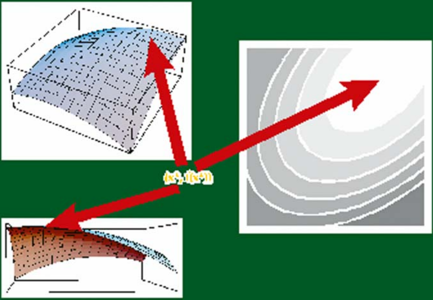
Α.Κ. Γεωργίου   Π.-Χ.Γ. Βασιλείου

Μη Γραμμικές Μέθοδοι Βελτιστοποίησης



Θεσσαλονίκη 1998
ΕΚΔΟΣΕΙΣ ΖΗΤΗ

Π ρ ό λ ο γ ο ς

Ο στόχος του παρόντος βιβλίου είναι η μεθοδική παρουσίαση περιγραφή και μελέτη κάποιων από τις σημαντικότερες μεθόδους βελτιστοποίησης που υπάγονται στη Θεωρία του Μη Γραμμικού Προγραμματισμού. Αποτελεί μία προσπάθεια καταγραφής του υλικού που διδάσκεται στο Τμήμα Μαθηματικών του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης την τελευταία πενταετία. Ταυτοχρόνως, πιστεύουμε ότι μπορεί να φανεί αρκετά χρήσιμο στον αναγνώστη που ενδιαφέρεται να έρθει σε μία αρχική επαφή με τον κλάδο αυτό της Επιχειρησιακής Ερευνας, είτε μεμονωμένα, είτε σε ένα γενικότερο πρόγραμμα εκμάθησης μεθόδων βελτιστοποίησης που πιθανόν να ξεκινάει από Γραμμικό Προγραμματισμό και να καταλήγει σε Μη Γραμμικές Μεθόδους περνώντας από Στοχαστικές Μεθόδους, Δυναμικό Προγραμματισμό κ.λ.π.

Υπάρχει πληθώρα αλγορίθμων οι οποίοι προτάθηκαν για την επίλυση γενικών μη γραμμικών προβλημάτων, δηλαδή προβλημάτων εντοπισμού μέγιστου ή ελάχιστου μίας μη γραμμικής αντικειμενικής συνάρτησης. Ο εντοπισμός της βέλτιστης λύσης μπορεί να γίνεται κάτω από περιορισμούς οι οποίοι εκφράζονται με γραμμικές ή μη γραμμικές εξισώσεις και ανισώσεις, ή χωρίς περιορισμούς στο χώρο των εφικτών λύσεων, που είναι και το αντικείμενο του παρόντος πρώτου μέρους. Ανάλογα με το είδος του προβλήματος που διαπραγματεύονται και τις πληροφορίες που χρειάζονται για να το επιλύσουν, οι Μη Γραμμικές Μέθοδοι διαχωρίζονται σε πολλές υποκατηγορίες. Εδώ, θα ασχοληθούμε με ένα πυρήνα αλγορίθμων που αφορούν προβλήματα χωρίς περιορισμούς, θεωρώντας ότι έτσι κλείνει μία ενότητα. Αλλωστε, για να καταγραφούν όλες οι παραλλαγές και οι υποκατηγορίες μεθόδων και τεχνικών θα χρειαστούν αρκετοί τόμοι ακόμα, εκτός από τον πρώτο.

Στο πρώτο κεφάλαιο εισάγεται το γενικευμένο μη γραμμικό πρόβλημα μέσα από παραδείγματα, ενώ στη συνέχεια δίνονται μερικά χρήσιμα αποτελέσματα από το Λογισμό και τη Θεωρία Πινάκων τα οποία θεωρούμε απαραίτητα για την κατανόηση του υλικού. Στο δεύτερο κεφάλαιο παρατίθεται μία ανάλυση των προβλημάτων που προκύπτουν κατά την ανάπτυξη ενός αλγορίθμου, καθώς επίσης και μία καταγραφή της θεωρίας που αφορά στο σημαντικό ερώτημα της ταχύτητας σύγκλισης των αλγορίθμων. Στο τρίτο κεφάλαιο περιέχονται μέθοδοι βελτιστοποίησης μονοδιάστατων μη γραμμικών προβλημάτων χωρίς περιορισμούς, διαφοροποιούμενες σε σχέση με την πληροφορία που υπάρχει, πρώτης ή δεύτερης τάξης. Στο τέταρτο κεφάλαιο ο αναγνώστης εισάγεται σε μεθόδους βελτιστοποίησης για πολυδιάστατα προβλήματα, όταν δηλαδή η αντικειμενική συνάρτηση είναι μία σημειακή συνάρτηση $f:\Re^n \rightarrow \Re$. Ταυτοχρόνως, παρουσιάζονται οι ικανές και αναγκαίες συνθήκες ύπαρξης βέλτιστης λύσης στα προβλήματα αυτά. Ακόμα, εξετάζεται το γενικό πρόβλημα επιλογής και μετακίνησης κατά μήκος μίας ευθείας μέσα στο χώρο προς εύρεση ικανοποιητικής επόμενης προσέγγισης βέλτιστου σημείου.

Προσπαθήσαμε για κάθε αλγόριθμο να δώσουμε αποτελέσματα που αφορούν στις συνθήκες κάτω από τις οποίες έχει ιδιότητες τοπικής ή σφαιρικής σύγκλισης αλλά και να εξετάσουμε και το ρυθμό της σύγκλισης συγκρίνοντας τις μεθόδους και μεταξύ τους. Πολλά από τα αποτελέσματα δίνονται σε μορφή θεωρημάτων μαζί με τις αντίστοιχες αποδείξεις, για να βοηθηθεί ο αναγνώστης στην κατανόηση της μεθοδολογίας που ακολουθείται στη θεωρία που στηρίζει την ανάπτυξη μίας επαναληπτικής μεθόδου.

Οπως τονίζουμε και στο κείμενο, μία μέθοδος βελτιστοποίησης ειδικά στον Μη Γραμμικό Προγραμματισμό, δεν αρκεί να συνοδεύεται από ικανοποιητικά θεωρητικά αποτελέσματα, θα πρέπει να αποδεικνύεται αρκετά αποτελεσματική και στην πράξη, δηλαδή να δίνει ικανοποιητικές λύσεις και στον υπολογιστή. Για κάθε μέθοδο παραθέτουμε τουλάχιστον ένα χαρακτηριστικό παράδειγμα αλλά και αντίστοιχα σχήματα και πίνακες αποτελεσμάτων.

Για την κατανόηση του παρόντος υλικού, απαιτούνται κάποιες γνώσεις Λογισμού και Θεωρίας Πινάκων οι οποίες στην πλειοψηφία τους έχουν συμπεριληφθεί στο πρώτο κεφάλαιο. Τέλος, ο αναγνώστης που ενδιαφέρεται να αποκτήσει ικανοποιητική ευχέρεια στη χρήση των μεθόδων, θα πρέπει να έχει πρόσβαση σε κάποιο υπολογιστικό σύστημα, έτσι ώστε να είναι σε θέση να προγραμματίσει του αλγορίθμους και να αντιμετωπίσει και τα πρακτικά προβλήματα που προκύπτουν σε μία διαδικασία βελτιστοποίησης.

Θεσσαλονίκη, Νοέμβριος 1993.

Π.-Χ.Γ. Βασιλείου   Α.Κ. Γεωργίου

Π ε ρ ι ε χ ό μ ε ν α

Πρόλογος 3

Συμβολισμοί 9

Κεφάλαιο Πρώτο

Εισαγωγή στα Μη Γραμμικά Προβλήματα 11

1.1 Εισαγωγή 11

1.2 Δομή των Προβλημάτων 13

1.3 Κατάταξη των Προβλημάτων 21

1.4 Μερικοί χρήσιμοι Ορισμοί και Αποτελέσματα 25

Κεφάλαιο Δεύτερο

Σύγκλιση των Αλγορίθμων 39

2.1 Αλγοριθμικά Προβλήματα 39

2.2 Σύγκλιση και Ρυθμός Σύγκλισης 46

Κεφάλαιο Τρίτο

Μονοδιάστατα Προβλήματα Χωρίς Περιορισμούς 53

3.1 Εισαγωγή 53

3.2 Η Μέθοδος Newton 55

3.3 Η βελτιωμένη μέθοδος Newton 66

3.4 Μέθοδοι χρήσης μόνο της Πρώτης Παραγώγου 77

3.5 Μέθοδοι χρήσης μόνο των τιμών της Συνάρτησης 93

Κεφάλαιο Τέταρτο

Πολυδιάστατα Προβλήματα Χωρίς Περιορισμούς 109

4.1 Εισαγωγή 109

4.2 Εμπειρικές Μέθοδοι 121

4.3 Η Μέθοδος της Μεγαλύτερης Αλλαγής 124

4.4 Η Πολυδιάστατη Μέθοδος Newton 149

4.5 Συζυγείς Διευθύνσεις 167

4.6 Μέθοδοι Τύπου Newton (Quasi Newton) 198

Βιβλιογραφία 227

Ευρετήριο Όρων 233

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΕΙΣΑΓΩΓΗ ΣΤΑ ΜΗ ΓΡΑΜΜΙΚΑ ΠΡΟΒΛΗΜΑΤΑ

1.1 ΕΙΣΑΓΩΓΗ

Η έννοια της Βελτιστοποίησης, όπως έχει ήδη αναφερθεί και σε άλλα μαθήματα του Τομέα Στατιστικής & Επιχειρησιακής Ερευνας του Τμήματος Μαθηματικών, μπορεί να οριστεί ως η διαδικασία η οποία με διάφορες μεθόδους εντοπίζει την καλύτερη λύση σε ένα μαθηματικό πρόβλημα όταν υπάρχει η δυνατότητα εναλλακτικών προτάσεων οι οποίες συνδέονται με κάποιο κόστος. Στην προηγούμενη πρόταση τις λέξεις "καλύτερη" και "κόστος" θα πρέπει να τις δούμε κάτω από ένα πολύ γενικό πρίσμα. Το καλύτερο συνδέεται πάντα με κάποιο κόστος και το κόστος δεν σημαίνει πάντα χρηματική απόδοση. Μέχρι το Β' Παγκόσμιο Πόλεμο οι μέθοδοι βελτιστοποίησης γενικά, που αφορούσαν συναρτήσεις πολλών μεταβλητών βρίσκονταν σε εμβρυϊκή μορφή. Η ανάπτυξη των Ηλεκτρονικών Υπολογιστών έδωσε την ώθηση που χρειάστηκε για την ανάπτυξη αλγορίθμων οι οποίοι θα μπορούσαν γρήγορα αλλά και με ακρίβεια η οποία σταθερά γινόταν μεγαλύτερη, να δώσουν προσεγγιστικές λύσεις σε τέτοια προβλήματα. Ετσι, υπήρξε ραγδαία ανάπτυξη των μεθόδων του Μαθηματικού Προγραμματισμού στις δεκαετίες του 40 και του 50 ενώ ακολούθησαν και άλλοι κλάδοι των μεθόδων βελτιστοποίησης όπως ο Δυναμικός Προγραμματισμός, ο Ακέραιος Προγραμματισμός, οι Μη γραμμικές Μέθοδοι κ.λ.π. Εδώ, θα πρέπει να σημειώσουμε ότι η λέξη "Προγραμματισμός" υποδηλώνει την έννοια "Βελτιστοποίηση" και δεν έχει άμεση σχέση με αυτό που έχει επικρατήσει, δηλαδή τον προγραμματισμό των Ηλεκτρονικών Υπολογιστών.

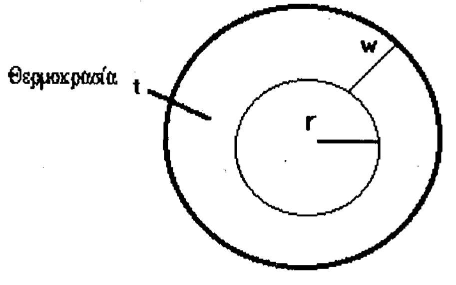
Στο παρόν, θα θέσουμε το γενικό πρόβλημα του Μη Γραμμικού Προγραμματισμού, δηλαδή το γενικό πρόβλημα βελτιστοποίησης μη γραμμικών προβλημάτων. Θα περιγράψουμε διάφορες πτυχές του αφού χωρίσουμε σε κατηγορίες τα διάφορα προβλήματα που έχουν μελετηθεί. Στη γενική του μορφή, το πρόβλημα είναι ήδη γνωστό στους φοιτητές του Τμήματος Μαθηματικών (Βασιλείου 1985), πως είναι η εύρεση ενός ακρότατου (μέγιστου ή ελάχιστου) μίας μη γραμμικής ως προς τις μεταβλητές (αντικειμενικής) συνάρτησης με τρόπο ώστε να ικανοποιούνται κάποιοι περιορισμοί που εκφράζονται σαν εξισώσεις ή ανισώσεις μεταξύ των μεταβλητών και οι οποίες μπορεί να είναι γραμμικές ή μη γραμμικές. Εχει επικρατήσει όταν αναφερόμαστε στο μη γραμμικό πρόβλημα προγραμματισμού, να μην περιέχονται ο ακέραιος μη γραμμικός προγραμματισμός, (δηλαδή όταν οι μεταβλητές και η αντικειμενική συνάρτηση παίρνουν μόνο ακέραιες τιμές), καθώς επίσης και το πρόβλημα του δυναμικού προγραμματισμού.

1.2 ΔΟΜΗ ΤΩΝ ΠΡΟΒΛΗΜΑΤΩΝ

Θα ξεκινήσουμε την παράγραφο αυτή με ένα παράδειγμα.

Παράδειγμα 1.1

Το ακόλουθο παράδειγμα είναι μια απλουστευμένη μορφή ενός προβλήματος που παρουσιάζεται στη μελέτη αντιδραστήρων πυρηνικής σύντηξης (Dennis & Schnabel 1983). Ενας αντιδραστήρας μπορεί να παρασταθεί όπως στο Σχήμα 1.1, δηλαδή από δύο ομόκεντρους κύκλους. Το πρόβλημα είναι να βρεθεί ο ιδανικός συνδυασμός της εσωτερικής ακτίνας r, του πλάτους w και της θερμοκρασίας t, έτσι ώστε να οδηγηθούμε σε παραγωγή ενέργειας με το ελάχιστο κόστος ανά μονάδα ενέργειας.



Σχήμα 1.1

Μετά από διάφορες εργαστηριακές μελέτες καθορίστηκε ότι το κόστος ανά μονάδα ενέργειας περιγράφεται από το ακόλουθο μοντέλο:

$$f(r,w,t)=10^6 t^2\left[c_1\left(r^2+\frac{2rw}{3}+\frac{w^2}{6}\right)+c_2\left(c_4+\frac{c_4^2}{r^2-c_4^2}\right)\right]+c_3\left(\frac{1}{2}+\frac{r}{w}\right)\left(10^3 t\right)^{-1.46}$$

(1.1)

όπου c_1, c_2, c_3, c_4 σταθερές. Οι **μεταβλητές** του προβλήματος είναι οι r, w και t ενώ οι c_1, c_2, c_3, c_4 είναι **παράμετροι**. Η λύση του πραγματικού φυσικού μη γραμμικού προβλήματος ανάγεται στον προσδιορισμό ενός ελαχίστου της πραγματικής συνάρτησης f η οποία εξαρτάται από την ακτίνα r, από το πλάτος w και από τη θερμοκρασία t. Βεβαίως, αν κανείς ήθελε να μελετήσει το πρόβλημα πιο αυστηρά, (πράγμα το οποίο προφανώς συμβαίνει στην πραγματικότητα), θα έπρεπε να θέσει και κάποιους περιορισμούς στις μεταβλητές, για παράδειγμα η ακτίνα και το πλάτος πρέπει να είναι θετικά. Το πρόβλημα στην πραγματική του μορφή είχε και πέντε περιορισμούς. Ετσι μπορεί να θεωρηθεί σαν ένα πρόβλημα μη γραμμικού προγραμματισμού με περιορισμούς.

Παράδειγμα 1.2

Ανάμεσα σε όλα τα τρίγωνα με βάση 8 εκατοστά και εμβαδόν τουλάχιστον 12 τετραγωνικά εκατοστά να βρεθεί αυτό με την ελάχιστη περίμετρο. (McCormick 1982).

Η βάση είναι σταθερή 8 cm οπότε θα θέλαμε να γνωρίζουμε τις τιμές από μία πλευρά και το ύψος (για παράδειγμα). Ετσι, αν x είναι μία πλευρά και y το ύψος, αυτές είναι οι μεταβλητές του προβλήματος και η βάση είναι παράμετρος. Από τα άπειρα x και y με τα οποία παράγονται τρίγωνα με εμβαδόν μεγαλύτερο ή ίσο των 12 τετραγωνικών εκατοστών, αναζητούμε τις τιμές x^*και y^* για τις οποίες η περίμετρος είναι ελάχιστη. Αφού $E = 8y/2 \geq 12$ έχουμε ότι $y \geq 3$. Είναι επίσης προφανές ότι $x \geq y$. Από το Πυθαγόρειο Θεώρημα προκύπτει το παρακάτω πρόβλημα:

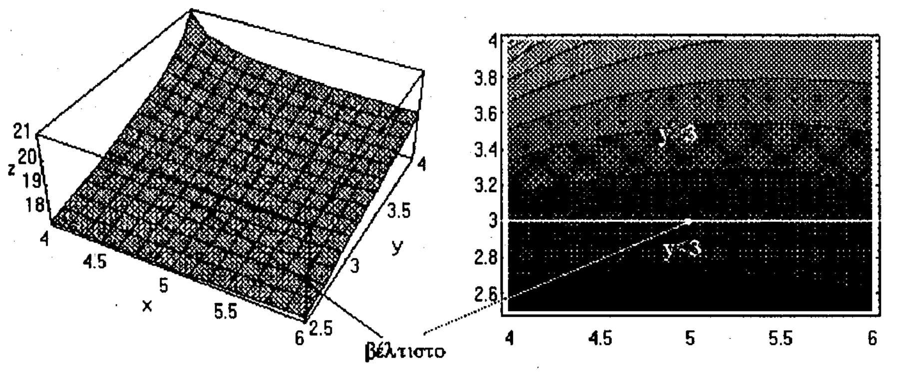
$$\underset{(x,y)}{\text{Minimize}} \;\; \mathrm{f}(x,y) = 8 + x + \left\{ y^2 + \left[8 - \left(x^2 - y^2 \right)^{1/2} \right]^2 \right\}^{1/2} \qquad \textbf{(1.2)}$$

με περιορισμούς:

$y \geq 3$ και

$x \geq y$

Η λύση του προβλήματος είναι το σημείο (5,3) που δίνει τιμή για τη συνάρτηση ίση με 18 (βλ. Σχήμα 1.2 τη συνάρτηση 1.2 στον $\Re^3$).



Σχήμα 1.2

Οπως είναι φανερό από τα παραπάνω παραδείγματα, η μελέτη των μη γραμμικών προβλημάτων και η κατασκευή ενός αντίστοιχου, μαθηματικού μοντέλου βελτιστοποίησης, εξαρτάται άμεσα από την αναγακαιότητα μεταφοράς συναρτησιακών ή ποιοτικών σχέσεων που ενυπάρχουν μεταξύ μετρήσιμων ή ποιοτικών χαρακτηριστικών, σε μαθηματικά σύμβολα. Αυτή η διαδικασία όπως ήδη γνωρίζετε ονομάζεται **μοντελοποίηση**. Κάθε μαθηματικό μοντέλο, ακόμα και το πιο απλό, σαν μεταφορά πραγματικών προβλημάτων σε μαθηματικούς τύπους, υπόκειται σε σφάλματα τα οποία προκύπτουν τουλάχιστον από τις παραδοχές που είμαστε υποχρεωμένοι να κάνουμε. Μία από τις πιο βασικές ικανότητες που πρέπει να αναπτύξει ένας επιχειρησιακός ερευνητής είναι η εύστοχη **απόρριψη** των ήσσονος σημασίας στοιχείων ενός προβλήματος και ταυτοχρόνως η **ανάδειξη** των σημαντικών στοιχείων και η προσκόλλησή τους στο μοντέλο.

Οπως προαναφέρθηκε, τα περισσότερα πραγματικά προβλήματα παρουσιάζουν καταστάσεις οι οποίες προκαλούν το πρόβλημα της μη ικανοποιητικής μοντελοποίησης τους σαν προβλήματα βελτιστοποίησης. Οι καταστάσεις που εμφανίζει ένα πραγματικό πρόβλημα χωρίζονται σε εννέα βασικές κατηγορίες (McCormick 1982).

1) Πολλοί αλληλοσυγκρουόμενοι μεταξύ τους στόχοι. Θα πρέπει να κατασκευάσουμε ένα κριτήριο βελτιστοποίησης, (μία αντικειμενική συνάρτηση), το οποίο να συμβιβάζει όσο γίνεται τους στόχους που θέτουμε.

2) Ασάφεια ως προς ποιες μεταβλητές αποτελούν τις μεταβλητές ελέγχου (εξηρτημένες, decision variables) και ποιες αποτελούν την είσοδο του προβλήματος δηλαδή οι ανεξάρτητες παράμετροι που βάζει ο ερευνητής. Είναι σημαντικό να καθοριστούν επακριβώς οι ανεξάρτητες μεταβλητές και οι μεταβλητές ελέγχου.

3) Αβεβαιότητα σε ό,τι αφορά τα φράγματα ή τους περιορισμούς των μεταβλητών ελέγχου και των συναρτήσεων τους. Πρέπει να καθορίζονται επακριβώς τα φράγματα στις μεταβλητές ή τις συναρτήσεις τους.

4) Ανάγκη να γνωρίζουμε τις υπαρκτές συναρτησιακές σχέσεις που υπάρχουν μεταξύ των ποσοτήτων που καθορίζουν ένα μοντέλο έτσι ώστε να τις μεταφέρουμε σωστά σε αλγεβρικές σχέσεις.

5) Οι συναρτησιακές σχέσεις μεταξύ των μεταβλητών μπορούν να είναι στοχαστικές ενώ ένα μη γραμμικό μοντέλο διαπραγματεύεται προσδιοριστικά ένα πρόβλημα. Ετσι θα πρέπει πρώτα να εκφράσουμε αυτές τις σχέσεις με αντίστοιχες μη στοχαστικές μορφές και στη συνέχεια να προχωρήσουμε στη βελτιστοποίηση.