταβλητών αρκεί για κάθε μεταβλητή να οριστεί μία συνάρτηση η οποία για αρνητικές τιμές της μεταβλητής να απειρίζεται αρνητικά, για θετικές τιμές να λαμβάνει πεπερασμένες τιμές και καθώς προσεγγίζει το μηδέν να τείνει προς $-\infty$. Η θεώρηση αυτή επιτρέπει την συνεχή μεταβολή της αντικειμενικής συνάρτησης και την ευχερέστερη μελέτη της.

Το πρόβλημα με εμπόδια ορίζεται ως εξής:

Μεγιστοποιείστε την συνάρτηση: $c^T x + \mu \sum_{i=1}^{n} \log x_i + \mu \sum_{j=1}^{m} \log w_j$

που υπόκειται στους περιορισμούς: $Ax + w = b$

Χωρίς τους περιορισμούς μη αρνητικότητας για τις μεταβλητές.

Το παραπάνω πρόβλημα είναι πρόβλημα μη γραμμικού προγραμματισμού καθόσον η αντικειμενική συνάρτηση είναι μη γραμμική συνάρτηση των μεταβλητών του προβλήματος.

Έχοντας κατά νου την γεωμετρική θεώρηση του προβλήματος του γραμμικού προγραμματισμού το οποίο στην γενικότερη διατύπωση του ορίζει ένα πολυεδρικό χώρο λύσεων κάθε έδρα του οποίου αντιστοιχεί στον μηδενισμό μίας μεταβλητής, η συνάρτηση με εμπόδια λαμβάνει τιμές μείον άπειρο επί των εδρών αυτών, ενώ λαμβάνει πεπερασμένες τιμές στο εσωτερικό χωρίο. Έτσι για διάφορες τιμές της παραμέτρου μ η μέγιστη τιμή της αντικειμενικής συνάρτησης ορίζει ένα εσωτερικό σημείο. Ο γεωμετρικός τόπος των σημείων αυτών για τιμές της παραμέτρου μ που μειώνονται μονοτονικά και τείνουν στο μηδέν διαγράφουν μία διαδρομή η οποία ορίζεται ως κεντρική διαδρομή η οποία συγκλίνει στη κορυφή του πολυέδρου που αποτελεί την βέλτιστη λύση.

Το πρόβλημα με εμπόδια στη διατύπωσή του με περιορισμούς ισότητας είναι δυνατόν να θεωρηθεί με βάση την θεωρία των πολλαπλασιαστών Lagrange ως εξής:

$$L(x,w,\lambda) = c^T x + \mu\sum_{i=1}^{n}\log x_i + \mu\sum_{j=1}^{m}\log w_j + y^T(b - Ax - w) \tag{8.1}$$

όπου y οι συντελεστές Lagrange.

Με βάση την παραπάνω συνάρτηση καθορίζονται οι συνθήκες βέλτιστης λύσης πρώτης τάξεως ως εξής:

$$\frac{\partial L}{\partial x_i} = c_i + \mu\frac{1}{x_j} - \sum_{j=1}^{m} y_j a_{ij} = 0, \qquad i = 1, 2, \ldots, n$$

$$\frac{\partial L}{\partial w_j} = \mu\frac{1}{w_j} - y_j = 0, \qquad j = 1, 2, \ldots, m$$

$$\frac{\partial L}{\partial y_j} = b_j - \sum_{i=1}^{n} a_{ij}x_i - w_j = 0, \qquad j = 1, 2, \ldots, m$$

Οι παραπάνω εξισώσεις μπορούν να γραφούν σε μητρωική μορφή ως εξής:

$$A^T y - \mu X^{-1} e = c$$

$$y = \mu W^{-1} e$$

$$Ax + w = b$$

όπου X υποδηλώνει το τετραγωνικό μητρώο με διαγώνιους όρους x_i και W το τετραγωνικό μητρώο με διαγώνιους όρους w_j, ενώ το διάνυσμα e υποδηλώνει το διάνυσμα με μονάδες σε όλες τις θέσεις.

Εισάγοντας επιπλέον το διάνυσμα $z = \mu X^{-1} e$ οι συνθήκες που εξασφαλίζουν την βέλτιστη λύση γράφονται ως εξής:

$$Ax + w = b$$

$$A^T y - z = c$$

$$z = \mu X^{-1} e$$

$$y = \mu W^{-1} e$$

και πολλαπλασιάζοντας την τρίτη και τέταρτη σχέση με X και W αντίστοιχα λαμβάνουμε τις σχέσεις:

$$Ax + w = b$$

$$A^T y - z = c$$

$$XZe = \mu e$$

$$YWe = \mu e$$

Παρατηρούμε ότι η πρώτη εξίσωση είναι η μητρωική εξίσωση των περιορισμών ισότητας του πρωτογενούς προβλήματος ενώ η δεύτερη αυτή του δυικού. Η τρίτη και τέταρτη εξίσωση μπορούν επίσης να γραφούν αναλυτικά ως εξής:

$$x_i z_i = \mu \qquad i = 1, 2, \ldots, n \tag{8.2α}$$

$$y_j w_j = \mu \qquad j = 1, 2, \ldots, m \tag{8.2β}$$

οι οποίες συνδέονται στενά με τις σχέσεις συμπληρωματικότητας (complementarity) στην βέλτιστη λύση και αναφέρονται ως σχέσεις μ συμπληρωματικότητας.

Οι παραπάνω σχέσεις ορίζουν ένα σύστημα $2n + 2m$ εξισώσεων ως προς $2n + 2m$ αγνώστους με μη γραμμική σύζευξη που ορίζουν η τρίτη και τέταρτη εξίσωση, ενώ οι δύο πρώτες είναι γραμμικές εξισώσεις.

Με βάση τις παραπάνω διαπιστώσεις είναι δυνατό να προκύψει ένας αλγόριθμος επίλυσης του προβλήματος του γραμμικού προγραμματισμού ακολουθώντας μία διαδρομή από εσωτερικά σημεία. Χαρακτηριστικό του αλγορίθμου ότι σε αντίθεση με τη μέθοδο Simplex που απαιτεί δύο φάσεις στη γενική περίπτωση ο προτεινόμενος αλγόριθμος είναι μία και μόνον φάσης.

Έτσι ξεκινώντας από οποιεσδήποτε θετικές τιμές για τις πρωτογενείς και τις δυικές μεταβλητές δηλ. $(x, w, y, z) > 0$ ακολουθώντας μία επαναληπτική διαδικασία καταλήγει στην βέλτιστη λύση.

Τα βήματα του αλγορίθμου είναι τα εξής:

1. Εκτίμηση της παραμέτρου μ

2. Υπολογισμός των κατευθύνσεων του βήματος $(\Delta x,\ \Delta w,\ \Delta y,\ \Delta z)$ που δείχνει προς το σημείο της κεντρικής διαδρομής $(x_\mu,\ w_\mu,\ y_\mu,\ z_\mu)$.

3. Υπολογισμός της παραμέτρου του μήκους του βήματος θ έτσι ώστε το νέο σημείο

$$\tilde{x} = x + \theta\ \Delta x$$

$$\tilde{w} = w + \theta\ \Delta w \quad \tilde{y} = y + \theta\ \Delta y \quad \tilde{z} = z + \theta\ \Delta z$$

να διατηρεί θετικούς όλους τους όρους

4. Αντικατάσταση του παλαιού σημείου $(x,\ w,\ y,\ z)$ με το νέο $(\tilde{x},\ \tilde{w},\ \tilde{y},\ \tilde{z})$.

Ο υπολογισμός των μεταβολών του δευτέρου βήματος γίνεται κατά τρόπο ώστε το νέο σημείο $(x + \Delta x,\ w + \Delta w,\ y + \Delta y,\ z + \Delta z)$ προσεγγίζει την πρωτογενή-δυική κεντρική διαδρομή στο σημείο $(x_\mu,\ w_\mu,\ y_\mu,\ z_\mu)$.

Οι εξισώσεις που ορίζουν την κεντρική διαδρομή είναι:

$$Ax + w = b$$

$$A^T y - z = c$$

$$XZe = \mu e$$

$$YWe = \mu e$$

Αν το νέο σημείο πρέπει να βρίσκεται ακριβώς επί της κεντρικής διαδρομής που αντιστοιχεί στην τιμή της παραμέτρου μ, τότε θα πρέπει να ισχύει:

$$A(x + \Delta x) + (w + \Delta w) = b$$

$$A^T(y + \Delta y) - (z + \Delta z) = c$$

$$(X + \Delta X)(Z + \Delta Z)e = \mu e$$

$$(Y + \Delta Y)(W + \Delta W)e = \mu e$$

Θεωρώντας τα (x, w, y, z), ως δεδομένα και τα $(\Delta x, \Delta w, \Delta y, \Delta z)$ ως αγνώστους οι παραπάνω εξισώσεις γράφονται ως εξής:

$$A\Delta x + \Delta w = b - Ax - w = \rho$$

$$A^T \Delta y - \Delta z = c - A^T y + z = \sigma$$

$$Z\Delta x + X\Delta z + \Delta X \Delta Z e = \mu e - XZe$$

$$W\Delta y + Y\Delta w + \Delta Y \Delta W e = \mu e - YWe$$

Οι παραπάνω εξισώσεις διατυπώνουν ένα σύστημα $2n + 2m$ μη γραμμικών εξισώσεων από το οποίο όμως αν παραλειφθούν οι μη γραμμικοί όροι ως προς τις μεταβλητές προκύπτει το παρακάτω γραμμικό σύστημα ως προς τις άγνωστες μεταβολές.

$$A\Delta x + \Delta w = b - Ax - w = \rho$$

$$A^T \Delta y - \Delta z = c - A^T y + z = \sigma$$

$$Z\Delta x + X\Delta z = \mu e - XZe$$

$$W\Delta y + Y\Delta w = \mu e - YWe$$

Οι παραπάνω εξισώσεις συμπίπτουν και με τη μέθοδο Newton για την επίλυση του συστήματος των μη γραμμικών εξισώσεων καθόσον:

$$F(\xi + \Delta\xi) \approx F(\xi) + F'(\xi)\Delta\xi + \ldots = 0$$

και άρα $F'(\xi)\Delta\xi = -F(\xi)$

οπότε για την περίπτωση των μη γραμμικών εξισώσεων (1.12) το διάνυσμα ξ είναι:

$$\xi = \begin{Bmatrix} x \\ w \\ y \\ z \end{Bmatrix} \text{ και } F(\xi) = \begin{Bmatrix} Ax + w - b \\ A^T y - z - c \\ XZe - \mu e \\ YWe - \mu e \end{Bmatrix}$$

από όπου παρατηρούμε ότι οι εξισώσεις που ορίζουν το σημείο της κεντρικής διαδρομής $(x_\mu, y_\mu, w_\mu, z_\mu)$ αποτελούν την ρίζα της διανυσματικής συνάρτησης F. Το μητρώο των παραγώγων της συνάρτησης ως προς τις μεταβλητές του προβλήματος είναι:

$$F'(\xi)=\begin{bmatrix} A & 0 & I & 0 \\ 0 & A^T & 0 & -I \\ Z & 0 & 0 & X \\ 0 & W & Y & 0 \end{bmatrix} \text{ και } \Delta\xi=\begin{Bmatrix} \Delta x \\ \Delta y \\ \Delta w \\ \Delta z \end{Bmatrix}$$

οπότε και η εξίσωση Newton γίνεται:

$$F'(\xi)\Delta\xi=\begin{bmatrix} A & 0 & I & 0 \\ 0 & A^T & 0 & -I \\ Z & 0 & 0 & X \\ 0 & W & Y & 0 \end{bmatrix}\begin{Bmatrix} \Delta x \\ \Delta y \\ \Delta w \\ \Delta z \end{Bmatrix}=-F(\xi)=-\begin{Bmatrix} Ax+w-b \\ A^T y-z-c \\ XZe-\mu e \\ YWe-\mu e \end{Bmatrix}$$

που συμπίπτει με την γραμμικοποιημένη σχέση (1.13).

8.1 Εκτίμηση της κατάλληλης τιμής της παραμέτρου μ

Αν το σημείο (x, w, y, z) ανήκει στην κεντρική διαδρομή τότε η τιμή της παραμέτρου μ που αντιστοιχεί σε αυτό προκύπτει από τις σχέσεις μ-συμπληρωματικότητας (1.11). Έτσι μπορεί να εκτιμηθεί από την σχέση $x_i z_i$ για συγκεκριμένο i, ή από την σχέση $y_j w_j$ για συγκεκριμένο δείκτη j. Επειδή όμως το σημείο (x, w, y, z) δεν βρίσκεται επί της κεντρικής διαδρομής για κάθε δείκτη i και j γενικά προκύπτει διαφορετική τιμή οπότε ως εκτιμώμενη τιμή μπορεί να χρησιμοποιηθεί ο μέσος όρος:

$$\mu=\frac{z^T x+y^T w}{n+m}$$

και επειδή ο αλγόριθμος αναζητά ένα σημείο πλησιέστερα προς το βέλτιστο σε σχέση με το εξεταζόμενο σημείο η τιμή της παραμέτρου δίνεται από τη σχέση:

$\mu=\delta\dfrac{z^T x+y^T w}{n+m}$ όπου δ καθαρός αριθμός μεταξύ μηδέν και ένα. Συνήθως χρησιμοποιείται η τιμή $\delta=1/10$ που δίνει καλά αποτελέσματα.

8.1.1 Επιλογή της παραμέτρου μήκους βήματος

Για τις ανάγκες της μεθόδου Newton το επόμενο σημείο της επαναληπτικής διαδικασίας μπορεί να είναι το $\left(x + \Delta x,\, w + \Delta w,\, y + \Delta y,\, z + \Delta z\right)$. Ενδέχεται όμως κάποιες ή και όλες οι μεταβλητές να μην διατηρήσουν το θετικό τους πρόσημο συνθήκη που πρέπει να διατηρείται κατά την επαναληπτική διαδικασία. Τούτο εξασφαλίζεται αν το επόμενο σημείο είναι το $\left(x + \theta\Delta x,\, w + \theta\Delta w,\, y + \theta\Delta y,\, z + \theta\Delta z\right)$ όπου θ μπορεί να είναι μικρότερο της μονάδος.

Για να προσδιοριστεί η τιμή του θ που να εξασφαλίζει θετικές τιμές για όλες τις μεταβλητές θα πρέπει να ισχύει:

$x_i + \theta\Delta x_i > 0, \qquad i = 1, 2, \ldots, n$ ή διαιρώντας δια $x_i > 0$ και $\theta > 0$

$\dfrac{1}{\theta} > -\dfrac{\Delta x_i}{x_i}, \qquad i = 1, 2, \ldots, n$ και άρα για όλες τις μεταβλητές να ισχύει το ίδιο οπότε:

$$\frac{1}{\theta} > \max_{i,j}\left\{-\frac{\Delta x_i}{x_i}, -\frac{\Delta w_j}{w_j}, -\frac{\Delta y_j}{y_j}, -\frac{\Delta z_i}{z_i},\right\}$$

Παρόλα αυτά το θ μπορεί να μη διατηρήσει την καθαρή ανισότητα και γιαυτό πρέπει να χρησιμοποιηθεί η παρακάτω σχέση

$$\theta > \min\left[r\left(\max_{i,j}\left\{-\frac{\Delta x_i}{x_i}, -\frac{\Delta w_j}{w_j}, -\frac{\Delta y_j}{y_j}, -\frac{\Delta z_i}{z_i},\right\}\right)^{-1},\ 1\right]$$

όπου r μία παράμετρος λίγο μικρότερη της μονάδος.

Σύμφωνα με τα παραπάνω ο αλγόριθμος που αντιστοιχεί στην μέθοδο εσωτερικής διαδρομής περιγράφεται στον εξής ψευδοκώδικα.

8.2 Αλγόριθμος Εσωτερικής διαδρομής

Αρχική επιλογή $(x,\ w,\ y,\ z) > 0$

While (μη βέλτιστη λύση) $\{$υπολόγισε:

$$\rho = b - Ax - w$$

$$\sigma = c - A^T y + z$$

$$\mu = \delta \frac{z^T x + y^T w}{n + m}$$

λύσε το σύστημα των εξισώσεων

$$A\Delta x + \Delta w = \rho$$

$$A^T \Delta y - \Delta z = \sigma$$

$$Z\Delta x + X\Delta z + \Delta X \Delta Z e = \mu e - XZe$$

$$W\Delta y + Y\Delta w + \Delta Y \Delta W e = \mu e - YWe$$

ως προς $(\Delta x,\ \Delta w,\ \Delta y,\ \Delta z)$

υπολόγισε: $\theta > \min\left[r\left(\max_{i,j}\left\{ -\frac{\Delta x_i}{x_i},\ -\frac{\Delta w_j}{w_j},\ -\frac{\Delta y_j}{y_j},\ -\frac{\Delta z_i}{z_i},\ \right\}\right)^{-1},\ 1 \right]$

αντικατέστησε:

$(x,\ w,\ y,\ z) \leftarrow (x + \theta\Delta x,\ w + \theta\Delta w,\ y + \theta\Delta y,\ z + \theta\Delta z)$

$\}$

8.3 Εξισώσεις Karush-Kuhn-Tucker – KKT

Το πλέον χρονοβόρο τμήμα του αλγορίθμου εσωτερικής διαδρομής είναι η επίλυση του συστήματος των εξισώσεων που προσδιορίζει τις επαυξήσεις.

$$A\Delta x + \Delta w = b - Ax - w = \rho$$

$$A^T\Delta y - \Delta z = c - A^T y + z = \sigma$$

$$Z\Delta x + X\Delta z = \mu e - XZe$$

$$W\Delta y + Y\Delta w = \mu e - YWe$$

Το σύστημα αυτό μπορεί να γραφεί σε μητρωική μορφή ως εξής:

$$\begin{bmatrix} -XZ^{-1} & 0 & -I & 0 \\ 0 & 0 & A & I \\ -I & A^T & 0 & 0 \\ 0 & I & 0 & YW^{-1} \end{bmatrix} \begin{Bmatrix} \Delta z \\ \Delta y \\ \Delta x \\ \Delta w \end{Bmatrix} = \begin{Bmatrix} -\mu Z^{-1}e + x \\ \rho \\ \sigma \\ \mu W^{-1}e - y \end{Bmatrix}$$

και καλείται σύστημα Karush-Kuhn-Tucker-KKT προς τιμήν των ερευνητών που ανεξάρτητα το διατύπωσαν. Το σύστημα αυτό είναι ένα συμμετρικό σύστημα $2n + 2m$ εξισώσεων με $2n + 2m$ αγνώστους. Η επίλυση του συστήματος παρέχει τις αυξήσεις των μεταβλητών κατά τα ενδιάμεσα βήματα της επαναληπτικής διαδικασίας που ορίζει ο αλγόριθμος της εσωτερικής διαδρομής.

Το KKT σύστημα εξισώσεων μπορεί να απλοποιηθεί περαιτέρω διατηρώντας την συμμετρία. Παρατηρούμε ότι η πρώτη και τελευταία σειρά επιλύονται άμεσα καθόσον τα μητρώα των συντελεστών είναι διαγώνια.

$$\Delta z = X^{-1}(\mu e - XZe - Z\Delta x)$$

$$\Delta w = Y^{-1}(\mu e - YWe - Y\Delta y)$$

και αντικαθιστώντας στις άλλες δύο λαμβάνουμε:

$$A\Delta x - Y^{-1}W\Delta y = \rho - \mu Y^{-1}e + w$$

$$A^T\Delta y + X^{-1}Z\Delta x = \sigma + \mu X^{-1}e - z$$

ή σε μητρωική μορφή:

$$\begin{bmatrix} -Y^{-1}W & A \\ A^T & X^{-1}Z \end{bmatrix} \begin{Bmatrix} \Delta y \\ \Delta x \end{Bmatrix} = \begin{Bmatrix} b - Ax - \mu Y^{-1}e \\ c - A^T y + \mu X^{-1}e \end{Bmatrix}$$

που αποτελεί την ανηγμένη έκφραση του συστήματος KKT και διατηρεί την συμμετρία.

8.4 Κανονικές Εξισώσεις KKT

Το σύστημα εξισώσεων KKT μπορεί να απλοποιηθεί περαιτέρω με δύο τρόπους. Σύμφωνα με τον πρώτο λύνουμε την πρώτη εξίσωση (1.13) ως προς Δx και αντικαθιστούμε στην δεύτερη εξίσωση (1.13β). Μπορούμε όμως να λύσουμε την δεύτερη ως προς Δy και να αντικαταστήσουμε στην πρώτη. Ακολουθώντας την δεύτερη διαδικασία λαμβάνουμε:

$$\Delta x = XZ^{-1}\left(c - A^T y + \mu X^{-1}e - A^T \Delta y\right)$$

$$-\left(Y^{-1}W + AXZ^{-1}A^T\right)\Delta y = b - Ax - \mu Y^{-1}e - AXZ^{-1}\left(c - A^T y + \mu X^{-1}e\right)$$

Το παραπάνω σύστημα είναι ένα σύστημα m εξισώσεων με m αγνώστους και αναφέρεται ως κανονικό σύστημα εξισώσεων πρωτογενούς μορφής. Ο πρώτος όρος του συντελεστή είναι ένα διαγώνιο μητρώο ενώ ο κύριος συντελεστής είναι ο δεύτερος όρος.

Αντίστοιχα άν λύσουμε ως προς Δy και αντικαταστήσουμε στην δεύτερη θα πάρουμε:

$$\Delta y = -YW^{-1}\left(b - Ax - \mu Y^{-1}e - A\Delta x\right)$$

και αντικαθιστώντας στην (1.13β) παίρνουμε:

$$\left(A^T YW^{-1}A + X^{-1}Z\right)\Delta x = c - A^T y + \mu X^{-1}e + A^T YW^{-1}\left(b - Ax - \mu Y^{-1}e\right)$$

που αποτελεί ένα σύστημα n εξισώσεων με n αγνώστους και αναφέρεται ως κανονικό σύστημα εξισώσεων σε δυική μορφή.

Στην πρωτογενή ή δυική μορφή αν το μητρώο A έχει μία στήλη πλήρη τότε το μητρώο των συντελεστών του πρωτογενούς συστήματος προκύπτει πλήρες, ενώ του δυικού παραμένει αραιό αν και μεγαλύτερων διαστάσεων. Γενικά γνωρίζουμε ότι η επίλυση ενός πλήρους μητρώου απαιτεί n^3 πράξεις για την διαγωνοποίηση ενώ ένα πολύ αραιό μητρώο μόνο n τάξεως πράξεις γεγονός που κάνει προτιμότερη την επιλογή του αραιού μητρώου δηλ. Των δυικών εξισώσεων.

8.5 Δευτεροβάθμιος Προγραμματισμός

Προβλήματα τα οποία ορίζονται μέσω αντικειμενικής δευτεροβάθμιας συνάρτησης που υπόκειται σε γραμμικούς περιορισμούς και μη αρνητικές τιμές των μεταβλητών καλούνται προβλήματα δευτεροβάθμιου προγραμματισμού (quadratic programming). Η γενική διατύπωση των προβλημάτων αυτών έχει ως εξής:

Ελαχιστοποιήστε την συνάρτηση $\min \quad c^T x + \frac{1}{2} x^T Q x$

Που υπόκειται στους περιορισμούς: $Ax \geq b$

$$x \geq 0$$

η διατύπωση του πρωτογενούς προβλήματος είναι για ελαχιστοποίηση με περιορισμούς μεγαλύτερο ίσο χωρίς να βλάπτεται η γενικότητα.

Για να συσχετιστεί το πρωτογενές με το δυικό πρόβλημα στο γραμμικό προγραμματισμό μελετήθηκε η αντικειμενική συνάρτηση με εμπόδια (barrier function). Στον δευτεροβάθμιο προγραμματισμό μπορεί να γίνει κάτι αντίστοιχο διατυπώνοντας το πρόβλημα με εμπόδια ως εξής:

$$\min\ c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \log x_i - \mu \sum_{j=1}^{m} \log w_j$$

που υπόκειται στους περιορισμούς:

$$Ax - w = b$$

για $x,\ w \geq 0$

Η συνάρτηση Lagrange που αντιστοιχεί στο παραπάνω πρόβλημα είναι:

$$f(x, w, y) = c^T x + \frac{1}{2} x^T Q x - \mu \sum_{i=1}^{n} \log x_i - \mu \sum_{j=1}^{m} \log w_j + y^T \left(b - Ax + w\right)$$

οι δε συνθήκες βελτιστοποίησης πρώτου βαθμού προκύπτουν με παραγώγιση της συνάρτησης Lagrange ως προς $x,\ w,$ και y ως εξής:

$$c + Qx - \mu X^{-1} e - A^T y = 0$$

$-\mu W^{-1}e + y = 0$

$b - Ax + w = 0$ αντίστοιχα.

Εισάγοντας το διάνυσμα $z = \mu X^{-1}e$

Προκύπτουν οι παρακάτω συνθήκες βελτιστοποίησης πρώτου βαθμού για το πρόβλημα ως εξής:

$$A^T y + z - Qx = c$$

$$Ax - w = b$$

$$XZe = \mu e$$

$$YWe = \mu e$$

όπου η πρώτη συνθήκη αντιστοιχεί στους περιορισμούς του δυικού προβλήματος. Η αντικειμενική συνάρτηση του δυικού προβλήματος προκύπτει ως εξής:

$$0 \le y^T w + z^T x = c^T x + x^T Qx - b^T y = c^T + \frac{1}{2}x^T Qx - \left(b^T y - \frac{1}{2}x^T Qx\right)$$

από όπου προκύπτει ότι το δυικό δευτεροβάθμιο πρόβλημα έχει την μορφή:

μεγιστοποιήσατε την συνάρτηση $b^T y - \frac{1}{2}x^T Qx$

που υπόκειται στους περιορισμούς: $A^T y + z - Qx = c$

και $y, z \ge 0$