ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Εθνικό Μετσόβιο Πολυτεχνείο

URL: http://www.semfe.ntua.gr

Προβλήματα Βελτιστοποίησης στον μη Γραμμικό Προγραμματισμό

Τσίγκος Δημήτριος eme19023

Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών

ΜΠΣΠ «Εφαρμοσμένες Μαθηματικές Επιστήμες»

Κατεύθυνση «Στατιστική και Πιθανότητες»

Επιβλέπων: Ιωάννης Κολέτσος, Αναπληρωτής Καθηγητής

Η τριμελής εξεταστική επιτροπή

Ι.Κολέτσος	**Β.Κοκκίνης**	**Π.Στεφανέας**
Αναπλ.Καθηγητής	**Αναπλ.Καθηγητής**	**Αναπλ.Καθηγητής**

Αθήνα, Ιούνιος 2022

Στους γονείς μου

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα εργασία αποτελεί διπλωματική εργασία στα πλαίσια του προγράμματος μεταπτυχιακών σπουδών στις Εφαρμοσμένες Μαθηματικές Επιστήμες με κατεύθυνση <<Στατιστική και Πιθανότητες>> της σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου.

Θα ήθελα να ευχαριστήσω θερμά τον κ. Ιωάννη Κολέτσο, επιβλέποντα καθηγητή μου, για την πολύτιμη βοήθεια και καθοδήγηση που μου προσέφερε, για την αμέριστη συμπαράσταση του καθ’ όλη τη διάρκεια εκπόνησης και για τη συμβολή του στην επίλυση των επιμέρους δυσκολιών μου. Θα ήθελα να επισημάνω ακόμη ότι οι επιστημονικές του γνώσεις βελτίωσαν σημαντικά το περιεχόμενο της εργασίας μου καθώς και το πνευματικό μου επίπεδο. Τον ευχαριστώ πολύ για την συνεργασία μας.

Ιδιαίτερες ευχαριστίες θέλω να εκφράσω προς τους γονείς μου Πολύδωρο και Αντιόπη για τη συνεχή συμπαράσταση τους και την υλική και ηθική στήριξη των επιλογών μου.

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή

1.1 Επιχειρησιακή έρευνα: Ορισμοί και ιστορία

Η επιχειρησιακή έρευνα (συχνά αναφέρεται ως επιστήμη διαχείρισης) είναι μια επιστημονική προσέγγιση στη λήψη αποφάσεων που επιδιώκει τον καλύτερο σχεδιασμό και λειτουργία ενός συστήματος, συνήθως υπό συνθήκες που απαιτούν την κατανομή μη ανεξάντλητων πόρων.

Με τον όρο σύστημα, εννοούμε μια οργάνωση αλληλοεξαρτώμενων στοιχείων που συνεργάζονται για την επίτευξη ενός στόχου. Για παράδειγμα, μια οποιαδήποτε επιχείρηση είναι ένα σύστημα του οποίου ο στόχος συνίσταται στη μεγιστοποίηση του κέρδους που μπορεί να αποκομίσει από την παραγωγή των προϊόντων της ή των υπηρεσιών της με την πλέον αποδοτική χρήση, οργάνωση, και συνεργασία μεταξύ, όλων των υλικών, άυλων και ανθρώπινων πόρων που έχει στη διάθεσή της.

Ο όρος επιχειρησιακή έρευνα επινοήθηκε στη Βρετανία κατά την διάρκεια του Β’ Παγκοσμίου Πολέμου όταν ζητήθηκε από επιστήμονες και μηχανικούς να αναλύσουν διάφορα στρατιωτικού τύπου προβλήματα όπως για παράδειγμα να αναπτύξουν αποτελεσματικές μεθόδους ώστε να χρησιμοποιήσουν ραντάρ τα οποία είχαν ανακαλυφθεί εκείνη την εποχή, να διαχειριστούν καλύτερα τις νηοπομπές και τα υποβρύχια, να διαχειριστούν τις αεροπορικές επιθέσεις με βόμβες και γενικά να βελτιστοποιήσουν τις στρατιωτικές επιχειρήσεις

Ο Watson Watt υπήρξε ένας από τους πρωτεργάτες της εποχής εκείνης και έδωσε τον εξής ορισμό:

> *Η Επιχειρησιακή Έρευνα αποσκοπεί στο να ερευνήσει ποσοτικά εάν ένας οργανισμός παίρνει από τη λειτουργία του εξοπλισμού του τη βέλτιστη δυνατή συνεισφορά σε σχέση με τον ολικό αντικειμενικό σκοπό του, ποιες αλλαγές σε εξοπλισμό και μεθόδους απαιτούνται για τη βελτίωση των αποτελεσμάτων με το μικρότερο δυνατό κόστος σε προσπάθεια και χρόνο και τέλος σε ποιο βαθμό μεταβολές στους επιμέρους αντικειμενικούς σκοπούς (τακτικοί αντικειμενικοί σκοποί) θα συνεισέφεραν στην πιο οικονομική και έγκαιρη εκτέλεση του ολικού στρατηγικού αντικειμενικού σκοπού.*

1.2 Στάδια επίλυσης προβλημάτων

Όταν η επιχειρησιακή έρευνα χρησιμοποιείται για την επίλυση του προβλήματος μιας επιχείρησης είναι καλό να ακολουθείται ένας αλγόριθμος που θα πρέπει να περιλαμβάνει τα ακόλουθα βήματα.

βήμα 1 **Διαπίστωση και σαφής προσδιορισμός του προβλήματος.** Διαπιστώνουμε ποιο ακριβώς είναι το ζητούμενο στο πρόβλημα που έχει ανατεθεί προς επίλυση. Αυτό μπορεί να γίνει με συζήτηση με τα στελέχη του οργανισμού με σύνταξη ερωτηματολογίων επιτόπια έρευνα σε τμήματα και παραγωγικές μονάδες της επιχείρησης ή και με οποιονδήποτε τρόπο κριθεί πρόσφορος για την συλλογή πληροφοριών. Ο στόχος σε αυτή τη φάση είναι να προσδιοριστεί ποια είναι η φύση του προβλήματος ποιες είναι οι επιλογές που υπάρχουν καθώς και να προσδιοριστούν οι φυσικοί ή οικονομικοί και πάσης φύσεως περιορισμοί.

βήμα 2 **Μοντελοποίηση του προβλήματος**. Σε αυτή τη φάση γίνεται προσπάθεια να εκφραστεί το σαφώς ορισμένο στην προηγούμενη φάση πρόβλημα σε μαθηματικές σχέσεις. Εάν η μαθηματική περιγραφή του προβλήματος μπορεί να γίνει με ένα από τα τυπικά μαθηματικά μοντέλα, όπως ο γραμμικός προγραμματισμός, μπορούμε συνήθως να φτάσουμε σε μια λύση χρησιμοποιώντας διαθέσιμους αλγόριθμους. Εναλλακτικά, εάν οι μαθηματικές σχέσεις είναι πολύ περίπλοκες για να επιτρέπουν τον προσδιορισμό μιας αποδεκτής λύσης, μπορούμε να επιλέξουμε την χρήση εναλλακτικών μαθηματικών μοντέλων (όπως θεωρία παιγνίων ή ανάλυση δικτύων), να χρησιμοποιήσουμε μια ευριστική προσέγγιση ή συνδυασμό από τις διαθέσιμες αυτές επιλογές.

βήμα 3 **Επαλήθευση και επανεκτίμηση της ορθότητας του μοντέλου που σχεδιάσαμε.** Έλεγχος της εγκυρότητας του μοντέλου που προσδιορίσαμε στην προηγούμενη φάση. χρησιμοποιώντας δεδομένα εισόδου παρελθόντων ετών και σύγκριση των πραγματικών δεδομένων εξόδου με αυτά που θα εξαχθούν από το μοντέλο. Σε περίπτωση που τα αποτελέσματα του ελέγχου δεν είναι ικανοποιητικά επιστρέφουμε στο προηγούμενο βήμα.

βήμα 4 **Παρουσίαση των αποτελεσμάτων και των συμπερασμάτων** της μελέτης στην Διοίκηση Σε αυτό το βήμα, παρουσιάζεται το μοντέλο και όλα τα αποτελέσματα που έχουν προκύψει στη διοίκηση της επιχείρησης. Μετά την παρουσίαση των αποτελεσμάτων της μελέτης είναι στη διακριτική ευχέρεια της Διοίκησης η έγκριση, η απόρριψη ή η επισήμανση τυχόν αστοχιών λόγω ελλιπούς πληροφόρησης ή λάθη εκτιμήσεων από τον μελετητή. Αυτό ίσως οδηγήσει σε επανεξέταση της διαδικασίας είτε στο βήμα 2 είτε στο βήμα 3.

βήμα 5 **Εφαρμογή και αξιολόγηση των προτεινόμενων από τη μελέτη λύσεων.** Εάν η Διοίκηση της επιχείρησης αποδεχθεί τη μελέτη, τότε ο αναλυτής βοηθά στην εφαρμογή των προτεινόμενων λύσεων. Το σύστημα πρέπει να παρακολουθείται συνεχώς (και να ενημερώνεται δυναμικά καθώς αλλάζει το περιβάλλον) για να διασφαλίζεται ότι οι συστάσεις επιτρέπουν στον οργανισμό να εκπληρώσει τους στόχους του.

1.3 Μέθοδοι που χρησιμοποιούνται στην επιχειρησιακή έρευνα

Ας δούμε όμως επιγραμματικά ποιες είναι οι κατηγορίες μεθόδων που χρησιμοποιούνται από την Επιχειρησιακή έρευνα

- **Δένδρα αποφάσεων**. Είναι ένα προσανατολισμένο δένδρο που αναπαριστά μια διαδικασία λήψης αποφάσεων. Οι κόμβοι συμβολίζουν χρονικά σημεία στα οποία:
    - Είτε πρέπει να ληφθεί κάποια απόφαση, Από κάθε κόμβο εκπορεύεται μια διακλάδωση εκπορεύεται για κάθε πιθανή απόφαση.

  - Είτε ο λήπτης αντιμετωπίζει κάποια φυσική κατάσταση. Από κάθε κόμβο εκπορεύεται μια διακλάδωση για πιθανές καταστάσεις της φύσης.
  - Είτε ολοκληρώνεται η διαδικασία.

  Τα δένδρα αποφάσεων είναι χρήσιμα για τον προσδιορισμό των βέλτιστων αποφάσεων για πολύπλοκες διαδικασίες.

- **Πολυκριτηριακή ανάλυση ή πολυκριτηριακή ιεράρχηση επιλογών.** Είναι εργαλείο λήψης απόφασης βάσει πολλών κριτηρίων, διαφορετικής βαρύτητας το καθένα. Βασικό στοιχείο της μεθόδου είναι ο καθορισμός των κριτηρίων, και η στάθμισή τους με όσο το δυνατόν αντικειμενικά κριτήρια έτσι ώστε το τελικό αποτέλεσμα να είναι το πραγματικά βέλτιστο.
- **Ανάλυση δικτύων.** Σαν δίκτυο ορίζουμε ένα σύνολο σημείων που ονομάζονται κόμβοι και ένα σύνολο από τόξα τα οποία συνδέουν συγκεκριμένα ζεύγη κόμβων. Κάποια προβλήματα είναι πιο εύκολο να επιλυθούν αν απεικονιστούν σε μορφή δικτύων. Τέτοια είναι τα προβλήματα εύρεσης συντομότερης διαδρομής, ελάχιστης κάλυψης μέγιστης ροής κλπ
- **Διαχείριση αποθεμάτων.** Με τον όρο αποθέματα εννοούμε περιουσιακά στοιχεία μιας επιχείρησης τα οποία προορίζονται είτε για πώληση, είτε για ανάλωση για την παραγωγή των προϊόντων που η επιχείρηση παράγει. Με την διαχείρισή τους προσπαθούμε να βρούμε την βέλτιστη ποσότητα αποθεμάτων που η επιχείρηση θα πρέπει να έχει έτσι ώστε να μην παρεμποδίζονται οι λειτουργίες της επιτυγχάνοντας έτσι το μικρότερο δυνατό κόστος που η διατήρηση αποθεμάτων επιφέρει
- **Ουρές αναμονής.** Μια διαδικασία ουράς αναμονής αποτελείται από πελάτες που προσέρχονται σε μια εγκατάσταση εξυπηρέτησης και περιμένουν σε σειρά αν όλοι οι εξυπηρετητές είναι απασχολημένοι. Χαρακτηρίζονται από πέντε συνιστώσες:
  - Το πρότυπο αφίξεων των πελατών,
  - το πρότυπο εξυπηρέτησης
  - αριθμό εξυπηρετητών
  - χωρητικότητα εγκατάστασης σε πελάτες και
  - τη σειρά με την οποία εξυπηρετούνται οι πελάτες

  Το ζητούμενο της έρευνας είναι η ελαχιστοποίηση του χρόνου αναμονής.

- **Θεωρία παιγνίων**. Με την θεωρία παιγνίων επιχειρούμε να μελετήσουμε μια κατάσταση κατά την οποία δύο ή περισσότεροι παίκτες ενεργούν με βάση προκαθορισμένους κανόνες ή και περιορισμούς έχοντας αντικρουόμενα ή και αλληλοεξαρτώμενα συμφέροντα. Κάθε παίκτης επιχειρεί τη μεγιστοποίηση της συνάρτησης ωφελείας του η οποία όμως περιέχει σα μεταβλητές της τις ενέργειες του ίδιου αλλά και των υπόλοιπων παικτών.
- **Προσομοίωση**. Με την προσομοίωση επιχειρείται η αναπαράσταση της λειτουργίας ενός συστήματος επιδιώκοντας από την μία να επισημανθούν κάποια προβλήματα στη λειτουργία του αφετέρου να γίνει μια εκτίμηση μελλοντικών αποτελεσμάτων της λειτουργίας του.
- **Μαθηματικός προγραμματισμός**. Με τον μαθηματικό προγραμματισμό επιχειρείται η χρήση μαθηματικών μοντέλων στην ανάλυση και επεξεργασία των προβλημάτων. Η επιστημονική προσέγγιση στη λήψη αποφάσεων

συνήθως περιλαμβάνει τη χρήση ενός ή περισσότερων μαθηματικών μοντέλων. Ένα μαθηματικό μοντέλο είναι μια μαθηματική αναπαράσταση μιας πραγματικής κατάστασης που μπορεί να χρησιμοποιηθεί για τη λήψη καλύτερων αποφάσεων ή απλώς για την καλύτερη κατανόηση της πραγματικής κατάστασης.

1.4 Μαθηματικά μοντέλα

1.4.1 Δομή μαθηματικών μοντέλων

Τα μαθηματικά μοντέλα που χρησιμοποιούνται στην επιχειρησιακή έρευνα περιγράφουν τη δομή και τη λειτουργία τους με μαθηματικές σχέσεις που περιλαμβάνουν δύο ομάδες στοιχείων.

- Παράμετροι ή μεταβλητές απόφασης οι οποίες ποσοτικοποιούν κάποιες αποφάσεις που πρέπει να ληφθούν.
- Ανισοεξισώσεις που προσπαθούν να παρουσιάσουν τις καθοριζόμενες από τον μελετητή σχέσεις μεταξύ των μεταβλητών και διακρίνονται σε δύο κατηγορίες.
  - Την Αντικειμενική συνάρτηση η οποία περιγράφει ένα κριτήριο ή μέτρο απόδοσης του συστήματος και της οποίας ζητείται η βελτιστοποίηση.
  - Περιορισμοί οι οποίοι μπορεί να προέρχονται από ανεπάρκεια είτε φυσικών πόρων είτε οικονομικών μέσων ή και οποιασδήποτε μορφής.

1.4.2 Διάκριση μαθηματικών μοντέλων

Τέλος τα μαθηματικά μοντέλα που χρησιμοποιούνται στην επιχειρησιακή έρευνα διακρίνονται ως προς:

- **Την δυναμική τους ως προς τον χρόνο (Στατικό ή δυναμικό μοντέλο)**
  - Ένα στατικό μοντέλο είναι αυτό στο οποίο οι μεταβλητές απόφασης δεν περιλαμβάνουν αλληλουχίες αποφάσεων σε πολλαπλές περιόδους. Σε ένα στατικό μοντέλο λύνουμε ένα πρόβλημα που έχουμε την εικόνα του σε μια συγκεκριμένη χρονική στιγμή υποθέτοντας ότι ο χρόνος δεν αποτελεί μεταβλητή αλλά σταθερά. Έτσι οι λύσεις που δίνονται ορίζουν τις βέλτιστες τιμές των μεταβλητών απόφασης σε όλα τα χρονικά σημεία.
  - Ένα δυναμικό μοντέλο είναι ένα μοντέλο στο οποίο οι μεταβλητές απόφασης περιλαμβάνουν αλληλουχίες αποφάσεων σε πολλαπλές περιόδους. Για παράδειγμα δυναμικού μοντέλου, σκεφτείτε μια εταιρεία ενοικίασης αυτοκινήτων που πρέπει να καθορίσει πώς να ελαχιστοποιήσει το κόστος της (έγκαιρης) κάλυψης της ζήτησης για ενοικιαζόμενα αυτοκίνητα κατά τη διάρκεια του επόμενου έτους λαμβάνοντας υπόψη την περιοδικότητα της ζήτησης. Μια τέτοια επιχείρηση θα πρέπει να καθορίσει την καλύτερη στρατηγική αυξομείωσης και ανανέωσης του στόλου της σε χρονικά σημεία που θα της φέρουν την μέγιστη απόδοση με τον μεγαλύτερο δυνατό βαθμό απασχόλησης του στόλου της. Οι αποφάσεις μιας τέτοιας επιχείρησης περιλαμβάνουν αποφάσεις που λαμβάνονται σε πολλαπλές περιόδους, επομένως ένα κατάλληλο μοντέλο του προβλήματος θα ήταν το δυναμικό μοντέλο.

- **την γραμμικότητα των μεταβλητών του προβλήματος: (Γραμμικό μη γραμμικό μοντέλο)**
    - **Γ**ραμμικότητα των μεταβλητών ενός προβλήματος έχουμε όταν η μεταβολή των ανεξάρτητων μεταβλητών προκαλεί αντίστοιχη μεταβολή στις εξαρτημένες μεταβλητές. Τέτοια μοντέλα είναι της μορφής y=$a \cdot x + b$ Ένα τέτοιο μοντέλο που όλες οι εμπλεκόμενες συναρτήσεις είναι γραμμικές είναι ένα γραμμικό μοντέλο.
    - Εάν σε κάποια από τις εμπλεκόμενες συναρτήσεις υπάρχει μη γραμμικότητα τότε το μοντέλο είναι μη γραμμικό. Γενικά, τα μη γραμμικά μοντέλα είναι πολύ πιο δύσκολο να επιλυθούν από τα γραμμικά μοντέλα.
- **την αριθμητική μορφή των μεταβλητών (Ακέραιο ή μη ακέραιο μοντέλο)**
    - Εάν μία ή περισσότερες μεταβλητές απόφασης πρέπει να είναι ακέραιες, τότε λέμε ότι ένα μοντέλο βελτιστοποίησης είναι ένα ακέραιο μοντέλο. . Εάν οι μεταβλητές απόφασης σε ένα μοντέλο αντιπροσωπεύουν τον αριθμό των εργαζομένων που ξεκινούν να εργάζονται κατά τη διάρκεια κάθε βάρδιας σε μια επιχείρηση, τότε σαφώς έχουμε ένα ακέραιο μοντέλο.
    - Εάν όλες οι μεταβλητές απόφασης είναι ελεύθερες να πάρουν κλασματικές τιμές, τότε το μοντέλο βελτιστοποίησης είναι ένα μη ακέραιο μοντέλο. Ο όγκος, η θερμοκρασία, η πίεση και η ποσοστιαία σύνθεση των εισροών μας μπορεί να λαμβάνουν όλες κλασματικές τιμές

Τα ακέραια μοντέλα είναι πολύ πιο δύσκολο να λυθούν από τα μη γραμμικά μοντέλα.

Κεφάλαιο 2^{ο}

Μη γραμμικός Προγραμματισμός

2.1 Ορισμός και ένα μικρό ιστορικό

Όπως και προηγουμένως αναφέραμε το κεντρικό πρόβλημα του μη γραμμικού προγραμματισμού είναι η ελαχιστοποίηση ή μεγιστοποίηση μιας δεδομένης συνάρτησης πολλών μεταβλητών η οποία και θα την αναφέρουμε είτε ως αντικειμενική συνάρτηση είτε ως συνάρτηση στόχο. Η αντικειμενική μπορεί να υπόκειται σε ένα πεπερασμένο σύνολο περιορισμών που εκφράζονται με συναρτήσεις μιας ή και περισσότερων μεταβλητών ανισότητας ή και ισότητας. Προϋπόθεση για να εντάσσεται το πρόβλημα στην κατηγορία του μη γραμμικού προγραμματισμού είναι τουλάχιστον μία εκ των συναρτήσεων (συνάρτηση στόχος ή συναρτήσεις που εκφράζουν τους περιορισμούς) να μην είναι γραμμική. Δηλαδή το πρόβλημα είναι της μορφής:

είτε:

$maxf(x_1, x_2, \dots, x_n)$ ή $minf(x_1, x_2, \dots, x_n)$

είτε:

$maxf(x_1, x_2, \dots, x_n)$		$minf(x_1, x_2, \dots, x_n)$
Με περιορισμούς		Με περιορισμούς
$g_1(x_1, x_2, \dots, x_n) \leq b_1$	ή	$g_1(x_1, x_2, \dots, x_n) \geq b_1$
$g_2(x_1, x_2, \dots, x_n) \leq b_2$		$g_2(x_1, x_2, \dots, x_n) \geq b_2$
.		.
.		.
$g_n(x_1, x_2, \dots, x_n) \leq b_n$		$g_n(x_1, x_2, \dots, x_n) \geq b_n$

Τα προβλήματα, τόσο θεωρητικά όσο και πρακτικά, τα οποία μπορούν να μεταφραστούν σε πρόβλημα μη γραμμικού προγραμματισμού, είναι στην πραγματικότητα αμέτρητα και προκύπτουν σε διαφορετικά επιστημονικά πεδία, όπως στην οικονομική επιστήμη στη θεωρία παιγνίων, στην επιχειρησιακή έρευνα, στην στατιστική, στη φυσική κ.λπ. Ο μη γραμμικός προγραμματισμός μπορεί να θεωρηθεί ως εκείνο το πεδίο της θεωρίας βελτιστοποίησης που αντιμετωπίζει προβλήματα στατικής και πεπερασμένων διαστάσεων βελτιστοποίησης με έμφαση στις υπολογιστικές πτυχές.

Επίσης χρησιμοποιείται ο όρος «μαθηματικός προγραμματισμός». Αυτός ο όρος εισήχθη για πρώτη φορά από τον Robert Dorfman το 1949, όπως αναφέρεται στο βιβλίο των J. K. LENSTRA, A. H. G. RINNOOY KAN and A. SCRIJVER « History of Mathematical Programming» που εκδόθηκε το 1991.

Ο όρος «μη γραμμικός προγραμματισμός» εμφανίζεται για πρώτη φορά στον τίτλο της διάσημης εργασίας των Kuhn και Tucker Nonlinear programming που δημοσιεύτηκε το 1951 σε έκδοση του UniversityofCaliforniaPress με τίτλο «Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability».

Ως αυτόνομο πεδίο έρευνας, ο μαθηματικός προγραμματισμός, που βασίζεται στη μαθηματική ανάλυση, την αριθμητική ανάλυση, τη γραμμική άλγεβρα, την επιχειρησιακή έρευνα κ.λπ. έχει χρησιμοποιηθεί ευρέως και έχει ενταχθεί στα εργαλεία πολλών επιστημών και σε πεδία έρευνας, όπως η θεωρία αποφάσεων, η θεωρία παιγνίων η μικροοικονομική θεωρία κ.λπ. Τα κύρια χαρακτηριστικά του καθιστούν εμφανή τη σημασία του για το σύνολο της οικονομικής και της χρηματοοικονομικής θεωρίας.

Η θεωρία βελτιστοποίησης πηγάζει, ιστορικά και λογικά, από την έρευνα μέγιστων και ελάχιστων σημείων σε συναρτήσεις μιας μεταβλητής.

Όπως είναι γνωστό η μελέτη συνάρτησης και δια μέσου αυτής ο έλεγχος για τα ακρότατα σημεία της συνιστά σημαντικό τμήμα στο βασικό μαθηματικό υπόβαθρο που διδάσκεται στο πρώτο έτος σπουδών οποιουδήποτε θετικής κατεύθυνσης πανεπιστημιακού τμήματος. Σαν συνέχεια της πρωτογενούς αυτής ανάλυσης ακολουθεί η μελέτη συναρτήσεων πολλών μεταβλητών.

Από την ως τώρα παρουσίαση προκύπτει ότι ένα τόσο σημαντικό εργαλείο το οποίο χρησιμοποιείται πλέον ευρύτατα από πληθώρα επιστημονικών πεδίων είναι σχετικά καινούριο. Όπως είναι γνωστό, ωστόσο, κάποιες πτυχές μερικών από τα μαθηματικά ερωτήματα που είναι κεντρικά για τον μαθηματικό προγραμματισμό, αντιμετωπίστηκαν σε μερικά από αυτά τα άλλα «συνδετικά» πεδία έρευνας πριν ο μαθηματικός προγραμματισμός εμφανιστεί ως αυτόνομος μαθηματικός κλάδος.

Είναι σαφές μέχρι τώρα ότι αναφερόμαστε σε μια θεωρία της οποίας η «επίσημη» γέννηση βρίσκεται στα μέσα του εικοστού αιώνα, και ακριβώς στα χρόνια που ακολούθησαν αμέσως μετά τον Δεύτερο Παγκόσμιο Πόλεμο. Ωστόσο, ξέρουμε ότι στον Μαθηματικό Προγραμματισμό χρησιμοποιούμε τον όρο «Πολλαπλασιαστές Lagrange», που πήρε το όνομά του από τον μαθηματικό που έζησε στα τέλη του $17^{ου}$ και στις αρχές του $18^{ου}$ αιώνα Joseph Louis Lagrange. Ο Joseph Louis Lagrange αντιμετώπισε το πρόβλημα της εύρεσης ακροτάτων σε μια συνάρτηση όταν οι μεταβλητές έχουν περιορισμούς της μορφής Ο Lagrange απέδειξε ότι τη λύση στο παραπάνω πρόβλημα τη δίνει η επίλυση του παρακάτω συστήματος εξισώσεων.

2.2 Βασικές μαθηματικές έννοιες

Θεωρώντας όμως αρκετά κατατοπιστική την ιστορική αναφορά και τα εισαγωγικά στοιχεία που δώσαμε για τον μη γραμμικό προγραμματισμό ας κάνουμε μια αναφορά σε βασικές μαθηματικές έννοιες που θα μας βοηθήσουν στη συνέχεια.

ορισμός 2.2.1 Έστω μια συνάρτηση συνεχής f σε ένα διάστημα Δ που αποτελεί αυτονόητα μέρος του συνόλου A που είναι το πεδίο ορισμού της συνάρτησης.

- Λέμε ότι η συνάρτηση έχει **τοπικό ελάχιστο** στο σημείο x_0 αν στο διάστημα αυτό ισχύει ότι:
  $f(x) > f(x_0)\ \forall\ x \in \Delta$
- Το σημείο αυτό ορίζεται σαν **ολικό ελάχιστο** αν για κάθε σημείο x του πεδίου ορισμού Α της συνάρτησης ισχύει ότι:
  $f(x) > f(x_0)\ \forall\ x \in A$
- Λέμε ότι η συνάρτηση έχει **τοπικό μέγιστο** στο σημείο x_0 αν στο διάστημα αυτό ισχύει ότι:
  $f(x) < f(x_0)\ \forall\ x \in \Delta$
- Το σημείο αυτό ορίζεται σαν **ολικό μέγιστο** αν για κάθε σημείο x του πεδίου ορισμού Α της συνάρτησης ισχύει ότι
  $f(x) < f(x_0)\ \forall\ x \in A$

ορισμός 2.2.2 Έστω συνάρτηση f που ορίζεται σε ένα διάστημα A (πεπερασμένο ή άπειρο). Τότε η συνάρτηση ορίζεται σαν:

- **κυρτή** όταν ισχύει ότι:

$$f(\alpha * x_1 + (1-a) * x_2) \leq \alpha * f(x_1) + (1-a) * f(x_2), 0 \leq a \leq 1$$

Θεώρημα 1. Έστω μια συνάρτηση f συνεχής σε ένα διάστημα Δ και δύο φορές παραγωγίσιμη στο εσωτερικό του Δ.

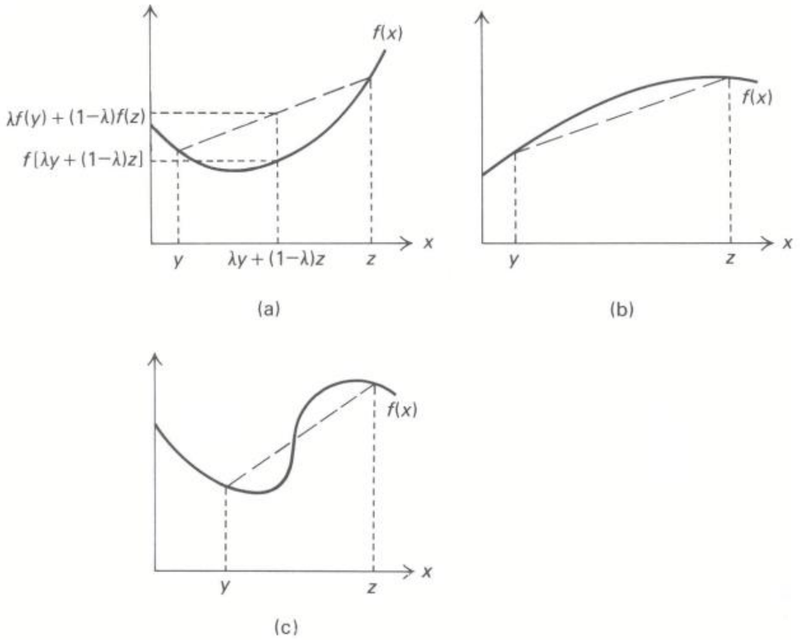
Αν $f''(x) > 0$ για κάθε εσωτερικό σημείο x του Δ, τότε η f είναι κυρτή στο Δ.

- **κοίλη** όταν ισχύει ότι:

$$f(a * x_1 + (1-\alpha) * x_2) \geq a * f(x_1) + (1-a) * f(x_2), 0 \leq a \leq 1$$

Θεώρημα 2. Έστω μια συνάρτηση f συνεχής σε ένα διάστημα Δ και δύο φορές παραγωγίσιμη στο εσωτερικό του Δ.

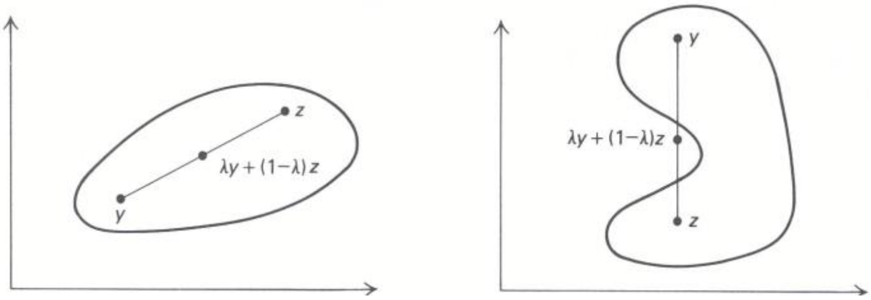
Αν $f''(x) < 0$ για κάθε εσωτερικό σημείο x του Δ, τότε η f είναι κοίλη στο Δ.



(a) Κυρτή συνάρτηση (b) Κοίλη συνάρτηση (c) Ούτε κυρτή ούτε κοίλη συνάρτηση

ορισμός 2.2.3 Έστω S υποσύνολο ενός πραγματικού γραμμικού χώρου $(S \dot{\subseteq} R^n)$. Τότε το **σύνολο καλείται κυρτό** αν όλα τα σημεία που ανήκει σε ευθεία που συνδέει δύο οποιαδήποτε σημεία που ανήκουν στο σύνολο S ανήκουν στο σύνολο S :

$$\forall\, x, y \in S, \forall\, \lambda \in [0,1]\ \tau\acute{o}\tau\varepsilon\ \lambda * x + (1-\lambda) * y \in S$$



Κυρτό και μη κυρτό σύνολο

ορισμός 2.2.4 Έστω διάνυσμα $\widehat{X}$ = $[\, x_1, x_2, \dots, x_n]^T$. Ορίζουμε ως **γειτονιά- *ε*** περί το $\widehat{X}$ το σύνολο όλων των διανυσμάτων $\boldsymbol{X}$ τέτοιων ώστε:

$$(X - \widehat{X}\,)^T * \left(X - \widehat{X}\right) = (x_1 - \widehat{x_1})^2 + (x_2 - \widehat{x_2})^2 + \ \dots + (x_n - \widehat{x_n})^2 \ < \ \varepsilon^2$$

ορισμός 2.2.5 Μία συνάρτηση $f(X)$ πολλών μεταβλητών έχει ένα **τοπικό μέγιστο** στο $\widehat{X}$ αν υπάρχει γειτονιά - ε περί το $\widehat{X}$ τέτοια ώστε να ισχύει:
$f(X) \leq f\left(\widehat{X}\,\right) \forall\ \widehat{X}$ της γειτονιάς – ε για το οποίο ορίζεται η συνάρτηση. Αν η συνθήκη ικανοποιείται για κάθε τιμή του ε, όσο μεγάλη και να είναι, τότε η συνάρτηση έχει ολικό μέγιστο στο $\widehat{X}$.
Μία συνάρτηση $f(X)$ έχει ένα **τοπικό ελάχιστο** στο $\widehat{X}$ αν υπάρχει γειτονιά- ε περί το $\widehat{X}$ έτοια ώστε να ισχύει: $f(X) \geq f\left(\widehat{X}\,\right) \forall\ \widehat{X}$ της γειτονιάς – ε για το

οποίο ορίζεται η συνάρτηση. Αν η συνθήκη ικανοποιείται για κάθε τιμή του ε, όσο μεγάλη και να είναι, τότε η συνάρτηση έχει ολικό ελάχιστο στο $\hat{X}$.

Θεώρημα 3. Αν η συνάρτηση $f(X)$ είναι συνεχής σε μια κλειστή και φραγμένη περιοχή τότε έχει ολικό μέγιστο και ολικό ελάχιστο σε αυτή την περιοχή.

ορισμός 2.2.6 Το διάνυσμα κλίσης ∇f που συσχετίζεται με συνάρτηση $f(x_1, x_2, \dots, x_n)$ η οποία έχει πρώτες μερικές παραγώγους ορίζεται από τον τύπο:

$\nabla f = [\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}]^T$ ορίζεται δηλαδή από τις μερικές παραγώγους ως προς κάθε μεταβλητή της συνάρτησης.

Η σημειογραφία $\nabla f_{\hat{X}}$ συμβολίζει την τιμή της κλίσης στο $\hat{X}$.

Θεώρημα 4. Αν η συνάρτηση $f(X)$ έχει τοπικό μέγιστο (ή τοπικό ελάχιστο) στο $\hat{X}$ και υπάρχει το διάνυσμα κλίσης ∇f σε γειτονιά-ε περί το $\hat{X}$ τότε $\nabla f_{\ \hat{X}} = 0$.

ορισμός 2.2.7 Αντίστροφος ενός τετραγωνικού πίνακα A είναι ο πίνακας (συμβολίζεται με A^{-1}) τέτοιος ώστε: πολλαπλασιαζόμενος επί τον αρχικό πίνακα να δίνει τον μοναδιαίο πίνακα ($A * \mathrm{A}^{-1} = I$).

ορισμός 2.2.8 Η μήτρα Hess που συσχετίζεται με μια συνάρτηση $f(x_1, x_2, \dots, x_n)$ η οποία έχει δεύτερες μερικές παραγώγους είναι η:

$H_f = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right] * (i, j = 1, 2, \dots, n)$ σε κάθε γραμμή της συνάρτησης ορίζεται η δεύτερη παράγωγος ως προς κάθε μεταβλητή και ως προς κάθε πρώτη παράγωγο που περιέχεται στο διάνυσμα κλίσης ∇f. Δηλαδή στη θέση i, j του πίνακα H περιέχεται η δεύτερη παράγωγος ως προς την j παράγωγο της x_i μεταβλητής της αρχικής συνάρτησης.

ορισμός 2.2.9 Μια συμμετρική μήτρα A διαστάσεων $n * n$ (τέτοια ώστε $A = A^T$) είναι αρνητικά (θετικά) ορισμένη αν το γινόμενο $\mathrm{X}^T A X$ είναι αρνητικό (θετικό) για κάθε μη μηδενικό διάνυσμα X n διαστάσεων με $X \neq 0$. Η σημειογραφία $\mathrm{H}f_{\hat{X}}$ συμβολίζει την τιμή της Η στο $\hat{X}$.

Θεώρημα 5. Έστω ότι $A = [a_{i,j}]$ πίνακας διαστάσεων $n * n$ και οι ορίζουσες:

$$A_1 = |a_{1,1}|,\ A_2 = -\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{1,3} & a_{1,4} \end{vmatrix},\ A_3 = +\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix}, \dots, A_n = (-1)^{n-1} * detA.$$

Ο πίνακας A είναι αρνητικά ορισμένος τότε και μόνον τότε αν όλες οι ορίζουσες $A_1, A_{2,\dots,}A_n$ είναι αρνητικές. Ο πίνακας A είναι αρνητικά ημιορισμένος αν και μόνον αν όταν οι ορίζουσες $A_1, A_{2,\dots,}A_r$ $(r < n)$ είναι αρνητικές και οι υπόλοιπες είναι ίσες με το 0.

Ισοδύναμα είναι αρνητικά ορισμένος αν όλες οι ιδιοτιμές του είναι αρνητικές.

Αντίστοιχα ο πίνακας A είναι θετικά ορισμένος αν:

$$A_1 = \left|a_{1,1}\right| , A_2 = \begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{1,3} & a_{1,4} \end{vmatrix} , A_3 = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} ,\ldots,\ A_n = detA$$

και $A_1, A_{2,\ldots,}A_n$ είναι όλες θετικές και θετικά ημιορισμένος αν $A_1, A_{2,\ldots,}A_r$ $(r < n)$ είναι θετικές και οι υπόλοιπες ίσες με το 0.
Ισοδύναμα είναι θετικά ορισμένος αν όλες οι ιδιοτιμές του είναι θετικές.

Θεώρημα 6. Αν η συνάρτηση $f(X)$ έχει μερικές παραγώγους σε γειτονιά-ε περί το $\hat{X}$, και αν $\nabla f_{\ \hat{X}} = 0$ και η μήτρα $\mathrm{H}f_{\hat{X}}$ είναι αρνητικά ορισμένη τότε η συνάρτηση $f(X)$ έχει τοπικό μέγιστο στο $\hat{X}$. Αντίστοιχα αν είναι θετικά ορισμένη τότε η συνάρτηση έχει τοπικό ελάχιστο. Τέλος αν δεν είναι ούτε αρνητικά ούτε θετικά ορισμένη τότε έχουμε σαγματικό σημείο.

ορισμός 2.2.10 Έστω ότι θέλουμε να βρούμε τα ακρότατα συνάρτησης $f(x, y)$ (που την ονομάζουμε αντικειμενική συνάρτηση) στα σημεία μιας καμπύλης που βρίσκεται στην περιοχή του πεδίου ορισμού της. Η καμπύλη αυτή περιγράφεται από μια εξίσωση $g(x, y) = c$ που ονομάζεται εξίσωση περιορισμού. Τέτοια ακρότατα ονομάζονται περιορισμένα ακρότατα. Όταν τα ακρότατα βρίσκονται εντός του πεδίου ορισμού της συνάρτησης ονομάζονται εσωτερικά ενώ όταν βρίσκονται στα σύνορα του πεδίου ορισμού ονομάζονται συνοριακά.

Θεώρημα 7. Αναγκαίες συνθήκες πρώτης τάξης για να έχει μια συνάρτηση περιορισμένο εσωτερικό ακρότατο είναι: $\frac{f_x}{f_y} = \frac{g_x}{g_y}$ και $g = c$. Δηλαδή η αντικειμενική συνάρτηση και η συνάρτηση περιορισμού να έχουν τον ίδιο ρυθμό υποκατάστασης και η συνάρτηση περιορισμού να είναι ισωτική.

ορισμός 2.2.11 Από την παραπάνω συνθήκη περιορισμένης στασιμότητας έχουμε: $\frac{f_x}{f_y} = \frac{g_x}{g_y} \Longrightarrow f_x * g_y = g_x * f_y \Leftrightarrow \frac{f_x}{g_x} = \frac{f_y}{g_y}$ όπου με τον συμβολισμό f_x και g_x ορίζονται οι μερικές παράγωγοι ως προς x. Ο λόγος που προέκυψε στην τελευταία ισότητα ονομάζεται πολλαπλασιαστής Lagrange και συμβολίζεται με το ελληνικό γράμμα λ. Από τον ορισμό του λ δηλαδή προκύπτει ότι: $f_x - \lambda * g_x = 0$.

ορισμός 2.2.12 Συνάρτηση Lagrange του προβλήματος περιορισμένου ακρότατου ονομάζεται η παρακάτω συνάρτηση τριών τουλάχιστον μεταβλητών $(x, y, \ldots, \lambda)$

$$g_1(x_1, x_2, \ldots, x_n) = c_1$$
$$g_2(x_1, x_2, \ldots, x_n) = c_2$$

$max/minf(x_1, x_2, \ldots, x_n)$ με

$$\ldots$$
$$g_m(x_1, x_2, \ldots, x_n) = c_m$$

$$L(x_i, \lambda_j) = f(x_i) - \sum_{j=1}^{m} \lambda_j * (\,g(x_j) - c_j),\ i = 1,2, \ldots, n,\ j = 1,2, \ldots, m$$

αν ζητάμε μεγιστοποίηση

$$L(x_i,\lambda_j) = f(x_i) + \sum_{j=1}^{m} \lambda_j * (\, g(x_j) - c_j), i = 1,2,\ldots,n, j = 1,2,\ldots,m$$

αν ζητάμε ελαχιστοποίηση

Τα κρίσιμα σημεία υπό περιορισμούς τα βρίσκουμε από την επίλυση των εξισώσεων:

$$\frac{\partial L}{\partial x_i} = 0 \,, i = 1,2,\ldots,n$$

$$\frac{\partial L}{\partial \lambda_j} = g_{j(X)} = 0 \,, j = 1,2,\ldots,m$$

Θεώρημα 8. Απαραίτητες προϋποθέσεις για τη λύση του προβλήματος βελτιστοποίησης της μορφής:

max $z = f(X)$

με περιορισμούς

$$\begin{gathered} g_1(X) \leq 0 \\ g_2(X) \leq 0 \\ \ldots \\ g_m(X) \leq 0 \\ X \geq 0 \end{gathered}$$

είναι

α. να ξαναγράψουμε τις μη αρνητικές συνθήκες με τη μορφή $X \geq 0$ έτσι ώστε η ομάδα των περιορισμών να περιέχει $m + n$ συνθήκες περιορισμών κάθε μία από τις οποίες έχει σχέση μικρότερου ή ίσου με.

β. Να προσθέσουμε χαλαρές μεταβλητές (με τον περιορισμό να είναι θετικές) στα αριστερά μέλη όλων των περιορισμών μετατρέποντας τις ανισότητες σε ισότητες.

Κεφάλαιο 3ο

Κατηγορίες και μέθοδοι Βελτιστοποίησης

3.1 Βελτιστοποίηση συναρτήσεων πολλών μεταβλητών χωρίς περιορισμούς

Σε αυτή την κατηγορία οι συναρτήσεις που προσπαθούμε να βελτιστοποιήσουμε έχουν τη μορφή

$$maxf(x_1, x_2, \dots, x_n) \text{ ή } minf(x_1, x_2, \dots, x_n)$$

3.1.1 Μεγιστοποίηση συνάρτησης στόχου με χρήση διαφορικού λογισμού Μέθοδος πρώτη

βήμα 1 Δημιουργούμε το διάνυσμα κλίσης παραγωγίζοντας την συνάρτηση ως προς κάθε μεταβλητή

βήμα 2 Εξισώνουμε κάθε συνάρτηση στο διάνυσμα κλίσης με το 0 και λύνουμε τις εξισώσεις.

βήμα 3 Οι τιμές που βρίσκουμε είναι οι τιμές των μεταβλητών που η συνάρτηση στόχος έχει ακρότατα ή είναι σαγματικό σημείο. Στη συνέχεια ελέγχουμε την συνάρτηση αν έχει ολικό μέγιστο ή τείνει στο άπειρο.

Παράδειγμα

Θέλουμε να μεγιστοποιήσουμε την συνάρτηση:

$$\max(x_1 * (x_2 - 4) + x_3 * ({x_3}^2 - 12)$$

Το διάνυσμα κλίσης (πρώτες μερικές παράγωγοι ως προς κάθε μεταβλητή) είναι
$\nabla f = (x_2 - 4),\ x_1\ , 3 * {x_3}^2 - 12$
εξισώνοντας κάθε παράγωγο με το 0 παίρνουμε:
$(x_2 - 4) = 0 \Longrightarrow x_2 = 4$
$x_1 = 0$
$3 * {x_3}^2 - 12 = 0 \Longrightarrow x_3 = 2$ ή $x_3 = -2$
που σημαίνει ότι έχουμε δύο διανύσματα σαν λύσεις.
$X_1 = [4{,}0{,}2]^T$ και $X_2 = [4{,}0,-2]^T$. Άρα έχουμε: $f(X_1) = -20$ και $f(X_2) = -12$
Ο Εσσιανός πίνακας που περιέχει τις μερικές παραγώγους ως προς το διάνυσμα που περιέχει τις μερικές παραγώγους της αρχικής συνάρτησης είναι:
$H = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 6x_3 \end{bmatrix}$ το οποίο βλέπουμε ότι η ορίζουσα του πίνακα A_{11} είναι 0 η ορίζουσα του δεύτερου πίνακα είναι $\det\left(H_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right) = -1.$ Άρα ο πίνακας θα είναι ή αρνητικά ημιορισμένος ή αόριστος γιατί θα έπρεπε να ισχύει $detH_2 \geq 0$ είτε αρνητικά ορισμένος γιατί $(-1)^1 * detH_2 = 1 > 0$. Άρα η συνάρτηση σε κάθε περίπτωση ή έχει τοπικό ελάχιστο ή τα κρίσιμα σημεία της είναι σαγματικά.

Παράδειγμα

Θέλουμε να μεγιστοποιήσουμε την συνάρτηση:

$$max(\eta\mu(x_1 + x_2) - \sigma\upsilon\nu(8 * x_1 * x_2))$$

Παρατηρούμε ότι λόγω των ιδιοτήτων των τριγωνομετρικών συναρτήσεων η αντικειμενική συνάρτηση αυτή δεν είναι κοίλη. Οπότε η γενική θεωρία για τα τοπικά βέλτιστα δεν εφαρμόζεται εδώ. Οπότε μία λύση είναι να βρούμε ένα άνω φράγμα για την αντικειμενική συνάρτηση και έπειτα να βρούμε τα σημεία (x_1, x_2) στα οποία η αντικειμενική συνάρτηση έχει την τιμή του φράγματος.
Έχουμε ότι

$$f(x_1, x_2) = \eta\mu(x_1 + x_2) - \sigma\upsilon\nu(8 * x_1 * x_2) \leq 2$$

Οι λύσεις του π.μ.γ.π θα είναι τα σημεία (x_1, x_2) με $f(x_1, x_2) = 2$.
Έχουμε ότι

$f(x_1, x_2) = 2 <=> \eta\mu(x_1 + x_2) = 1 \, , \sigma\upsilon\nu(8 * x_1 * x_2) = -1$

$<=> x_1 + x_2 = 2 * k * \pi + \left(\frac{\pi}{2}\right), \quad k \in Z$

$\quad 8 * x_1 * x_2 = 2 * \lambda * \pi + \pi, \quad \lambda \in Z$

Οπότε έχουμε άπειρες λύσεις του π.μ.γτ.π.
Π.χ. για $k = \lambda = 0$ έχουμε

$$(x_1, x_2) = (\frac{\pi}{8} + \frac{1}{2} * \sqrt{\frac{\pi^2}{8} - \pi}, \frac{\pi}{8} - \frac{1}{2} * \sqrt{\frac{\pi^2}{8} - \pi})$$

Το οποίο προκύπτει ως λύση του συστήματος $x_1 + x_2 = \frac{\pi}{2}$ και $8 * x_1 * x_2 = \pi$.

3.1.2 Τετραγωνικός προγραμματισμός. (Συνάρτηση που στην γενική της μορφή μπορεί να γίνει η περιγραφή της από την μορφή: $X^T * A * X - B * X + C$

Μια συνάρτηση πολλών μεταβλητών που είναι τετραγωνική δηλαδή της μορφής: $X^T * A * X - B * X + C$ η οποία έχει n μεταβλητές και είναι διπλά παραγωγίσιμη έχει παράγωγο την $2 * A * X + B$. Αυτό συμβαίνει γιατί κάθε μερική παράγωγος της αρχικής συνάρτησης θα έχει την μορφή:

$f'(x) = 2 * a * x + b$ και θα σχηματιστεί έτσι ένα διάνυσμα που θα αποτελείται από n μερικές παραγώγους, (όσες και οι μεταβλητές) με τους όρους που παραγωγίζουν την τετραγωνική μεταβλητή να πολλαπλασιάζονται επί δύο λόγω της γνωστής ιδιότητας της παραγώγου $f'(\alpha * x^k) = k * a * x^{k-1}$ άρα για $k = 2$ ισχύει ότι : $f'(\alpha * x^2) = 2 * a * x$.

Αν παραγωγίσουμε για δεύτερη φορά ως προς όλες τις μεταβλητές θα σχηματιστεί ο συμμετρικός εσσιανός πίνακας H για τον οποίο εξ' ορισμού θα πρέπει να ισχύει ότι $H * X - B = 0$ έχοντας όμως $H = 2 * A$.

Ταυτόχρονα ισχύει ότι για να έχει ακρότατα η συνάρτηση αυτή θα πρέπει η πρώτη παράγωγος να ισούται με το 0. Άρα θα ισχύει ότι : $H * X - B = 0 \iff 2 * A * X = B$.

Από την τελευταία σχέση διαπιστώνουμε ότι η λύση του προβλήματος ελαχιστοποίησης μιας συνάρτησης πολλών μεταβλητών που έχει την μορφή: $min\frac{1}{2} * X^T * H * X - B * X + C$ έγκειται τελικά στην επίλυση του προβλήματος $A * X = B = 0 \Leftrightarrow A * X = B \Leftrightarrow X = A^{-1} * B$ και από την σχέση $A = \frac{1}{2} * H \Longrightarrow X = \frac{1}{2} * H * B$ με αναγκαία συνθήκη ο πίνακας H να είναι θετικά ορισμένος ή θετικά ημιορισμένος.

Αλγόριθμος επίλυσης

βήμα 1 Αν ζητάμε μεγιστοποίηση της συνάρτησης την πολλαπλασιάζουμε επί (-1) και μετατρέπουμε το πρόβλημα σε πρόβλημα εύρεσης ελαχίστου.

βήμα 2 Παραγωγίζουμε δύο φορές ως προς όλες τις μεταβλητές της συνάρτησης.

βήμα 3 Ελέγχουμε αν ο Εσσσιανός πίνακας Η είναι θετικά ορισμένος ή θετικά ημιορισμένος ή αρνητικά ορισμένος ή αρνητικά ημιορισμένος ή αόριστος.

α. Αν ο πίνακας είναι αόριστος ή αρνητικά ορισμένος ή ημιορισμένος τέλος του αλγορίθμου χωρίς λύση.

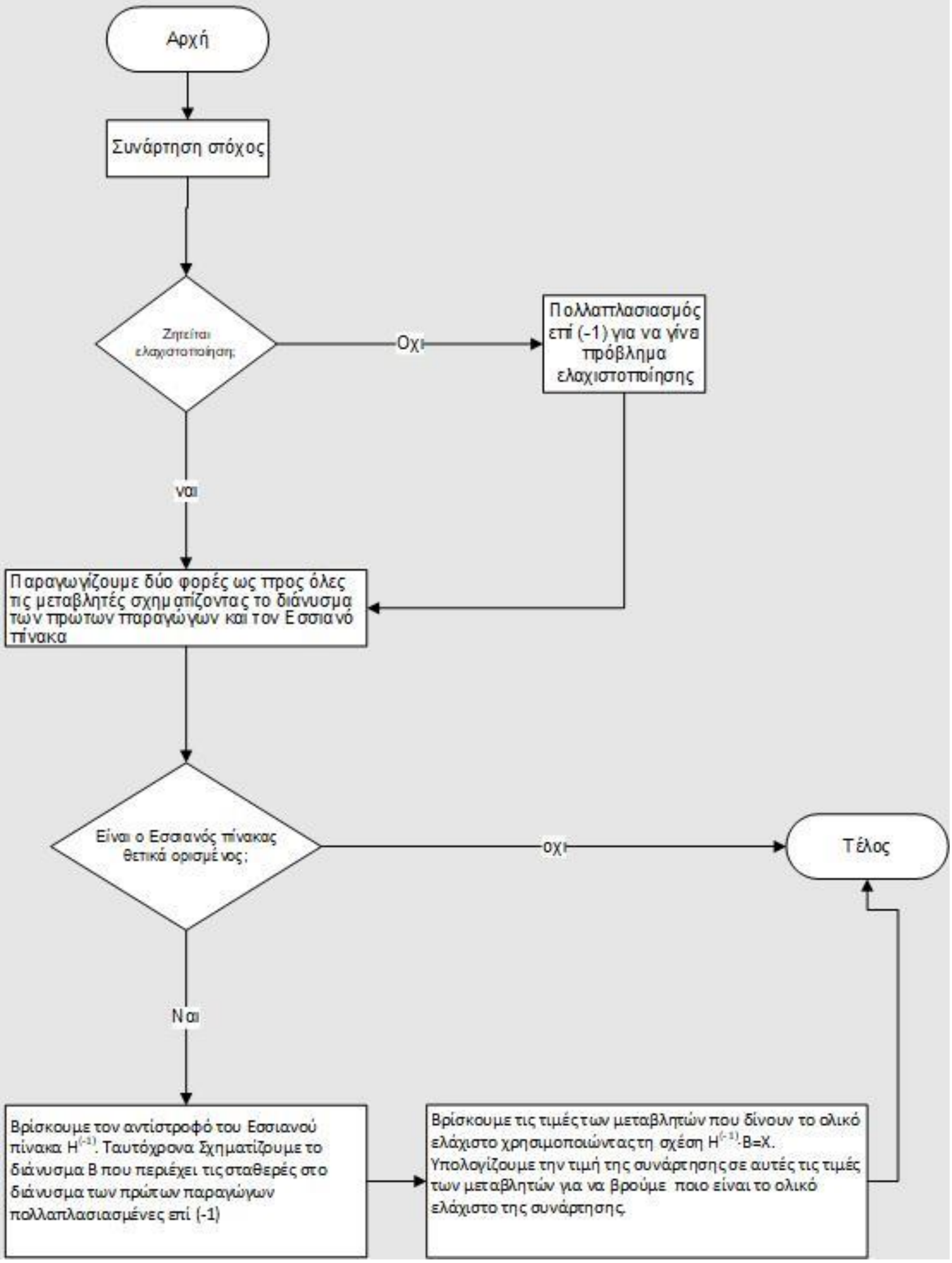
β. Αν ο πίνακας είναι θετικά ορισμένος ή ημιορισμένος η συνάρτηση έχει ολικό ελάχιστο.

βήμα 4 Βρίσκουμε τον αντίστροφο του Εσσιανού πίνακα όπου συνηθέστερα χρησιμοποιούμε την μέθοδο απαλοιφής του Gauss. Σχηματίζουμε το διάνυσμα Β που περιέχει τις σταθερές στο διάνυσμα των πρώτων παραγώγων πολλαπλασιασμένες επί (-1).

βήμα 5 Βρίσκουμε τις τιμές των μεταβλητών που δίνουν το ολικό ελάχιστο χρησιμοποιώντας τη σχέση $H^{-1} * B = X$. Υπολογίζουμε την τιμή της συνάρτησης σε αυτές τις τιμές των μεταβλητών για να βρούμε ποιο είναι το ολικό ελάχιστο της συνάρτησης.

βήμα 6 Τέλος.

Διάγραμμα ροής της μεθόδου τετραγωνικού προγραμματισμού χωρίς περιορισμούς



Παράδειγμα

Συνάρτηση στόχος:

$$\min(3 * x^2 + 2 * y^2 - 4 * x - 5 * y - 4 * x * y)$$

βήμα 1 Ζητάμε ελαχιστοποίηση άρα πάμε στο

βήμα 2 Παραγωγίζουμε δύο φορές.
Διάνυσμα πρώτης παραγώγου:

$$\nabla f = \begin{bmatrix} 6x - 4y - 4 \\ -4x + 4y - 5 \end{bmatrix}$$

Εσσιανός πίνακας (πίνακας δεύτερης παραγώγου)

$$H := Hessian(f, [x, y]) = \begin{bmatrix} 6 & -4 \\ -4 & 4 \end{bmatrix}$$

βήμα 3 Έλεγχος αν ο πίνακας είναι θετικά ορισμένος:

$$H_{(1,1)} = 6 > 0$$
$$\det(H) = 6 * 4 - (-4) * (-4) = 24 - 16 = 8 > 0$$

άρα η συνάρτηση έχει ολικό ελάχιστο και προχωράμε στο

βήμα 4 Βρίσκουμε τον αντίστροφο του Εσσιανού πίνακα.

$$H^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

Οι συντελεστές των πρωτοβάθμιων μεταβλητών είναι

$$B = \begin{bmatrix} -4 \\ -5 \end{bmatrix} * (-1) = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$$

βήμα 5 Πολλαπλασιάζουμε τον αντίστροφο του εσσιανού με το διάνυσμα B για να βρούμε τις τιμές των μεταβλητών που ελαχιστοποιούν τη συνάρτηση.

$$\begin{bmatrix} x \\ y \end{bmatrix} = H^{-1} * B = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{4} \end{bmatrix} * \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{9}{2} \\ \frac{23}{4} \end{bmatrix}$$

Αντικαθιστώντας τις τιμές στη συνάρτηση στόχο παίρνουμε:

$$f = 3 * x^2 + 2 * y^2 - 4 * x - 5 * y - 4 * x * y = -\frac{187}{8}$$

βήμα 6 Τέλος του αλγόριθμου.

3.1.3 Μεγιστοποίηση συνάρτησης στόχου με χρήση αριθμητικών μεθόδων

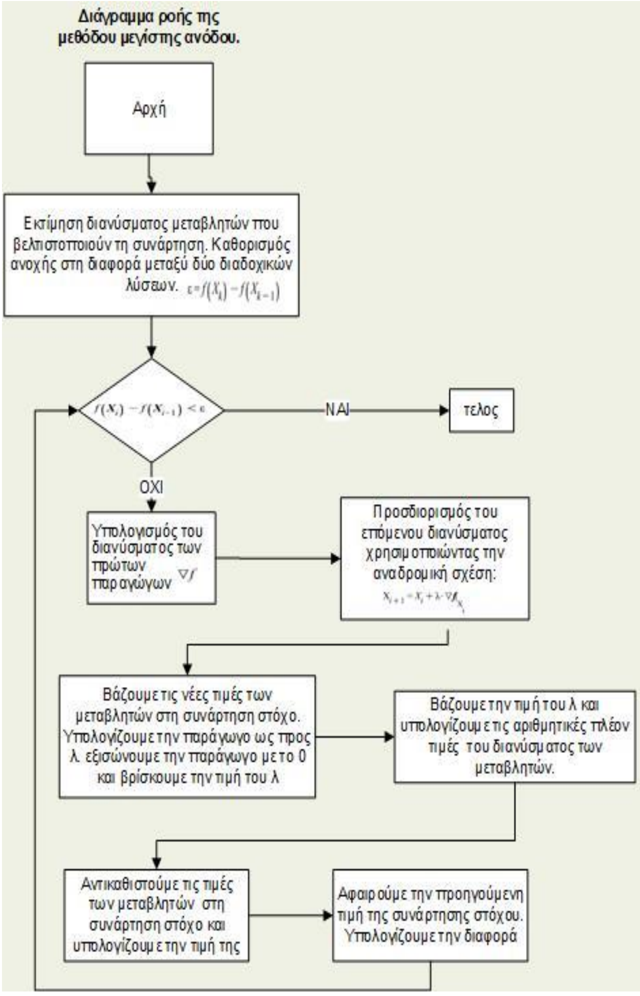
Μέθοδος μέγιστης ανόδου

βήμα 1 Επιλογή ενός αρχικού διανύσματος X_0 εκτιμώντας κατά προσέγγιση την πιθανή θέση του ολικού μέγιστου. Ταυτόχρονα καθορίζουμε ανοχή $\varepsilon = f(X_k) - f(X_{k-1})$ που θα γίνει αποδεκτό το διάνυσμα X_k σαν λύση του προβλήματος.

βήμα 2 Στη συνέχεια υπολογίζουμε το διάνυσμα ∇f που περιέχει τις πρώτες μερικές παραγώγους της συνάρτησης.

βήμα 3 Αρχίζουμε να προσδιορίζουμε διανύσματα $X_{1,} X_{2,\ldots}$ χρησιμοποιώντας την αναδρομική σχέση $X_{i+1} = X_i + \lambda * \nabla f_{X_i}$ όπου λ είναι μια ποσότητα όπου αντιπροσωπεύει ένα ολικό μέγιστο της συνάρτησης $f(x + \lambda * \nabla f_{X_i})$.

βήμα 4 Το τελευταίο διάνυσμα το οποίο θα δίνει την καθορισμένη στο βήμα 1 ανοχή θεωρείται ότι είναι το διάνυσμα που προσεγγίζει πιο πολύ το ολικό μέγιστο.



Παράδειγμα: Μεγιστοποίηση της συνάρτησης $f \coloneqq -(2*x-5)^2 - (y-3)^2$
Εκτιμάμε ότι το ολικό μέγιστο της συνάρτησης βρίσκεται στο διάνυσμα $X_0 = (2,1)^T$ όπου η συνάρτηση παίρνει την τιμή: -5. Θεωρούμε ότι μια ανεχτή διαφορά μεταξύ των διαδοχικών τιμών των συναρτήσεων είναι 0.05.
Βρίσκουμε το διάνυσμα των πρώτων παραγώγων της συνάρτησης:
$\frac{\partial}{\partial x}(f) = -8*x+20$ και $\frac{\partial}{\partial y}(f) = -2*y+6$

Δοκιμή πρώτη:

Προσθέτουμε στο διάνυσμα των εκτιμώμενων τιμών των μεταβλητών τις τιμές των παραγώγων στο σημείο εκτίμησης πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \lambda * \begin{bmatrix} -8*2+20 \\ -2*1+6 \end{bmatrix}$$

εφαρμόζουμε τις τιμές στην συνάρτηση στόχο:

$$-(2*x_1-5)^2 - (y_1-3)^2 = -(2*(4*\lambda+2)-5)^2 - (4*\lambda+1-3)^2 =$$
$$= -(8*\lambda-1)^2 - (4*\lambda-2)^2$$

Βρίσκουμε την πρώτη παράγωγο και την εξισώνουμε με το 0 για να βρούμε την τιμή του λ που βελτιστοποιεί την συνάρτηση:
$\frac{d}{d\lambda}(f_1) = -160*\lambda+32 = 0 \Rightarrow \lambda = \frac{1}{5}$

Στη συνέχεια εφαρμόζουμε την τιμή στο διάνυσμα που βρήκαμε για να βρούμε τις νέες τιμές των μεταβλητών:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \left(\frac{1}{5}\right) * \begin{bmatrix} -8 * 2 + 20 \\ -2 * 1 + 6 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ \frac{9}{5} \end{bmatrix}$$

Εφαρμόζοντας στην συνάρτηση στόχο τις νέες τιμές παίρνουμε την παρακάτω τιμή

$$-(2 * x - 5)^2 - (y - 3)^2 = -\left(2 * \frac{14}{5} - 5\right)^2 - \left(\frac{9}{5} - 3\right)^2 = -\frac{9}{25} - \frac{36}{25} = -\frac{45}{25} =$$
$$= -\frac{9}{5}$$

Η νέα τιμή είναι μεγαλύτερη από την τιμή της συνάρτησης στην οποία είχαμε εφαρμόσει το αρχικό διάνυσμα $X_0 = (2,1)^T$. Άρα θεωρούμε ότι οι νέες τιμές προσεγγίζουν πιο πολύ σε αυτές που βελτιστοποιούν τη συνάρτηση και συνεχίζουμε με την

Δοκιμή δεύτερη

Προσθέτουμε στο διάνυσμα των τιμών των μεταβλητών από την προηγούμενη δοκιμή τις τιμές των παραγώγων στο σημείο αυτό πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ \frac{9}{5} \end{bmatrix} + \lambda * \begin{bmatrix} -(8 * \frac{14}{5}) + 20 \\ -2 * \frac{9}{5} + 6 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} - \frac{12 * \lambda}{5} \\ \frac{9}{5} + \frac{12 * \lambda}{5} \end{bmatrix}$$

εφαρμόζουμε τις τιμές στη συνάρτηση στόχο

$$-(2 * x - 5)^2 - (y - 3)^2 = -\left(-\frac{24 * \lambda}{5} + \frac{3}{5}\right)^2 - \left(\frac{12 * k}{5} - \frac{6}{5}\right)^2$$

Βρίσκουμε την πρώτη παράγωγο και την εξισώνουμε με το 0 για να βρούμε την τιμή του λ που βελτιστοποιεί την συνάρτηση:

$$\frac{d}{d\lambda}\left(-\left(-\frac{24 * \lambda}{5} + \frac{3}{5}\right)^2 - \left(\frac{12 * \lambda}{5} - \frac{6}{5}\right)^2\right) = -\frac{288 * \lambda}{5} + \frac{288}{25} = 0 \Rightarrow \lambda = \frac{1}{5}$$

Στη συνέχεια εφαρμόζουμε την τιμή στο διάνυσμα που βρήκαμε για να βρούμε τις νέες τιμές των μεταβλητών:

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{14}{5} \\ \frac{9}{5} \end{bmatrix} + \frac{1}{5} * \begin{bmatrix} -(8 * \frac{14}{5}) + 20 \\ -2 * \frac{9}{5} + 6 \end{bmatrix} = \begin{bmatrix} \frac{58}{25} \\ \frac{57}{25} \end{bmatrix}$$

Εφαρμόζοντας στην συνάρτηση στόχο τις νέες τιμές παίρνουμε την παρακάτω τιμή

$-(2 * x - 5)^2 - (y - 3)^2 = -\left(2 * \frac{58}{25} - 5\right)^2 - \left(\frac{57}{25} - 3\right)^2 = -\frac{81}{125}$ η διαφορά από την πρώτη δοκιμή είναι $-\frac{81}{125} - \left(-\frac{9}{5}\right) = \frac{144}{25}$ και δεν έχει φτάσει ακόμα στο ανεχτό όριο. Επομένως προχωράμε στην

Δοκιμή τρίτη

Προσθέτουμε στο διάνυσμα των τιμών των μεταβλητών από την προηγούμενη δοκιμή τις τιμές των παραγώγων στο σημείο αυτό πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{58}{25} \\ \frac{57}{25} \end{bmatrix} + \lambda * \begin{bmatrix} -(8 * \frac{58}{25}) + 20 \\ -2 * \frac{57}{25} + 6 \end{bmatrix} = \begin{bmatrix} \frac{58}{25} + \frac{36 * \lambda}{25} \\ \frac{57}{25} + \frac{36 * \lambda}{25} \end{bmatrix}$$

εφαρμόζουμε τις τιμές στη συνάρτηση στόχο

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\left(-\frac{9}{25} + \frac{72 * \lambda}{25}\right)^2 - \left(-\frac{18}{25} + \frac{36 * \lambda}{25}\right)^2$$

Βρίσκουμε την πρώτη παράγωγο και την εξισώνουμε με το 0 για να βρούμε την τιμή του λ που βελτιστοποιεί την συνάρτηση:

$$\frac{d}{d\lambda}\left(-\left(-\frac{9}{25} + \frac{72 * \lambda}{25}\right)^2 - \left(-\frac{18}{25} + \frac{36 * \lambda}{25}\right)^2\right) = \frac{2592}{625} - \frac{2592 * \lambda}{125}$$

Επιλύοντας έχουμε: $\frac{2592}{625} - \frac{2592*\lambda}{125} = 0 \Longrightarrow \lambda = \frac{1}{5}$ και εφαρμόζουμε την τιμή που βρήκαμε στο διάνυσμα:

$$\begin{bmatrix} x_3 \\ y_3 \end{bmatrix} = \begin{bmatrix} \frac{58}{25} \\ \frac{57}{25} \end{bmatrix} + \frac{1}{5} * \begin{bmatrix} -(8 * \frac{58}{25}) + 20 \\ -2 * \frac{57}{25} + 6 \end{bmatrix} = \begin{bmatrix} \frac{326}{125} \\ \frac{321}{125} \end{bmatrix}$$

εφαρμόζοντας στη συνάρτηση στόχο παίρνουμε την τιμή:

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\left(2 * \frac{326}{125} - 5\right)^2 - \left(\frac{321}{125} - 3\right)^2 = -\frac{729}{3125} =$$
$$= -0.23328$$

Η διαφορά από την προηγούμενη εκτίμηση είναι

$-\frac{729}{3125}-\left(-\frac{5881}{125}\right)=\frac{1296}{3125}=0.41472$ η οποία εκτιμάται ότι είναι και πάλι εκτός των ορίων ανοχής. Άρα συνεχίζουμε με την

Δοκιμή τέταρτη

Προσθέτουμε στο διάνυσμα των τιμών των μεταβλητών από την προηγούμενη δοκιμή τις τιμές των παραγώγων στο σημείο αυτό πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} = \begin{bmatrix} \frac{326}{125} \\ \frac{321}{125} \end{bmatrix} + \lambda * \begin{bmatrix} -8 * \frac{326}{125} + 20 \\ -2 * \frac{321}{125} + 6 \end{bmatrix} = \begin{bmatrix} \frac{326}{125} - \frac{108 * \lambda}{125} \\ \frac{321}{125} + \frac{108 * \lambda}{125} \end{bmatrix}$$

εφαρμόζουμε τις τιμές στη συνάρτηση στόχο

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\left(-\frac{216}{125} * \lambda + \frac{27}{125}\right)^2 - \left(\frac{108}{125} * \lambda - \frac{54}{125}\right)^2$$

Βρίσκουμε την πρώτη παράγωγο και την εξισώνουμε με το 0 για να βρούμε την τιμή του λ που βελτιστοποιεί την συνάρτηση:

$$\frac{d}{d\lambda}\left(-\left(-\frac{216}{125} * \lambda + \frac{27}{125}\right)^2 - \left(\frac{108}{125} * \lambda - \frac{54}{125}\right)^2\right) = -\frac{23328}{3125} * \lambda + \frac{23328}{15625} = 0$$
$$\Rightarrow \lambda = \frac{1}{5}$$

αντικαθιστώντας την τιμή του λ στο διάνυσμα παίρνουμε:

$$\begin{bmatrix} x_4 \\ y_4 \end{bmatrix} = \begin{bmatrix} \frac{326}{125} \\ \frac{321}{125} \end{bmatrix} + \frac{1}{5} * \begin{bmatrix} -8 * \frac{326}{125} + 20 \\ -2 * \frac{321}{125} + 6 \end{bmatrix} = \begin{bmatrix} \frac{1522}{625} \\ \frac{1713}{625} \end{bmatrix}$$

Και η τιμή της συνάρτησης f γίνεται:
$f := -(2 * x - 5)^2 - (y - 3)^2 = -\frac{6561}{78125} = -0.083980$. Η διαφορά από την προηγούμενη τιμή που προσεγγίσαμε είναι:

$$-\frac{6561}{78125} - \left(-\frac{729}{3125}\right) = \frac{11664}{78125} = 0.149299 > 0.05 = \varepsilon$$

Άρα προχωράμε στην

Δοκιμή πέμπτη

Προσθέτουμε στο διάνυσμα των τιμών των μεταβλητών από την προηγούμενη δοκιμή τις τιμές των παραγώγων στο σημείο αυτό πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_5 \\ y_5 \end{bmatrix} = \begin{bmatrix} \frac{1522}{625} \\ \frac{1713}{625} \end{bmatrix} + \lambda * \begin{bmatrix} -8 * \frac{1522}{625} + 20 \\ -2 * \frac{321}{625} + 6 \end{bmatrix} = \begin{bmatrix} \frac{324}{625} * \lambda + \frac{1522}{625} \\ \frac{3108}{625} * \lambda + \frac{1713}{625} \end{bmatrix}$$

εφαρμόζουμε τις τιμές στη συνάρτηση στόχο

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\left(\frac{642 * \lambda}{625} - \frac{81}{625}\right)^2 - \left(\frac{324 * \lambda}{625} - \frac{162}{625}\right)^2$$

Παραγωγίζουμε:

$$\frac{d}{d\lambda}\left(-\left(\frac{648}{625} * \lambda - \frac{81}{625}\right)^2 - \left(\frac{324}{625} * \lambda - \frac{162}{625}\right)^2\right) = -\frac{209952}{78125} * \lambda + \frac{209952}{390625}$$

Εξισώνουμε την παράγωγο με το 0 και βρίσκουμε την τιμή του λ

$$\frac{209952}{390625} - \frac{209952}{78125} * \lambda \Longrightarrow \lambda = \frac{1}{5}$$

αντικαθιστώντας την τιμή του λ στο διάνυσμα παίρνουμε:

$$\begin{bmatrix} x_5 \\ y_5 \end{bmatrix} = \begin{bmatrix} \frac{1522}{625} \\ \frac{1713}{625} \end{bmatrix} + \frac{1}{5} * \begin{bmatrix} -\left(8 * \frac{1522}{625}\right) + 20 \\ -2 * \frac{1713}{625} + 6 \end{bmatrix} = \begin{bmatrix} \frac{7934}{3125} \\ \frac{8889}{3125} \end{bmatrix} = \begin{bmatrix} 2.53888 \\ 2.84448 \end{bmatrix}$$

και η τιμή της συνάρτησης f γίνεται:

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\frac{59049}{1953125} = -0.030233$$

Η διαφορά με την προηγούμενη τιμή της συνάρτησης είναι:

$$\left(-\frac{59049}{1953125}\right) - \left(-\frac{11664}{78125}\right) = 0.1190611 > 0.05 = \varepsilon$$

Άρα συνεχίζουμε στην

Δοκιμή έκτη

Προσθέτουμε στο διάνυσμα των τιμών των μεταβλητών από την προηγούμενη δοκιμή τις τιμές των παραγώγων στο σημείο αυτό πολλαπλασιασμένες με άγνωστο συντελεστή λ.

$$\begin{bmatrix} x_6 \\ y_6 \end{bmatrix} = \begin{bmatrix} \frac{7934}{3125} \\ \frac{8889}{3125} \end{bmatrix} + \lambda * \begin{bmatrix} -\left(8 * \frac{7934}{3125}\right) + 20 \\ -2 * \frac{8889}{3125} + 6 \end{bmatrix} = \begin{bmatrix} \frac{7934}{3125} - \frac{972}{3125} * \lambda \\ \frac{8889}{3125} + \frac{972}{3125} * \lambda \end{bmatrix}$$

εφαρμόζουμε τις τιμές στη συνάρτηση στόχο

$$f(\lambda) = -\left(-\frac{1944}{3125} * \lambda - \frac{243}{3125}\right)^2 - \left(\frac{972}{3125} * \lambda - \frac{486}{3125}\right)^2$$

παραγωγίζουμε και εξισώνουμε με το 0 για να βρούμε το λ που δίνει κρίσιμο σημείο της συνάρτησης

$$\frac{df}{d\lambda}\left(-\left(-\frac{1944}{3125} * \lambda + \frac{243}{3125}\right)^2 - \left(\frac{972}{3125} * \lambda - \frac{486}{3125}\right)^2\right) =$$
$$= -\frac{1889568}{1953125} * \lambda + \frac{1889568}{9765625} \Longrightarrow \lambda = \frac{1}{5}$$

αντικαθιστούμε την τιμή του λ στο διάνυσμα:

$$\begin{bmatrix} x_6 \\ y_6 \end{bmatrix} = \begin{bmatrix} \frac{7934}{3125} \\ \frac{8889}{3125} \end{bmatrix} + \frac{1}{5} * \begin{bmatrix} -\left(8 * \frac{7934}{3125}\right) + 20 \\ -2 * \frac{8889}{3125} + 6 \end{bmatrix} = \begin{bmatrix} \frac{38698}{15625} \\ \frac{45417}{15625} \end{bmatrix} = \begin{bmatrix} 2.477 \\ 2.906 \end{bmatrix}$$

Εφαρμόζουμε τις τιμές του διανύσματος στη συνάρτηση και παίρνουμε:

$$f := -(2 * x - 5)^2 - (y - 3)^2 = -\frac{531441}{48828125} = -0.010884$$

Η διαφορά από την προηγούμενη δοκιμή είναι:

$$-\frac{531441}{48828125} - \left(-\frac{59049}{1953125}\right) = \frac{944784}{48828125} = 0.019349 < 0.05 = \varepsilon$$

άρα η διαδικασία που έχουμε ορίσει ολοκληρώθηκε.

Σημείωση: Είναι φανερό ότι το ολικό μέγιστο της συνάρτησης είναι το 0 με
$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2.5 \\ 3 \end{bmatrix}$
Θεωρούμε όμως ότι με αυτό το παράδειγμα δείξαμε την αποτελεσματικότητα της μεθόδου αυτής καθώς επίσης ότι με συνεχείς επαναλήψεις θα βελτιώναμε την εκτίμηση για το διάνυσμα που δίνει την βέλτιστη τιμή. Θέλουμε όμως να δείξουμε και συνέπεια προς τον αλγόριθμο ως προς την τιμή ε της ανοχής.

Μέθοδος Newton-Raphson

βήμα 1 Επιλογή ενός αρχικού διανύσματος X_0 εκτιμώντας κατά προσέγγιση την πιθανή θέση του ολικού μέγιστου.. Ταυτόχρονα καθορίζουμε ανοχή $\varepsilon = f(X_k) - f(X_{k-1})$ που θα γίνει αποδεκτό το διάνυσμα X_0 σαν λύση του προβλήματος.

βήμα 2 Εάν $f(X_{i+1}) - f(X_i) < \varepsilon$ τερματίζεται. Διαφορετικά:

βήμα 3 Υπολογίζουμε το διάνυσμα ∇f που περιέχει τις πρώτες μερικές παραγώγους της συνάρτησης.

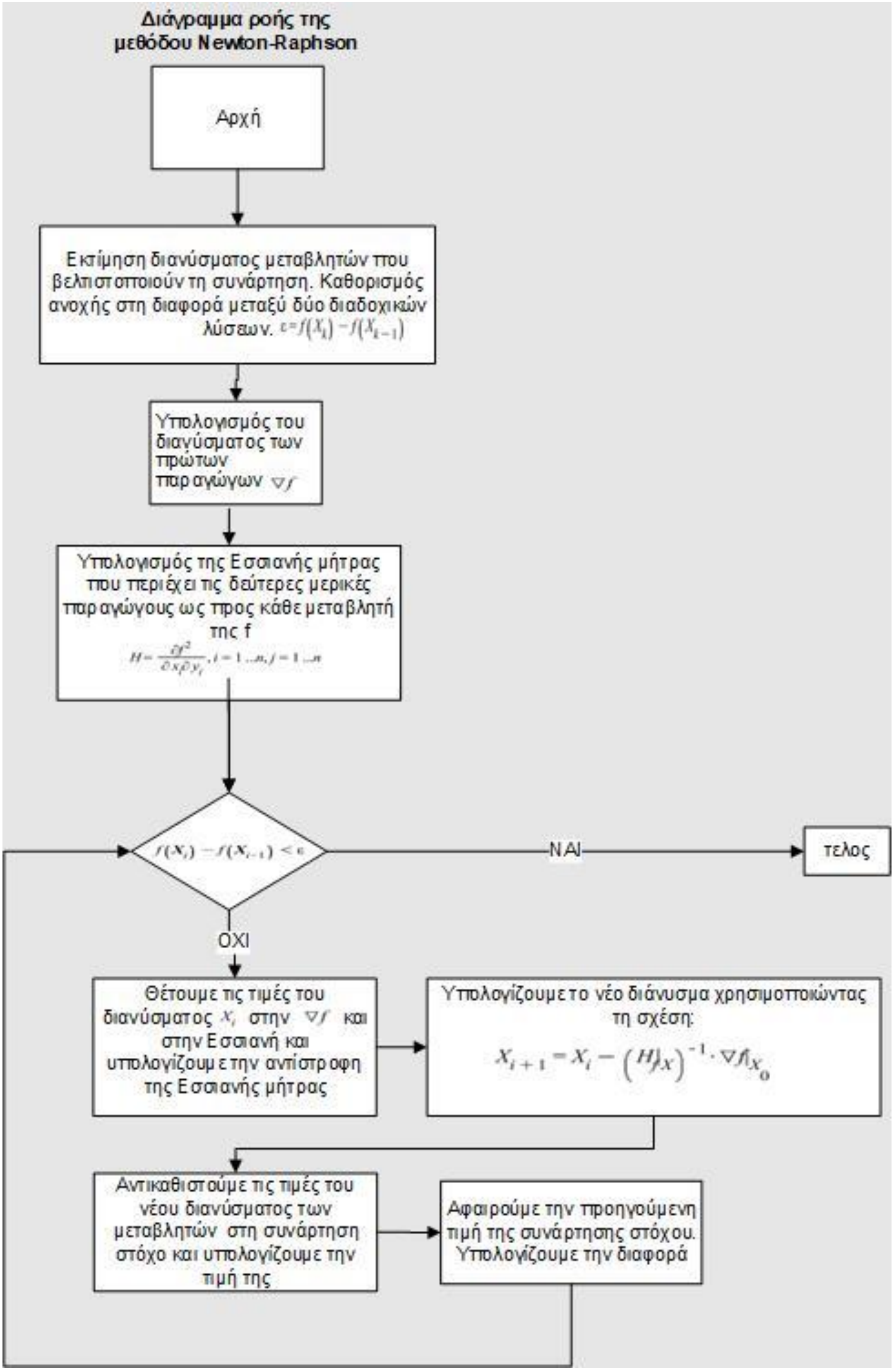
βήμα 4 Υπολογίζουμε την μήτρα Hess που περιέχει τις δεύτερες μερικές παραγώγους της συνάρτησης.

βήμα 5 Εφαρμόζουμε τις τιμές του διανύσματος X_i και στο διάνυσμα ∇f και στην μήτρα Hess. Δηλαδή τα $\nabla f_{\hat{X}}$ και $H_{f_{X_n}}$. Υπολογίζουμε την αντίστροφη της μήτρας Hess $\left(H_{f_{X_n}}\right)^{-1}$.

βήμα 6 Προσδιορίζουμε το επόμενο διάνυσμα X_{i+1} με τον τύπο $X_{i+1} = X_i -$ $- \left(H_{f_{X_i}}\right)^{-1} * \nabla_{f_{X_i}}$.

βήμα 7 Υπολογίζουμε την τιμή της $f(X_{i+1})$ και αφαιρούμε από αυτήν την τιμή της $f(X_i)$.

βήμα 8 Πάμε στο βήμα 2.


Διάγραμμα ροής της μεθόδου Newton-Raphson
Αρχή
Εκτίμηση διανύσματος μεταβλητών που βελτιστοποιούν τη συνάρτηση. Καθορισμός ανοχής στη διαφορά μεταξύ δύο διαδοχικών λύσεων. $\varepsilon = f(X_i) - f(X_{i-1})$
Υπολογισμός του διανύσματος των πρώτων παραγώγων ∇f
Υπολογισμός της Εσσιανής μήτρας που περιέχει τις δεύτερες μερικές παραγώγους ως προς κάθε μεταβλητή της f
$H = \frac{\partial f^2}{\partial x_i \partial y_j}, i = 1 \dots n, j = 1 \dots n$
$f(X_i) - f(X_{i-1}) < \varepsilon$
ΝΑΙ
τελος
ΟΧΙ
Θέτουμε τις τιμές του διανύσματος X_i στην ∇f και στην Εσσιανή και υπολογίζουμε την αντίστροφη της Εσσιανής μήτρας
Υπολογίζουμε το νέο διάνυσμα χρησιμοποιώντας τη σχέση:
$X_{i+1} = X_i - \left(H_f x\right)^{-1} \cdot \nabla f_{X_0}$
Αντικαθιστούμε τις τιμές του νέου διανύσματος των μεταβλητών στη συνάρτηση στόχο και υπολογίζουμε την τιμή της
Αφαιρούμε την προηγούμενη τιμή της συνάρτησης στόχου. Υπολογίζουμε την διαφορά

Παράδειγμα:
Μεγιστοποίηση της συνάρτησης:

$$f := -2*(x+3)^2 - (y+4)^2 - x^2 - y^2 - (z-3)^2 - z^2 + xz + yx - 10$$

Καθορίζουμε σαν αρχικό διάνυσμα το $X_0 := \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$ και ανοχή ίση με 0.05.
Υπολογίζοντας την τιμή της συνάρτησης με τις τιμές του διανύσματος X_0 βρίσκουμε ότι είναι -91.
Υπολογισμός του διανύσματος των πρώτων μερικών παραγώγων της μήτρας Hess και της αντίστροφής της:

$$\nabla f = \begin{bmatrix} -6*x - 12 + z + y \\ -4*y - 8 + x \\ -4*z + 6 + x \end{bmatrix}, H = \begin{bmatrix} -6 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & -4 \end{bmatrix}, H^{-1} = \begin{bmatrix} -\frac{4}{23} & -\frac{1}{23} & 0 \\ -\frac{1}{23} & -\frac{6}{23} & 0 \\ 0 & 0 & -\frac{1}{4} \end{bmatrix}$$

Δοκιμή πρώτη:
Από τις τιμές του κατ’ εκτίμηση διανύσματος, αφαιρούμε το γινόμενο της αντίστροφης μήτρας Hess επί το διάνυσμα των πρώτων παραγώγων εφαρμόζοντας τις τιμές των μεταβλητών στο διάνυσμα ∇f και στη μήτρα Hess (αν χρειάζεται).

$$X_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} -\frac{4}{23} & -\frac{1}{23} & 0 \\ -\frac{1}{23} & -\frac{6}{23} & 0 \\ 0 & 0 & -\frac{1}{4} \end{bmatrix} * \begin{bmatrix} -20 \\ -10 \\ -4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} - \begin{bmatrix} \frac{90}{23} \\ \frac{80}{23} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{44}{23} \\ -\frac{57}{23} \\ 2 \end{bmatrix}$$

εφαρμόζοντας τις τιμές του διανύσματος X_1 στην f παίρνουμε ότι $f = -28.565$ ενώ η διαφορά από την προηγούμενη τιμή της συνάρτησης είναι: $-28.565 - (-91) = 62.435 > 0.05 = \varepsilon$ άρα συνεχίζουμε στην
Δοκιμή δεύτερη.
Το διάνυσμα $X_1 = \begin{bmatrix} -\frac{44}{23} \\ -\frac{57}{23} \\ 2 \end{bmatrix}$ είναι το νέο διάνυσμα με το οποίο θα συνεχίσουμε την προσπάθεια βελτιστοποίησης. Επαναλαμβάνουμε την παραπάνω διαδικασία και έχουμε:

$$X_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} -\frac{44}{23} \\ -\frac{57}{23} \\ 2 \end{bmatrix} - \begin{bmatrix} -\frac{4}{23} & -\frac{1}{23} & 0 \\ -\frac{1}{23} & -\frac{6}{23} & 0 \\ 0 & 0 & -\frac{1}{4} \end{bmatrix} * \begin{bmatrix} -1 \\ 0 \\ -\frac{90}{23} \end{bmatrix} = \begin{bmatrix} -\frac{44}{23} \\ -\frac{57}{23} \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{4}{23} \\ \frac{1}{23} \\ \frac{45}{46} \end{bmatrix} = \begin{bmatrix} -\frac{48}{23} \\ -\frac{58}{23} \\ \frac{47}{46} \end{bmatrix}$$

αντικαθιστώντας τις τιμές στην f έχουμε:

$$f := -2 * (x+3)^2 - (y+4)^2 - x^2 - y^2 - (z-3)^2 - z^2 + xz + yx - 10 =$$

$$= -\frac{27925}{1058} = -26.394$$

Ενώ η διαφορά από την προηγούμενη τιμή της διαμορφώνεται σε:

$$-\frac{657}{23} - \left(-\frac{27925}{1058}\right) = \frac{2297}{1058} = 2.171 > \varepsilon = 0.05$$

άρα προχωράμε στην
Δοκιμή Τρίτη:

Το διάνυσμα $X_2 = \begin{bmatrix} -\frac{48}{23} \\ -\frac{58}{23} \\ \frac{47}{46} \end{bmatrix}$ είναι το νέο διάνυσμα με το οποίο θα συνεχίσουμε την προσπάθεια βελτιστοποίησης. Επαναλαμβάνουμε την παραπάνω διαδικασία και έχουμε:

$$X_3 = \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} -\frac{48}{23} \\ -\frac{58}{23} \\ \frac{47}{46} \end{bmatrix} - \begin{bmatrix} -\frac{4}{23} & -\frac{1}{23} & 0 \\ -\frac{1}{23} & -\frac{6}{23} & 0 \\ 0 & 0 & -\frac{1}{4} \end{bmatrix} * \begin{bmatrix} -\frac{45}{46} \\ 0 \\ -\frac{4}{23} \end{bmatrix} = \begin{bmatrix} -\frac{48}{23} \\ -\frac{58}{23} \\ \frac{47}{46} \end{bmatrix} - \begin{bmatrix} \frac{90}{529} \\ \frac{45}{1058} \\ \frac{1}{23} \end{bmatrix} =$$

$$= \begin{bmatrix} -\frac{1194}{529} \\ -\frac{2713}{1058} \\ \frac{45}{46} \end{bmatrix}$$

αντικαθιστώντας τις τιμές στην f έχουμε:

$$f := -2 * (x+3)^2 - (y+4)^2 - x^2 - y^2 - (z-3)^2 - z^2 + xz + yx - 10 =$$
$$= -\frac{27925}{1058} = -26.394$$

$-\frac{27925}{1058} - \left(-\frac{27925}{1058}\right) = 0 < 0.05 = \varepsilon.$ Η διαδικασία ολοκληρώθηκε. Η κατά προσέγγιση με αριθμητική μέθοδο της μέγιστης τιμής της συνάρτησης είναι $-\frac{27925}{1058}$ με τιμές των μεταβλητών $X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -\frac{1194}{529} \\ -\frac{2713}{1058} \\ \frac{45}{46} \end{bmatrix}$.

3.2 Βελτιστοποίηση τετραγωνικών συναρτήσεων πολλών μεταβλητών με περιορισμούς

Όπως και πιο πάνω είπαμε εδώ έχουμε συνάρτηση πολλών μεταβλητών (συνάρτηση στόχο) που οι τιμές των μεταβλητών της υπόκεινται σε σειρά περιορισμών και αναζητάμε τις τιμές εκείνες των μεταβλητών που δίνουν την βέλτιστη τιμή στη συνάρτηση. Οι μορφές που μπορεί να έχει είναι:

3.2.1 Με εξισωτικούς περιορισμούς

Μέθοδος Newton-Raphson

Ο αλγόριθμος είναι παρόμοιος με την ομώνυμη μέθοδο για βελτιστοποίηση χωρίς μεταβλητές με τη διαφορά ότι η συνάρτηση που παραγωγίζεται είναι η συνάρτηση Lagrange.

Μέθοδος Newton-Raphson

βήμα 1 Σχηματίζουμε την συνάρτηση Lagrange $L(X, \lambda)$

βήμα 2 Σχηματίζουμε το διάνυσμα των πρώτων μερικών παραγώγων ως προς κάθε μεταβλητή και ως προς κάθε πολλαπλασιαστή Lagrange.

βήμα 3 Σχηματίζουμε τον Εσσιανό πίνακα που περιέχει τις δεύτερες παραγώγους ως προς κάθε μεταβλητή και ως προς τους πολλαπλασιαστές Lagrange και βρίσκουμε τον αντίστροφό του.

βήμα 4 Επιλογή ενός αρχικού διανύσματος X_0 εκτιμώντας κατά προσέγγιση την πιθανή θέση του ολικού μέγιστου.. Ταυτόχρονα καθορίζουμε ανοχή $\varepsilon = f(X_k) - f(X_{k-1})$ που θα γίνει αποδεκτό το διάνυσμα X_k σαν λύση του προβλήματος.

βήμα 5 Εάν $f(X_{i+1}) - f(X_i) < \varepsilon$ τερματίζεται. Διαφορετικά:

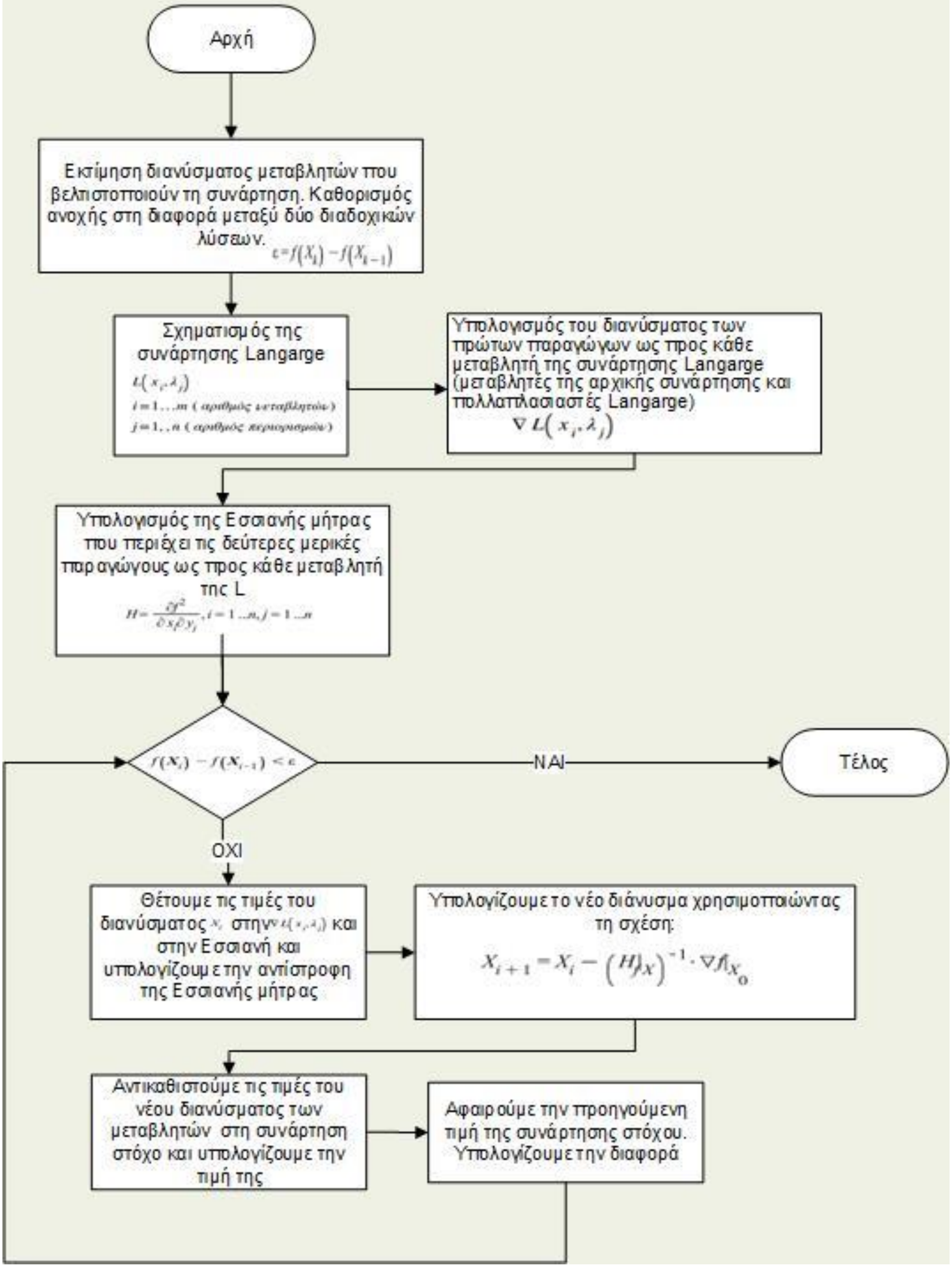
βήμα 6 Εφαρμόζουμε τις τιμές του διανύσματος X_i και στο διάνυσμα ∇L και στην αντίστροφη μήτρα Hess. Δηλαδή τα $\nabla L_{\bar{X}}$ και $\left(H_{L_{X_n}}\right)^{-1}$.

βήμα 7 Προσδιορίζουμε το επόμενο διάνυσμα X_{i+1} με τον τύπο $X_{i+1} = X_i - \left(H_{L_{X_i}}\right)^{-1} * \nabla \mathrm{L} f_{X_i}$

βήμα 8 Υπολογίζουμε την τιμή της $f(X_{i+1})$ και αφαιρούμε από αυτήν την τιμή της $f(X_i)$

βήμα 9 Πάμε στο βήμα 5.

Διάγραμμα ροής της μεθόδου Newton-Raphson για βελτιστοποίηση με περιορισμούς



Παράδειγμα

Μεγιστοποίηση της συνάρτησης:

$f := -(2*(x-2)^2 + 2*x*y + 3*(y-1)^2)$

με περιορισμούς

$4*x + 3*y - 10 = 0$
$3*\mathrm{x} + \mathrm{y} - 6 = 0$

βήμα 1 Σχηματίζουμε την συνάρτηση Lagrange
$L(x,y,k,m) = (4*x + 3*y - 10)*k - (3*\mathrm{x} + \mathrm{y} - 6)*\mathrm{m} -$
$-2(x-2)^2 - 2*x*y - 3*(y-1)^2$

όπου (k,m) οι πολλαπλασιαστές Lagrange

βήμα 2 Υπολογίζουμε το διάνυσμα των πρώτων μερικών παραγώγων ως προς κάθε μεταβλητή της συνάρτησης Lagrange

$$\begin{bmatrix} \frac{\partial f}{\partial x}(Lag) \\ \frac{\partial f}{\partial y}(Lag) \\ \frac{\partial f}{\partial k}(Lag) \\ \frac{\partial f}{\partial m}(Lag) \end{bmatrix} = \begin{bmatrix} -4*k - 3*m - 4*x + 8 - 2*y \\ -3*k - m - 2*x - 6*y + 6 \\ -4*x - 3*y + 10 \\ 6 - 3*x - y \end{bmatrix}$$

βήμα 3 Υπολογίζουμε τον εσσιανό πίνακα που περιέχει τις δεύτερες μερικές παραγώγους ως προς όλες τις μεταβλητές της συνάρτησης Lagrange και βρίσκουμε τον αντίστροφό του.

$$\begin{bmatrix} \frac{\partial^2}{\partial x^2}(Lag) & \frac{\partial^2}{\partial x \partial y}(Lag) & \frac{\partial^2}{\partial x \partial k}(Lag) & \frac{\partial^2}{\partial x \partial m}(Lag) \\ \frac{\partial^2}{\partial y \partial x}(Lag) & \frac{\partial^2}{\partial y \partial y}(Lag) & \frac{\partial^2}{\partial y \partial k}(Lag) & \frac{\partial^2}{\partial y \partial m}(Lag) \\ \frac{\partial^2}{\partial k \partial x}(Lag) & \frac{\partial^2}{\partial k \partial y}(Lag) & \frac{\partial^2}{\partial k \partial k}(Lag) & \frac{\partial^2}{\partial k \partial m}(Lag) \\ \frac{\partial^2}{\partial m \partial x}(Lag) & \frac{\partial^2}{\partial m \partial y}(Lag) & \frac{\partial^2}{\partial m \partial k}(Lag) & \frac{\partial^2}{\partial m \partial m}(Lag) \end{bmatrix} =$$

$$= \begin{bmatrix} -4 & -2 & -4 & -3 \\ -2 & -6 & -3 & -1 \\ -4 & -3 & 0 & 0 \\ -3 & -1 & 0 & 0 \end{bmatrix}$$

ο αντίστροφος πίνακας είναι:

$$H^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ \frac{1}{5} & -\frac{3}{5} & \frac{46}{25} & -\frac{58}{25} \\ -\frac{3}{5} & \frac{4}{5} & -\frac{58}{25} & \frac{84}{25} \end{bmatrix}$$

βήμα 4 Επιλέγουμε σαν αρχικό διάνυσμα για την πιθανή θέση των μεταβλητών το διάνυσμα:

$$X = \begin{bmatrix} x \\ y \\ k \\ m \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

Ορίζουμε την διαφορά όπου θα κρίνουμε ότι έχουμε πλησιάσει τη βέλτιστη τιμή της συνάρτησης $f(X_{i+1}) - f(X_i) < \varepsilon = 0.05$

βήμα 5 Η τιμή της συνάρτησης με βάση την αρχική εκτίμηση είναι:

$$f := -(2 * (x-2)^2 + 2 * x * y + 3 * (y-1)^2) = -17$$

βήμα 6 Εφαρμόζουμε τις τιμές του διανύσματος στον πίνακα που περιέχει τις πρώτες παραγώγους

$$\nabla L_{\hat{X}} = \begin{bmatrix} -4 * k - 3 * m - 4 * x + 8 - 2 * y \\ -3 * k - m - 2 * x - 6 * y + 6 \\ -4 * x - 3 * y + 10 \\ 6 - 3 * x - y \end{bmatrix} = \begin{bmatrix} -15 \\ -16 \\ -8 \\ -5 \end{bmatrix}$$

ο αντίστροφος του Εσσιανού πίνακα είναι ο

$$H^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ \frac{1}{5} & -\frac{3}{5} & \frac{46}{25} & -\frac{58}{25} \\ -\frac{3}{5} & \frac{4}{5} & -\frac{58}{25} & \frac{84}{25} \end{bmatrix}$$

βήμα 7 Προσδιορίζουμε το επόμενο διάνυσμα X_{i+1} με τον τύπο $X_{i+1} = X_i - \left(H_{L_{X_i}}\right)^{-1} * \nabla \mathrm{L} f_{X_i}$

$$X_1 = \begin{bmatrix} x_1 \\ y_1 \\ k_1 \\ m_1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & \frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ \frac{1}{5} & -\frac{3}{5} & \frac{46}{25} & -\frac{58}{25} \\ -\frac{3}{5} & \frac{4}{5} & -\frac{58}{25} & \frac{84}{25} \end{bmatrix} * \begin{bmatrix} -15 \\ -16 \\ -8 \\ -5 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} \frac{7}{5} \\ \frac{4}{5} \\ \frac{87}{25} \\ -\frac{51}{25} \end{bmatrix} = \begin{bmatrix} \frac{8}{5} \\ \frac{6}{5} \\ -\frac{62}{25} \\ \frac{76}{25} \end{bmatrix}$$

βήμα 8 Υπολογίζουμε την τιμή της αρχικής συνάρτησης με τις τιμές του νέου διανύσματος

$$f_{\hat{X}} = -(2 * (x-2)^2 + 2 * x * y + 3 * (y-1)^2) = -4.28$$

βήμα 9 Επιστροφή στο βήμα 5

$$\nabla \mathrm{L}_{X_1} = \begin{bmatrix} \frac{\partial f}{\partial x}(Lag) \\ \frac{\partial f}{\partial y}(Lag) \\ \frac{\partial f}{\partial k}(Lag) \\ \frac{\partial f}{\partial m}(Lag) \end{bmatrix} = \begin{bmatrix} -4 * k - 3 * m - 4 * x + 8 - 2 * y \\ -3 * k - m - 2 * x - 6 * y + 6 \\ -4 * x - 3 * y + 10 \\ 6 - 3 * x - y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

βήμα 5 $f(X_{i+1}) - f(X_i) = -4.28 - (-17) = 12.72 > \varepsilon = 0.05$ άρα πάμε με τις τιμές στου νέου διανύσματος στο

βήμα 6 Εφαρμόζουμε τις τιμές του διανύσματος στον πίνακα που περιέχει τις πρώτες παραγώγους και στον Εσσιανό πίνακα

$$H^{-1} = \begin{bmatrix} 0 & 0 & \frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ \frac{1}{5} & -\frac{3}{5} & \frac{46}{25} & -\frac{58}{25} \\ -\frac{3}{5} & \frac{4}{5} & -\frac{58}{25} & \frac{84}{25} \end{bmatrix}$$

βήμα 7 Προσδιορίζουμε το επόμενο διάνυσμα X_{i+1} με τον τύπο
$X_{i+1} = X_i - \left(H_{L_{X_i}}\right)^{-1} * \nabla L f_{X_i}$

$$X_2 = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & \frac{1}{5} & -\frac{3}{5} \\ 0 & 0 & -\frac{3}{5} & \frac{4}{5} \\ \frac{1}{5} & -\frac{3}{5} & \frac{46}{25} & -\frac{58}{25} \\ -\frac{3}{5} & \frac{4}{5} & -\frac{58}{25} & \frac{84}{25} \end{bmatrix} * \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{8}{5} \\ \frac{6}{5} \\ -\frac{62}{25} \\ \frac{76}{25} \end{bmatrix}$$

βήμα 8 Βλέπουμε ότι το διάνυσμα $X_2 = X_1$ είναι το ίδιο άρα η τιμή της συνάρτησης θα είναι ή ίδια άρα πηγαίνοντας στο

βήμα 5 Θα είναι $f(X_{i+1}) - f(X_i) = -4.28 - (-4.28) = 0 < \varepsilon = 0.05$ άρα ο αλγόριθμος θα τερματίσει. Βέλτιστη τιμή της συνάρτησης είναι $f_{\hat{X}} = -(2 * (x-2)^2 + 2 * x * y + 3 * (y-1)^2) = -4.28$ με τιμές των μεταβλητών $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \frac{8}{5} \\ \frac{6}{5} \end{bmatrix} = \begin{bmatrix} 1.6 \\ 1.2 \end{bmatrix}$

3.2.2 Με ανισωτικούς περιορισμούς

Όπως είδαμε και στην κατηγορία της βελτιστοποίησης χωρίς περιορισμούς ο τετραγωνικός προγραμματισμός περιλαμβάνει συναρτήσεις της μορφής:

$$X^T * A * X - B * X + C$$

ή όπως είδαμε στην ίδια ενότητα της μορφής:

$$\frac{1}{2} * X^T * H * X - B * X + C$$

όπου H είναι ο εσσιανός πίνακας της συνάρτησης.

Τώρα θα εξετάσουμε την περίπτωση όπου η συνάρτηση στόχος αυτής της μορφής υπόκειται σε γραμμικούς ανισωτικούς ή και εξισωτικούς περιορισμούς. Για να μπορέσουμε να λύσουμε το πρόβλημα θα το θέσουμε στην τυπική του μορφή με την προϋπόθεση ότι η συνάρτηση στόχος είναι κυρτή.

$$\min\left(\frac{1}{2} * x^T * H * x - B * x + C\right)$$

$$G^T * x \geq g$$
$$D^T * x = d$$
$$x \geq 0$$

Για να έχει λύση το συγκεκριμένο πρόβλημα απαραίτητη προϋπόθεση είναι η κυρτότητα της συνάρτησης στόχου που σημαίνει ότι ο Εσσιανός πίνακας της συνάρτησης θα πρέπει να είναι είτε θετικά ορισμένος (οπότε και υπάρχει ολικό ελάχιστο) είτε θετικά ημιορισμένος που σημαίνει ότι μπορεί το ελάχιστο που θα βρεθεί να μην είναι ολικό ή εάν είναι ολικό αυτό να προκύπτει εξαιτίας των περιορισμών.. Οι πίνακες G και D είναι διαστάσεων $m * n_1$ και $m * n_2$ αντίστοιχα όπου n_1 είναι ο αριθμός των ανισωτικών περιορισμών και n_2 ο αριθμός των εξισωτικών περιορισμών. Τέλος τα διανύσματα g και d είναι οι αντίστοιχες δεξιές πλευρές των περιορισμών.
Αφού θέσουμε το πρόβλημα στην τυπική του μορφή, μπορούμε πλέον να προχωρήσουμε στον σχηματισμό της συνάρτησης Lagrange και στην εφαρμογή των συνθηκών Karush-Kuhn-Tucker.

Συνάρτηση Lagrange:
$$L(x, \lambda, \kappa) = \frac{1}{2} * x^T * H * x - B * x + C - \lambda * (G^T * x - g) - \kappa * (D^T * x - d)$$
Και οι συνθήκες KKT είναι:

$$\nabla_x L(x, \lambda, \kappa) = 2 * \tfrac{1}{2} * H * x - B + G * \lambda - D * \kappa = 0$$

$$\nabla_x L(x, \lambda, \kappa) = -G^T * x + g = 0$$

$$\nabla_x L(x, \lambda, \kappa) = -D^T * x + d = 0$$

Στη γενική συνθήκη η λύση του πρωτοβάθμιου συστήματος εξισώσεων που προκύπτει δίνει μια αποδεκτή λύση του προβλήματος ή (εάν οι περιορισμοί συντελούν σε αυτό) δεν υπάρχει εφικτή λύση.

Παράδειγμα: έχουμε την παρακάτω συνάρτηση με περιορισμούς για βελτιστοποίηση

$$\min(2 * x + 5 * y + 3 * x^2 + 2 * y^2 + (z - 1)^2 + z * x)$$
$$\mathrm{g}_1 \coloneqq x - y \geq 2$$
$$\mathrm{g}_2 \coloneqq x + y \geq 6$$
$$\mathrm{g}_3 \coloneqq 2 * z + x + 3 * y \geq 16$$

Διαπιστώνουμε ότι έχει την τυπική μορφή και ότι η συνάρτηση στόχος μπορεί να γραφτεί:

$$2*x+5*y+3*x^2+2*y^2+z^2-2*z+1+z*x=$$
$$=[x \quad y \quad z]*\begin{bmatrix}3 & 0 & \frac{1}{2}\\ 0 & 2 & 0\\ \frac{1}{2} & 0 & 1\end{bmatrix}*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+[2 \quad 5 \quad -2]*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+1$$
$$=\frac{1}{2}*[x \quad y \quad z]*\begin{bmatrix}6 & 0 & 1\\ 0 & 4 & 0\\ 1 & 0 & 2\end{bmatrix}*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+[2 \quad 5 \quad -2]*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+1$$

Η συνάρτηση Lagrange είναι:

$$L(x,y,z,\lambda,\kappa,\mu)=\frac{1}{2}*[x \quad y \quad z]*\begin{bmatrix}6 & 0 & 1\\ 0 & 4 & 0\\ 1 & 0 & 2\end{bmatrix}*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+[2 \quad 5 \quad -2]*\begin{bmatrix}x\\ y\\ z\end{bmatrix}+1-$$
$$-\lambda*(x-y-2)-\kappa*(x+y-6)-\mu*(x+3*y+2*z-16)$$

Οι συνθήκες ΚΚΤ είναι:

$$\frac{\partial}{\partial x}(L)=6*x+z+2-\lambda-\kappa-\mu=0$$
$$\frac{\partial}{\partial y}(L)=4*y+5+\lambda-\kappa-3*\mu=0$$
$$\frac{\partial}{\partial z}(L)=x+2*z-2-2*\mu=0$$
$$\frac{\partial}{\partial \mu}(L)=16-2*z-x-3*y=0$$
$$\frac{\partial}{\partial \lambda}(L)=2-x+y=0$$
$$\frac{\partial}{\partial \kappa}(L)=6-x-y=0$$

Και επιλύοντας το σύστημα των εξισώσεων έχουμε:

$$\frac{\partial}{\partial x}(L)=6*x+z+2-\lambda-\kappa-\mu=0 \qquad (1)$$
$$\frac{\partial}{\partial y}(L)=4*y+5+\lambda-\kappa-3*\mu=0 \qquad (2)$$
$$\frac{\partial}{\partial z}(L)=x+2*z-2-2*\mu=0 \qquad (3)$$
$$\frac{\partial}{\partial \mu}(L)=16-2*z-x-3*y=0 \qquad (4)$$
$$\frac{\partial}{\partial \lambda}(L)=2-x+y=0 \qquad (5)$$
$$\frac{\partial}{\partial \kappa}(L)=6-x-y=0 \qquad (6)$$

Από την (5) $\Longrightarrow x=y+2$ (7)
Από (6) και (7) $\Longrightarrow 6-(y+2)-y=0 \Longrightarrow 4-2*y=0 \Longrightarrow y=2$ (8)

Από (7) και (8) $\Longrightarrow x = 2 + 2 = 4$ (9)
Από (4), (8) και (9) $\Longrightarrow 16 - 2 * z - 4 - 3 * 2 = 0 \Longrightarrow 2 * z = 6 \Longrightarrow z = 3$ (10)
Από (3), (9) και (10) $\Longrightarrow x + 2 * z - 2 - 2 * \mu = 0 \Longrightarrow 4 + 6 - 2 = 2 * \mu \Longrightarrow \mu = 4$ (11)
Αντικαθιστούμε τις τιμές των μεταβλητών που βρήκαμε στις (8), (9), (10), (11) στις (1) και (2)
$6 * x + z + 2 - \lambda - \kappa - \mu = 0 \Longrightarrow 24 + 3 + 2 - \lambda - \kappa - 4 = 0 \Longrightarrow 25 - \lambda - -\kappa = 0$ (12)
$4 * y + 5 + \lambda - \kappa - 3 * \mu = 0 \Longrightarrow 8 + 5 + \lambda - \kappa - 12 = 0 \Longrightarrow \lambda - \kappa - 1 = 0 \Longrightarrow \Longrightarrow \lambda = \kappa + 1$ (13)
Από (12) και (13) $\Longrightarrow 25 - (\kappa + 1) - \kappa = 0 \Longrightarrow 24 - 2 * \kappa = 0 \Longrightarrow \kappa = 12$ (14)
Από (13) και (14) $\Longrightarrow \lambda = 13$
Και αντικαθιστώντας στην συνάρτηση στόχο παίρνουμε
$2 * x + 5 * y + 3 * x^2 + 2 * y^2 + (z - 1)^2 + z * x = 2 * 4 + 5 * 2 + 3 * 16 + +2 * 4 + 4 + 2 * 3 = 8 + 10 + 54 + 8 + 4 + 6 = 90$
άρα η βέλτιστη (ελάχιστη) τιμή της συνάρτησης είναι 90 με τιμές των μεταβλητών

$$\begin{Bmatrix} x \\ y \\ z \end{Bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 3 \end{bmatrix}$$

Παράδειγμα

Έστω ότι έχουμε την παρακάτω συνάρτηση με περιορισμούς για βελτιστοποίηση

$$max(-{x_1}^2 - {x_2}^2 + 2 * x_1 + 3 * x_2)$$

$$x \geq 0$$

Εξετάζω αν η αντικειμενική συνάρτηση είναι κοίλη ($f(x) = -{x_1}^2 - {x_2}^2 + 2 * * x_1 + 3 * x_2$)

$$H = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

Οπότε $-H = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

Έχουμε $D_1 = |2| = 2 > 0$ και $D_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4 > 0$, οπότε ο πίνακας $-H$ είναι θετικά ορισμένος, άρα ο πίνακας H είναι αρνητικά ορισμένος, οπότε η συνάρτηση f είναι κοίλη.
Έχουμε ότι

$$\frac{\partial f}{\partial x_1}(x^*) = 0 <=> -2 * x_1 + 2 = 0 <=> x_1 = 1 > 0$$
$$\frac{\partial f}{\partial x_2}(x^*) = 0 <=> -2 * x_2 + 3 = 0 <=> x_2 = \frac{3}{2} > 0$$

Άρα $x^* = \left(1, \frac{3}{2}\right)$ είναι η λύση του π.μ.γ.π και έχει ολικό μέγιστο $maxf = \frac{13}{4}$

Τέλος, θα παραθέσουμε άλλο ένα παραδειγμα επίλυσης ενός π.μ.γ.π με την τροποποιημένη μέθοδο Simplex.

Παράδειγμα

Να λυθεί το π.μ.γ.π

$$max(4*x_1+6*x_2-2*{x_1}^2-2*x_1*x_2-2*{x_2}^2)$$

με τους περιορισμούς

$$x_1+2*x_2\leq 2$$

$$x_1,x_2>0$$

Έχουμε ότι $D=\begin{bmatrix}-2 & -1\\ -1 & -2\end{bmatrix}$ και $-D=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$ με $D_1=|2|=2>0$ και $D_2=\begin{vmatrix}2 & 1\\ 1 & 2\end{vmatrix}=3>0.$

Παρατηρούμε ότι $-D$ είναι θετικά ορισμένος, οπότε ο D είναι αρνητικά ορισμένος. Οπότε η συνάρτηση f $(f(x)=4*x_1+6*x_2-2*{x_1}^2-2*x_1*x_2-2*{x_2}^2)$ είναι κοίλη, οπότε έχουμε πρόβλημα τετραγωνικού προγραμματισμού δηλαδή τροποποιημένη Simplex με $c=(4{,}6)^t$, $b=2$ και $A=(1.2)$.

Εφαρμόζοντας τις σχέσεις

$$-2*D*x+A^t*\mu-y=c$$
$$A*x+\lambda=b$$
$$x,y,\lambda,\mu\geq 0$$
$$x_j*y_j=0,\mu_i*\lambda_i=0,j=1,\dots,n,i=1,\dots,m$$

στο πρόβλημά μας έχουμε

$$-2*\begin{bmatrix}-2 & -1\\ -1 & -2\end{bmatrix}*\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\begin{bmatrix}1\\ 2\end{bmatrix}*\mu_1-\begin{bmatrix}y_1\\ y_2\end{bmatrix}=\begin{bmatrix}4\\ 6\end{bmatrix}$$

και $$(1\quad 2)*\begin{bmatrix}x_1\\ x_2\end{bmatrix}+\lambda_1=2$$

Οπότε έχουμε

$$4*x_1+2*x_2+\mu_1-y_1=4$$
$$2*x_1+4*x_2+2*\mu_1-y_2=6$$
$$x_1+2*x_2+\lambda_1=2$$
$$x,y,\lambda_1,\mu_1\geq 0$$
$$x_j*y_j=0,\ \mu_i*\lambda_i=0$$

Έχουμε λοιπόν τον πίνακα A ο οποίος είναι

$$A=\begin{pmatrix}P_1 & P_2 & P_3 & P_4 & P_5 & P_6\\ 4 & 2 & 1 & -1 & 0 & 0\\ 2 & 4 & 2 & 0 & -1 & 0\\ 1 & 2 & 0 & 0 & 0 & 1\end{pmatrix}$$

Παρατηρούμε ότι ο A δεν περιέχει τον μοναδιαίο πίνακα, άρα εφαρμόζουμε τη M-μέθοδο.

$$max(M * z_1 + M * z_2), \mathrm{M} << 0$$

$$4 * x_1 + 2 * x_2 + \mu_1 - y_1 + z_1 = 4$$
$$2 * x_1 + 4 * x_2 + 2 * \mu_1 - y_2 + z_2 = 6$$
$$x_1 + 2 * x_2 + \lambda_1 = 2$$
$$x, y, \lambda_1, \mu_1, z \geq 0$$

Οπότε

$$\tilde{A} = \begin{array}{c} \begin{matrix} P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & P_8 \end{matrix} \\ \begin{pmatrix} 4 & 2 & 1 & -1 & 0 & 0 & 1 & 0 \\ 2 & 4 & 2 & 0 & -1 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \end{array}$$

Εφαρμόζουμε τροποποιημένη Simplex όπου κατά την επιλογή της μεταβλητής που θα μπει στη βάση σε κάθε επανάληψη Simplex απαλοίφουμε τις μεταβλητές των οποίων οι συμπληρωματικές είναι βασικές, δηλαδή ποτέ στη βάση x_j, y_j και μ_i, λ_i.
Συμπληρωματικές στήλες είναι οι $(P_1, P_4), (P_2, P_5)$ και (P_3, P_6).

Οπότε ο αρχικός πίνακας tableau είναι ο εξής :

			0	**0**	**0**	**0**	**0**	**0**	**M**	**M**		
B	**C_b**	**b**	**P_1**	**P_2**	**P_3**	**P_4**	**P_5**	**P_6**	**P_7**	**P_8**	**θ**	
P_7	M	4	4	2	1	-1	0	0	1	0	1	$Γ_1$
P_8	M	6	2	4	2	0	-1	0	0	1	3	$Γ_2$
P_6	0	2	1	2	0	0	0	1	0	0	2	$Γ_3$
	z	10M	6M	6M	~~3M~~	-M	-M	0	0	0		$Γ_4$

Διαλέγω το 6M του P_1 τυχαία. Διαγράφω την P_3 ως συμπληρωματική της P_6 που είναι ήδη στη βάση.

Διαπιστώνουμε ότι στο πρώτο tableau φεύγει η P_7 από τη βάση και αντικαθίσταται στο δεύτερο tableau από την P_1.

			0	0	0	0	0	0	M	M		
B	C_b	b	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	θ	
P_1	0	1	*1*	$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	*0*	*0*	$\frac{1}{4}$	*0*	$\frac{4}{2}$	Γ΄₁
P_8	M	4	*0*	*3*	$\frac{3}{2}$	2	*-1*	*0*	$-\frac{1}{2}$	*-1*	$\frac{4}{3}$	Γ΄₂
P_6	0	1	*0*	$\boxed{\frac{3}{2}}$	$-\frac{1}{4}$	$\frac{1}{4}$	*0*	*1*	$-\frac{1}{4}$	*0*	$\frac{2}{3}$	Γ΄₃
	Z	$\frac{4}{9}$	0	3M	$\cancel{\frac{3}{2}M}$	$\cancel{\frac{1}{2}M}$	-M	0	$-\frac{6}{4}M$	0		Γ΄₄

Διαγράφω την P_3 και P_4 ως συμπληρωματικές των P_6 και P_1 αντίστοιχα που είναι ήδη στη βάση.

Διαπιστώνουμε ότι στο δεύτερο tableau φεύγει η P_6 από τη βάση και αντικαθίσταται στο τρίτο tableau από την P_2.

			0	0	0	0	0	0	M	M		
B	C_b	b	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	θ	
P_1	0	$\frac{2}{3}$	*1*	*0*	$\frac{1}{3}$	$-\frac{1}{3}$	*0*	$-\frac{1}{3}$	$\frac{1}{3}$	*0*	*2*	Γ΄΄₁
P_8	M	2	*0*	*0*	$\boxed{2}$	0	*-1*	*-2*	0	*1*	*1*	Γ΄΄₂
P_2	0	$\frac{2}{3}$	*0*	*1*	$-\frac{1}{6}$	$\frac{1}{6}$	*0*	$\frac{2}{3}$	$-\frac{1}{6}$	*0*	-	Γ΄΄₃
	Z	$2M$	*0*	0	$2M$	$\cancel{0}$	$\cancel{-M}$	-2M	$-M$	0		Γ΄΄₄

Διαγράφω την P_4 και P_5 ως συμπληρωματικές των P_1 και P_2 αντίστοιχα που είναι ήδη στη βάση.

Διαπιστώνουμε ότι στο τρίτο tableau φεύγει η P_8 από τη βάση και αντικαθίσταται στο τέταρτο tableau από την P_3.

			0	0	0	0	0	0	M	M		
B	**C_b**	**b**	**P_1**	**P_2**	**P_3**	**P_4**	**P_5**	**P_6**	**P_7**	**P_8**	**θ**	
P_1	0	$\frac{1}{3}$										Γ'''_1
P_3	0	1	*0*	*0*	*1*	0	$-\frac{1}{2}$	*-1*	0	$\frac{1}{2}$		Γ'''_2
P_2	0	$\frac{5}{6}$										Γ'''_3
	z	*0*	*0*	0	*0*	0	0	0	$-M$	$-M$		Γ'''_4

Οπότε η άριστη λύση του ΠΜΓΠ είναι $(x_1, x_2, \mu_1, y_1, y_2, \lambda_1, z_1, z_2) = (\frac{1}{3}, \frac{5}{6}, 1,0,0,0,0,0)$

Με $(x^*{}_1, x^*{}_2) = (\frac{1}{3}, \frac{5}{6})$ και η αντίστοιχη τιμή της αντικειμενικής συνάρτησης είναι $maxf = \frac{150}{36}$.

Βιβλιογραφία

Bronson R.- Naadimuu G. Επιχειρησιακή Έρευνα Επιστημονική Επιμέλεια Ελληνικής έκδοσης: Σαμαράς Ν. Εκδόσεις Κλειδάριθμος Αθήνα 1997 (Τίτλος πρωτοτύπου: Schaum's outline of Theory and Problems of OPERATIONS RESEARCH Second Edition

Φακίνος Δ. Οικονόμου Α. Εισαγωγή στην Επιχειρησιακή Έρευνα Εκδόσεις Συμμετρία Αθήνα 2003

Nocedal J. Wright S. Numerical Optimization Εκδόσεις Springer New York 2006

Taha H. Operation Research an introduction Eighth Edition. Pearson Education New Jersey 2007

Hillier F. Lieberman G. Introduction to Operation Research Tenth Edition Published by McGraw-Hill Education. New York 2015

Winston W. Operation Research Applications and Algorithms Fourth Edition Thomson Books/Cole 2004

Giorgio Giorgi Tinne Hoff Kjeldsen Traces and Emergence of Nonlinear Programming Springer New York 2014