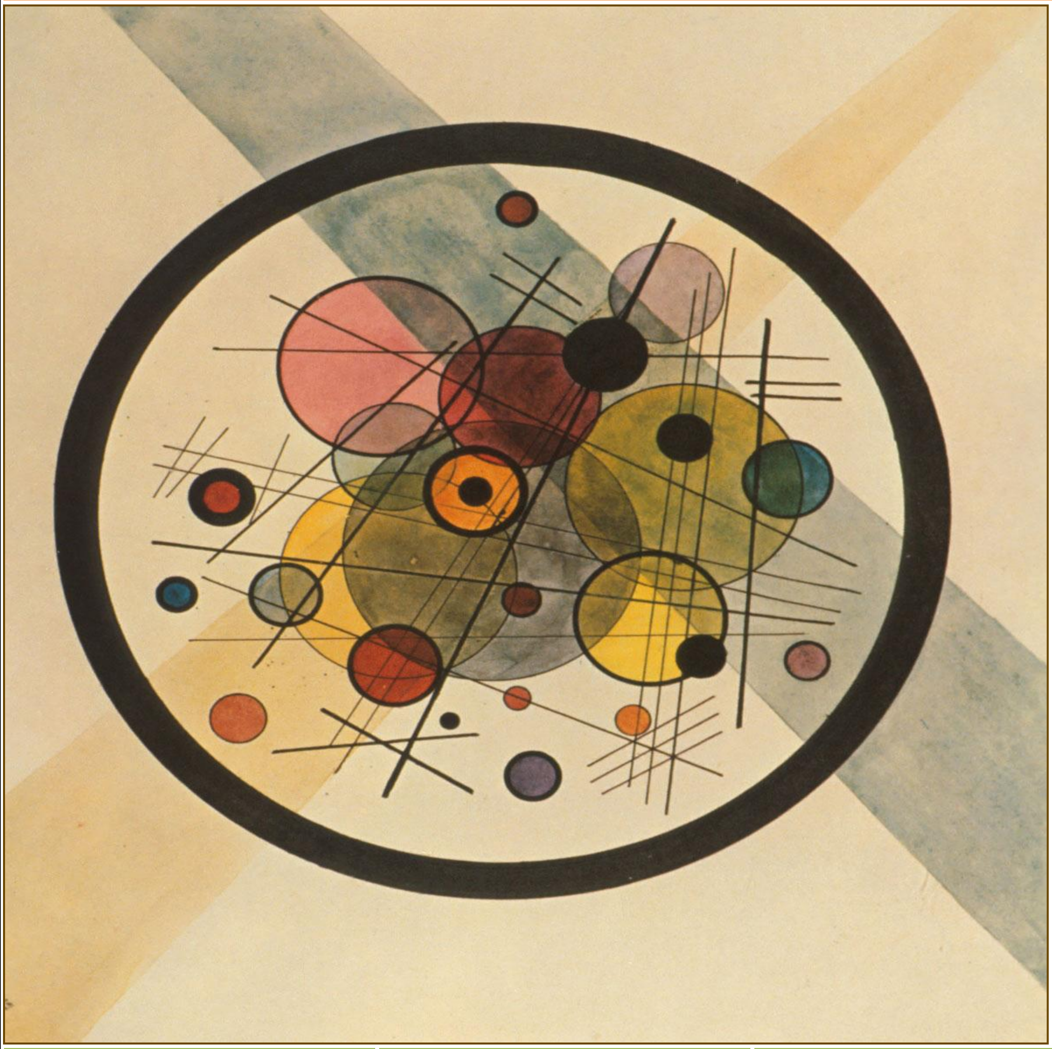
ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
«ΕΦΑΡΜΟΣΜΕΝΕΣ ΜΑΘΗΜΑΤΙΚΕΣ ΕΠΙΣΤΗΜΕΣ»



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΣΤΗΝ ΕΠΙΧΕΙΡΗΣΙΑΚΗ ΕΡΕΥΝΑ

«ΕΙΣΑΓΩΓΗ ΣΤΟ ΜΗ ΓΡΑΜΜΙΚΟ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟ»

(Τετραγωνικός, Κυρτός, Διαχωρίσιμος, Μη Κυρτός, Γεωμετρικός και Κλασματικός Προγραμματισμός)

ΧΑΡΙΤΟΣ ΝΤΑΚΟΛΙΑ

Α.Μ.:09410020

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ

Επιβλέπων Καθηγητής : Επίκουρος Καθηγητής Ιωάννης Κολέτσος

Καθηγητής Ιωάννης Τσινιάς

Καθηγητής Ίων Χρυσοβέργης

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ & ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Στην αδερφή μου, Νάντια…

Πίνακας περιεχομένων

Ιστορικό σημείωμα

Οι ρίζες της Επιχειρησιακής Έρευνας βρίσκονται πολλές δεκαετίες πίσω, όταν έγιναν οι πρώτες προσπάθειες να χρησιμοποιηθεί μια επιστημονική προσέγγιση στη διαχείριση των επιχειρήσεων.

Ωστόσο, η αρχή της δραστηριότητας που ονομάζεται Επιχειρησιακή Έρευνα έχει γενικώς συνδεθεί με τις στρατιωτικές υπηρεσίες στις αρχές του Β΄ Παγκοσμίου Πολέμου.[1]

Συγκεκριμένα, παρατηρούμε ότι στην ιστορία της ανθρωπότητας συναντάμε συχνά συνεργασίες επιστημόνων με στρατιωτικούς αρχηγούς προς μια κοινή κατεύθυνση, προκειμένου να βρεθεί η βέλτιστη απόφαση στο πεδίο της μάχης.

Επομένως, πολλοί ειδικοί θεωρούν ως αρχή της Επιχειρησιακής Έρευνας τον 3ο αιώνα π.Χ., κατά τη διάρκεια του Β′ Καρχηδονικού Πολέμου (218 – 201 π.Χ.), όπου ο Αρχιμήδης (287 – 212 π.Χ.) βάσει προσωπικών αναλύσεων, έδωσε τη λύση για την άμυνα της πόλης των Συρακουσών από τους Ρωμαίους. Μεταξύ των εφευρέσεών του συναντώνται ο καταπέλτης και ένα σύστημα με καθρέφτες, που έθετε φωτιά στα εχθρικά πλοία με τη βοήθεια των ηλιακών ακτινών.

Το 1503 ο Leonardo da Vinci (1452 – 1519) συμμετείχε ως μηχανικός της Cesena στον πόλεμο κατά της Prisa, λόγω των γνώσεών του σε τεχνικές βομβαρδισμού, κατασκευής πλοίων, εξοπλισμένων οχημάτων, κανονιών, καταπελτών και άλλων πολεμικού τύπου μηχανών.

Άλλο ένα ιστορικό παράδειγμα χρήσης της Επιχειρησιακής Έρευνας αποτελεί αυτό του F. W. Lanchester (1868 – 1946), ο οποίος έκανε μια μαθηματική μελέτη πάνω στη βαλλιστική ισχύ των αντιπάλων και ανέπτυξε, από ένα σύστημα

[1] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1.

διαφορικών εξισώσεων, τον τετραγωνικό νόμο Lanchester[2], με τον οποίο μπορεί να προσδιοριστεί η έκβαση μιας πολεμικής μάχης.

Επιπλέον, ο Thomas Edison (1847 – 1931) χρησιμοποίησε την Επιχειρησιακή Έρευνα συνεισφέροντας στον αντιυποβρύχιο πόλεμο με τις ιδέες του, όπως αυτή των ασπίδων κατά των τορπιλών για τα πλοία.

Από μαθηματικής άποψης, στον 17ο και 18ο αιώνα, οι Isaac Newton (1642 – 1727), Gottfried Leibniz (1646 – 1716), Daniel Bernoulli (1700 – 1782) και Joseph-Louis Lagrange (1736 – 1813) εργάστηκαν με σκοπό την εύρεση συνθηκών μεγίστου και ελαχίστου συγκεκριμένων συναρτήσεων, ενώ ο Γάλλος μαθηματικός Jean Baptiste Joseph Fourier (1768 – 1830) σχεδίασε μεθόδους γραμμικού προγραμματισμού. Επίσης, στα τέλη του 18ου αιώνα, ο Gaspard Monge (1746 – 1818) εδραίωσε τις προγενέστερες μελέτες σχετικά με τη γραφική μέθοδο χάρη στην ανάπτυξη της περιγραφικής γεωμετρίας.

Αξίζει, ακόμη, να αναφερθεί ότι ο John von Neumann (1903 – 1957) με το επιστημονικό του έργο πάνω στη θεωρία παιγνίων και στη θεωρία διαφόρων μορφών τελεστών, συνεισέφερε σημαντικά στην ανάπτυξη του γραμμικού προγραμματισμού. Συγκεκριμένα, το 1947 διέκρινε την ομοιότητα ανάμεσα στα προβλήματα γραμμικού προγραμματισμού και τη θεωρία πινάκων που ανέπτυξε ο ίδιος.

Το 1939, ο Ρώσος μαθηματικός Leonid Vitaliyevich Kantorovich (1912 – 1986) σε συνεργασία με τον Ολλανδό μαθηματικό Tjalling Charles Koopmans (1910 – 1985) ανέπτυξε τη μαθηματική θεωρία του γραμμικού προγραμματισμού, η οποία τιμήθηκε με το βραβείο Νόμπελ στις οικονομικές επιστήμες (1975).

Επιπλέον, ο George Joseph Stigler (1911 – 1991) παρουσίασε ένα ιδιαίτερο πρόβλημα, γνωστό ως το *πρόβλημα της δίαιτας*. Το πρόβλημα αυτό δημιουργήθηκε λόγω της ανησυχίας του αμερικάνικου στρατού να διατηρήσει

[2] Lanchester's Square Law: Αν δύο αντίπαλες ομάδες στη μάχη πυροβολούν ο ένας ενάντια του άλλου από απόσταση, τότε μπορούν να επιτεθούν σε πολλαπλούς στόχους αλλά και να δεχθούν πυρά από πολλαπλές κατευθύνσεις. Ο τετραγωνικός νόμος του Lanchester δηλώνει ότι ο ρυθμός φθοράς σε μια τέτοια κατάσταση εξαρτάται μόνο από τον αριθμό των όπλων που πυροβολούν. Συγκεκριμένα, η ισχύς μιας τέτοιας δύναμης είναι αναλογική με τον αριθμό τον μονάδων στο τετράγωνο. Βλέπε αναλυτικά στο Trevor N. Dupuy, *Numbers, Predictions and War*, Indianapolis, 1979.

σε χαμηλό κόστος τις διατροφικές ανάγκες των στρατευμάτων. Αν και το πρόβλημα αυτό είχε επιλυθεί διαισθητικά, η λύση του διέφερε μόλις λίγα εκατοστά από τη βέλτιστη λύση που δόθηκε αργότερα με βάση τη μέθοδο simplex (1947).[3]

Το 1941 και 1942, οι L. V. Kantorovich και T. C. Koopmans μελέτησαν ανεξάρτητα ο ένας από τον άλλον το *πρόβλημα μεταφοράς*[4] για πρώτη φορά. Για το λόγο αυτό, προβλήματα τέτοιας μορφής είναι γνωστά και ως *Koopmans – Kantorovich προβλήματα*. Για τη λύση του, χρησιμοποίησαν γεωμετρικές μεθόδους οι οποίες σχετίζονται με τη θεωρία κυρτότητας του Minkowski.[5]

Ωστόσο, οι επιστημονικές ανακαλύψεις και η προσφορά των παραπάνω μαθηματικών θεωριών στις επιχειρήσεις δεν ήταν αρκετές για τη δημιουργία μιας νέας επιστήμης ονόματι Επιστημονική Έρευνα, μέχρι την έναρξη του Β΄ Παγκοσμίου Πολέμου (1939 – 1945). Κατά τη διάρκεια του εναέριου πολέμου στην Αγγλία, η βρετανική κυβέρνηση αναζητούσε μια μέθοδο για να υπερασπιστεί τη χώρα της, επιστρατεύοντας πολλούς επιστήμονες διαφορετικών ειδικοτήτων με σκοπό την εύρεση μιας λύσης για τη βέλτιστη χρήση των διαθέσιμων ραντάρ. Χάρη στην κατάλληλη τοποθέτηση των κεραιών, κατάφεραν να αποκτήσουν τη βέλτιστη διανομή των σημάτων και, κατά συνέπεια, το διπλασιασμό της αποτελεσματικότητας του εναέριου αμυντικού συστήματος.

Το εγχείρημα της Αγγλίας ακολούθησαν και οι Ηνωμένες Πολιτείες όταν συμμετείχαν στον Πόλεμο το 1942, όπου δημιούργησαν το πρόγραμμα SCOOP[6] στο οποίο εργαζόταν ο George Dantzig[7], δημιουργός του αλγορίθμου simplex[8] το 1947.

---

[3] Η Μέθοδος Simplex αποτελεί μια γενική διαδικασία για την επίλυση προβλημάτων γραμμικού προγραμματισμού. Αναπτύχθηκε από τον George Dantzig (1914 – 2005) το 1947. Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.109.
[4] Το πρόβλημα μεταφοράς (Transportation Problem) έγινε γνωστό από το Γάλλο μαθηματικό G. Monge το 1781. Βλέπε αναλυτικά στο G. Monge, *Mémoire sur la théorie des déblais et des remblais, Histoire de l'Académie Royale des Sciences de Paris* 1781.
[5] Hermann Minkowski (1864 – 1909).
[6] SCOOP: Scientific Computation Of Optimum Procurement.
[7] Το 1942 ο G. Dantzig εγκατέλειψε το διδακτορικό του πρόγραμμα στο Πανεπιστήμιο της Καλιφόρνιας, Berkeley για να ενταχθεί στο U.S. Air Force Office of Statistical Control.
[8] Αλγόριθμος Simplex ή Μέθοδος Simplex.

Κατά τη διάρκεια του Ψυχρού Πολέμου (1946 – 1991), η Σοβιετική Ένωση (URRS), η οποία απέρριψε το σχέδιο Marshall[9] το 1947, ήθελε να ελέγξει τις εδαφικές επικοινωνίες, συμπεριλαμβανομένων των ποτάμιων διαδρομών από το Βερολίνο. Με σκοπό την αποφυγή της κατάκτησης της πόλης και την καθολική υποταγή της στη γερμανική κομμουνιστική ζώνη, η Αγγλία και οι Ηνωμένες Πολιτείες αποφάσισαν να εφοδιάσουν την πόλη με τα απαραίτητα για την προστασία της. Επομένως, για την εύρεση της βέλτιστης ποσότητας εφοδιασμού χρησιμοποίησαν μεθόδους Επιχειρησιακής Έρευνας και την ομάδα έρευνας SCOOP. Αυτή τους η επιλογή αποθάρρυνε τελικώς τα σχέδια της Σοβιετικής Ένωσης για την κατάκτηση ολόκληρης της πόλης.

Μετά το τέλος του Β΄ Παγκοσμίου Πολέμου, οι Αρχές των Πόρων των Ηνωμένων Πολιτειών χρησιμοποίησαν μοντέλα βελτιστοποίησης και την επίλυσή τους με γραμμικό προγραμματισμό, για τη σωστή κατανομή των πόρων της.

Παράλληλα με την Επιχειρησιακή Έρευνα, οι υπολογιστικές τεχνικές, καθώς και οι ηλεκτρονικοί υπολογιστές αναπτύχθηκαν και, κατ’ επέκταση, ο χρόνος επίλυσης μαθηματικών προβλημάτων μειώθηκε σημαντικά.

Κατά τη διάρκεια των δεκαετιών του ‘50 και ‘60, οι εφαρμογές της Επιχειρησιακής Έρευνας στο εμπόριο και στις βιομηχανίες ενέτειναν το ενδιαφέρον για την Επιχειρησιακή Έρευνα, καθώς και την ανάπτυξή της ως επιστήμη.

Στην εποχή μας, η Επιχειρησιακή Έρευνα χρησιμοποιείται όχι μόνο σε στρατιωτικές επιχειρήσεις αλλά και σε άλλους κλάδους, όπως αυτός της γεωργίας, της βιομηχανίας, καθώς και σε προβλήματα μεταφοράς, δίαιτας, ουρών προτεραιότητας κ.ά.

Στη συνέχεια, ο *πίνακας 1*[10] φανερώνει πραγματικές περιπτώσεις επιχειρήσεων, οι οποίες χρησιμοποίησαν την Επιχειρησιακή Έρευνα με σκοπό τη μεγιστοποίηση του κέρδους τους.[11]

---

[9] Επίσημη ονομασία: European Recovery Program, ERP.

[10] Οι πληροφορίες αντλήθηκαν από την ιστοσελίδα www.phpsimplex.com (25/07/2011) και Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.4.

Επιχείρηση	Εφαρμογή	Έτος	Ετήσια Αποταμίευση
The Netherlands Rijkswaterstaat	Ανάπτυξη της εθνικής πολιτικής της διαχείρισης των υδάτων, περιλαμβανομένων των νέων εγκαταστάσεων, των επιχειρησιακών διαδικασιών και τη χρηματοδότηση.	1985	\$15 εκατομ.
Monsanto Corp.	Βελτιστοποίηση της διαδικασίας παραγωγής με σκοπό την ελαχιστοποίηση του κόστους.	1985	\$2 εκατομ.
Weyerhauser Co.	Βελτιστοποίηση της κοπής των δέντρων στα ξύλινα προϊόντα με σκοπό τη μεγιστοποίηση της παραγωγής.	1986	\$15 εκατομ.
Electrobas/CEPAL Brasil	Βέλτιστη κατανομή των υδραυλικών και θερμικών πόρων στο εθνικό σύστημα παραγωγής ενέργειας.	1986	\$43 εκατομ.
United Airlines	Προγραμματισμός της εργασίας στα γραφεία κρατήσεων και στα αεροδρόμια με σκοπό την κάλυψη των αναγκών του πελάτη με το ελάχιστο κόστος.	1986	\$6 εκατομ.
Citgo Petroleum Corp.	Βελτιστοποίηση της συσκευασίας, της προσφοράς, της διανομής και της εμπορίας των προϊόντων επιχειρήσεων.	1987	\$70 εκατομ.
SANTOS, Ltd., Australia	Βελτιστοποίηση των κεφαλαιακών επενδύσεων για την παραγωγή φυσικού	1987	\$3 εκατομ.

[11] Οι πληροφορίες του ιστορικού σημειώματος αντλήθηκαν από την ιστοσελίδα www.phpsimplex.com/en/history.htm (25/07/2011).

	αερίου κατά τη διάρκεια 25 χρόνων.		
Electric Power Research Institute	Η χορήγηση των αποθεμάτων του πετρελαίου και του άνθρακα για την ηλεκτρική υπηρεσία με στόχο την εξισορρόπηση των εξόδων αποθεμάτων και του ρίσκου υπολειμμάτων.	1989	$59 εκατομ.
San Francisco Police Department	Βελτιστοποίηση του προγραμματισμού και της εκχώρησης των υπαλλήλων της Patrol με ένα υπολογιστικό σύστημα.	1989	$11 εκατομ.
Texaco Inc.	Βελτιστοποίηση του διαθέσιμου μείγματος συστατικών προκειμένου τα προϊόντα βενζίνης να πληρούν τα αιτήματα των πωλήσεων και της ποιότητας.	1989	$30 εκατομ.
IBM	Ενσωμάτωση ενός εθνικού δικτύου αποθήκευσης εξαρτημάτων για τη βελτίωση της υποστήριξης των υπηρεσιών.	1990	$20 εκατομ. και $250 εκατομ. Σε μικρότερο απόθεμα.
U.S. Military Airlift Command	Ταχύτητα στο συντονισμό αεροπλάνων, πληρώματος, φορτίου και επιβατών για την εκκένωση από αέρα στο πρόγραμμα “Desert Storm” στη Μέση Ανατολή.	1992	Νίκη
American Airlines	Σχεδιασμός της τιμολόγησης, υπερβολικής ζήτησης και συντονισμού του συστήματος πτήσεων για τη βέλτιστη χρήση του.	1992	$500 εκατομ. από πρόσθετα έσοδα.
Yellow Freight System, Inc.	Βελτιστοποίηση του σχεδιασμού ενός εθνικού δικτύου μεταφοράς και του προγραμματισμού	1992	$17.3 εκατομ

	των δρομολογίων μεταφοράς.		
New Haven Health Dept.	Σχεδιασμός ενός αποτελεσματικού προγράμματος της αλλαγής των βελονών για την καταπολέμηση της εξάπλωσης του AIDS.	1993	33% μείωση μετάδοσης
AT&T	Ανάπτυξη ενός συστήματος που βασίζεται σ' έναν ηλεκτρονικό υπολογιστή για να οδηγεί τους πελάτες της επιχείρησης στο σχεδιασμό του κέντρου κλήσεων.	1993	$750 εκατομ
Delta Airlines	Η μεγιστοποίηση των κερδών από την εκχώρηση των τύπων αεροσκαφών σε 2.500 εθνικές πτήσεις.	1994	$100 εκατομ
Digital Equipment Corp.	Αναδιοργάνωση της αλυσίδας όλων των προμηθευτών μεταξύ των προμηθευτών, των εγκαταστάσεων, των κέντρων διανομής, πιθανές θέσεις και τομείς της αγοράς.	1995	$800 εκατομ
China	Επιλογή και βέλτιστο προγραμματισμό μαζικών έργων για να υπακούν με τις μελλοντικές ενεργειακές ανάγκες της χώρας.	1995	$425 εκατομ
Cuerpo de defensa de Sudáfrica	Βέλτιστο επανασχεδιασμό του μεγέθους και της μορφής του Εθνικού Σώματος Άμυνας και του οπλικού του συστήματος.	1997	$1.100 εκατομ
Procter and Gamble	Επανασχεδιασμός της παραγωγής της Βόρειας Αμερικής και του συστήματος διανομής για να μειώσει το κόστος και να βελτιώσει την ταχύτητα εισχώρηση στην αγορά.	1997	$200 εκατομ

Taco Bell	Βέλτιστο προγραμματισμό των εργαζόμενων για να παρέχουν την κατάλληλη υπηρεσία στους επιθυμητούς πελάτες με ελάχιστο κόστος.	1998	$13 εκατομ
Hewlett-Packard	Επανασχεδιασμός της ασφάλειας του μεγέθους και της τοποθεσίας των αποθεμάτων στη γραμμή παραγωγής του εκτυπωτή για να υπακούσει στους στόχους παραγωγής.	1998	$280 εκατομ από πρόσθετα έσοδα.

Πίνακας 1 – πραγματικές περιπτώσεις επιχειρήσεων, οι οποίες χρησιμοποίησαν την Επιχειρησιακή Έρευνα.

Εισαγωγικές Έννοιες

Αρχικά, πριν προβούμε στην παρουσίαση της θεωρίας του μη – γραμμικού προγραμματισμού, θεωρήθηκε αναγκαία η εισαγωγή του αναγνώστη σε βασικές μαθηματικές έννοιες, οι οποίες θα χρησιμοποιηθούν σε επόμενα κεφάλαια.

Τέτοια έννοια αποτελεί και η *κυρτότητα,* η οποία συναντάται συχνά στη θεωρία του μη – γραμμικού προγραμματισμού.

Κυρτότητα

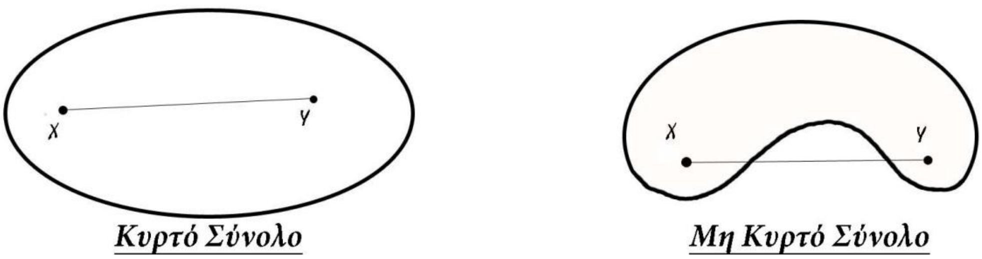
Στην ενότητα αυτή θα δοθεί ο ορισμός ενός κυρτού συνόλου, μιας κυρτής συνάρτησης, καθώς και θεωρήματα που σχετίζονται άμεσα με την έννοια της κυρτότητας.

ΟΡΙΣΜΟΣ 2.1[12]

Έστω S ένα υποσύνολο ενός πραγματικού γραμμικού χώρου ($S \subseteq \Re^n$). Το σύνολο S καλείται κυρτό αν

$$\boldsymbol{\lambda x} + (\mathbf{1} - \boldsymbol{\lambda})\boldsymbol{y} \in \boldsymbol{S}\,, \qquad \forall x, y \in S, \forall \lambda \in [0,1].$$

Παρακάτω δίνεται ένα παράδειγμα ενός κυρτού και ενός μη κυρτού συνόλου.



Γράφημα 1 - Παραδείγματα κυρτότητας.

ΟΡΙΣΜΟΣ 2.2[13]

Έστω S ένα μη κενό και κυρτό υποσύνολο ενός πραγματικού γραμμικού χώρου ($S \subseteq \Re^n$). Η συνάρτηση $f\,:S \to \mathbb{R}$ καλείται κυρτή αν

[12] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin ³2007, σελ.11.
[13] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin ³2007, σελ.11.

$$\boldsymbol{f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)}, \qquad \forall x, y \in S, \forall \lambda \in [0,1].$$

Ομοίως, η συνάρτηση f λέγεται αυστηρά κυρτή αν

$$\boldsymbol{f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)}, \qquad \forall x, y \in S, \forall \lambda \in [0,1].$$ [14]

ΟΡΙΣΜΟΣ 2.3[15]

Έστω S ένα μη κενό και κυρτό υποσύνολο ενός πραγματικού γραμμικού χώρου ($S \subseteq \Re^n$). Η συνάρτηση $f : S \to \mathbb{R}$ καλείται κοίλη αν η συνάρτηση $-f$ είναι κυρτή, ή αλλιώς με βάση τον ορισμό 2.2, αν

$$\boldsymbol{f(\lambda x + (1-\lambda)y) \geq \lambda f(x) + (1-\lambda)f(y)}, \qquad \forall x, y \in S, \forall \lambda \in [0,1].$$

Ομοίως, η συνάρτηση f λέγεται αυστηρά κοίλη αν

$$\boldsymbol{f(\lambda x + (1-\lambda)y) > \lambda f(x) + (1-\lambda)f(y)}, \qquad \forall x, y \in S, \forall \lambda \in [0,1].$$ [16]

Ερμηνεύοντας γεωμετρικά την έννοια της κυρτότητας, παρατηρούμε ότι αν η συνάρτηση $f(x)$ έχει δεύτερη παράγωγο παντού, τότε η $f(x)$ είναι κυρτή αν και μόνο αν $\frac{d^2 f(x)}{d^2 x} \geq 0$ για όλες τις πιθανές τιμές της μεταβλητής x.[17]

<u>**Έλεγχος κυρτότητας για συναρτήσεις μιας μεταβλητής**</u>[18]

Θεωρούμε μια συνάρτηση μιας μεταβλητής $f(x)$, η οποία έχει δεύτερη παράγωγο για όλες τις τιμές της μεταβλητής x. Τότε, η $f(x)$ είναι

- ***Κυρτή*** αν και μόνο αν $\frac{d^2 f(x)}{d^2 x} \geq 0$, για όλες τις πιθανές τιμές της μεταβλητής x.
- ***Αυστηρά κυρτή*** αν και μόνο αν $\frac{d^2 f(x)}{d^2 x} > 0$, για όλες τις πιθανές τιμές της μεταβλητής x.

---

[14] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1159.

[15] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.11.

[16] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1159.

[17] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1159.

[18] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1159.

- ***Κοίλη*** αν και μόνο αν $\frac{d^2 f(x)}{d^2 x} \leq 0$, για όλες τις πιθανές τιμές της μεταβλητής x.
- ***Αυστηρά κοίλη*** αν και μόνο αν $\frac{d^2 f(x)}{d^2 x} < 0$, για όλες τις πιθανές τιμές της μεταβλητής x.

Έλεγχος κυρτότητας για συναρτήσεις πολλών μεταβλητών[19]

Έστω S ένα μη κενό και κυρτό υποσύνολο ενός πραγματικού γραμμικού χώρου ($S \subseteq \Re^n$) και έστω η συνάρτηση $f : S \rightarrow \mathbb{R}$ δύο φορές συνεχώς παραγωγίσιμη, τότε η f είναι

- ***Κυρτή*** αν και μόνο αν ο $n \times n$ Hessian πίνακας[20] είναι θετικά ημιορισμένος[21].
- ***Αυστηρά κυρτή*** αν και μόνο αν ο $n \times n$ Hessian πίνακας είναι θετικά ορισμένος.
- ***Κοίλη*** αν και μόνο αν ο $n \times n$ Hessian πίνακας είναι αρνητικά ημιορισμένος.
- ***Αυστηρά κοίλη*** αν και μόνο αν ο $n \times n$ Hessian πίνακας είναι αρνητικά ορισμένος.

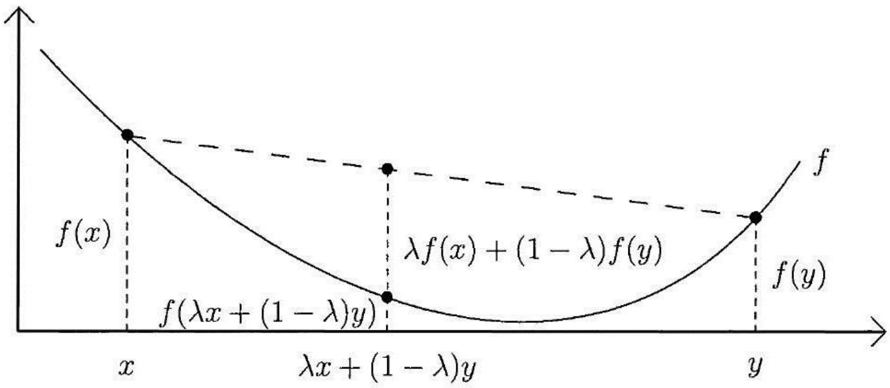
Παρακάτω δίνονται παραδείγματα κυρτότητας διαφόρων συναρτήσεων.

---

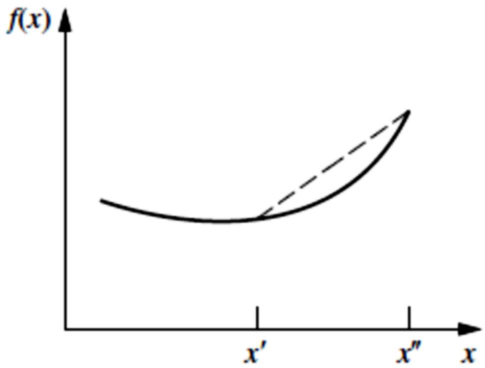
[19] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1162.

[20] Εσσιανός πίνακας (Hessian) ονομάζεται ο πίνακας που φτιάχνεται από τις δεύτερες μερικές παραγώγους της f. Ο πίνακας αυτός μπορεί να θεωρηθεί σαν τη δεύτερη παράγωγο της f και ορίζεται ως ο $n \times n$ συμμετρικός πίνακας με στοιχεία: $H_{i,j}(x) = [\nabla^2 f(x)]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x), 1 \leq i, j \leq n$. John H. Hubbard & Barbara Burke Hubbard, *Διανυσματικός Λογισμός, Γραμμική Άλγεβρα και Διαφορικές Μορφές – Μια ενοποιημένη Προσέγγιση*, μετάφραση των Β. Μεταφτσή και Α. Τσολομύτη, Πάτρα 2006, σελ.421 και Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods – Nonlinear Programming*, NY 2006, σελ.23.

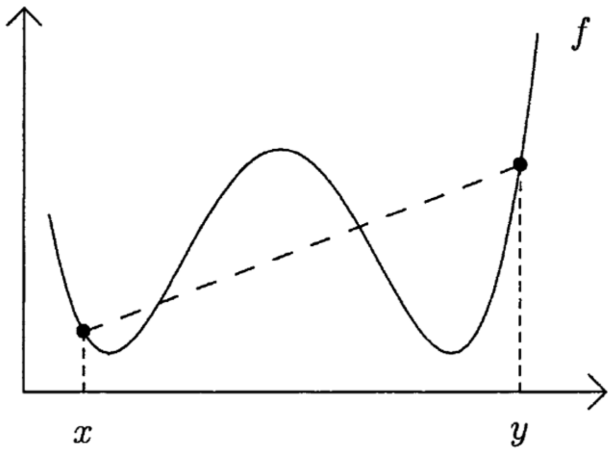
[21] Μια τετραγωνική μορφή Q είναι θετικά (ημι)ορισμένη αν και μόνο αν $Q(x)(\geq) > 0$, όταν $x \neq 0$. Μια τετραγωνική μορφή Q είναι αρνητικά (ημι)ορισμένη αν και μόνο αν $Q(x)(\leq) < 0$, όταν $x \neq 0$. John H. Hubbard & Barbara Burke Hubbard, *Διανυσματικός Λογισμός, Γραμμική Άλγεβρα και Διαφορικές Μορφές – Μια ενοποιημένη Προσέγγιση*, μετάφραση των Β. Μεταφτσή και Α. Τσολομύτη, Πάτρα 2006, σελ.410.



Γράφημα 2 - Παράδειγμα κυρτής συνάρτησης.[22]

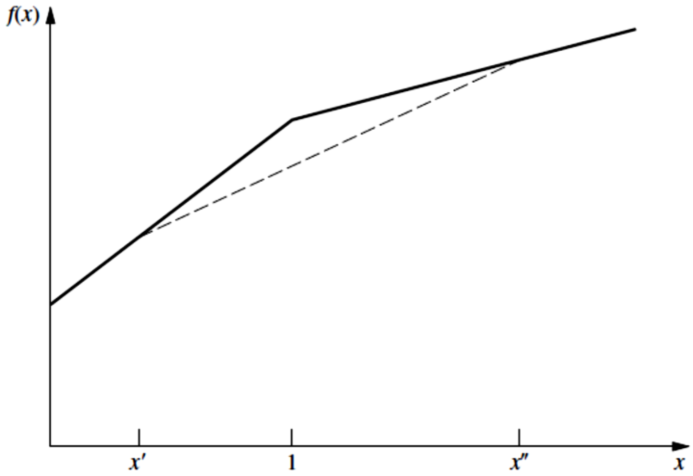


Γράφημα 3 - Παράδειγμα αυστηρά κυρτής συνάρτησης.[23]

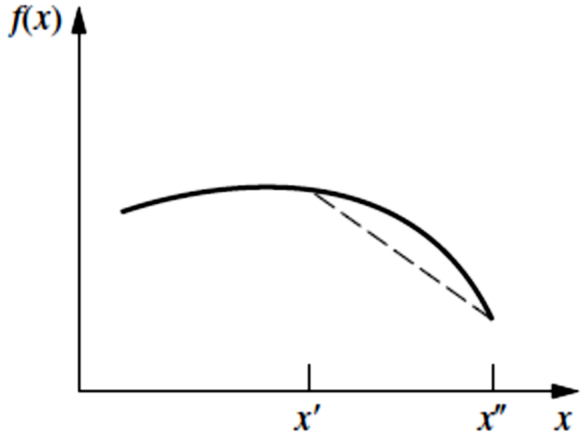


Γράφημα 4 - Παράδειγμα μη κυρτής συνάρτησης.[24]

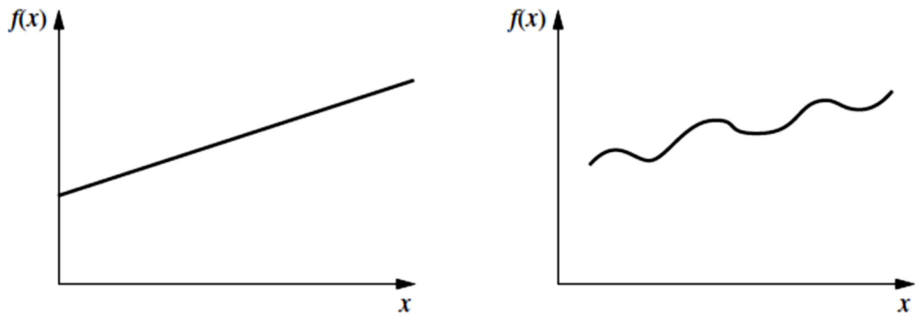
[22] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.12.
[23] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1160.
[24] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.13.



Γράφημα 5 - Παράδειγμα κοίλης συνάρτησης.[25]



Γράφημα 6 - Παράδειγμα αυστηρά κοίλης συνάρτησης.[26]



Γράφημα 7 - Παράδειγμα συνάρτησης που είναι κυρτή και κοίλη, συνάρτησης που δεν είναι ούτε κυρτή ούτε κοίλη.[27]

Ωστόσο, η κυρτότητα ενός συναρτησιακού μπορεί να χαρακτηριστεί με τη βοήθεια του επιγραφήματος[28].

[25] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1161.
[26] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1161.
[27] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1161.

Στο σημείο αυτό, είναι αναγκαία η διατύπωση της παρακάτω υπόθεσης για την παρουσίαση της ακόλουθης θεωρίας.

Υποθεση 2.1[29]

Έστω ότι $(X, \|\cdot\|)$ είναι ένας πραγματικός χώρος εφοδιασμένος με μια νόρμα. Έστω, επίσης, ότι S είναι ένα μη κενό υποσύνολο του χώρου ($\emptyset \neq S \subseteq X$) και έστω $f : S \to \mathbb{R}$ μια δοθείσα συνάρτηση.

Ορισμος 2.4[30]

Έστω ότι ικανοποιείται η υπόθεση 2.1 και έστω ότι το σύνολο S είναι κυρτό. Τότε τα παρακάτω είναι ισοδύναμα:

Η συνάρτηση f είναι κυρτή

$\Leftrightarrow$ *epif* ή $E(f)$ *είναι κυρτό*

$\Leftrightarrow \forall a \in \mathbb{R}, \text{το σύνολο } S_a \coloneqq \{x \in S \mid f(x) \leq a\} \text{ είναι κυρτό}.$

Αποδειξη

Αρχικά θα εργαστούμε με σκοπό να αποδείξουμε την πρώτη ισοδυναμία του θεωρήματος, δηλαδή, αν η συνάρτηση f είναι κυρτή, τότε το $E(f)$ είναι κυρτό.

Αν η συνάρτηση f είναι κυρτή, τότε για τα τυχαία σημεία $(x, a), (y, b) \in E(f)$ και για ένα τυχαίο $k \in [0, 1]$ ισχύει

$$f(kx + (1-k)y) \leq kf(x) + (1-k)f(y) \Leftrightarrow$$

$$\Leftrightarrow f(kx + (1-k)y) \leq ka + (1-k)b$$

Καταλήγοντας ότι

$$k(x, a) + (1-k)(y, b) \in E(f)$$

Επομένως, το επιγράφημα της f είναι κυρτό.

---

[28] Επιγράφημα μιας συνάρτησης $f: S \subset \mathbb{R}^n \to \mathbb{R}$, λέγεται το σύνολο των σημείων που κείτονται πάνω στο γράφημα ή πάνω από το γράφημα της f και συμβολίζεται με $E(f)$ ή $epif = \{(x, a) | f(x) \leq a, x \in S, a \in \mathbb{R}\}$. Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods – Nonlinear Programming*, NY 2006, σελ.38.

[29] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.7.

[30] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.12.

Στη συνέχεια, θα αποδείξουμε τη δεύτερη ισοδυναμία του θεωρήματος, δηλαδή, αν $E(f)$ είναι κυρτό, τότε $\forall a \in \mathbb{R}$, το σύνολο $S_a \coloneqq \{x \in S \mid f(x) \leq a\}$ είναι κυρτό.

Θεωρούμε, λοιπόν, ότι το $E(f)$ είναι κυρτό και επιλέγουμε ένα $a \in \mathbb{R}$ για το οποίο το σύνολο S_a είναι μη κενό (η περίπτωση $S_a = \emptyset$ είναι τετριμμένη). Για τυχαία $x, y \ \in S_a$ έχουμε ότι $(x, a) \in E(f)$ και $(y, a) \in E(f)$, άρα για τυχαίο $k \in [0, 1]$ παίρνουμε

$$k(x, a) + (1 - k)(y, a) \in E(f).$$

Αυτό σημαίνει ότι

$$(kx + (1 - k)y) \ \leq \ kf(x) + (1 - k)f(y) \Leftrightarrow$$

$$(kx + (1 - k)y) \ \leq \ \kappa\alpha + (1 - \kappa)\alpha = \alpha$$

και

$$kx + (1 - k)y \in S_a.$$

Επομένως, το σύνολο S_a είναι κυρτό.

Τέλος, απομένει να αποδείξουμε την τρίτη ισοδυναμία του θεωρήματος, δηλαδή, αν $\forall a \in \mathbb{R}$, το σύνολο $S_a \coloneqq \{x \in S \mid f(x) \leq a\}$ είναι κυρτό, τότε η συνάρτηση f είναι κυρτή.

Για τυχαία $x, y \ \in S$ και ένα τυχαίο $k \in [0, 1]$ παίρνουμε

$$k\big(x, f(x)\big) + (1 - k)(y, f(y)) \in E(f)$$

το οποίο υποδηλώνει ότι

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

Επομένως, η συνάρτηση f είναι κυρτή.

---*τέλος απόδειξης*---

Στη συνέχεια παρατίθεται ο ορισμός της έννοιας *οιονεί – κυρτότητα*.

Ορισμος 2.5[31]

Έστω ότι ικανοποιείται η υπόθεση 2.1 και έστω ότι το σύνολο S είναι κυρτό. Αν $\forall a \in \mathbb{R}$, το σύνολο $S_a \coloneqq \{x \in S \mid f(x) \leq a\}$ είναι κυρτό, τότε η συνάρτηση f είναι οιονεί – κυρτή.

Κλασικές Μέθοδοι Βελτιστοποίησης

Στην ενότητα αυτή θα επικεντρωθούμε στην παρουσίαση κλασικών μεθόδων βελτιστοποίησης του λογισμού για την εύρεση μιας λύσης, η οποία μεγιστοποιεί ή ελαχιστοποιεί μια συνάρτηση μιας μεταβλητής, πολλών μεταβλητών και πολλών μεταβλητών με περιορισμούς στις τιμές των μεταβλητών. Επιπλέον, στην παρούσα ενότητα, θεωρούμε ότι όλες οι συναρτήσεις στις οποίες θα αναφερθούμε, έχουν συνεχείς πρώτες και δεύτερες παραγώγους, καθώς και μερικές παραγώγους παντού.

Αρχικά, παρατίθενται οι ορισμοί του τοπικού και ολικού ελαχίστου και μεγίστου, του στάσιμου σημείου, καθώς και της καθοδικής κατεύθυνσης μιας συνάρτησης.

Ορισμος 2.6[32]

Ένα σημείο $\vec{x}^*$ ονομάζεται τοπικό ελάχιστο αν υπάρχει $\delta > 0$ τέτοιο ώστε $f(\vec{x}^*) \leq f(\vec{x})$ για όλα τα $\vec{x} \in \mathbb{R}^n$, ικανοποιώντας τη σχέση $\|\vec{x} - \vec{x}^*\| < \delta$. Ομοίως, ένα σημείο $\vec{x}^*$ ονομάζεται αυστηρά τοπικό ελάχιστο αν υπάρχει $\delta > 0$ τέτοιο ώστε $f(\vec{x}^*) < f(\vec{x})$ για όλα τα $\vec{x} \in \mathbb{R}^n$ με $\vec{x} \neq \vec{x}^*$, ικανοποιώντας τη σχέση $\|\vec{x} - \vec{x}^*\| < \delta$.

Ορισμος 2.7[33]

Ένα σημείο $\vec{x}^*$ ονομάζεται τοπικό μέγιστο αν υπάρχει $\delta > 0$ τέτοιο ώστε $f(\vec{x}^*) \geq f(\vec{x})$ για όλα τα $\vec{x} \in \mathbb{R}^n$, ικανοποιώντας τη σχέση $\|\vec{x} - \vec{x}^*\| < \delta$. Ομοίως, ένα σημείο $\vec{x}^*$ ονομάζεται αυστηρά τοπικό μέγιστο αν υπάρχει $\delta > 0$ τέτοιο

[31] Johannes Jahn, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007, σελ.14.
[32] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 58.
[33] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 58.

ώστε $f(\vec{x}^*) > f(\vec{x})$ για όλα τα $\vec{x} \in \mathbb{R}^n$ με $\vec{x} \neq \vec{x}^*$, ικανοποιώντας τη σχέση $\|\vec{x} - \vec{x}^*\| < \delta$.

Ορισμος 2.8[34]

Ένα σημείο $\vec{x}^* \in \mathbb{R}^n$ ονομάζεται στάσιμο ή κρίσιμο σημείο για μια διαφορίσιμη συνάρτηση f αν $\boldsymbol{\nabla f(\vec{x}^*) = 0}$.

Επίσης, αξίζει να σημειωθεί ότι αν ένα στάσιμο σημείο $\vec{x}^*$ δεν είναι ούτε ελάχιστο ούτε μέγιστο, τότε το σημείο αυτό ονομάζεται σαγματικό σημείο.

Ορισμος 2.9[35]

Έστω μια συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$, η οποία είναι διαφορίσιμη στο $\vec{x} \in \mathbb{R}^n$. Αν στο σημείο αυτό υπάρχει ένα διάνυσμα $d \in \mathbb{R}^n$ τέτοιο ώστε

[36] $$\boldsymbol{\langle \nabla f(\vec{x}), \vec{d} \rangle < 0},$$

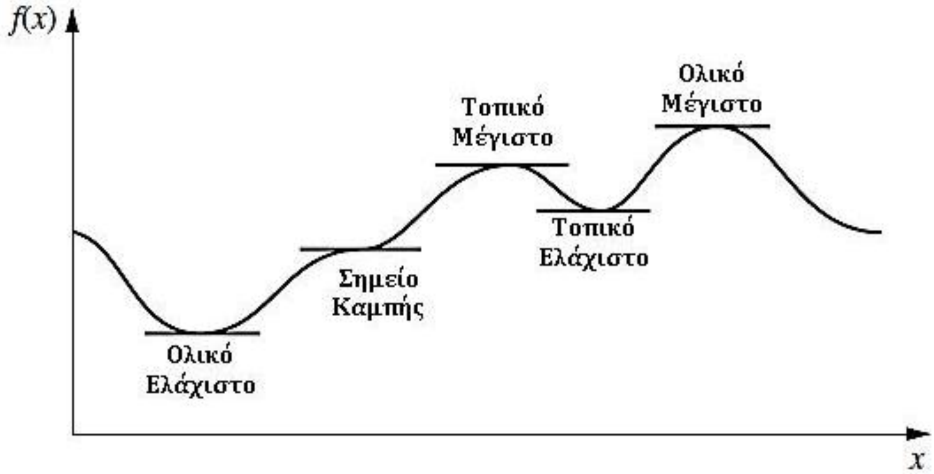
τότε το διάνυσμα $\vec{d}$ καλείται καθοδική κατεύθυνση της f στο $\vec{x}$.

Αφού παρουσιάστηκαν οι ορισμοί ελαχίστου και μεγίστου, στο σημείο αυτό θα ασχοληθούμε με τις συνθήκες βελτιστοποίησης χωρίς περιορισμούς για συναρτήσεις μιας μεταβλητής.

Βελτιστοποίηση χωρίς περιορισμούς για συναρτήσεις μιας μεταβλητής[37]

Θεωρούμε μια συνάρτηση μιας μεταβλητής, τέτοια όπως απεικονίζεται στο παρακάτω γράφημα.

[34] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 62.
[35] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 58.
[36] $\langle *,* \rangle$: συμβολίζει το εσωτερικό γινόμενο
[37] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1165.



Γράφημα 8 - Συνάρτηση μιας μεταβλητής.[38]

Μια αναγκαία συνθήκη για μια μερική λύση $x = x^*$ να είναι είτε ελάχιστο είτε μέγιστο είναι η ακόλουθη

$$\left.\frac{\boldsymbol{df}(\boldsymbol{x})}{\boldsymbol{dx}}\right|_{\boldsymbol{x}\,=\,\boldsymbol{x}^*} = \mathbf{0} \quad (1^{\eta\varsigma}\text{ Τάξης Αναγκαία Συνθήκη})$$

Στην Γράφημα 8 υπάρχουν πέντε λύσεις, οι οποίες ικανοποιούν αυτές τις συνθήκες. Για να εξασφαλίσουμε περισσότερες πληροφορίες σχετικά με τα πέντε αυτά κρίσιμα σημεία[39], είναι αναγκαία η εξέταση της δεύτερης παραγώγου.

Επομένως, παρατηρούμε ότι αν

$$\left.\frac{\boldsymbol{d}^2\boldsymbol{f}(\boldsymbol{x})}{\boldsymbol{dx}^2}\right|_{\boldsymbol{x}\,=\,\boldsymbol{x}^*} > \mathbf{0} \quad (2^{\eta\varsigma}\text{ Τάξης Αναγκαία Συνθήκη})$$

τότε το x^* πρέπει να είναι τουλάχιστον ένα σημείο τοπικού ελαχίστου. Δηλαδή, για μια περιοχή γύρω από το x^*, $U(x^*, \delta)$, ισχύει $f(x^*) \leq f(x)$.

Επίσης, παρατηρούμε ότι αν

[38] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1166.

[39] Έστω $U \subset \mathbb{R}^n$ ανοικτό και $f: U \to \mathbb{R}$ μια διαφορίσιμη συνάρτηση. Ένα κρίσιμο σημείο της f είναι ένα σημείο όπου η παράγωγος μηδενίζεται. Η τιμή της f σε ένα κρίσιμο σημείο είναι μια κρίσιμη τιμή. John H. Hubbard & Barbara Burke Hubbard, *Διανυσματικός Λογισμός, Γραμμική Άλγεβρα και Διαφορικές Μορφές – Μια ενοποιημένη Προσέγγιση*, μετάφραση των Β. Μεταφτσή και Α. Τσολομύτη, Πάτρα 2006, σελ.418.

$$\left.\boldsymbol{\frac{d^2 f(x)}{dx^2}}\right|_{\boldsymbol{x = x^*}} < \boldsymbol{0} \text{ (2}^{\text{ης}}\text{ Τάξης Αναγκαία Συνθήκη)}$$

τότε το x^* πρέπει να είναι τουλάχιστον ένα σημείο τοπικού μεγίστου. Δηλαδή, για μια περιοχή γύρω από το x^*, $U(x^*, \delta)$, ισχύει $f(x^*) \geq f(x)$.

Ερμηνεύοντας την παραπάνω θεωρία με βάση την έννοια της κυρτότητας, μπορούμε να πούμε ότι το x^* είναι τοπικό ελάχιστο αν η $f(x)$ είναι αυστηρά κυρτή μέσα σε μια γειτονιά του x^*. Ομοίως, το x^* είναι τοπικό μέγιστο αν η $f(x)$ είναι αυστηρά κοίλη μέσα σε μια γειτονιά του x^*.

Στο σημείο αυτό, αξίζει να αναφέρουμε ότι στην ειδική περίπτωση, όπου η δεύτερη παράγωγος ισούται με μηδέν, δηλαδή

$$\left.\boldsymbol{\frac{d^2 f(x)}{dx^2}}\right|_{\boldsymbol{x = x^*}} = \boldsymbol{0}$$

τότε το σημείο x^* αποτελεί σημείο καμπής και καθίσταται αναγκαία η περαιτέρω μελέτη ανώτερης τάξης παραγώγων.

Επιπλέον, για τον προσδιορισμό ενός ολικού ελαχίστου ή μεγίστου, δηλαδή ενός σημείου x^* τέτοιου ώστε $f(x^*) \leq f(x)$ ή $f(x^*) \geq f(x)$ αντιστοίχως, για όλα τα x που ανήκουν στο πεδίο ορισμού της συνάρτησης f, είναι αναγκαία η σύγκριση όλων των τοπικών ελαχίστων ή μεγίστων και να επιλεγεί αυτό με τη μικρότερη ή τη μεγαλύτερη τιμή της f, αντιστοίχως, καθώς το x παίρνει όλες τις τιμές στο πεδίο ορισμού.

Ωστόσο, αν η συνάρτηση f είναι κυρτή ή κοίλη, η ανάλυση γίνεται πολύ πιο εύκολη. Συγκεκριμένα, αν η συνάρτηση f είναι κυρτή τότε κάθε σημείο x^*, το οποίο μηδενίζει την πρώτη παράγωγο της f, αποτελεί ολικό ελάχιστο. Δηλαδή, η παρακάτω συνθήκη είναι ικανή και αναγκαία για την εύρεση ενός ολικού ελαχίστου μιας κυρτής συνάρτησης

$$\left.\boldsymbol{\frac{df(x)}{dx}}\right|_{\boldsymbol{x = x^*}} = \boldsymbol{0} \quad \text{(Ικανή και Αναγκαία Συνθήκη)}$$

Αξίζει να σημειωθεί, ότι αν η συνάρτηση f είναι αυστηρά κυρτή, τότε το σημείο x^* ολικού ελαχίστου είναι και μοναδικό.

Ομοίως, αν η συνάρτηση f είναι κοίλη τότε κάθε σημείο x^*, το οποίο μηδενίζει την πρώτη παράγωγο της f, αποτελεί ολικό μέγιστο. Δηλαδή, η παρακάτω συνθήκη είναι ικανή και αναγκαία για την εύρεση ενός ολικού μεγίστου μιας κοίλης συνάρτησης

$$\left.\frac{\boldsymbol{df(x)}}{\boldsymbol{dx}}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = \mathbf{0} \quad \text{(Ικανή και Αναγκαία Συνθήκη)}$$

Αξίζει να σημειωθεί, ότι αν η συνάρτηση f είναι αυστηρά κοίλη, τότε το σημείο x^* ολικού μεγίστου είναι και μοναδικό.

Στη συνέχεια, αφού ολοκληρώσαμε την παρουσίαση της θεωρίας που σχετίζεται με τη βελτιστοποίηση χωρίς περιορισμούς για συναρτήσεις μιας μεταβλητής, θα μεταβούμε στην ανάλυση της βελτιστοποίησης χωρίς περιορισμούς για συναρτήσεις πολλών μεταβλητών.

Βελτιστοποίηση χωρίς περιορισμούς για συναρτήσεις πολλών μεταβλητών

Η ανάλυση για μια συνάρτηση πολλών μεταβλητών $f(\vec{x})$, όπου $\vec{x} = (x_1, \dots, x_n)$, είναι ανάλογη με αυτή της προηγούμενης περίπτωσης.

Θεωρημα 2.1(1ης Ταξησ Αναγκαια Συνθηκη)[40]

Έστω $f : D \subset \mathbb{R}^n \to \mathbb{R}$ μια συνεχώς διαφορίσιμη συνάρτηση σε ένα ανοιχτό σύνολο D. Αν $\vec{x}^* \in D$ είναι ένα τοπικό ελάχιστο ή τοπικό μέγιστο της συνάρτησης f, τότε

$$\boldsymbol{\nabla f(\vec{x}^*) = 0}$$

Αποδειξη

[40] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 59.

Η απόδειξη γίνεται για το σημείο $\vec{x}^*$ να είναι τοπικό ελάχιστο, όμοια είναι η απόδειξη και για το τοπικό μέγιστο.

Α' Τροπος

Έστω ότι το σημείο $\vec{x}^*$ είναι ένα τοπικό ελάχιστο. Θεωρούμε την ακολουθία

$$\vec{x}_k = \vec{x}^* - a_k \nabla f(\vec{x}^*)\,, a_k > 0.$$

Από το ανάπτυγμα του Taylor, για k αρκετά μεγάλο, παίρνουμε

$$0 \leq f(\vec{x}_k) - f(\vec{x}^*) = -a_k \nabla f(\vec{\eta}_k)^T \nabla f(\vec{x}^*)\,,$$

όπου $\vec{\eta}_k$ είναι ένας κυρτός συνδυασμός των $\vec{x}_k$ και $\vec{x}^*$. Διαιρώντας με a_k και παίρνοντας το όριο, και επειδή $f \in \mathbb{C}^1$ συνεπάγεται

$$^{41}\; 0 \leq -\|\nabla f(\vec{x}^*)\|^2$$

Πράγμα το οποίο σημαίνει ότι $\nabla f(\vec{x}^*) = 0$.

Β' Τροπος (Εις Άτοπο Απαγωγη)

Θεωρούμε ότι $\nabla f(\vec{x}^*) \neq 0$. Παίρνοντας ότι $\vec{d} = -\nabla f(\vec{x}^*)$ έχουμε ότι

$$\vec{d}^T \nabla f(\vec{x}^*) = -\|\nabla f(\vec{x}^*)\|^2 < 0.$$

Επομένως, $\vec{d}$ είναι μια καθοδική κατεύθυνση και υπάρχει $\delta > 0$ τέτοιο ώστε

$$f\left(\vec{x}^* + a\vec{d}\right) < f(\vec{x}^*), \forall a \in (0, \delta)$$

το οποίο έρχεται σε αντίφαση με την υπόθεση ότι το $\vec{x}^*$ είναι ένα τοπικό ελάχιστο.

Γ' Τροπος

---

[41] Έστω X διανυσματικός χώρος. Μια απεικόνιση $\|\cdot\|: X \to \mathbb{R}$ λέγεται νόρμα αν ικανοποιεί τις ακόλουθες ιδιότητες:

i) $\|x\| \geq 0\,, \forall x \in X \;και\; \|x\| = 0 \;αν\; και\; μόνο\; αν\; x = 0.$
ii) $\|ax\| = |a|\|x\|, \forall a \in \mathbb{R} \;και\; \forall x \in X.$
iii) $\|x + y\| \leq \|x\| + \|y\|, \;\; \forall x, y \in X.$

Σπύρος Αργυρός, *Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης*, Αθήνα ²2004, σελ.29.
Lloyd N. Trefethen & David Bau III, *Numerical Linear Algebra*, Philadelphia 1997, σελ.17.
Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ.3.

Έστω $\vec{x}^*$ ένα τοπικό ελάχιστο, τότε υπάρχει $\delta > 0$, τέτοιο ώστε $f(\vec{x}) \geq f(\vec{x}^*)$ για κάθε $\vec{x}$ με $\|\vec{x} - \vec{x}^*\| < \delta$. Από το ανάπτυγμα Taylor

$$f(\vec{x}) = f(\vec{x}^*) + \nabla f(\vec{x}^*)^T(\vec{x} - \vec{x}^*) + o(\|\vec{x} - \vec{x}^*\|) \geq f(\vec{x}^*).$$

Διαιρώντας με $\|\vec{x} - \vec{x}^*\|$ και επειδή $\vec{x} \longrightarrow \vec{x}^*$ συνεπάγεται

$$\nabla f(\vec{x}^*)^T \frac{(\vec{x} - \vec{x}^*)}{\|\vec{x} - \vec{x}^*\|} \geq 0.$$

Θέτοντας $s = (\vec{x} - \vec{x}^*)/\|\vec{x} - \vec{x}^*\|$, η παραπάνω ανισότητα γίνεται

$$\nabla f(\vec{x}^*)^T s \geq 0\,, \nabla s\ \mu\varepsilon\ \|s\| = 1.$$

Επιλέγοντας $s = \pm e_i, (i = 1, \dots, n)$, παίρνουμε ότι $\nabla f(\vec{x}^*) = 0$.

--*τέλος απόδειξης*--

Όταν μια μεταβλητή x_j έχει έναν μη αρνητικό περιορισμό $x_j \geq 0$, η ικανή και (ίσως) αναγκαία συνθήκη αλλάζει ελαφρώς στην ακόλουθη

$$\left.\frac{\partial \mathbf{f}}{\partial \mathbf{x_j}}\right|_{\vec{\mathbf{x}} = \vec{\mathbf{x}}^*} \begin{cases} \leq \mathbf{0} \ \ \boldsymbol{\alpha\nu}\ \mathbf{x_j^*} = \mathbf{0} \\ = \mathbf{0} \ \ \boldsymbol{\alpha\nu}\ \mathbf{x_j^*} > \mathbf{0} \end{cases} \text{, για κάθε j.}$$

Προβλήματα τέτοιας μορφής, όπου περιέχουν μη αρνητικούς περιορισμούς αλλά όχι συναρτησιακούς περιορισμούς, αποτελούν μια ειδική περίπτωση ($m = 0$) της επόμενης ενότητας.[42]

Θεωρημα 2.2 (2ης Ταξης Αναγκαια Συνθηκη)[43]

Έστω $f : D \subset \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση δύο φορές συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο D. Αν $\vec{x}^* \in D$ είναι ένα τοπικό ελάχιστο ή τοπικό μέγιστο της συνάρτησης f, τότε $\nabla f(\vec{x}^*) = 0$ και $\nabla^2 f(\vec{x}^*)$ είναι θετικά ημιορισμένο ή αρνητικά ημιορισμένο, αντιστοίχως.

[42] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.665.
[43] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 60.

Αποδειξη

Η απόδειξη γίνεται για το σημείο $\vec{x}^*$ να είναι τοπικό ελάχιστο, όμοια είναι η απόδειξη και για το τοπικό μέγιστο.

Α' Τροπος

Από το θεώρημα 2.1 γνωρίζουμε ότι $\nabla f(\vec{x}^*) = 0$, άρα αρκεί να αποδείξουμε μόνο ότι $\nabla^2 f(\vec{x}^*)$ είναι θετικά ημιορισμένο. Θεωρούμε την ακολουθία

$$\vec{x}_k = \vec{x}^* + a_k \vec{d}\,, a_k > 0,$$

όπου $\vec{d}$ είναι αυθαίρετο. Αφού $f \in \mathbb{C}^2$ και $\nabla f(\vec{x}^*) = 0$, τότε από το ανάπτυγμα του Taylor, για k αρκετά μεγάλο, παίρνουμε

$$0 \leq f(\vec{x}_k) - f(\vec{x}^*) = \frac{1}{2} a_k^2 \vec{d}^T \nabla^2 f(\vec{\eta}_k) \vec{d}\,,$$

όπου $\vec{\eta}_k$ είναι ένας κυρτός συνδυασμός των $\vec{x}_k$ και $\vec{x}^*$. Διαιρώντας με $\frac{1}{2} a_k^2$ και παίρνοντας το όριο, έχουμε

$$\vec{d}^T \nabla^2 f(\vec{\eta}_k) \vec{d} \geq 0\,, \forall \vec{d} \in \mathbb{R}^n.$$

Β' Τροπος (Εις Άτοπο Απαγωγη)

Θεωρούμε ότι $\nabla^2 f(\vec{x}^*)$ δεν είναι θετικά ημιορισμένο, μπορούμε να διαλέξουμε ένα $\vec{d} \in \mathbb{R}^n$ τέτοιο ώστε $\vec{d}^T \nabla^2 f(\vec{x}^*) \vec{d} < 0$. Αφού $f \in \mathbb{C}^2$, υπάρχει $\delta > 0$ και μπορούμε να επιλέξουμε ένα $\varepsilon > 0$, τέτοιο ώστε $\vec{x}^* + \varepsilon \vec{d} \in B(\vec{x}^*, \delta)$ και $\vec{d}^T \nabla^2 f\left(\vec{x}^* + \varepsilon \vec{d}\right) \vec{d} < 0$.

Λαμβάνοντας υπόψη ότι $\nabla f(\vec{x}^*) = 0$, συνεπάγεται ότι

$$f\left(\vec{x}^* + \varepsilon \vec{d}\right) = f(\vec{x}^*) + \frac{1}{2} \varepsilon^2 \vec{d}^T \nabla^2 f\left(\vec{x}^* + \theta \varepsilon \vec{d}\right) \vec{d},$$

όπου $0 \leq \theta \leq 1$. Επομένως, $f\left(\vec{x}^* + \varepsilon \vec{d}\right) < f(\vec{x}^*)$, το οποίο έρχεται σε αντίφαση με την υπόθεση ότι το $\vec{x}^*$ είναι ένα τοπικό ελάχιστο.

---*τέλος απόδειξης*---

Θεωρημα 2.3 (2ης Ταξης Ικανη Συνθηκη)[44]

Έστω $f : D \subset \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση δύο φορές συνεχώς διαφορίσιμη σε ένα ανοιχτό σύνολο D. Αν $\nabla f(\vec{x}^*) = 0$ και $\nabla^2 f(\vec{x}^*)$ είναι θετικά ορισμένο ή αρνητικά ορισμένο, τότε το $\vec{x}^* \in D$ είναι ένα αυστηρά τοπικό ελάχιστο ή μέγιστο, αντιστοίχως.

Αποδειξη

Η απόδειξη γίνεται για $\nabla^2 f(\vec{x}^*)$ είναι θετικά ορισμένο, όμοια είναι η απόδειξη και για $\nabla^2 f(\vec{x}^*)$ είναι αρνητικά ορισμένο.

Α' Τροπος

Υποθέτουμε ότι $\nabla f(\vec{x}^*) = 0$ και ότι $\nabla^2 f(\vec{x}^*)$ είναι θετικά ορισμένο. Από το ανάπτυγμα του Taylor, για κάθε διάνυσμα $\vec{d} \in \mathbb{R}^n$, τέτοιο ώστε $\vec{x}^* + \vec{d}$ να βρίσκεται σε μία περιοχή του $\vec{x}^*$ στην οποία $\nabla^2 f\left(\vec{x}^* + \vec{d}\right)$ είναι θετικά ορισμένο, επομένως έχουμε

$$f\left(\vec{x}^* + \vec{d}\right) = f(\vec{x}^*) + \frac{1}{2}\vec{d}^T \nabla^2 f\left(\vec{x}^* + \theta \vec{d}\right)\vec{d},$$

όπου $\theta \in (0,1)$. Τότε, μπορούμε να διαλέξουμε ένα $\delta > 0$, τέτοιο ώστε $\vec{x}^* + \vec{d} \in B(\vec{x}^*, \delta)$ και $\vec{d}^T \nabla^2 f\left(\vec{x}^* + \theta \vec{d}\right)\vec{d} > 0$. Συνεπώς,

$$f\left(\vec{x}^* + \vec{d}\right) > f(\vec{x}^*).$$

Β' Τροπος (Εις Άτοπο Απαγωγη)

Υποθέτουμε ότι το $\vec{x}^*$ δεν είναι αυστηρά τοπικό ελάχιστο, τότε υπάρχει μια ακολουθία $\{x_k\} \subset D$ με $x_k \neq \vec{x}^*$, $\forall k$, τέτοια ώστε $f(x_k) \leq f(\vec{x}^*)$ για k αρκετά μεγάλο. Χρησιμοποιώντας το ανάπτυγμα του Taylor

$$0 \geq f(x_k) - f(\vec{x}^*) = \nabla f(\vec{x}^*)^T(\vec{x}_k - \vec{x}^*) + \frac{1}{2}(\vec{x}_k - \vec{x}^*)^T \nabla^2 f(\eta_k)(\vec{x}_k - \vec{x}^*),$$

[44] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 61.

όπου η_k είναι ένας κυρτός συνδυασμός των $\vec{x}_k$ και $\vec{x}^*$. Λαμβάνοντας υπόψη ότι $\nabla f(\vec{x}^*) = 0$ και διαιρώντας με $\frac{1}{2}\|\vec{x}_k - \vec{x}^*\|^2$ και παίρνοντας το όριο, έχουμε

$$0 \geq \bar{e}^T \nabla^2 f(\vec{x}^*)\bar{e} \quad \boxed{1},$$

όπου $\bar{e}$ είναι σημείο συσσώρευσης (ή οριακό σημείο) μιας ομοιόμορφης φραγμένης ακολουθίας $\{(\vec{x}_k - \vec{x}^*)/\|\vec{x}_k - \vec{x}^*\|\}$ και $\|\bar{e}\| = 1$. Προφανώς, η (1) έρχεται σε αντίφαση με τη θετική ορισιμότητα του $\nabla^2 f(\vec{x}^*)$.

---*τέλος απόδειξης*---

Θεωρημα 2.4[45]

Έστω $S \subset \mathbb{R}^n$ ένα μη κενό κυρτό σύνολο και $f : \mathbb{R}^n \to \mathbb{R}$. Έστω $\vec{x}^* \in S$ ένα τοπικό ελάχιστο της f στο S.

1. Αν η f είναι κυρτή, τότε το $\vec{x}^*$ είναι επίσης ένα ολικό ελάχιστο.
2. Αν η f είναι αυστηρά κυρτή, τότε το $\vec{x}^*$ είναι το μοναδικό ολικό ελάχιστο.

Αποδειξη

(1) Έστω ότι η συνάρτηση f είναι κυρτή και το σημείο $\vec{x}^*$ είναι ένα τοπικό ελάχιστο, τότε υπάρχει μια δ-γειτονιά $B(\vec{x}^*, \delta)$, τέτοια ώστε

$$f(\vec{x}) \geq f(\vec{x}^*), \forall x \in S \cap B(\vec{x}^*, \delta) \quad \boxed{1}.$$

Από εις άτοπο απαγωγή, υποθέτουμε ότι το $\vec{x}^*$ δεν είναι ένα ολικό ελάχιστο. Τότε, μπορούμε να βρούμε κάποιο $\vec{x}^* \in S$, τέτοιο ώστε $f(\hat{\vec{x}}) < f(\vec{x}^*)$. Από κυρτότητα της f, έχουμε για $\alpha \in (0,1)$

$$f\left(\alpha\hat{\vec{x}} + (1-\alpha)\vec{x}^*\right) \leq \alpha f\left(\hat{\vec{x}}\right) + (1-\alpha)f(\vec{x}^*) \Rightarrow$$

$$f\left(\alpha\hat{\vec{x}} + (1-\alpha)\vec{x}^*\right) < \alpha f(\vec{x}^*) + (1-\alpha)f(\vec{x}^*) \Rightarrow$$

$$f\left(\alpha\hat{\vec{x}} + (1-\alpha)\vec{x}^*\right) < f(\vec{x}^*) \quad \boxed{2}$$

[45] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 62.

Ωστόσο, για αρκετά μικρό $\alpha > 0, \alpha\hat{\vec{x}} + (1-\alpha)\vec{x}^* \in S \cap B(\vec{x}^*, \delta)$. Επομένως, η (2) έρχεται σε αντίφαση με την (1). Αυτή η αντίφαση αποδεικνύει το πρώτο συμπέρασμα.

(2) Από το πρώτο μέρος, έχουμε ότι το $\vec{x}^*$ είναι ένα ολικό ελάχιστο, επειδή η αυστηρή κυρτότητα ισοδυναμεί με κυρτότητα. Επομένως, αρκεί να αποδείξουμε τη μοναδικότητα.

Από εις άτοπο απαγωγή, υποθέτουμε ότι το $\vec{x}^*$ δεν είναι το μοναδικό ολικό ελάχιστο, άρα μπορούμε να βρούμε ένα $\vec{x} \in S, \vec{x} \neq \vec{x}^*$, τέτοιο ώστε $f(\vec{x}) = f(\vec{x}^*)$.

Από αυστηρή κυρτότητα της f

$$f\left(\frac{1}{2}\vec{x} + \frac{1}{2}\vec{x}^*\right) < \frac{1}{2}f(\vec{x}) + \frac{1}{2}f(\vec{x}^*) = f(\vec{x}^*)$$

Αξίζει να σημειωθεί ότι από την κυρτότητα του συνόλου S έχουμε ότι $\frac{1}{2}\vec{x} + \frac{1}{2}\vec{x}^* \in S$. Επομένως, η παραπάνω σχέση έρχεται σε αντίφαση με το γεγονός ότι το $\vec{x}^*$ είναι ένα ολικό ελάχιστο.

---*τέλος απόδειξης*---

Θεωρημα 2.5

Έστω $S \subset \mathbb{R}^n$ ένα μη κενό κυρτό σύνολο και $f : \mathbb{R}^n \to \mathbb{R}$. Έστω $\vec{x}^* \in S$ ένα τοπικό μέγιστο της f στο S.

1. Αν η f είναι κοίλη, τότε το $\vec{x}^*$ είναι, επίσης, ένα ολικό μέγιστο.
2. Αν η f είναι αυστηρά κοίλη, τότε το $\vec{x}^*$ είναι το μοναδικό ολικό μέγιστο.

Η απόδειξη του θεωρήματος παραλείπεται διότι είναι ανάλογη της απόδειξης του θεωρήματος 2.4.

Θεωρημα 2.6[46]

Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια διαφορίσιμη κυρτή ή κοίλη συνάρτηση. Τότε, το $\vec{x}^*$ είναι ένα ολικό ελάχιστο ή μέγιστο, αντιστοίχως, αν και μόνο αν $\nabla f(\vec{x}^*) = 0$.

[46] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 63.

Απόδειξη

Η απόδειξη γίνεται για το σημείο $\vec{x}^*$ να είναι ολικό ελάχιστο, όμοια είναι η απόδειξη και για το ολικό μέγιστο.

Αναγκαιότητα

Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια διαφορίσιμη κυρτή συνάρτηση στο $\mathbb{R}^n$ και $\nabla f(\vec{x}^*) = 0$, τότε

$$f(\vec{x}) \geq f(\vec{x}^*) + \nabla f(\vec{x}^*)(\vec{x} - \vec{x}^*) = f(\vec{x}^*)\,, \forall \vec{x} \in \mathbb{R}^n$$

το οποίο δηλώνει ότι το $\vec{x}^*$ είναι ένα ολικό ελάχιστο της f.

Ικανότητα

Είναι προφανές ότι το ολικό ελάχιστο είναι και τοπικό ελάχιστο, καθώς επίσης στάσιμο σημείο.

---*τέλος απόδειξης*---

Τέλος, θα ολοκληρώσουμε την ανάλυση αυτής της ενότητας με την παρουσίαση της θεωρίας για τη βελτιστοποίηση με περιορισμούς ισότητας.

Δεσμευμένη βελτιστοποίηση με περιορισμούς ισότητας[47]

Στη συνέχεια, θεωρούμε το πρόβλημα εύρεσης ελαχίστου ή μεγίστου μιας συνάρτησης $f(\vec{x})$, υπό τον περιορισμό ότι η μεταβλητή $\vec{x}$ πρέπει να ικανοποιεί όλες τις ισότητες

$$g_1(\vec{x}) = b_1$$

$$g_2(\vec{x}) = b_2$$

$$\vdots$$

$$g_m(\vec{x}) = b_m$$

[47] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.1166.

όπου $m < n$.

Μια κλασική μέθοδος για την επίλυση ενός τέτοιου προβλήματος είναι η μέθοδος των πολλαπλασιαστών Lagrange. Αρχικά, ορίζουμε την Λαγκραζιανή συνάρτηση

$$h(\vec{x},\vec{\lambda}) = f(\vec{x}) - \sum_{i=1}^{m} \lambda_i [g_i(\vec{x}) - b_i]$$

όπου οι νέες μεταβλητές $\vec{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_m)$ ονομάζονται Λαγκραζιανοί πολλαπλασιαστές. Παρατηρούμε ότι για τις εφικτές τιμές της μεταβλητής $\vec{x}$ ισχύει

$$g_i(\vec{x}) - b_i = 0\,, για\ όλα\ τα\ i$$

και άρα, $h(\vec{x},\vec{\lambda}) = f(\vec{x})$. Επομένως, μπορεί εύκολα να γίνει αντιληπτό ότι αν $(\vec{x},\vec{\lambda}) = (\vec{x}^*,\vec{\lambda}^*)$ είναι ένα τοπικό ή ολικό ελάχιστο ή μέγιστο για τη χωρίς περιορισμούς συνάρτηση $h(\vec{x},\vec{\lambda})$, τότε το $\vec{x}^*$ είναι ένα αντίστοιχο κρίσιμο σημείο για το αρχικό πρόβλημα. Κατά συνέπεια, η μέθοδος περιορίζεται στην ανάλυση της $h(\vec{x},\vec{\lambda})$ με τη μέθοδο που αναπτύχθηκε στην ενότητα για τη βελτιστοποίηση χωρίς περιορισμούς. Έτσι, οι $n + m$ μερικές παράγωγοι τίθενται ίσες με το μηδέν

$$\frac{\boldsymbol{\partial h(\vec{x},\vec{\lambda})}}{\boldsymbol{\partial x_j}} = \frac{\boldsymbol{\partial f(\vec{x})}}{\boldsymbol{\partial x_j}} - \sum_{i=1}^{m} \boldsymbol{\lambda_i} \frac{\boldsymbol{\partial g_i(\vec{x})}}{\boldsymbol{\partial x_j}} = \mathbf{0}\,, για\ j = 1, 2, \dots, n \quad \boxed{1}$$

$$\frac{\boldsymbol{\partial h(\vec{x},\vec{\lambda})}}{\boldsymbol{\partial \lambda_i}} = -\boldsymbol{g_i(\vec{x})} + \boldsymbol{b_i} = \mathbf{0}\,, για\ i = 1, 2, \dots, m \quad \boxed{2}$$

τότε τα κρίσιμα σημεία βρίσκονται λύνοντας αυτές τις εξισώσεις για τα $(\vec{x},\vec{\lambda})$. Στο σημείο αυτό αξίζει να σημειωθεί, ότι οι εξισώσεις της μορφής (2) είναι ισοδύναμες με τους περιορισμούς του αρχικού προβλήματος, κατ' επέκταση, μόνο οι εφικτές λύσεις λαμβάνονται υπόψη. Ύστερα από περεταίρω ανάλυση για την εύρεση του ολικού ελαχίστου ή μεγίστου της συνάρτησης $h(\vec{x},\vec{\lambda})$, η τιμή του $\vec{x}$, η οποία προκύπτει από την επίλυση, είναι η επιθυμητή λύση του αρχικού προβλήματος.

Ωστόσο, από υπολογιστική σκοπιά, η μέθοδος των πολλαπλασιαστών Lagrange δεν είναι μια ιδιαίτερα ισχυρή διαδικασία. Συχνά, είναι ουσιωδώς αδύνατον να επιλυθούν οι εξισώσεις, ώστε να βρεθούν τα κρίσιμα σημεία. Επιπλέον, ακόμα κι αν τα κρίσιμα σημεία μπορούν να προσδιοριστούν, ο αριθμός των κρίσιμων σημείων μπορεί να είναι τόσο μεγάλος (συχνά άπειρος) που δεν καθίσταται πρακτική η προσπάθεια εύρεσης του ολικού ελαχίστου ή μεγίστου. Παρ' όλα αυτά, για συγκεκριμένες περιπτώσεις μικρών προβλημάτων, αυτή η μέθοδος μπορεί, μερικές φορές, να χρησιμοποιηθεί επιτυχώς.

Μη Γραμμικός Προγραμματισμός

Μια βασική υπόθεση στον γραμμικό προγραμματισμό είναι ότι τόσο η αντικειμενική συνάρτηση όσο και οι συναρτήσεις των περιορισμών είναι γραμμικές. Ωστόσο, αυτή η υπόθεση συνήθως δεν μπορεί να εφαρμοστεί σε αρκετές περιπτώσεις. Επομένως, κρίνεται αναγκαία η απευθείας αντιμετώπιση προβλημάτων μη γραμμικού προγραμματισμού. Για το λόγο αυτό, στο παρόν κεφάλαιο θα παρουσιάσουμε διάφορους τύπους προβλημάτων μη γραμμικού προγραμματισμού, καθώς και τεχνικές επίλυσής τους.[48]

Αρχικά, θα ορίσουμε τη γενική μορφή ενός προβλήματος μη γραμμικού προγραμματισμού. Σε μια γενική μορφή, το πρόβλημα μη γραμμικού προγραμματισμού έγκειται στην εύρεση ενός διανύσματος $\vec{x} = (x_1, x_2, \dots, x_n)$, τέτοιο ώστε

$$maximize\ \ f(\vec{x}) \quad \boxed{1}$$

υπό τους περιορισμούς

$$g_i(\vec{x}) \leq b_i\ , \qquad \text{για } i = 1, \dots, m \quad \boxed{2}$$

και

$$\vec{x} \geq \vec{0} \quad \boxed{3}$$

όπου $f(\vec{x})$ και $g_i(\vec{x})$ δοθείσες συναρτήσεις των μεταβλητών απόφασης[49].

Στη συνέχεια, με βάση την κανονική μορφή ενός προβλήματος θα δώσουμε την ορολογία για τις λύσεις που προκύπτουν από τις μεθόδους επίλυσης της επιχειρησιακής έρευνας για προβλήματα γραμμικού και μη γραμμικού προγραμματισμού.

[48] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.654.

[49] Στο συγκεκριμένο κεφάλαιο, για λόγους ευκολίας, θεωρούμε ότι όλες οι συναρτήσεις είναι είτε παντού διαφορίσιμες είτε μερικώς γραμμικές συναρτήσεις.

- ***Λύση*** ενός προβλήματος μη γραμμικού προγραμματισμού καλείται κάθε λύση του συστήματος (1). Συγκεκριμένα, κάθε διάνυσμα $\vec{x} = (x_1, x_2, \dots, x_n)$ το οποίο ικανοποιεί το σύστημα (1).
- ***Δυνατή*** ή ***εφικτή λύση*** ενός προβλήματος μη γραμμικού προγραμματισμού καλείται κάθε λύση για την οποία όλοι οι περιορισμοί ικανοποιούνται. Συγκεκριμένα, κάθε λύση του συστήματος (2) που ικανοποιεί τους περιορισμούς (3).
- ***Αδύνατη ή ανέφικτη λύση*** ενός προβλήματος μη γραμμικού προγραμματισμού καλείται κάθε λύση για την οποία τουλάχιστον ένας περιορισμός δεν ικανοποιείται.
- ***Βέλτιστη εφικτή λύση*** ενός προβλήματος μη γραμμικού προγραμματισμού καλείται κάθε εφικτή λύση αυτού που βελτιστοποιεί την αντικειμενική συνάρτηση του προβλήματος. [50]

[50] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.34.

Γραφική Απεικόνιση Προβλημάτων Μη Γραμμικού Προγραμματισμού

Στην περίπτωση που ένα πρόβλημα μη γραμμικού προγραμματισμού έχει μόνο μια ή δύο μεταβλητές, μπορεί να αναπαρασταθεί γραφικά. Για την ευκολότερη κατανόηση της παρούσας ενότητας, η θεωρία θα αναπτυχθεί με τη βοήθεια ενός παραδείγματος. Αρχικά, θα παρουσιάσουμε τη γραφική απεικόνιση ενός προβλήματος γραμμικού προγραμματισμού με δύο μεταβλητές και στη συνέχεια, με βάση το πρόβλημα του γραμμικού προγραμματισμού, θα εξετάσουμε τέσσερις περιπτώσεις μη γραμμικού προγραμματισμού.

Παρακάτω δίνεται το πρόβλημα γραμμικού προγραμματισμού με δύο μεταβλητές το οποίο και επιλύεται γραφικά

$$\max z = 3x_1 + 5x_2$$

υπό τους περιορισμούς:

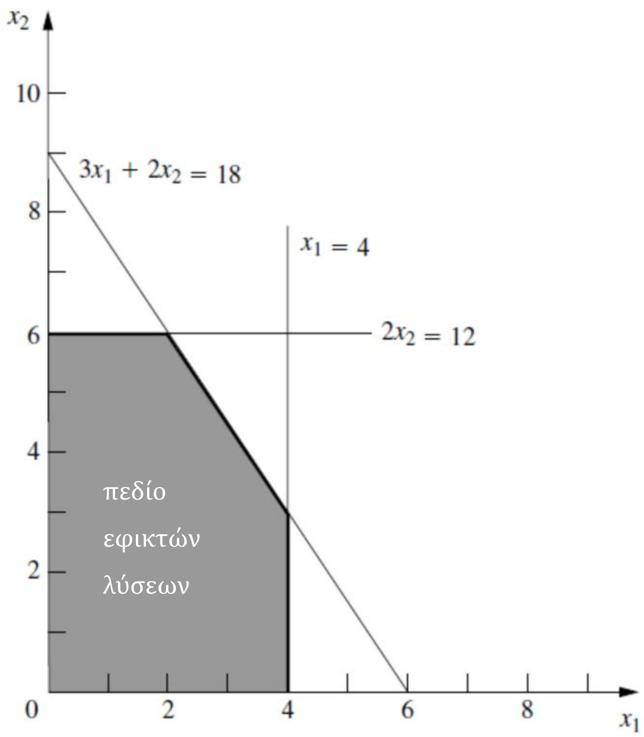
$$\begin{aligned} x_1 \qquad &\leq 4 \\ 2x_2 &\leq 12 \\ 3x_1 + 2x_2 &\leq 18 \end{aligned}$$

με περιορισμούς μη αρνητικότητας:

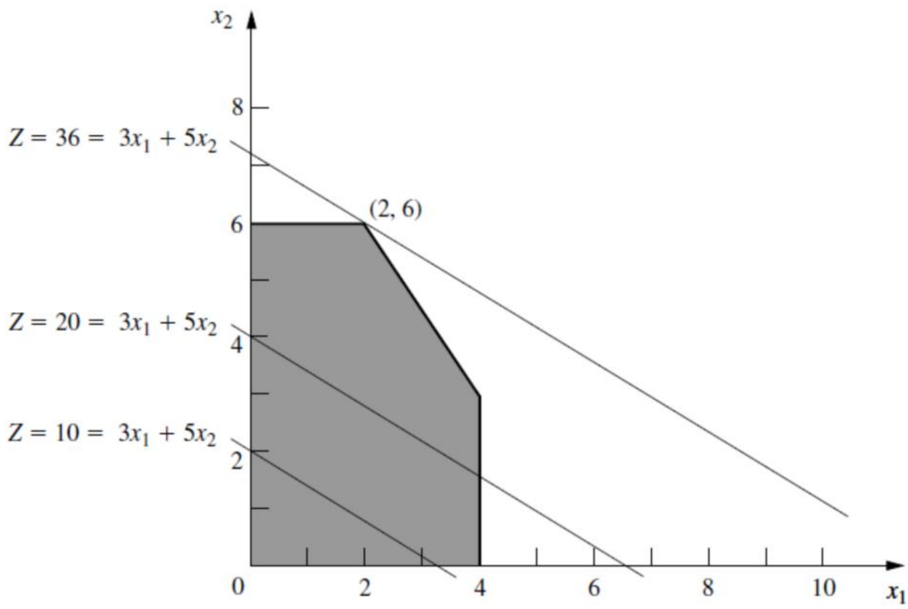
$$x_1, x_2 \geq 0$$

Από τη γραφική επίλυση (Γραφ. 10) παρατηρούμε ότι το γωνιακό σημείο του πεδίου εφικτών λύσεων, το οποίο βελτιστοποιεί την αντικειμενική συνάρτηση του προβλήματος είναι το $(2, 6)$ με μέγιστη τιμή $\max z = 36$.[51]

[51] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.659.



Γράφημα 9 - Γραφική Επίλυση, πεδίο εφικτών λύσεων.[52]



Γράφημα 10 - Γραφική Επίλυση, βέλτιστη λύση.[53]

[52] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.28.
[53] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.29.

Στη συνέχεια θα παρουσιάσουμε περιπτώσεις μη γραμμικού προγραμματισμού και θα επιχειρήσουμε να τις επιλύσουμε γραφικά.

1η περίπτωση

Στην περίπτωση αυτή θα μελετήσουμε το αρχικό μας πρόβλημα αλλάζοντας το δεύτερο και τρίτο περιορισμό με τον ακόλουθο μη γραμμικό περιορισμό

$$9x_1^2 + 5x_2^2 \leq 216$$

Επομένως, έχουμε το πρόβλημα μη γραμμικού προγραμματισμού

$$\mathbf{max}\, \boldsymbol{z} = \mathbf{3}\boldsymbol{x}_\mathbf{1} + \mathbf{5}\boldsymbol{x}_\mathbf{2}$$

υπό τους περιορισμούς:

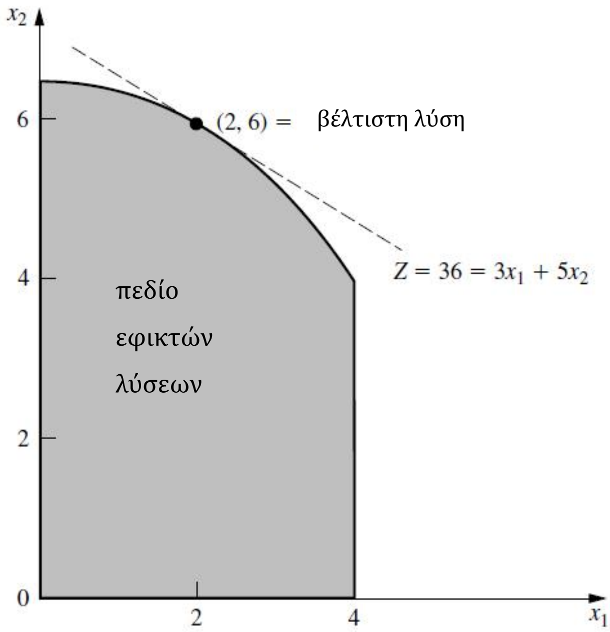
$$\boldsymbol{x}_\mathbf{1} \leq \mathbf{4}$$
$$\mathbf{9}\boldsymbol{x}_\mathbf{1}^\mathbf{2} + \mathbf{5}\boldsymbol{x}_\mathbf{2}^\mathbf{2} \leq \mathbf{216}$$

με περιορισμούς μη αρνητικότητας:

$$\boldsymbol{x}_\mathbf{1}, \boldsymbol{x}_\mathbf{2} \geq \mathbf{0}$$

Από τη γραφική επίλυση (Γραφ. 11) παρατηρούμε ότι το σημείο του πεδίου εφικτών λύσεων, το οποίο βελτιστοποιεί την αντικειμενική συνάρτηση του προβλήματος είναι το $(2, 6)$ με μέγιστη τιμή $\max z = 36$, δηλαδή ίδιο με το αρχικό πρόβλημα, ωστόσο το σημείο αυτό ενώ βρίσκεται στο σύνορο του πεδίου των εφικτών λύσεων, δεν είναι γωνιακό σημείο.[54]

[54] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.659.



Γράφημα 11 - Γραφική επίλυση, 1η περίπτωση.[55]

2η περίπτωση

Στην περίπτωση αυτή θα μελετήσουμε το αρχικό μας πρόβλημα αλλάζοντας την αντικειμενική συνάρτηση με την ακόλουθη μη γραμμική

$$z = 126x_1 - 9x_1^2 + 182x_2 - 13x_2^2$$

Επομένως, έχουμε το πρόβλημα μη γραμμικού προγραμματισμού

$$\max z = 126x_1 - 9x_1^2 + 182x_2 - 13x_2^2$$

υπό τους περιορισμούς:

$$\begin{aligned} x_1 \quad\quad &\leq 4 \\ 2x_2 &\leq 12 \\ 3x_1 + 2x_2 &\leq 18 \end{aligned}$$

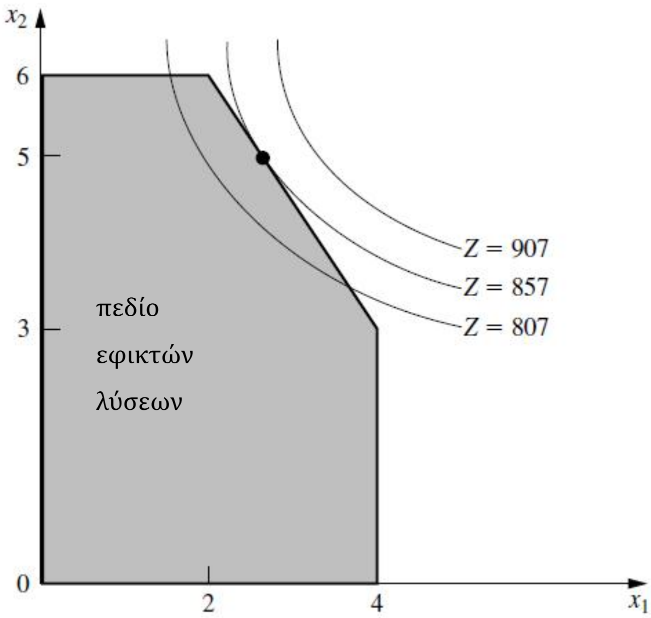
με περιορισμούς μη αρνητικότητας:

$$x_1, x_2 \geq 0$$

Από τη γραφική επίλυση (Γραφ. 12) παρατηρούμε ότι το σημείο του πεδίου εφικτών λύσεων, το οποίο βελτιστοποιεί την αντικειμενική συνάρτηση του

[55] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.660.

προβλήματος είναι το $(\frac{8}{3}, 5)$ με μέγιστη τιμή $\max z = 857$. Παρατηρούμε ότι το σημείο αυτό, ενώ βρίσκεται στο σύνορο του πεδίου των εφικτών λύσεων, δεν είναι γωνιακό σημείο και επίσης, ο τόπος όλων των σημείων με $z = 857$ τέμνει το πεδίο εφικτών λύσεων μόνο σε αυτό το σημείο, ενώ ο τόπος των σημείων με μεγαλύτερη τιμή z δεν τέμνεται με το πεδίο εφικτών λύσεων. [56]



Γράφημα 12 - Γραφική Επίλυση, 2η περίπτωση.[57]

3η περίπτωση

Στην περίπτωση αυτή θα μελετήσουμε το αρχικό μας πρόβλημα αλλάζοντας την αντικειμενική συνάρτηση με την ακόλουθη μη γραμμική

$$z = 54x_1 - 9x_1^2 + 78x_2 - 13x_2^2$$

Επομένως, έχουμε το πρόβλημα μη γραμμικού προγραμματισμού

$$\mathbf{\max z = 54x_1 - 9x_1^2 + 78x_2 - 13x_2^2}$$

υπό τους περιορισμούς:

$$\begin{aligned} \mathbf{x_1} \quad\quad & \mathbf{\leq 4} \\ \mathbf{2x_2} & \mathbf{\leq 12} \end{aligned}$$

[56] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York 72001, σελ.661.

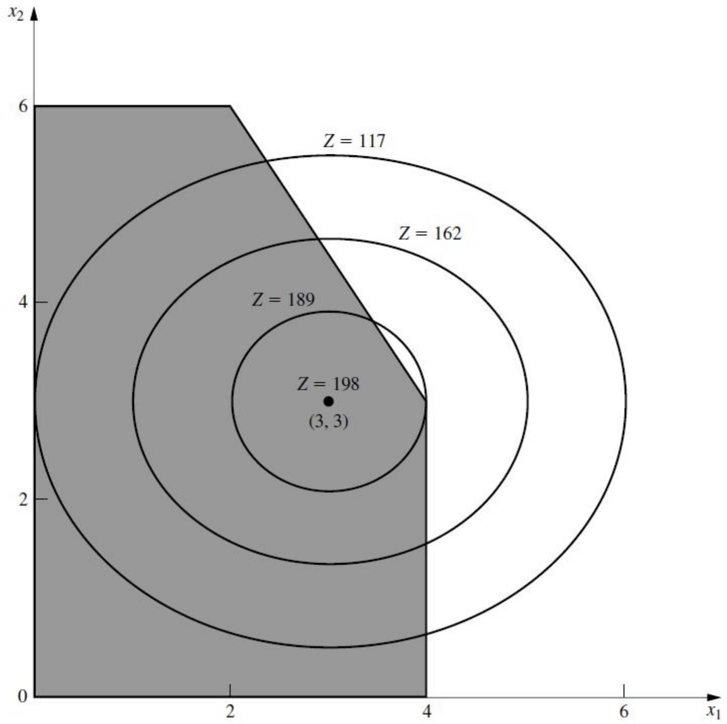
[57] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York 72001, σελ.660.

$$3x_1 + 2x_2 \leq 18$$

με περιορισμούς μη αρνητικότητας:

$$x_1, x_2 \geq 0$$

Από τη γραφική επίλυση (Γραφ. 13) παρατηρούμε ότι το σημείο του πεδίου εφικτών λύσεων, το οποίο βελτιστοποιεί την αντικειμενική συνάρτηση του προβλήματος είναι το $(3, 3)$ με μέγιστη τιμή $\max z = 252$. Παρατηρούμε ότι το σημείο αυτό βρίσκεται μέσα στο πεδίο των εφικτών λύσεων, δηλαδή, πρέπει να ληφθούν υπόψη όλες οι λύσεις στο πεδίο εφικτών λύσεων και όχι μόνο αυτές που βρίσκονται στο σύνορο.[58]



Γράφημα 13 - Γραφική Επίλυση, 3η περίπτωση.[59]

4η περίπτωση

Στην περίπτωση αυτή θα μελετήσουμε το αρχικό μας πρόβλημα αλλάζοντας τον δεύτερο και τρίτο περιορισμό με τον ακόλουθο κυρτό περιορισμό

[58] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.661.
[59] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.661.

$$8x_1 - x_1^2 + 14x_2 - x_2^2 \leq 49$$

Η συνάρτηση $g_2(\vec{x}) = 8x_1 - x_1^2 + 14x_2 - x_2^2$ είναι κυρτή, αφού $8x_1 - x_1^2$ και $14x_2 - x_2^2$ είναι κυρτές συναρτήσεις.

Επομένως, έχουμε το πρόβλημα μη γραμμικού προγραμματισμού

$$\mathbf{max}\, \boldsymbol{z} = \mathbf{3}\boldsymbol{x_1} + \mathbf{5}\boldsymbol{x_2}$$

υπό τους περιορισμούς:

$$\begin{aligned} \boldsymbol{x_1} \qquad\qquad\qquad\qquad & \leq \mathbf{4} \\ \mathbf{8}\boldsymbol{x_1} + \mathbf{14}\boldsymbol{x_2} - \boldsymbol{x_1^2} - \boldsymbol{x_2^2} & \leq \mathbf{49} \end{aligned}$$

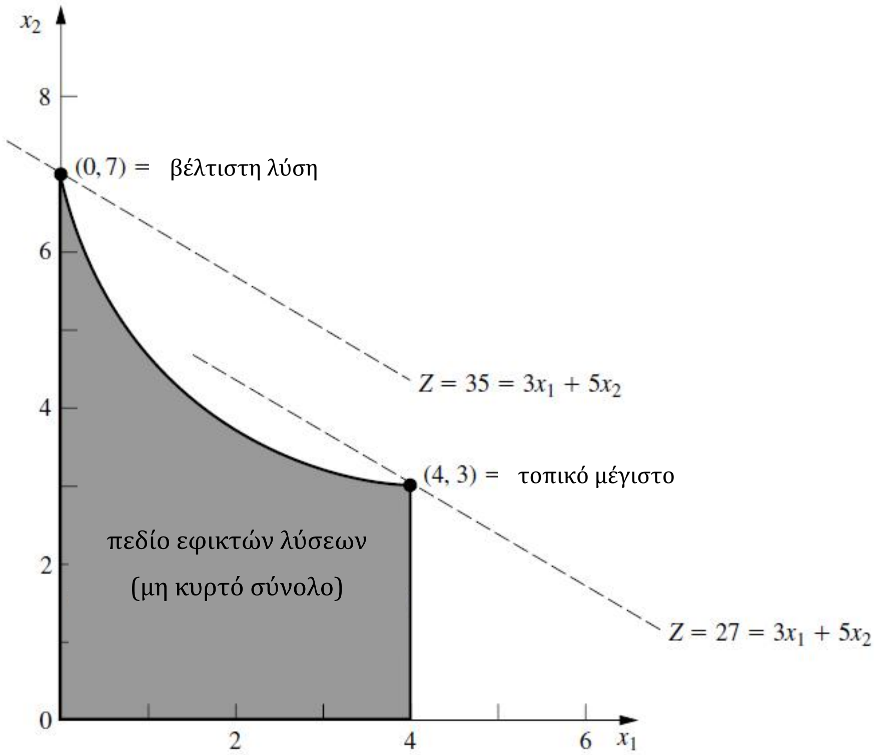
με περιορισμούς μη αρνητικότητας:

$$\boldsymbol{x_1}, \boldsymbol{x_2} \geq \mathbf{0}$$

Από τη γραφική επίλυση (Γραφ. 14) παρατηρούμε ότι το σημείο του πεδίου εφικτών λύσεων, το οποίο βελτιστοποιεί την αντικειμενική συνάρτηση του προβλήματος είναι το $(0, 7)$ με μέγιστη τιμή $\max z = 35$. Παρατηρούμε ότι το πεδίο εφικτών λύσεων δεν είναι κυρτό σύνολο και περιέχει τα σημεία τοπικού μεγίστου $(0, 7)$ και $(4, 3)$, επομένως δεν μπορούμε να εγγυηθούμε ότι τα σημεία αυτά είναι και σημεία ολικού μεγίστου. Πράγματι, από τα δύο αυτά σημεία μόνο το σημείο $(0, 7)$ είναι ολικό μέγιστο.

Καταλήγουμε, λοιπόν, στο συμπέρασμα ότι για να είναι ένα τοπικό μέγιστο και ολικό για ένα πρόβλημα μη γραμμικού προγραμματισμού με περιορισμούς $g_i(\vec{x}) \leq b_i$ $(i = 1,2, \ldots, m)$ και $\vec{x} \geq 0$, η αντικειμενική συνάρτηση $f(\vec{x})$ πρέπει να είναι κοίλη συνάρτηση και κάθε $g_i(\vec{x})$ πρέπει να είναι κυρτή συνάρτηση. Τέτοιο πρόβλημα λέγεται πρόβλημα κυρτού προγραμματισμού. [60]

[60] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.663.



Γράφημα 14 - Γραφική Επίλυση, 4η περίπτωση.[61]

[61] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.664.

Τύποι Προβλημάτων Μη Γραμμικού Προγραμματισμού

Στο κεφάλαιο αυτό, θα παρουσιάσουμε αναλυτικά τους τύπους προβλημάτων που συναντάμε στον μη γραμμικό προγραμματισμό. Γενικώς, τα προβλήματα μη γραμμικού προγραμματισμού εμφανίζονται σε πολλές διαφορετικές μορφές και επομένως, δεν υπάρχει ένας συγκεκριμένος αλγόριθμος, ο οποίος να επιλύει όλα τα προβλήματα. Ωστόσο, για συγκεκριμένες μορφές προβλημάτων έχουν αναπτυχθεί αλγόριθμοι επίλυσης, οι οποίοι παρουσιάζονται διεξοδικά στη συνέχεια.

Βελτιστοποίηση Χωρίς Περιορισμούς

Τα προβλήματα βελτιστοποίησης χωρίς περιορισμούς έχουν την ακόλουθη μορφή,

$$maximise\ f(\vec{x})$$

ή

$$minimise\ f(\vec{x})$$

πάνω σε όλες τις τιμές της μεταβλητής $\vec{x} = (x_1, x_2, \dots, x_n)$. Όπως έχουμε ήδη αναφέρει στο κεφάλαιο *Κλασικές Μέθοδοι Βελτιστοποίησης*, η αναγκαία συνθήκη για μια μερική λύση $\vec{x} = \vec{x}^*$ να είναι η βέλτιστη όταν $f(\vec{x})$ είναι μια διαφορίσιμη συνάρτηση είναι

$$\nabla f(\vec{x})\Big|_{\vec{x}\,=\,\vec{x}^*} = 0$$

Ωστόσο, για μη γραμμικές συναρτήσεις $f(\vec{x})$, αυτές οι εξισώσεις είναι επίσης μη γραμμικές και κατ' επέκταση είναι σχεδόν αδύνατον να λυθούν αναλυτικά για την εύρεση της λύσης. Στην παράγραφο αυτή, λοιπόν, θα παρουσιαστούν αλγοριθμικές διαδικασίες αναζήτησης για την εύρεση του $\vec{x}^*$, αρχικά για μονοδιάστατο χώρο και στη συνέχεια για $n > 1$. Επιπλέον, αυτές οι τεχνικές παίζουν σημαντικό ρόλο στην επίλυση αρκετών τύπων προβλημάτων, που θα παρουσιαστούν εκτενώς στη συνέχεια. Το γεγονός αυτό έγκειται στο ότι πολλοί αλγόριθμοι για προβλήματα με περιορισμούς έχουν σχεδιαστεί με σκοπό να εστιάζουν στην περίπτωση ενός προβλήματος χωρίς περιορισμούς κατά τη διάρκεια μιας επανάληψης.[62]

Αρχικά, λοιπόν, θα εξετάσουμε την περίπτωση της βελτιστοποίησης μιας μεταβλητής χωρίς περιορισμούς και μάλιστα, για λόγους ευκολίας, στις παρακάτω ενότητες του παρόντος κεφαλαίου, θα αναπτύξουμε τη θεωρία για μια διαφορίσιμη κοίλη συνάρτηση $f(x)$, η οποία πρέπει να μεγιστοποιηθεί.

Βελτιστοποίηση μιας μεταβλητής χωρίς περιορισμούς

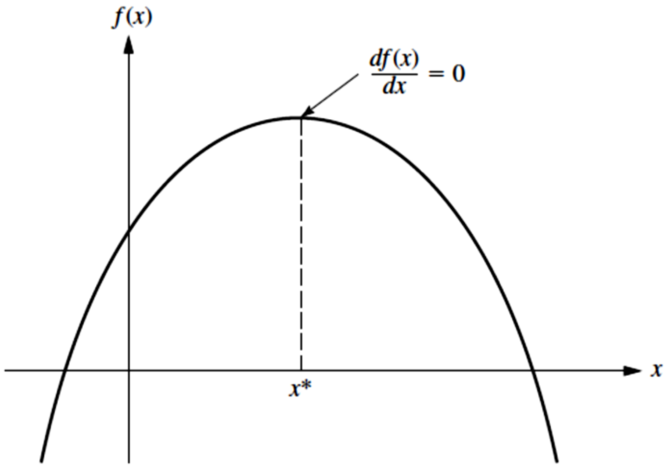
Η περίπτωση αυτή αποτελεί και την πιο εύκολη που μπορούμε να συναντήσουμε στη θεωρία του μη γραμμικού προγραμματισμού.

Επειδή η συνάρτηση f είναι κοίλη και διαφορίσιμη, η ικανή και αναγκαία συνθήκη για μια μερική λύση $x = x^*$ να είναι βέλτιστη (ολικό μέγιστο) είναι

$$\left.\frac{df}{dx}\right|_{x\,=\,x^*} = 0$$

όπως φαίνεται και στο Γράφημα 14. Αν η εξίσωση αυτή μπορεί να επιλυθεί απευθείας για x^*, η διαδικασία έχει ολοκληρωθεί. Ωστόσο, αν η συνάρτηση δεν είναι ιδιαίτερα απλή, ώστε η παράγωγος να μην είναι απλά μια γραμμική ή τετραγωνική συνάρτηση, ίσως να μην είναι εφικτή η επίλυσή της αναλυτικά. Στην περίπτωση αυτή, η *μονοδιάστατη διαδικασία αναζήτησης* παρέχει μια αριθμητική επίλυση του προβλήματος.

[62] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.665.



Γράφημα 14 – Πρόβλημα μιας μεταβλητής χωρίς περιορισμούς όταν η συνάρτηση είναι κοίλη.

Μονοδιάστατη Διαδικασία Αναζήτησης

Όπως άλλες διαδικασίες αναζήτησης στον μη γραμμικό προγραμματισμό, η *μονοδιάστατη διαδικασία αναζήτησης* βρίσκει μια ακολουθία δοκιμαστικών λύσεων, η οποία οδηγεί σε μια βέλτιστη λύση. Σε κάθε επανάληψη, ξεκινάμε από την τρέχουσα δοκιμαστική λύση για τη διεξαγωγή μιας συμμετρικής αναζήτησης, η οποία ολοκληρώνεται με την αναγνώριση μιας νέας βελτιωμένης δοκιμαστικής λύσης.

Η ιδέα στην οποία βασίζεται η μονοδιάστατη διαδικασία αναζήτησης είναι διαισθητική, συγκεκριμένα αν η κλίση (παράγωγος) είναι θετική ή αρνητική σε μια δοκιμαστική λύση φανερώνει αν η βελτίωση βρίσκεται άμεσα στα δεξιά ή αριστερά, αντιστοίχως. Επομένως, αν η εκτιμώμενη παράγωγος σε μια συγκεκριμένη τιμή του x είναι θετική, τότε το x^* πρέπει να είναι μεγαλύτερο από το x (βλέπε Γράφημα 14). Έτσι αυτό το x γίνεται ένα κάτω φράγμα στις δοκιμαστικές λύσεις που πρέπει να θεωρηθούν στη συνέχεια. Αναλόγως, αν η εκτιμώμενη παράγωγος σε μια συγκεκριμένη τιμή του x είναι αρνητική, τότε το x^* πρέπει να είναι μικρότερο από το x (βλέπε Γράφημα 14), έτσι αυτό το x γίνεται ένα άνω φράγμα στις δοκιμαστικές λύσεις που πρέπει να θεωρηθούν στη συνέχεια. Ακολούθως, αφού αναγνωριστούν τα άνω και κάτω όρια, κάθε νέα επιλεγόμενη δοκιμαστική λύση ανάμεσα στα τρέχοντα όρια παρέχει ένα νέο πιο αυστηρό άνω ή κάτω φράγμα. Ως αποτέλεσμα, καταλήγουμε σε στενότερο πεδίο αναζήτησης για τη βέλτιστη λύση. Εφόσον ένας λογικός κανόνας χρησιμοποιείται για την επιλογή κάθε δοκιμαστικής λύσης, η ακολουθία που προκύπτει από τις δοκιμαστικές λύσεις πρέπει να οδηγεί στο x^*.

Στην πράξη αυτό σημαίνει ότι η διαδικασία θα συνεχιστεί μέχρι το κενό μεταξύ του άνω και κάτω φράγματος να είναι αρκετά μικρό, έτσι ώστε η επόμενη δοκιμαστική λύση να είναι μέσα σε ένα προκαθορισμένο διάστημα ανοχής σφάλματος του x^*. Δηλαδή, στόχος είναι ο προσδιορισμός ενός σημείου όπου η παράγωγος στο σημείο αυτό να είναι (ουσιωδώς) μηδενική.

Η παραπάνω διαδικασία συνοψίζεται ως εξής:

Έστω

$$x' = \text{τρέχουσα δοκιμαστική λύση}$$
$$\overline{x} = \text{τρέχον άνω φράγμα}$$
$$\underline{x} = \text{τρέχον κάτω φράγμα}$$
$$\epsilon = \text{ανοχή σφάλματος για το } x^*$$

Αν και υπάρχουν πολλοί λογικοί κανόνες για την επιλογή κάθε καινούριας δοκιμαστικής λύσης, αυτός που χρησιμοποιήθηκε στην ακόλουθη διαδικασία είναι ο *κανόνας μέσου*[63], βάσει του οποίου επιλέγουμε το μέσο μεταξύ των δύο τρεχόντων φραγμάτων.

Αλγοριθμος

Αρχικοποίηση:

1. Επιλογή του ϵ.
2. Εύρεση των αρχικών $\overline{x}$ και $\underline{x}$ με έλεγχο (ή βρίσκοντας μια τυχαία τιμή του x, στην οποία η παράγωγος είναι θετική και μετά αρνητική).
3. Επιλογή μιας αρχικής δοκιμαστικής λύσης

$$x' = \frac{\overline{x} + \underline{x}}{2}.$$

Επανάληψη:

1. Υπολογισμός

[63] Παραδοσιακά ονομάζεται σχέδιο αναζήτησης Bolzano.

$$\left.\frac{df(x)}{dx}\right|_{x\,=\,x'}.$$

2. Αν

$$\left.\frac{df(x)}{dx}\right|_{x\,=\,x'} \geq 0$$

τότε επαναπροσδιορισμός

$$\underline{x} = x'.$$

3. Αν

$$\left.\frac{df(x)}{dx}\right|_{x\,=\,x'} \leq 0$$

τότε επαναπροσδιορισμός

$$\overline{x} = x'.$$

4. Επιλογή ενός καινούριου

$$x' = \frac{\overline{x} + \underline{x}}{2}.$$

Κανόνας τερματισμού ή διακοπής:

Αν

$$\overline{x} - \underline{x} \leq 2\epsilon$$

έτσι ώστε το καινούριο x' να είναι μέσα στο ϵ του x^*, σταματάμε. Αλλιώς, εκτελούμε ακόμα μια επανάληψη.[64]

Βελτιστοποίηση πολλών μεταβλητών χωρίς περιορισμούς

Στην ενότητα αυτή, θεωρούμε το πρόβλημα μεγιστοποίησης μιας κοίλης συνάρτησης $f(\vec{x})$ πολλών μεταβλητών $\vec{x} = (x_1, x_2, \dots, x_n)$ όταν δεν επιδέχονται περιορισμούς στις επιτρεπόμενες τιμές. Σε αντίθεση με την περίπτωση, όπου έχουμε μόνο μια μεταβλητή, στη συγκεκριμένη υπάρχουν αναρίθμητες πιθανές κατευθύνσεις στις οποίες μπορούμε να κινηθούμε. Κάθε κατεύθυνση αντιστοιχεί

---

[64] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.670 – 673.

στους αναλογικούς ρυθμούς με τους οποίους οι σχετικές μεταβλητές μπορούν να αλλάξουν. Στόχος είναι να προσδιοριστεί το σημείο, όπου όλες οι μερικές παράγωγοι είναι (ουσιωδώς) μηδενικές. Επομένως, η επέκταση της *μονοδιάστατης διαδικασίας αναζήτησης* απαιτεί τη χρήση των τιμών των μερικών παραγώγων για να επιλεγεί μια συγκεκριμένη κατεύθυνση, στην οποία πρέπει να κινηθούμε. Αυτή η επιλογή εμπλέκει τη χρήση της κλίσης[65] της αντικειμενικής συνάρτησης, όπως περιγράφεται ακολούθως.

Επειδή η αντικειμενική συνάρτηση $f(\vec{x})$ θεωρείται διαφορίσιμη, ορίζεται το ανάδελτά της, $\nabla f(\vec{x})$, σε κάθε σημείο $\vec{x}$. Συγκεκριμένα, το ανάδελτα σε ένα συγκεκριμένο σημείο $\vec{x} = \vec{x}'$ είναι ένα διάνυσμα, του οποίου τα στοιχεία είναι οι αντίστοιχες μερικές παράγωγοι εκτιμώμενες στο σημείο $\vec{x} = \vec{x}'$, έτσι ώστε

$$\nabla f(\vec{x})\bigg|_{\vec{x}\,=\,\vec{x}'} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right).$$

Η σπουδαιότητα του ανάδελτα είναι ότι η (απειροελάχιστη) αλλαγή στο $\vec{x}$, η οποία μεγιστοποιεί το ρυθμό με τον οποίο η $f(\vec{x})$ αυξάνεται, είναι η αλλαγή που είναι ανάλογη με το $\nabla f(\vec{x})$. Για να εκφράσουμε αυτή την ιδέα γεωμετρικά, η ‘κατεύθυνση’ της κλίσης $\nabla f(\vec{x}')$ ερμηνεύεται ως η κατεύθυνση του κατευθυνόμενου ευθυγράμμου τμήματος (βέλος) από την αρχή $(0, 0, \dots, 0)$ στο σημείο $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}\right)$, οπού $\frac{\partial f}{\partial x_j}$ εκτιμάται στο $x_j = x_j'$. Επομένως, μπορεί να ειπωθεί ότι ο ρυθμός με τον οποίο αυξάνεται η $f(\vec{x})$, μεγιστοποιείται αν (απειροελάχιστες) αλλαγές στο $\vec{x}$ είναι στην κατεύθυνση του ανάδελτα $\nabla f(\vec{x})$. Επειδή ο στόχος είναι να βρεθεί μια εφικτή λύση η οποία να μεγιστοποιεί την $f(\vec{x})$, σκοπός μας είναι να επιχειρήσουμε να κινηθούμε στην κατεύθυνση του ανάδελτα όσο το δυνατόν περισσότερο.

<u>Διαδικασία Αναζήτησης Κλίσης</u>

Το γεγονός ότι το συγκεκριμένο πρόβλημα δεν επιδέχεται περιορισμούς, σε συνδυασμό με τη χρήση του ανάδελτα, υποδηλώνει ότι μια αποτελεσματική

[65] Κλίση ή ανάδελτα (gradient) της συνάρτησης $f(\vec{x}) : \nabla f(\vec{x})$.

διαδικασία αναζήτησης θα πρέπει να κινείται συνεχώς προς την κατεύθυνση της κλίσης μέχρι να φτάσει (ουσιωδώς) μια βέλτιστη λύση $\vec{x}^*$, όπου $\nabla f(\vec{x}^*) = 0$. Ωστόσο, μια συνεχόμενη εναλλαγή των $\vec{x}$ προς την κατεύθυνση της κλίσης, $\nabla f(\vec{x})$, δεν αποτελεί μια πρακτική επίλυση του προβλήματος, καθώς οι επαναλαμβανόμενες αλλαγές απαιτούν τον συνεχή υπολογισμό των ποσοτήτων $\frac{\partial f}{\partial x_j}$ και κατ' επέκταση την αλλαγή κατεύθυνσης του μονοπατιού. Επομένως, μια καλύτερη και αποτελεσματικότερη προσέγγιση είναι να συνεχίσουμε να κινούμαστε προς μια σταθερή κατεύθυνση από την τρέχουσα δοκιμαστική λύση, μέχρι να σταματήσει η $f(\vec{x})$ να αυξάνεται. Αυτό το σημείο διακοπής θα γίνει η επόμενη δοκιμαστική λύση, υπολογίζοντας ξανά το ανάδελτα, ώστε να αποφασιστεί η καινούργια κατεύθυνση στην οποία θα κινηθούμε. Με αυτή την προσέγγιση, κάθε επανάληψη επαναπροσδιορίζει την τρέχουσα δοκιμαστική λύση $\vec{x}'$, ως εξής:

Επαναπροσδιορισμός $\vec{x}' = \vec{x}' + t^*\nabla f(\vec{x}')$,

όπου t^* είναι η θετική τιμή του t, η οποία μεγιστοποιεί την $f\big(\vec{x}' + t\nabla f(\vec{x}')\big)$, δηλαδή,

$$f\big(\vec{x}' + t^*\nabla f(\vec{x}')\big) = \max_{t\geq 0} f\big(\vec{x}' + t\nabla f(\vec{x}')\big).$$

Να σημειωθεί ότι $f\big(\vec{x}' + t\nabla f(\vec{x}')\big)$ είναι απλώς η $f(\vec{x})$, όπου

$$x_j = x_j' + t\left(\frac{\partial f}{\partial x_j}\right)_{\vec{x}=\vec{x}'} \text{, για } j = 1,2, \dots, n$$

και επειδή οι εκφράσεις για τα x_j περιέχουν μόνο σταθερές και τη μεταβλητή t, η $f(x)$ γίνεται συνάρτηση της μιας μόνο μεταβλητής t.

Οι επαναλήψεις της *διαδικασίας αναζήτησης κλίσης* συνεχίζονται μέχρι $\nabla f(\vec{x}) = 0$ εντός μιας μικρής ανοχής ϵ, δηλαδή, μέχρι

$$\left|\frac{\partial f}{\partial x_j}\right| \leq \epsilon \text{, για } j = 1,2, \dots, n$$ [66]

[66] Αυτός ο κανόνας διακοπής γενικώς δίνει μια λύση $\vec{x}$, η οποία βρίσκεται κοντά στη βέλτιστη λύση $\vec{x}^*$ και με την τιμή της $f(\vec{x})$ να είναι πολύ κοντά στην $f(\vec{x}^*)$. Ωστόσο, αυτό δεν μπορεί να εγγυηθεί αφού είναι δυνατόν η συνάρτηση να διατηρεί μια πολύ μικρή θετική κλίση ($\leq \epsilon$) πάνω σε μια σημαντική απόσταση από το $\vec{x}$ στο $\vec{x}^*$.

Το δυσκολότερο μέρος της διαδικασίας αναζήτησης κλίσης συνήθως είναι η εύρεση του t^*, δηλαδή την τιμή της μεταβλητής t η οποία μεγιστοποιεί την fστην κατεύθυνση της κλίσης, σε κάθε επανάληψη. Επίσης, επειδή τα $\vec{x}$ και $\nabla f(\vec{x})$ έχουν σταθερές τιμές για τη μεγιστοποίηση και επειδή η f είναι κοίλη, το παρόν πρόβλημα μπορεί να θεωρηθεί ως πρόβλημα μεγιστοποίησης μιας κοίλης συνάρτησης μιας μεταβλητής t. Επομένως, μπορεί να επιλυθεί με βάση τη *μονοδιάστατη διαδικασία αναζήτησης* της προηγούμενης παραγράφου, όπου το αρχικό κάτω φράγμα στο t πρέπει να είναι μη αρνητικό λόγω του περιορισμού $t \geq 0$. Εναλλακτικώς, αν η f είναι μια απλή συνάρτηση, τότε μπορεί να προσδιοριστεί η αναλυτική λύση θέτοντας την παράγωγο ως προς t ίση με το μηδέν και να επιλυθεί.

Η παραπάνω διαδικασία συνοψίζεται ως εξής:

Έστω

$x' =$ τρέχουσα δοκιμαστική λύση

$t^* =$ η θετική τιμή του t, η οποία μεγιστοποιεί την $f\left(\vec{x}' + t\nabla f(\vec{x}')\right)$

$\epsilon =$ ανοχή σφάλματος για το x^*

Αλγοριθμος

Αρχικοποίηση:

1. Επιλογή του ϵ.
2. Επιλογή μιας αρχικής δοκιμαστικής λύσης x'.
3. Μετάβαση στον κανόνα διακοπής.

Επανάληψη:

1. Εκφράζουμε την $f\left(\vec{x}' + t\nabla f(\vec{x}')\right)$ ως συνάρτηση του t θέτοντας
$$x_j = x_j' + t\left(\frac{\partial f}{\partial x_j}\right)_{\vec{x}=\vec{x}'}, \text{για } j = 1,2,\ldots,n.$$
και υποκαθιστούμε τις εκφράσεις αυτές στην $f(\vec{x})$.
2. Χρήση της μονοδιάστατης διαδικασίας αναζήτησης (ή υπολογιστικά) για την εύρεση του $t = t^*$, το οποίο μεγιστοποιεί την ποσότητα $f\left(\vec{x}' + t\nabla f(\vec{x}')\right)$ με $t \geq 0$.

3. Επαναπροσδιορισμός

$$\vec{x}' = \vec{x}' + t\nabla f(\vec{x}').$$

4. Μετάβαση στον κανόνα διακοπής.

Κανόνας τερματισμού ή διακοπής:

Υπολογισμός του $\nabla f(\vec{x}')$ στο $\vec{x} = \vec{x}'$. Ελέγχουμε αν

$$\left|\frac{\partial f}{\partial x_j}\right| \leq \epsilon \text{ , για } j = 1,2, \dots, n$$

Αν ισχύει, το καινούριο x' αποτελεί την επιθυμητή προσέγγιση μιας βέλτιστης λύσης x^* και σταματάμε. Αλλιώς, εκτελούμε ακόμα μια επανάληψη.[67]

Βελτιστοποίηση με Γραμμικούς Περιορισμούς

Τα προβλήματα τέτοιας μορφής χαρακτηρίζονται από περιορισμούς, οι οποίοι ταιριάζουν σε γραμμικό προγραμματισμό. Δηλαδή, όλοι οι συναρτησιακοί περιορισμοί $g_i(\vec{x})$ είναι γραμμικοί αλλά η αντικειμενική συνάρτηση $f(\vec{x})$ είναι μη γραμμική. Το πρόβλημα απλουστεύεται σε σημαντικό βαθμό, καθώς υπάρχει μια μόνο μη γραμμική συνάρτηση που πρέπει να ληφθεί υπόψη σε συνδυασμό με το πεδίο εφικτών λύσεων του γραμμικού προγραμματισμού που προκύπτει από τους γραμμικούς περιορισμούς. Τέλος για την επίλυση τέτοιων προβλημάτων έχουν αναπτυχθεί πολλοί αλγόριθμοι, οι οποίοι επεκτείνουν τη μέθοδο simplex για μη γραμμική αντικειμενική συνάρτηση.[68] Ένας τέτοιος αλγόριθμος επίλυσης είναι και ο ακόλουθος.

Συνθήκες Karush – Kuhn – Tucker (KKT) για βελτιστοποίηση με περιορισμούς

Στην παρούσα ενότητα θα αναπτυχθεί η θεωρία που σχετίζεται με την βελτιστοποίηση όταν υπάρχουν περιορισμοί. Οι περιορισμοί για μια γενική

[67] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.673 – 676.

[68] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.665.

περίπτωση ονομάζονται ***Karush – Kuhn – Tucker συνθήκες*** (ή ΚΚΤ συνθήκες), επειδή εμπνεύστηκαν ανεξάρτητα από τον William Karush[69] (1917 – 1997) και από τους Harold William Kuhn (1925) και Albert William Tucker[70] (1905 – 1995). Το βασικό τους αποτέλεσμα ενσωματώνεται στο ακόλουθο θεώρημα.

ΘΕΩΡΗΜΑ 3.1[71]

Θεωρούμε ότι οι $f(\vec{x}), g_1(\vec{x}), g_2(\vec{x}), \dots, g_m(\vec{x})$ είναι διαφορίσιμες συναρτήσεις, οι οποίες ικανοποιούν συγκεκριμένες συνθήκες κανονικότητας. Τότε το

$$\vec{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$$

μπορεί να είναι μια βέλτιστη λύση για ένα πρόβλημα μη γραμμικού προγραμματισμού μόνο αν υπάρχουν m αριθμοί $u_1, u_2, \dots, u_m$ τέτοιοι ώστε όλες οι ακόλουθες ΚΚΤ συνθήκες ικανοποιούνται:

$$\left.\begin{array}{ll} 1. & \frac{\partial f}{\partial x_j} - \sum_{i=1}^m u_i \frac{\partial g_i}{\partial x_j} \le 0 \\ 2. & x_j^* \left(\frac{\partial f}{\partial x_j} - \sum_{i=1}^m u_i \frac{\partial g_i}{\partial x_j}\right) = 0 \end{array}\right\} \quad \text{στο } \vec{x} = \vec{x}^*, \text{για } j = 1, 2, \dots, n.$$

$$\left.\begin{array}{ll} 3. & g_i(\vec{x}^*) - b_i \le 0 \\ 4. & u_i[g_i(\vec{x}^*) - b_i] = 0 \end{array}\right\} \quad \text{για } i = 1, 2, \dots, m.$$

$$5. \quad x_j^* \ge 0 \qquad \text{για } j = 1, 2, \dots, n.$$

$$6. \quad u_i \ge 0 \qquad \text{για } i = 1, 2, \dots, m.$$

Παρατηρούμε ότι οι συνθήκες 2 και 4 απαιτούν το γινόμενο των δύο ποσοτήτων να ισούται με μηδέν. Επομένως, κάθε μια από αυτές τις συνθήκες δηλώνουν ότι τουλάχιστον μια από τις δύο ποσότητες πρέπει να είναι μηδέν. Ακολούθως, η συνθήκη 4 μπορεί να συνδυαστεί με τη συνθήκη 3, εκφράζοντας τις συνθήκες αυτές με μια ισοδύναμη μορφή, όπως

$$(3,4) \quad g_i(\vec{x}^*) - b_i = 0$$

$$(\text{ή} \le 0 \ \text{αν } u_i = 0), \text{για } i = 1, 2, \dots, m.$$

[69] W. Karush , *Minima of Functions of Several Variables with Inequalities as Side Conditions*, M.S. thesis, Department of Mathematics, University of Chicago 1939.
[70] H.W. Kuhn & A.W. Tucker, *Nonlinear Programming,* Berkeley 1951, σελ. 481 – 492.
[71] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.679.

Ομοίως, η συνθήκη 2 μπορεί να συνδυαστεί με τη συνθήκη 1 ως

$$(1,2) \quad \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} = 0$$

$$(\text{ή} \leq 0 \ \text{αν}\ x_j^* = 0), \text{για}\ j = 1, 2, \dots, n.$$

Αξίζει να σημειωθεί ότι για $m = 0$, το παραπάνω άθροισμα δεν υφίσταται και επομένως η συνθήκη (1 ,2) παίρνει τη γνωστή μορφή περιορισμού όπως έχουμε ήδη περιγράψει στην περίπτωση που το πρόβλημα μη γραμμικού προγραμματισμού επιδέχεται μόνο περιορισμούς μη αρνητικότητας (όπως φαίνεται και στον Πίνακα 2).

Στις συνθήκες 1, 2, 4 και 6 οι αριθμοί u_i εμφανίζονται στη μαθηματική παραγώγιση ως πολλαπλασιαστές Lagrange[72]. Επίσης, οι συνθήκες 3 και 5 εξασφαλίζουν ότι η λύση είναι εφικτή. Οι υπόλοιπες συνθήκες αποκλείουν τις περισσότερες εφικτές λύσεις να είναι η βέλτιστη.

Πρόβλημα	Αναγκαίες Συνθήκες για Βελτιστότητα	Επίσης Ικανή αν:
Μιας μεταβλητής χωρίς περιορισμούς	$\frac{df}{dx} = 0$	$f(x)$ κοίλη
Πολλών μεταβλητών χωρίς περιορισμούς	$\frac{\partial f}{\partial x_j} = 0 \quad (j = 1, 2, \dots, n)$	$f(\mathbf{x})$ κοίλη
Μόνο με περιορισμούς μη αρνητικότητας	$\frac{\partial f}{\partial x_j} = 0 \quad (j = 1, 2, \dots, n)$ (or ≤ 0 if $x_j = 0$)	$f(\mathbf{x})$ κοίλη
Γενικό πρόβλημα με περιορισμούς	Karush-Kuhn-Tucker συνθήκες	$f(\mathbf{x})$ κοίλη και $g_i(\mathbf{x})$ κυρτή $(i = 1, 2, \dots, m)$

Πίνακας 2 - Αναγκαίες και Ικανές Συνθήκες για Βελτιστότητα.

Ωστόσο, η ικανοποίηση αυτών των συνθηκών δεν εξασφαλίζει ότι η λύση είναι βέλτιστη. Όπως φαίνεται και στον Πίνακα 2, οι υποθέσεις κυρτότητας είναι υποχρεωτικές για την εξασφάλιση της βέλτιστης λύσης. Οι υποθέσεις αυτές παρουσιάζονται παρακάτω ως επέκταση του θεωρήματος.

Πορισμα 3.1

Υποθέτουμε ότι η $f(\vec{x})$ είναι μια κοίλη συνάρτηση και οι $g_1(\vec{x}), g_2(\vec{x}), \dots, g_m(\vec{x})$ είναι κυρτές συναρτήσεις (δηλαδή το πρόβλημα είναι πρόβλημα κυρτού προγραμματισμού), όπου όλες οι συναρτήσεις ικανοποιούν τις συνθήκες

[72] Βλέπε ενότητα *Δεσμευμένη Βελτιστοποίηση με Περιορισμούς Ισότητας*, σελ. 28.

κανονικότητας. Τότε το $\vec{x}^* = (\vec{x}_1^*,\ \vec{x}_2^*, \dots, \vec{x}_n^*)$ είναι μια βέλτιστη λύση αν και μόνο αν όλες οι συνθήκες του θεωρήματος 3.1 ικανοποιούνται.

Επιπλέον, για πιο περίπλοκα προβλήματα ίσως είναι δύσκολο αν όχι αδύνατον να βρεθεί μια βέλτιστη λύση κατ' ευθείαν από τις ΚΚΤ συνθήκες. Ωστόσο, αυτές οι συνθήκες παρέχουν πολύτιμα συμπεράσματα για την εύρεση μιας βέλτιστης λύσης, ενώ επίσης μας επιτρέπουν να ελέγξουμε αν μια προτεινόμενη λύση μπορεί να είναι βέλτιστη.

Τέλος, υπάρχουν αρκετές χρήσιμες έμμεσες εφαρμογές των ΚΚΤ συνθηκών. Μια απ' αυτές αποτελεί η θεωρία δυισμού, η οποία έχει αναπτυχθεί για το μη γραμμικό προγραμματισμό.

Τετραγωνικός Προγραμματισμός

Τα προβλήματα τετραγωνικού προγραμματισμού έχουν γραμμικούς περιορισμούς αλλά η αντικειμενική συνάρτηση $f(\vec{x})$ είναι τετραγωνική. Επομένως, η μόνη διαφορά μεταξύ του γραμμικού και του τετραγωνικού προγραμματισμού είναι ότι κάποιοι όροι της αντικειμενικής συνάρτησης περιλαμβάνουν μεταβλητές στο τετράγωνο x_j^2 ή γινόμενο δύο μεταβλητών $x_i x_j$ $(i \neq j)$.

Πολλοί αλγόριθμοι έχουν αναπτυχθεί για αυτήν την περίπτωση υπό την επιπλέον υπόθεση ότι η συνάρτηση f είναι μια κοίλη[73] συνάρτηση. Ένας τέτοιος αλγόριθμος, ο οποίος περιλαμβάνει μια άμεση επέκταση της μεθόδου simplex είναι αυτός που θα παρουσιαστεί στη συνέχεια.

Το πρόβλημα του τετραγωνικού προγραμματισμού συνίσταται στην εύρεση του $\vec{x}$, ώστε να μεγιστοποιεί την f υπό συγκεκριμένους περιορισμούς. Δηλαδή,

$$maximize\ f(\vec{x}\,) = \vec{c}\vec{x} - \frac{1}{2}\vec{x}^T Q\vec{x},$$

[73] Ένας τρόπος για να εξετάσουμε αν η αντικειμενική συνάρτηση είναι κοίλη είναι να εξετάσουμε αν $\vec{x}^T Q\vec{x} \geq 0$ για όλα τα $\vec{x}$, δηλαδή, ο Q να είναι θετικά ημιορισμένος πίνακας. Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.684.

υπό τους περιορισμούς

$$A\vec{x} \leq \vec{b}$$

και

$$\vec{x} \geq 0$$

όπου $\vec{c}$ είναι ένα διάνυσμα γραμμή, $\vec{x}$ και $\vec{b}$ είναι διανύσματα στήλη, Q και A είναι πίνακες και το σύμβολο T δηλώνει τον ανάστροφο. Επίσης, για τα στοιχεία του Q ισχύει $q_{ij} = q_{ji}$.

Εκφράζοντας τη συνάρτηση f με εναλλακτικό τρόπο χρησιμοποιώντας τα στοιχεία των πινάκων και των διανυσμάτων έχουμε

$$f(\vec{x}\,) = \vec{c}\vec{x} - \frac{1}{2}\vec{x}^T Q\vec{x} = \sum_{j=1}^{n} c_j x_j - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} q_{ij}x_i x_j.$$

Στη συνέχεια, θα παρουσιάσουμε την τροποποιημένη μέθοδο simplex. Το κλειδί σε αυτή τη μέθοδο είναι η κατασκευή των ΚΚΤ συνθηκών της προηγούμενης ενότητας και ακολούθως η αναδιατύπωση τους σε μια βολική μορφή, η οποία δεν θα διαφέρει πολύ από αυτή του γραμμικού προγραμματισμού.

Οι ΚΚΤ συνθήκες για τον τετραγωνικό προγραμματισμό

Για κάθε πρόβλημα τετραγωνικού προγραμματισμού οι ΚΚΤ συνθήκες μπορούν να περιοριστούν στην ακόλουθη βολική μορφή, περιλαμβάνοντας μόνο περιορισμούς γραμμικού προγραμματισμού και έναν περιορισμό συμπληρωματικότητας

$$Q\vec{x} + A^T\vec{u} - \vec{y} = \vec{c}^T,$$

$$A\vec{x} + \vec{v} = \vec{b},$$

$$\vec{x} \geq 0\,, \vec{u} \geq 0\,, \vec{y} \geq 0\,, \vec{v} \geq 0$$

$$\vec{x}^T\vec{y} + \vec{u}^T\vec{v} = 0,$$

όπου τα στοιχεία του διανύσματος στήλης $\vec{u}$ είναι τα u_i της προηγούμενης ενότητας και τα στοιχεία του διανύσματος γραμμής $\vec{y}$ και $\vec{v}$ είναι χαλαρές μεταβλητές.

Επειδή η αντικειμενική συνάρτηση του αρχικού προβλήματος έχει θεωρηθεί ότι είναι κοίλη και επειδή οι περιορισμοί είναι γραμμικοί και κατ' επέκταση κυρτοί, μπορεί να εφαρμοστεί το θεώρημα 3.1. Επομένως, το $\vec{x}$ είναι βέλτιστο αν και μόνο αν υπάρχουν τιμές των $\vec{y}$, $\vec{v}$ και $\vec{u}$ τέτοιες ώστε τα τέσσερα διανύσματα να ικανοποιούν όλες αυτές τις συνθήκες. Το αρχικό πρόβλημα έχει επομένως περιοριστεί στο ισοδύναμο πρόβλημα της εύρεσης μιας εφικτής λύσης για τους συγκεκριμένους περιορισμούς.

Η τροποποιημένη μέθοδος simplex

Η τροποποιημένη μέθοδος simplex εκμεταλλεύεται το γεγονός ότι οι ΚΚΤ συνθήκες στη βολική μορφή που παρουσιάστηκαν παραπάνω, εξαιρουμένου του περιορισμού συμπληρωματικότητας, αποτελούν περιορισμούς γραμμικού προγραμματισμού. Επιπλέον, ο περιορισμός συμπληρωματικότητας δηλώνει ότι δεν επιτρέπονται και οι δύο συμπληρωματικές μεταβλητές για κάθε ζεύγος να είναι βασικές μεταβλητές[74] στις βασικές εφικτές λύσεις. Επομένως, το πρόβλημα περιορίζεται στην εύρεση μιας αρχικής βασική λύσης σε οποιοδήποτε πρόβλημα γραμμικού προγραμματισμού το οποίο επιδέχεται αυτούς τους περιορισμούς, υπό τον επιπλέον περιορισμό στην ταυτοποίηση των βασικών μεταβλητών. (Αυτή η αρχική βασική εφικτή λύση ίσως είναι η μόνη εφικτή λύση σε αυτή την περίπτωση.)

Όπως είναι γνωστό, η εύρεση μιας αρχικής εφικτής λύσης είναι σχετικά μια απλή και ευθύς διαδικασία. Στην απλή περίπτωση όπου $\vec{c}^T \leq 0$ και $\vec{b} \geq 0$, οι αρχικές βασικές μεταβλητές είναι τα στοιχεία των διανυσμάτων $\vec{y}$ και $\vec{v}$, έτσι ώστε η επιθυμητή λύση να είναι $\vec{x} = 0, \vec{u} = 0, \vec{y} = -\vec{c}^T, \vec{v} = \vec{b}$. Διαφορετικά, πρέπει να αναθεωρήσουμε το πρόβλημα εισάγοντας μια τεχνητή μεταβλητή σε κάθε εξίσωση όπου $c_j > 0$ ή $b_i > 0$, ώστε να χρησιμοποιηθούν αυτές οι τεχνητές

[74] Δηλαδή η λύση να μην είναι εκφυλισμένη.

μεταβλητές ως αρχικές βασικές μεταβλητές για το αναθεωρημένο πρόβλημα. Αξίζει να σημειωθεί ότι αυτή η επιλογή των αρχικών βασικών μεταβλητών ικανοποιεί τον περιορισμό συμπληρωματικότητας, διότι ως μη βασικές μεταβλητές $\vec{x} = 0$ και $\vec{u} = 0$ αυτομάτως.

Στη συνέχεια, εφαρμόζουμε την παρακάτω μέθοδο για την εύρεση μιας βασικής εφικτής λύσης για το πραγματικό πρόβλημα,

χρησιμοποιώντας τη μέθοδο simplex (με κάποια τροποποίηση) για το ακόλουθο πρόβλημα γραμμικού προγραμματισμού,

$$minimize\ Z = \Sigma\ τεχνητών\ μεταβλητών,$$

υπό τους περιορισμούς του γραμμικού προγραμματισμού τους οποίους πήραμε από τις ΚΚΤ συνθήκες περιλαμβανομένων των τεχνητών μεταβλητών.

Η τροποποίηση στη μέθοδο simplex είναι η ακόλουθη:

<u>Κανόνας Περιορισμένης – Εισόδου</u>: όταν διαλέγουμε ένα βασικό διάνυσμα εισόδου, εξαιρούμε από την εξέταση κάθε μη βασική μεταβλητή της οποίας η συμπληρωματική μεταβλητή είναι ήδη μια βασική μεταβλητή. Η επιλογή πρέπει να γίνει από τις υπόλοιπες μη βασικές μεταβλητές σύμφωνα με το συνηθισμένο κριτήριο για τη μέθοδο simplex.

Αυτός ο κανόνας εξασφαλίζει την ικανοποίηση του περιορισμού συμπληρωματικότητας στον αλγόριθμο. Όταν μια βέλτιστη λύση

$$\vec{x}^*, \vec{u}^*, \vec{y}^*, \vec{v}^* \text{ με όλες τις τεχνητές μεταβλητές ίσες με μηδέν,}$$

εξασφαλιστεί για την πρώτη φάση του προβλήματος, το $\vec{x}^*$ είναι μια επιθυμητή βέλτιστη λύση για το αρχικό τετραγωνικό πρόβλημα προγραμματισμού.[75]

[75] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.683 – 690.

Κυρτός Προγραμματισμός

Ο κυρτός προγραμματισμός καλύπτει μια ευρεία κατηγορία προβλημάτων, η οποία συμπεριλαμβάνει ως ειδικές περιπτώσεις όλους τους προηγούμενους τύπους όταν η $f(\vec{x})$ είναι μια κοίλη συνάρτηση. Οι υποθέσεις είναι ότι,

- η $f(\vec{x})$ είναι μια κοίλη συνάρτηση.
- Κάθε $g_i(\vec{x})$ είναι μια κυρτή συνάρτηση.

Όπως έχει ήδη αναφερθεί σε προηγούμενες ενότητες, αυτές οι υποθέσεις είναι αρκετές για να εξασφαλίσουν ότι ένα τοπικό μέγιστο είναι ένα ολικό μέγιστο. [76]

Ωστόσο, δεν υπάρχει ένας συγκεκριμένος αλγόριθμος για όλες τις περιπτώσεις επίλυσης προβλημάτων κυρτού προγραμματισμού. Πολλοί διαφορετικοί αλγόριθμοι έχουν αναπτυχθεί, καθένας με τα πλεονεκτήματα και τα μειονεκτήματά τους, οι οποίοι στις περισσότερες των περιπτώσεων υπόκεινται σε μία από τις παρακάτω κατηγορίες.

Αλγόριθμοι κλίσης

Στην κατηγορία αυτή χρησιμοποιείται η διαδικασία αναζήτησης κλίσης, η οποία τροποποιείται μερικώς για τη διατήρηση του μονοπατιού αναζήτησης μέσα στο σύνορο των περιορισμών. Μια τέτοια διαδεδομένη μέθοδο κλίσης αποτελεί και η *γενικευμένη μειωμένη κλίση* (GRC), της οποίας ο αλγόριθμος παρουσιάζεται αναλυτικά.[77]

Γενικευμένη μειωμένη κλίση

Αλγοριθμος[78]

1. Δοθέντος ενός σημείου $\vec{x}_1 \in X$, $\epsilon \geq 0$, $\bar{\epsilon} > 0$, θετικού ακεραίου M και για $k := 1$.
2. Υπολογίζουμε

[76] [76] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ.667.
[77] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 697.
[78] Wenyu Sun & Ya-Xiang Yuan, *Optimization Theory and Methods - Nonlinear Programming*, NY 2006, σελ. 510.

$$\nabla c^T(\vec{x}_k) = \begin{bmatrix} A_B \\ A_N \end{bmatrix},$$

όπου η διαμέριση ικανοποιεί ότι ο $A_B \in \Re^{m\times m}$ δεν είναι ιδιάζων πίνακας.

Υπολογίζουμε το λ από τη σχέση,

$$\frac{\partial f(\vec{x})}{\partial \vec{x}_B} = \frac{\partial c^T(\vec{x})}{\partial \vec{x}_B}\lambda$$

και το $\tilde{g}_k$ από τη σχέση,

$$\tilde{g}(\vec{x}_N) = \frac{\partial}{\partial \vec{x}_N}[f(\vec{x}) - \lambda^T c(\vec{x})]$$

3. Αν $\|\tilde{g}_k\| \le \epsilon$, τότε σταματάμε.

   Θέτω $\bar{d}_k = -\tilde{g}_k$ και $\alpha = \alpha_k^{(0)} > 0$.
4. $\vec{x}_N = (\vec{x}_k)_N + \alpha \bar{d}_k$

   $\vec{x}_B = (\vec{x}_k)_B$, $j \coloneqq 0$.
5. $\vec{x}_B = \vec{x}_B - A_B^{-T} c(\vec{x}_B, \vec{x}_N)$.

   Υπολογίζουμε $c(\vec{x}_B, \vec{x}_N)$.

   Αν $\|c(\vec{x}_B, \vec{x}_N)\| \le \bar{\epsilon}$, τότε πηγαίνουμε στο βήμα 7.

   $j \coloneqq j + 1$.

   Αν $j < M$, πηγαίνουμε στο βήμα 5.
6. $\alpha := \alpha/2$, πηγαίνουμε στο βήμα 4.
7. Αν $f(\vec{x}_B, \vec{x}_N) \ge f(\vec{x}_K)$, τότε πηγαίνουμε στο βήμα 6.

   $\vec{x}_{k+1} = (\vec{x}_B, \vec{x}_N)$, $k \coloneqq k + 1$, πηγαίνουμε στο βήμα 2.

Ακολουθιακοί αλγόριθμοι χωρίς περιορισμούς

Η δεύτερη κατηγορία περιέχει μια συνάρτηση ποινής, καθώς και μεθόδους με συναρτήσεις φράγματα. Αυτοί οι αλγόριθμοι μετατρέπουν το αρχικό πρόβλημα βελτιστοποίησης με περιορισμούς σε μια ακολουθία προβλημάτων βελτιστοποίησης χωρίς περιορισμούς, των οποίων οι βέλτιστες λύσεις συγκλίνουν στη βέλτιστη λύση του αρχικού προβλήματος. Κάθε ένα από αυτά τα προβλήματα βελτιστοποίησης χωρίς περιορισμούς μπορούν να επιλυθούν με τη διαδικασία αναζήτησης κλίσης. Αυτή η τροποποίηση επιτυγχάνεται ενσωματώνοντας τους περιορισμούς σε μια συνάρτηση ποινής (ή συνάρτηση φράγμα), η οποία είναι αφαιρούμενη από την αντικειμενική συνάρτηση, ώστε να επιβληθούν μεγάλες ποινές για παραβιάσεις περιορισμών.[79]

Ακολουθιακοί – προσεγγιστικοί αλγόριθμοι

Η τρίτη κατηγορία περιλαμβάνει μεθόδους γραμμικής προσέγγισης και τετραγωνικής προσέγγισης. Αυτοί οι αλγόριθμοι αντικαθιστούν τη γραμμική αντικειμενική συνάρτηση με μια διαδοχή από γραμμικές ή τετραγωνικές προσεγγίσεις. Για προβλήματα βελτιστοποίησης με γραμμικούς περιορισμούς, αυτές οι προσεγγίσεις επιτρέπουν επαναλαμβανόμενη εφαρμογή των αλγορίθμων γραμμικού ή τετραγωνικού προγραμματισμού.

Ένας τέτοιος αλγόριθμος είναι και ο Frank – Wolfe αλγόριθμος για την περίπτωση κυρτού προγραμματισμού με γραμμικούς περιορισμούς, τον οποίο θα παρουσιάσουμε στη συνέχεια.

Frank – Wolfe αλγόριθμος

Αλγοριθμος

Αρχικοποίηση:

Επιλογή μιας εφικτής αρχικής δοκιμαστικής λύσης $\vec{x}^{(0)}$, εφαρμόζοντας διαδικασίες γραμμικού προγραμματισμού για την εύρεση μια αρχικής βασικής εφικτής λύσης.

Θέτω $k = 1$.

Επανάληψη:

1. Για $j = 1, 2, \dots, n$, υπολογίζουμε

[79] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 697.

$$\left.\frac{\partial f(\vec{x})}{\partial x_j}\right|_{\vec{x}\,=\,\vec{x}^{(k-1)}}$$

Θέτουμε το c_j ίσο με αυτή την τιμή.

2. Εύρεση μιας βέλτιστης λύσης $\vec{x}_{LP}^{(k)}$ για το ακόλουθο πρόβλημα γραμμικού προγραμματισμού.

$$maximize \quad g(\vec{x}) = \sum_{j=1}^{n} c_j x_j$$

υπό τους περιορισμούς

$$A\vec{x} \leq \vec{b} \quad \text{και} \quad \vec{x} \geq 0$$

3. Για τη μεταβλητή t $(0 \leq t \leq 1)$, θέτω

$h(t) = f(\vec{x})$ για $\vec{x} = \vec{x}^{(k-1)} + t\left(\vec{x}_{LP}^{(k)} - \vec{x}^{(k-1)}\right)$,

έτσι ώστε η $h(t)$ να δίνει την τιμή της $f(\vec{x})$ στο ευθύγραμμο τμήμα μεταξύ των $\vec{x}^{(k-1)}$ (για $t = 0$) και $\vec{x}_{LP}^{(k)}$ (για $t = 1$). Χρήση κάποιας διαδικασίας όπως αυτή της μονοδιάστατης διαδικασίας αναζήτησης για την βελτιστοποίηση της $h(t)$ στο $0 \leq t \leq 1$, και θέτουμε $\vec{x}^{(k)}$ ίσο με το αντίστοιχο $\vec{x}$.

Μετάβαση στον κανόνα διακοπής.

Κανόνας τερματισμού ή διακοπής:

Αν $\vec{x}^{(k-1)}$ και $\vec{x}^{(k)}$ είναι ικανοποιητικά κοντά, σταματάμε και χρησιμοποιούμε το $\vec{x}^{(k)}$ (ή κάποια επέκταση των $\vec{x}^{(0)}, \vec{x}^{(1)}$, ..., $\vec{x}^{(k)}$) ως εκτίμηση μιας βέλτιστης λύσης. Αλλιώς, θέτουμε $k = k + 1$ και εκτελούμε άλλη μια επανάληψη.[80]

Διαχωρίσιμος προγραμματισμός

Ο διαχωρίσιμος προγραμματισμός αποτελεί μια ειδική περίπτωση του κυρτού προγραμματισμού, όπου προστίθεται και μια επιπλέον υπόθεση. Συγκεκριμένα έχουμε,

- η $f(\vec{x})$ είναι μια κοίλη συνάρτηση.

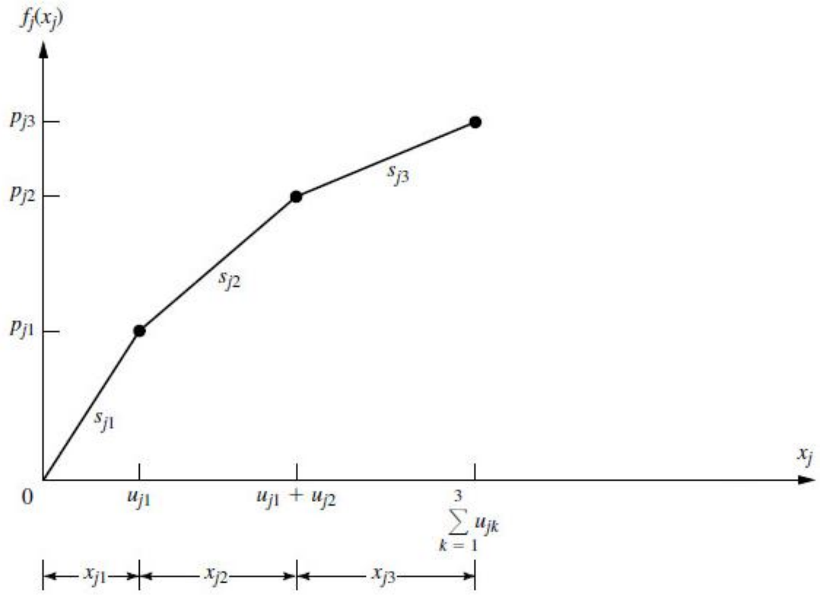
[80] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 697 – 699.

- Κάθε $g_i(\vec{x})$ είναι μια κυρτή συνάρτηση.
- Οι συναρτήσεις $f(\vec{x})$ και $g_i(\vec{x})$ είναι διαχωρίσιμες συναρτήσεις.

Διαχωρίσιμη συνάρτηση ονομάζεται η συνάρτηση, όπου κάθε όρος εμπλέκει μόνο μια μεταβλητή, ώστε η συνάρτηση να είναι χωρισμένη σε ένα άθροισμα από συναρτήσεις μιας μεταβλητής. Δηλαδή είναι της μορφής[81]

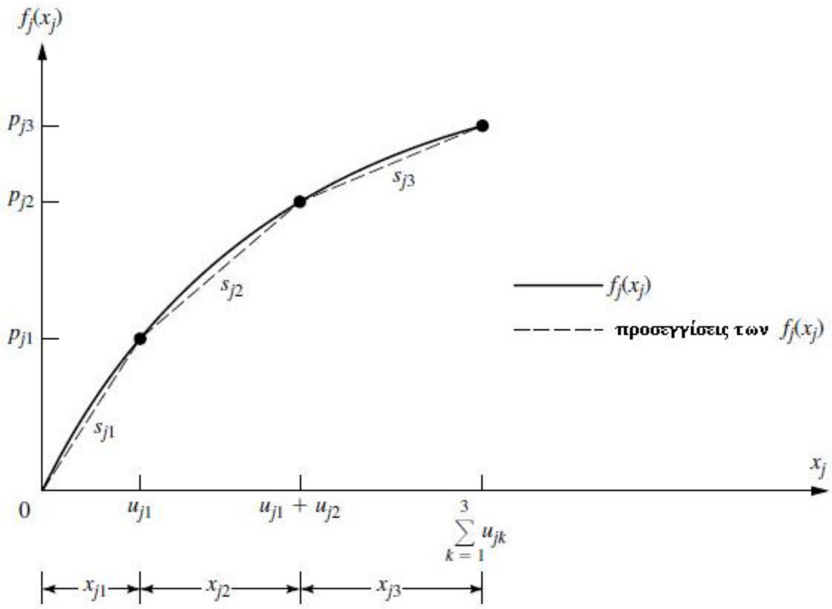
$$f(\vec{x}) = \sum_{j=1}^{n} f_j(x_j).$$

έτσι ώστε κάθε $f_j(x_j)$ να είναι της μορφής του γραφήματος 15 ή 16 σε μια εφικτή έκταση των τιμών του x_j.



Γράφημα 15 – 1η περίπτωση: $f_j(x_j)$ είναι κοίλη και κατά τμήματα γραμμική.

[81] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 667.



Γράφημα 16 – 2η περίπτωση: $f_j(x_j)$ είναι μόνο κοίλη.

Αυτή η προσέγγιση είναι πολύ βολική, διότι μια κατά τμήματα γραμμική συνάρτηση μιας μεταβλητής μπορεί να γραφτεί ως μια γραμμική συνάρτηση πολλών μεταβλητών, με έναν ειδικό περιορισμό στις τιμές των μεταβλητών, όπως θα περιγραφεί στη συνέχεια.

Αναδιατύπωση ως ένα πρόβλημα γραμμικού προγραμματισμού

Το κλειδί για τη μετατροπή μιας κατά τμήματα γραμμικής συνάρτησης σε γραμμική είναι η χρήση μιας διαφορετικής μεταβλητής για κάθε γραμμικό τμήμα.

Έστω n_j το πλήθος των γραμμικών τμημάτων στην $f_j(x_j)$, έτσι ώστε τα

$$x_j = \sum_{k=1}^{n_j} x_{jk}$$

να αντικαθιστούν στο αρχικό μοντέλο και οι

$$f_j(x_j) = \sum_{k=1}^{n_j} s_{jk} x_{jk}$$

να αντικαθιστούν στην αντικειμενική συνάρτηση για $j = 1, 2, \dots, n$. Το μοντέλο που προκύπτει είναι

$$maximise \quad Z = \sum_{j=1}^{n} \left(\sum_{k=1}^{n_j} s_{jk} x_{jk} \right),$$

υπό τους περιορισμούς

$$\sum_{j=1}^{n} a_{ij} \left(\sum_{k=1}^{n_j} x_{jk} \right) \leq b_i \text{ , για } i = 1, 2, \dots, m$$

$$x_{jk} \leq u_{jk} \text{ , για } k = 1, 2, \dots, n_j \text{ και } j = 1, 2, \dots, n$$

και

$$x_{jk} \geq 0 \text{ , για } k = 1, 2, \dots, n_j \text{ και } j = 1, 2, \dots, n.$$

Αν κάποια αρχική μεταβλητή x_j δεν έχει άνω φράγμα, τότε $u_{jn_j} = \infty$, ώστε ο περιορισμός που περιέχει αυτή την ποσότητα να διαγραφεί.

Δυστυχώς, κάποιες κατά τμήματα γραμμικές συναρτήσεις δεν μπορούν να μετατραπούν σε μορφή γραμμικού προγραμματισμού. Ωστόσο, οι $f_j(x_j)$ υποθέτονται κοίλες, ώστε $s_{j1} > s_{j2} > \dots$, έτσι ώστε ο αλγόριθμος μεγιστοποίησης της $f(\vec{x})$ να δίνει αυτόματα μέγιστη προτεραιότητα στη χρήση του x_{j1} όταν αυξάνεται το x_j από το μηδέν και ούτω καθεξής, χωρίς να συμπεριληφθεί ο ειδικός περιορισμός άμεσα στο μοντέλο. Αυτή η παρατήρηση οδηγεί στην ακόλουθη σημαντική ιδιότητα.

Σημαντική ιδιότητα του διαχωρίσιμου προγραμματισμού:

Όταν οι $f(\vec{x})$ και $g_i(\vec{x})$ ικανοποιούν τις υποθέσεις του διαχωρίσιμου προγραμματισμού, και όταν οι προκύπτουσες κατά τμήματα γραμμικές συναρτήσεις μετατρέπονται σε γραμμικές συναρτήσεις, διαγράφοντας τον ειδικό περιορισμό, καταλήγουμε σε ένα μοντέλο γραμμικού προγραμματισμού, του οποίου η βέλτιστη λύση αυτόματα ικανοποιεί τον ειδικό περιορισμό.

Τέλος, ένας αποτελεσματικός τρόπος επίλυσης αυτού του μοντέλου αποτελεί η χρήση της μεθόδου simplex. Αφού βρεθεί η βέλτιστη λύση αυτού του μοντέλου, θα πρέπει να υπολογιστούν οι ποσότητες

$$x_j = \sum_{k=1}^{n_j} x_{jk} \text{ για } j = 1, 2, \dots, n,$$

για τον προσδιορισμό της βέλτιστης λύσης του αρχικού προβλήματος διαχωρίσιμου προγραμματισμού.[82]

Μη κυρτός προγραμματισμός

Ο μη κυρτός προγραμματισμός περικλείει όλα τα προβλήματα του μη γραμμικού προγραμματισμού, τα οποία δεν ικανοποιούν τις υποθέσεις του κυρτού προγραμματισμού. Στην περίπτωση αυτή, ακόμα κι αν βρεθεί επιτυχώς ένα σημείο τοπικού μεγίστου, δεν υπάρχει εγγύηση ότι το σημείο αυτό αποτελεί ολικό μέγιστο. Επομένως, δεν υπάρχει ένας συγκεκριμένος αλγόριθμος ο οποίος να εξασφαλίζει την εύρεση μιας βέλτιστης λύσης. Ωστόσο, υπάρχουν διάφοροι αλγόριθμοι που εφαρμόζονται για την εύρεση τοπικού μεγίστου.[83]

Μια συνηθισμένη προσέγγιση για την επίλυση τέτοιων προβλημάτων είναι η εφαρμογή μιας αλγοριθμικής διαδικασίας αναζήτησης, η οποία θα σταματήσει όταν βρεθεί ένα τοπικό μέγιστο και στη συνέχεια θα επανεκκινήσει για συγκεκριμένο αριθμό επαναλήψεων από διάφορες αρχικές δοκιμαστικές λύσεις, ώστε να βρεθούν όσο το δυνατόν περισσότερα διακεκριμένα τοπικά μέγιστα. Τέλος, το βέλτιστο τοπικό μέγιστο επιλέγεται για εφαρμογή.

Μια τέτοια διαδικασία αναζήτησης, η οποία χρησιμοποιείται ευρέως από την δημιουργία της στη δεκαετία του 1960, αποτελεί η ακολουθιακή τεχνική ελαχιστοποίησης χωρίς περιορισμούς (SUMT). Υπάρχουν δύο εκδοχές της SUMT, η μια βασίζεται στον αλγόριθμο με εξωτερικό σημείο και η άλλη στον αλγόριθμο

[82] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 690 – 694.

[83] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 668.

με εσωτερικό σημείο, η οποία αντιμετωπίζει άμεσα εφικτές λύσεις ενώ χρησιμοποιεί μια συνάρτηση – φράγμα για τη συγκράτηση της λύσης μέσα στο εφικτό πεδίο λύσεων. Στη συνέχεια της ενότητας, θα παρουσιαστεί η δεύτερη εκδοχή της SUMT μετατρέποντάς την, ωστόσο, σε μέθοδο μεγιστοποίησης, ώστε να συνάδει με τις αλγοριθμικές μεθόδους μεγιστοποίησης που παρουσιάστηκαν σε προηγούμενες ενότητες.

Ακολουθιακή τεχνική μεγιστοποίησης χωρίς περιορισμούς (SUMT)

Όπως υποδηλώνει και το όνομά της, η μέθοδος αυτή αντικαθιστά το αρχικό πρόβλημα με μια ακολουθία προβλημάτων χωρίς περιορισμούς, των οποίων οι λύσεις συγκλίνουν σε μια λύση (τοπικό μέγιστο) του αρχικού προβλήματος. Σε κάθε ένα από τα προβλήματα χωρίς περιορισμούς σε αυτή την ακολουθία, θα επιλεγεί μια (ικανοποιητικά μικρή) γνήσια θετική τιμή ενός βαθμωτού r και εν συνεχεία θα επιλυθεί για $\vec{x}$ τέτοιο ώστε:

$$maximize \quad P(\vec{x}; r) = f(\vec{x}) - rB(\vec{x}).$$

Όπου η $B(\vec{x})$ είναι μια συνάρτηση – φράγμα, η οποία έχει τις ακόλουθες ιδιότητες (για εφικτές τιμές του $\vec{x}$ για το αρχικό πρόβλημα):

- Η $B(\vec{x})$ είναι μικρή όταν το $\vec{x}$ είναι μακριά από το σύνορο του πεδίου εφικτών λύσεων.
- Η $B(\vec{x})$ είναι μεγάλη όταν το $\vec{x}$ είναι κοντά στο σύνορο του πεδίου εφικτών λύσεων.
- Η $B(\vec{x}) \to \infty$, καθώς η απόσταση από το (κοντινότερο) σύνορο του πεδίου εφικτών λύσεων $\to 0$.

Η πιο συνηθισμένη επιλογή της $B(\vec{x})$ είναι η

$$B(\vec{x}) = \sum_{i=1}^{m} \frac{1}{b_i - g_i(\vec{x})} + \sum_{j=1}^{n} \frac{1}{x_j}$$

SUMT αλγόριθμος

Αλγοριθμος

Αρχικοποίηση:

Επιλογή μιας εφικτής αρχικής δοκιμαστικής λύσης $\vec{x}^{(0)}$, η οποία δεν είναι πάνω στο σύνορο του πεδίου εφικτών λύσεων.

Θέτω $k = 1$ και επιλέγω κατάλληλες γνήσια θετικές τιμές για το αρχικό r και για $\theta < 1$.

Επανάληψη:

Ξεκινώντας από το $\vec{x}^{(k-1)}$, εφαρμόζουμε τη διαδικασία αναζήτησης κλίσης, όπως αυτή έχει περιγραφεί αναλυτικά σε προηγούμενη ενότητα, ή κάποια παρόμοια μέθοδο, για την εύρεση ενός τοπικού μεγίστου $\vec{x}^{(k)}$ του

$$P(\vec{x};r) = f(\vec{x}) - r\left[\sum_{i=1}^{m} \frac{1}{b_i - g_i(\vec{x})} + \sum_{j=1}^{n} \frac{1}{x_j}\right]$$

Κανόνας τερματισμού ή διακοπής:

Αν η αλλαγή από $\vec{x}^{(k-1)}$ σε $\vec{x}^{(k)}$ είναι αμελητέα, σταματάμε και χρησιμοποιούμε το $\vec{x}^{(k)}$ (ή κάποια επέκταση των $\vec{x}^{(0)}, \vec{x}^{(1)}$, ..., $\vec{x}^{(k)}$) ως εκτίμηση ενός τοπικού μεγίστου του αρχικού προβλήματος. Αλλιώς, θέτουμε $k = k + 1$ και $r = \theta r$ και εκτελούμε άλλη μια επανάληψη.

Όταν οι υποθέσεις του κυρτού προγραμματισμού δεν ικανοποιούνται, αυτός ο αλγόριθμος πρέπει να επαναληφθεί αρκετές φορές ξεκινώντας από διαφορετικές αρχικές δοκιμαστικές εφικτές λύσεις. Το καλύτερο από τα τοπικά μέγιστα θα χρησιμοποιηθεί ως η καλύτερη προσέγγιση για το ολικό μέγιστο.

Τέλος, αξίζει να σημειωθεί ότι η μέθοδος SUMT μπορεί να επεκταθεί, ώστε να αντιμετωπίζει και ισότητες της μορφής $g_i(\vec{x}) = b_i$. Ένας συνηθισμένος τρόπος είναι ο ακόλουθος. Για κάθε περιορισμό ισότητας, η ποσότητα $\frac{-[b_i - g_i(\vec{x})]^2}{\sqrt{r}}$

αντικαθίσταται με $\frac{-r}{b_i - g_i(\vec{x})}$ στην έκφραση της $P(\vec{x}; r)$, και στη συνέχεια χρησιμοποιείται η ίδια διαδικασία.[84]

Γεωμετρικός προγραμματισμός

Αρχικά, θα ορίσουμε ένα μονώνυμο ως συνάρτηση $f : R_+^n \to R$

$$f(\vec{x}) = d x_1^{a^{(1)}} x_2^{a^{(2)}} \dots x_n^{a^{(n)}}$$

όπου η πολλαπλασιαστική σταθερά $d \geq 0$ και οι εκθετικές σταθερές $a_{(i)} \in \mathbb{R}\,, i = 1,2,\dots,n$. Το άθροισμα των μονωνύμων, με συμβολισμό ένα δείκτη k, ονομάζονται ποζυνώμια,

$$\sum_{k=1}^{K} d_k x_1^{a_k^{(1)}} x_2^{a_k^{(2)}} \dots x_n^{a_k^{(n)}}$$

όπου $d_k \geq 0, k = 1, 2, \dots, K$ και $a_{kj} \in \mathbb{R}\,, j = 1, 2, \dots, n, k = 1, 2, \dots, K$.[85]

Προβλήματα γεωμετρικού προγραμματισμού ονομάζονται τα προβλήματα μη γραμμικού προγραμματισμού, των οποίων η αντικειμενική συνάρτηση και οι περιορισμοί συχνά έχουν την ακόλουθη μορφή,

$$f(\vec{x}) = \sum_{k=1}^{K} d_k P_k(\vec{x}),$$

όπου

$$P_k(\vec{x}) = x_1^{a_k^{(1)}} x_2^{a_k^{(2)}} \dots x_n^{a_k^{(n)}}, \qquad \gamma\iota\alpha\ k = 1,2,\dots,K.$$

Αυτές οι συναρτήσεις γενικά δεν είναι ούτε κυρτές ούτε κοίλες, ώστε να χρησιμοποιηθούν οι τεχνικές κυρτού προγραμματισμού για την επίλυση τέτοιων προβλημάτων. Ωστόσο, για μια συγκεκριμένη περίπτωση το πρόβλημα γεωμετρικού προγραμματισμού μπορεί να μετασχηματιστεί σε ένα ισοδύναμο

[84] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 702 – 705.
[85] Mung Chiang, *Geometric Programming for Communication Systems*, MA 2005, σελ. 09 – 10.

πρόβλημα κυρτού προγραμματισμού. Στην περίπτωση αυτή, όλοι οι συντελεστές d_k σε κάθε συνάρτηση είναι αυστηρά θετικές, ώστε οι συναρτήσεις να είναι γενικευμένα ποζυνώμια και η αντικειμενική συνάρτηση πρέπει να ελαχιστοποιηθεί. [86]

Η κανονική μορφή ενός προβλήματος γεωμετρικού προγραμματισμού είναι η ακόλουθη

$$minimize \quad f_0(\vec{x})$$

υπό τους περιορισμούς

$$f_i(\vec{x}) \leq 1, \qquad i = 1, 2, \dots, m$$

$$h_l(\vec{x}) = 1, \qquad l = 1, 2, \dots, M$$

όπου τα f_i , $i = 0, 1, 2, \dots, m$ είναι ποζυνώμια της μορφής

$$f_i(\vec{x}) = \sum_{k=1}^{K_i} d_{ik} x_1^{a_{ik}^{(1)}} x_2^{a_{ik}^{(2)}} \dots x_n^{a_{ik}^{(n)}},$$

και τα h_l , $l = 1, 2, \dots, M$ είναι μονώνυμα της μορφής

$$h_l(\vec{x}) = d_l x_1^{a_l^{(1)}} x_2^{a_l^{(2)}} \dots x_n^{a_l^{(n)}}.$$ [87]

Για ένα δοθέν πρόβλημα γεωμετρικού προγραμματισμού στην κανονική του μορφή, μπορούμε να σχηματίσουμε τον πίνακα A, όπου κάθε γραμμή αποτελείται από τις εκθετικές σταθερές, οι οποίες σχετίζονται με κάθε όρο του μονωνύμου που παρουσιάζεται στην αντικειμενική συνάρτηση και τις σταθερές, και ένα διάνυσμα d που αποτελείται από όλες τις πολλαπλασιαστικές σταθερές. Κάθε πρόβλημα γεωμετρικού προγραμματισμού μπορεί να οριστεί με μοναδικό τρόπο υπό την ακόλουθη μορφή δεδομένων: A, d και ένα αναγνωριστικό για το ποιες γραμμές στα A και d ανήκουν στην αντικειμενική συνάρτηση και ποιες στις συναρτήσεις των περιορισμών.

[86] Stephen Boyd & Seung-Jean Kim & Lieven Vandenberghe & Arash Hassibi, *A Tutorial On Geometric Programming*, Published online 2007, σελ. 70.
[87] Stephen Boyd & Seung-Jean Kim & Lieven Vandenberghe & Arash Hassibi, *A Tutorial On Geometric Programming*, Published online 2007, σελ. 70.
Mung Chiang, *Geometric Programming for Communication Systems*, MA 2005, σελ.10.

Όπως προαναφέρθηκε, το πρόβλημα γεωμετρικού προγραμματισμού στην κανονική του μορφή δεν κατατάσσεται στην κατηγορία προβλημάτων κυρτού προγραμματισμού, επειδή τα ποζυνώμια δεν είναι κυρτές συναρτήσεις. Ωστόσο ένας λογαριθμικός μετασχηματισμός της μορφής $y_j = \log x_j$, $b_{ik} = logd_{ik}$, $b_l = logd_l$, μετατρέπει το αρχικό πρόβλημα σ' ένα αντίστοιχο κυρτού προγραμματισμού.

Αναλυτικά, το ισοδύναμο πρόβλημα κυρτού προγραμματισμού με βάση τον παραπάνω μετασχηματισμό και με μεταβλητές απόφασης $y_1, y_2, \dots, y_n$, σχηματίζεται θέτοντας

$$x_j = e^{y_j} \ , \qquad για\ j = 1,2, \dots, n$$

στο αρχικό πρόβλημα, ώστε τώρα να μπορεί να χρησιμοποιηθεί αλγόριθμος επίλυσης προβλημάτων κυρτού προγραμματισμού. [88]

Επομένως, το ισοδύναμο πρόβλημα[89] έχει τη μορφή

$$minimize \quad \log f_0(\vec{x})$$

υπό τους περιορισμούς

$$\log f_i\left(e^{\vec{y}}\right) \leq 0, \qquad i = 1, 2, \dots, m$$

$$\log h_l\left(e^{\vec{y}}\right) = 0, \qquad l = 1, 2, \dots, M.$$

Παρουσιάζοντας το παραπάνω πρόβλημα με μορφή δεδομένων: A, d, έχουμε

$$minimize \quad \sum_{k=1}^{K_0} e^{\left(\vec{a}_{0k}^T \vec{y} + b_{ok}\right)}$$

υπό τους περιορισμούς

$$\sum_{k=1}^{K_i} e^{\left(\vec{a}_{ik}^T \vec{y} + b_{ik}\right)} \leq 1, \qquad i = 1, 2, \dots, m$$

[88] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 668.
[89] Stephen Boyd & Seung-Jean Kim & Lieven Vandenberghe & Arash Hassibi, *A Tutorial On Geometric Programming*, Published online 2007, σελ. 73.

$$\vec{a}_l^T \vec{y} + b_l = 0, \qquad l = 1, 2, \dots, M$$

το οποίο είναι ισοδύναμο με το ακόλουθο πρόβλημα γεωμετρικού προγραμματισμού σε μορφή κυρτού προγραμματισμού

$$minimize \;\; p_0(\vec{y}) = log\left(\sum_{k=1}^{K_0} e^{(\vec{a}_{0k}^T \vec{y} + b_{ok})}\right)$$

υπό τους περιορισμούς

$$p_i(\vec{y}) = log\left(\sum_{k=1}^{K_i} e^{(\vec{a}_{ik}^T \vec{y} + b_{ik})}\right) \leq 0, \qquad i = 1, 2, \dots, m$$

$$q_l(\vec{y}) = \vec{a}_l^T \vec{y} + b_l = 0, \qquad l = 1, 2, \dots, M.$$ [90]

Λημμα 3.1[91]

Το $log - sum - \exp function\; f(\vec{x}) = \log \sum_{i=1}^{n} e^{x_i}$ είναι κυρτό στο $\vec{x}$.

Αποδειξη

Έστω η ακόλουθη $log - sum$ ανισότητα,

$$\sum_{i=1}^{n} a_i \log \frac{a_i}{b_i} \geq \left(\sum_{i=1}^{n} a_i\right) \log \frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}, \qquad \boxed{1}$$

όπου $a, b \succcurlyeq 0$.

Δοθείσας μιας συνάρτησης $f: \mathbb{R}^n \to \mathbb{R}$, η συνάρτηση $f^*: \mathbb{R}^n \to \mathbb{R}$ ορισμένη ως

$$f^*(\vec{y}) = \sup_{\vec{x} \subset dom f} \left(\vec{y}^T \vec{x} - f(\vec{x})\right), \quad \boxed{2}$$

ονομάζεται συζυγής συνάρτηση της f. Αφού η f^* είναι κατά σημείο ελάχιστο άνω φράγμα μιας οικογένειας αφφινικών συναρτήσεων του $\vec{y}$, είναι πάντα κυρτή συνάρτηση.

Έστω $\hat{b}_i = \log b_i$ και $\sum_{i=1}^{n} a_i = 1$ στην (1). Παίρνουμε,

---

[90] Mung Chiang, *Geometric Programming for Communication Systems*, MA 2005, σελ.11 – 12.
[91] Mung Chiang, *Geometric Programming for Communication Systems*, MA 2005, σελ.12 – 13.

$$\log\left(\sum_{i=1}^{n} e^{\hat{b}_i}\right) \geq \vec{a}^T\hat{\vec{b}} - \sum_{i=1}^{n} a_i \log a_i,$$

με ισότητα αν και μόνο αν $a_i = \frac{e^{\hat{b}_i}}{\sum_{j=1} e^{\hat{b}_j}}$. Αυτό εξ' ορισμού δείχνει ότι η $log - sum - exp$ συνάρτηση είναι η συζυγής συνάρτηση της αρνητικής εντροπίας. Αφού όλες οι συζυγείς συναρτήσεις είναι κυρτές, τότε και η $log - sum - exp$ συνάρτηση είναι κυρτή.

---*τέλος απόδειξης*---

Τέλος, υπάρχουν τουλάχιστον δύο σημαντικές προσεγγίσεις για την επίλυση ενός προβλήματος γεωμετρικού προγραμματισμού. Η μια είναι η *μέθοδος εσωτερικού σημείου* και η άλλη είναι ένας μη εφικτός αλγόριθμος.

Στην παρούσα ενότητα θα παρουσιάσουμε τον αλγόριθμο για την πρώτη μέθοδο. Η βασιζόμενη σε φράγματα μέθοδος εσωτερικού σημείου για κυρτό προγραμματισμό μπορεί να εφαρμοστεί στα προβλήματα γεωμετρικού προγραμματισμού με ορθό και εμπρός τρόπο.

Η κεντρική ιδέα του αλγορίθμου είναι η επίλυση μιας ακολουθίας από προβλήματα χωρίς περιορισμούς, τα οποία απορροφούν τους περιορισμούς σε μια καινούρια αντικειμενική συνάρτηση, η οποία είναι σταθμικό άθροισμα της αρχικής αντικειμενικής συνάρτησης και μια συνάρτηση – φράγμα ϕ των περιορισμών. Καθώς το βάρος t στην αρχική αντικειμενική συνάρτηση γίνεται μεγαλύτερο, το πρόβλημα χωρίς περιορισμούς γίνεται μια πιο στενή προσέγγιση του αρχικού προβλήματος.

Αλγόριθμος για τη μέθοδο – φράγμα

Αλγοριθμος

Αρχικοποίηση:

Έστω ένα δοθέν αυστηρώς εφικτό σημείο $\vec{x}$, το οποίο μπορεί να βρεθεί είτε επαληθεύοντας ότι ένα δοθέν $\vec{x}$ είναι αυστηρώς εφικτό ή επιλύοντας ένα εφικτό πρόβλημα γεωμετρικού προγραμματισμού, και $t := t^{(0)} > 0$, $\mu > 1$, και ανοχή σφάλματος $\epsilon > 0$.

Επανάληψη:

1. Κεντρικό βήμα: υπολογισμός του $\vec{x}^*(t)$ ελαχιστοποιώντας την ποσότητα $tf_0(\vec{x}) + \phi(\vec{x})$ ξεκινώντας από το $\vec{x}$. Αυτή είναι μια χωρίς περιορισμούς, λεία, κυρτή ελαχιστοποίηση, η οποία μπορεί να υλοποιηθεί από διάφορες επαναληπτικές μεθόδους, όπως η *μέθοδος καθοδικής κλίσεως* ή η *μέθοδος Newton*.
2. Ανανέωση: $\vec{x} := \vec{x}^*(t)$.

Κανόνας τερματισμού ή διακοπής:

Αν $\frac{m}{t} \leq \epsilon$, σταματάμε, αλλιώς αυξάνουμε το t θέτοντας $t := \mu t$ και εκτελούμε άλλη μια επανάληψη.[92]

Κλασματικός προγραμματισμός

Στη συγκεκριμένη ενότητα θεωρούμε ότι η αντικειμενική συνάρτηση είναι σε κλασματική μορφή

$$maximize\ \ f(\vec{x}) = \frac{f_1(\vec{x})}{f_2(\vec{x})}.$$

Για προβλήματα τέτοιας μορφής έχουν αναπτυχθεί ειδικές διαδικασίες επίλυσης όταν οι συναρτήσεις $f_1(\vec{x})$ και $f_2(\vec{x})$ έχουν συγκεκριμένες μορφές.

Όταν αυτό είναι εφικτό, η πιο ορθή προσέγγιση για την επίλυση ενός προβλήματος κλασματικού προγραμματισμού είναι η μετατροπή του σε ένα ισοδύναμο πρόβλημα ενός κανονικού τύπου για τον οποίο αποτελεσματικές διαδικασίες επίλυσης είναι ήδη διαθέσιμες.

Υποθέτουμε, λοιπόν, ότι η συνάρτηση $f(\vec{x})$ είναι σε μορφή γραμμικού κλασματικού προγραμματισμού

[92] Mung Chiang, *Geometric Programming for Communication Systems*, MA 2005, σελ.35 – 36.

$$f(\vec{x}) = \frac{\vec{c}\vec{x} + c_0}{\vec{d}\vec{x} + d_0},$$

όπου $\vec{c}$ και $\vec{d}$ είναι διανύσματα – γραμμές, $\vec{x}$ είναι διάνυσμα – στήλη και c_0 και d_0 είναι βαθμωτά. Επίσης, υποθέτουμε ότι οι συναρτήσεις, $g_i(\vec{x})$, των περιορισμών είναι γραμμικοί, έτσι ώστε οι περιορισμοί σε μορφή πινάκων να είναι $\vec{A}\vec{x} \leq \vec{b}$ και $\vec{x} \geq \vec{0}$.

Υπό ήπιες επιπλέον υποθέσεις, μπορούμε να μετασχηματίσουμε το πρόβλημα σε ένα ισοδύναμο πρόβλημα γραμμικού προγραμματισμού θέτοντας

$$\vec{y} = \frac{\vec{x}}{\vec{d}\vec{x} + d_0} \quad \text{και} \quad t = \frac{1}{\vec{d}\vec{x} + d_0},$$

έτσι ώστε $\vec{x} = \frac{\vec{y}}{t}$. Επομένως, παίρνουμε

$$maximize \quad Z = \vec{c}\vec{y} + c_0 t$$

υπό τους περιορισμούς

$$\vec{A}\vec{y} - \vec{b}t \leq \vec{0}\,,$$

$$\vec{d}\vec{y} + d_0 t = 1\,,$$

και

$$\vec{y} \geq \vec{0}\,, \qquad t \geq 0$$

το οποίο μπορεί να επιλυθεί με τη μέθοδο simplex. Γενικότερα, ανάλογος μετασχηματισμός μπορεί να εφαρμοστεί για τη μετατροπή ενός προβλήματος κλασματικού προγραμματισμού με κοίλη συνάρτηση $f_1(\vec{x})$, κυρτή $f_2(\vec{x})$ και κυρτές $g_i(\vec{x})$, σε ένα ισοδύναμο πρόβλημα κυρτού προγραμματισμού.[93]

[93] Federich S. Hillier & Gerald J. Liebermann, *Introduction to Operations Research*, New York [7]2001, σελ. 668 – 669.

Εφαρμογές

Στην συγκεκριμένη ενότητα, θα εργαστούμε με σκοπό την παρουσίαση δύο εφαρμογών του μη γραμμικού προγραμματισμού.

Πρώτη Εφαμογή

Έστω, λοιπόν, το πρόβλημα

$$\boldsymbol{maximize} \;\; \boldsymbol{f(x_1, x_2) = 7x_1 + 14x_2 + 3x_1x_2 - 2x_1^2 - 6x_2^2}$$

υπό τους περιορισμούς

$$\boldsymbol{x_1 + 4x_2 \leq 3}$$

$$\boldsymbol{x_1 \geq 0}\,, \boldsymbol{x_2 \geq 0} \;\; \text{(περιορισμοί μη αρνητικότητας).}$$

Παρατηρούμε ότι η αντικειμενική συνάρτηση είναι τετραγωνικής μορφής με γραμμικούς περιορισμούς, επομένως το πρόβλημα αποτελεί ένα πρόβλημα μη γραμμικού τετραγωνικού προγραμματισμού.

Με βάση τη θεωρία του τετραγωνικού προγραμματισμού, όπως αυτή παρουσιάστηκε σε προηγούμενη ενότητα[94], για τη συγκεκριμένη εφαρμογή έχουμε

$\vec{c} = \begin{bmatrix} 7 & 14 \end{bmatrix}$

$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$Q = \begin{bmatrix} 4 & -3 \\ -3 & 12 \end{bmatrix}$

$A = \begin{bmatrix} 1 & 4 \end{bmatrix}$

$\vec{b} = [3]$

Γενική μορφή τετραγωνικού προγραμματισμού:

$$maximise\; f(\vec{x}\,) = \vec{c}\vec{x} - \frac{1}{2}\vec{x}^T Q \vec{x},$$

υπό τους περιορισμούς

$$A\vec{x} \leq \vec{b}$$

και

$$\vec{x} \geq 0$$

[94] Βλέπε στο παρόν σελ. 52.

Αφού έχουμε πρόβλημα τετραγωνικού προγραμματισμού με γραμμικούς περιορισμούς θα χρησιμοποιήσουμε τις συνθήκες ΚΚΤ για τετραγωνικό προγραμματισμό.

<u>Οι ΚΚΤ συνθήκες για τον τετραγωνικό προγραμματισμό</u>

Παρατηρούμε ότι $m = 1$ αφού έχουμε μόνο έναν περιορισμό.

ΚΚΤ συνθήκες:
$Q\vec{x} + A^T\vec{u} - \vec{y} = \vec{c}^T,$
$A\vec{x} + \vec{v} = \vec{b},$
$\vec{x} \geq 0\,, \vec{u} \geq 0\,, \vec{y} \geq 0\,, \vec{v} \geq 0$
$\vec{x}^T\vec{y} + \vec{u}^T\vec{v} = 0$

1. $\left(\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j} \leq 0\right)$

   <u>$\text{για } j = 1$</u>

   $7 + 3x_2 - 4x_1 - u_1 \leq 0$

   <u>$\text{για } j = 2$</u>

   $14 + 3x_1 - 12x_2 - 4u_1 \leq 0$

2. $\left(x_j^* \left[\frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} u_i \frac{\partial g_i}{\partial x_j}\right] = 0\right)$

   <u>$\text{για } j = 1$</u>

   $x_1(7 + 3x_2 - 4x_1 - u_1) = 0$

   <u>$\text{για } j = 2$</u>

   $x_2(14 + 3x_1 - 12x_2 - 4u_1) = 0$

3. $(g_i(\vec{x}^*) - b_i \leq 0)$

   $x_1 + 4x_2 - 3 \leq 0$

4. $(u_1[g_i(\vec{x}^*) - b_i] = 0)$

   $u_1(x_1 + 4x_2 - 3) = 0$

5. $(x_j^* \geq 0)$

   $x_1 \geq 0\,, x_2 \geq 0$

6. $(u_i \geq 0)$

   $u_1 \geq 0$

Για να εκφράσουμε σε πιο βολική μορφή τις συνθήκες, μεταφέρουμε τις σταθερές στις συνθήκες $1.j = 1$, $1.j = 2$ και 3 στο δεξί μέρος και στη συνέχεια εισάγουμε μη αρνητικές χαλαρές μεταβλητές y_1, y_2 και v_1, αντίστοιχα, για να μετατρέψουμε αυτές τις ανισότητες σε ισότητες.

1. <u>$\text{για } j = 1$</u>

   $3x_2 - 4x_1 - u_1 + y_1 = -7$

$\underline{\text{για } j = 2}$

$3x_1 - 12x_2 - 4u_1 + y_2 = -14$

3. $x_1 + 4x_2 + v_1 = 3$

Επίσης, παρατηρούμε ότι η συνθήκη $2.j = 1$ μπορεί να εκφραστεί ως μια απλή απαίτηση είτε $x_1 = 0$ είτε $y_1 = 0$ αφού για $y_1 = 0$ παίρνουμε $3x_2 - 4x_1 - u_1 = -7$, δηλαδή

2. $\underline{\text{για } j = 1}$

$x_1 y_1 = 0$

ομοίως μπορούν να εκφραστούν οι συνθήκες $2.j = 2$ και 4

2. $\underline{\text{για } j = 2}$

$x_2 y_2 = 0$

4. $u_1 v_1 = 0.$

Οι τρεις παραπάνω συνθήκες μπορούν να συνδυαστούν σε έναν περιορισμό

$x_1 y_1 + x_2 y_2 + u_1 v_1 = 0$ (*περιορισμός συμπληρωματικότητας*).

Αφού πολλαπλασιάσουμε με -1 τους περιορισμούς $1.j = 1$ και $1.j = 2$ για να εξασφαλίσουμε μη αρνητικά δεξιά μέρη, καταλήγουμε σε μια βολική μορφή για τις συνθήκες

$$
\begin{aligned}
4x_1 - 3x_2 + u_1 - y_1 &= 7 \\
-3x_1 + 12x_2 + 4u_1 - y_2 &= 14 \\
x_1 + 4x_2 + v_1 &= 3 \\
x_1 y_1 + x_2 y_2 + u_1 v_1 &= 0
\end{aligned}
$$

$$x_1 \geq 0 , x_2 \geq 0, u_1 \geq 0, y_1 \geq 0, y_2 \geq 0, v_1 \geq 0$$

Επομένως, το παραπάνω πρόβλημα παίρνει την ακόλουθη μορφή

$$maximize \quad f(x_1, x_2) = 7x_1 + 14x_2 + 3x_1 x_2 - 2x_1^2 - 6x_2^2$$

υπό τους περιορισμούς

$$
\begin{aligned}
4x_1 - 3x_2 + u_1 - y_1 &= 7 \\
-3x_1 + 12x_2 + 4u_1 - y_2 &= 14
\end{aligned}
$$

$$x_1 \ + 4x_2 \qquad\qquad\qquad + v_1 = 3$$
$$x_1y_1 + x_2y_2 + u_1v_1 = 0$$

$$x_1 \geq 0\,, x_2 \geq 0, u_1 \geq 0, y_1 \geq 0, y_2 \geq 0, v_1 \geq 0$$

Από το Πόρισμα 3.1, το $\vec{x}^* = (\vec{x}_1^*,\ \vec{x}_2^*, \dots, \vec{x}_n^*)$ είναι μια βέλτιστη λύση αν και μόνο αν όλες οι συνθήκες του θεωρήματος 3.1 ικανοποιούνται.

Λύνοντας το παραπάνω σύστημα καταλήγουμε στο συμπέρασμα ότι η λύση

$\vec{x}^* = (1{,}280000082 \quad 0{,}430000008)$

είναι μια βέλτιστη λύση αφού υπάρχει $u_1 = 3{,}170000038$, ώστε οι συνθήκες ΚΚΤ για τετραγωνικό προγραμματισμό να ικανοποιούνται για το $\vec{x}^*$. Μάλιστα η μέγιστη τιμή της συνάρτησης f που προκύπτει είναι 12,24500036.

Δεύτερη Εφαρμογή

Στην εφαρμογή αυτή θα παρουσιάσουμε ένα πρόβλημα διαχωρίσιμου προγραμματισμού.

Έστω ότι μια εταιρία διαθέτει 3 μηχανήματα για συγκεκριμένες ώρες λειτουργίας για την παραγωγή δύο προϊόντων, τα οποία επιφέρουν συγκεκριμένο κέρδος, όπως φαίνεται αναλυτικά στον παρακάτω πίνακα

Μηχάνημα	**Χρόνος παραγωγής / παρτίδα (ώρες)**		**Διαθέσιμος χρόνος λειτουργίας / βδομάδα (ώρες)**
	Προϊόν		
	1	**2**	
1	1	0	4
2	0	2	12
3	3	2	18
Κέρδος/παρτίδα	3.000 €	5.000 €	

Πίνακας 3 - Δεδομένα για την εταιρία.

Μοντελοποιηση Προβληματος

Το παραπάνω πρόβλημα αποτελεί πρόβλημα γραμμικού προγραμματισμού με στόχο τη μεγιστοποίηση του κέρδους και η μοντελοποίηση του είναι η ακόλουθη.

Έστω οι μεταβλητές απόφασης

$$x_1 = ο\ αριθμός\ των\ παρτίδων\ του\ προϊόντος\ 1\ ανά\ εβδομάδα$$

$$x_2 = ο\ αριθμός\ των\ παρτίδων\ του\ προϊόντος\ 2\ ανά\ εβδομάδα$$

και η αντικειμενική συνάρτηση

$$maximize\ \ Z = 3x_1 + 5x_2$$

υπό τους περιορισμούς

Περιορισμοί διαθέσιμου χρόνου λειτουργίας

$$x_1 \leq 4$$
$$2x_2 \leq 12$$
$$3x_1 + 2x_2 \leq 18$$

Περιορισμοί μη αρνητικότητας

$$x_1 \geq 0$$

$$x_2 \geq 0$$

Λύνοντας το παραπάνω πρόβλημα με τη μέθοδο simplex με τη βοήθεια του λογισμικού προγράμματος excel – solver παίρνουμε την βέλτιστη λύση

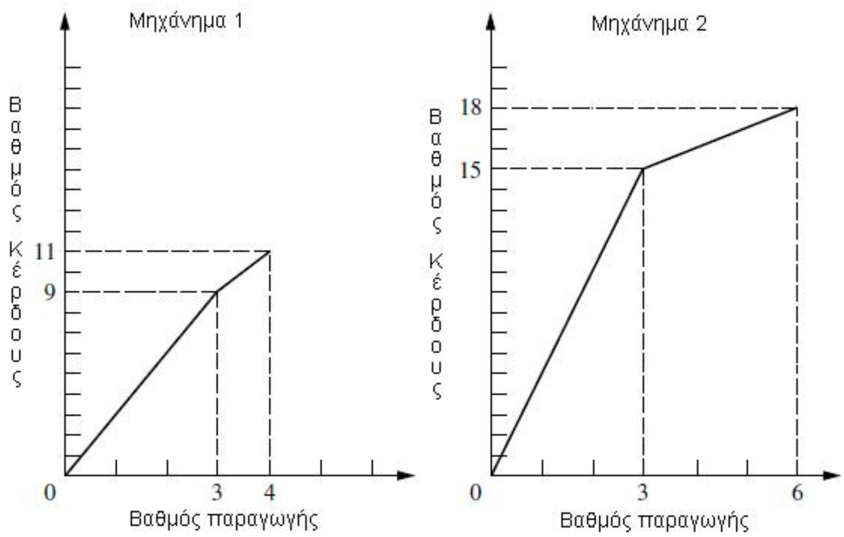
$$x_1 = 2$$

$$x_2 = 6$$

με μέγιστο κέρδος 36.000 €.

Στη συνέχεια, η εταιρία δέχεται μια ειδική παραγγελία για χειροποίητα προϊόντα τα οποία θα υλοποιηθούν με τη βοήθεια των μηχανημάτων 1 και 2 στους επόμενους τέσσερεις μήνες.

Για την υλοποίηση της παραγγελίας θα χρειαστεί να απασχοληθούν κάποιοι εργαζόμενοι από το υπαλληλικό προσωπικό που χρησιμοποιείται για την παραγωγή των δύο βασικών προϊόντων της εταιρίας. Συγκεκριμένα, για την παραγωγή του καινούργιου προϊόντος θα χρειαστεί η χρήση του 25% της παραγωγικής δυνατότητας του μηχανήματος 1 για την παραγωγή του προϊόντος 1 και η χρήση του 50% της παραγωγικής δυνατότητας του μηχανήματος 2 του προϊόντος 2. Η υπερωριακή αυτή εργασία θα μειώσει το κέρδος για κάθε εμπλεκόμενη μονάδα από 3€ σε 2€ για το προϊόν 1 και από 5€ σε 1€ για το προϊόν 2, παίρνοντας την ακόλουθη καμπύλη κέρδους για κάθε προϊόν.



Γράφημα 17 - Δεδομένα κέρδους της εταιρίας για τους επόμενους τέσσερεις μήνες.

Για την υλοποίηση της καινούργιας παραγγελίας αποφασίστηκε να χρησιμοποιηθούν οι ήδη υπάρχοντες υπάλληλοι. Ωστόσο, θα πρέπει το υπαλληλικό προσωπικό να χρησιμοποιηθεί πλήρως στον κανονικό χρόνο πριν γίνει χρήση της υπερωρίας. Επιπλέον, η παρούσα αναλογία παραγωγής $x_1 = 2$ για το προϊόν 1 και $x_2 = 6$ για το προϊόν 2 πρέπει να αλλάξει προσωρινά ώστε να βελτιωθεί το ολικό κέρδος της εταιρίας. Το καινούργιο πλάνο παραγωγής φαίνεται στον επόμενο πίνακα.

Μηχάνημα	Χρόνος παραγωγής / παρτίδα (ώρες)				Διαθέσιμος χρόνος λειτουργίας / βδομάδα (ώρες)
	Προϊόν				
	1Κ	1Υ	2Κ	2Υ	
1	1	1	0	0	4
2	0	0	2	2	12
3	3	3	2	2	18
Κέρδος/παρτίδα	3.000 €	2.000 €	5.000 €	1.000 €	

Πίνακας 4 - Δεδομένα για την εταιρία στους επόμενους 4 μήνες.

ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΠΡΟΒΛΗΜΑΤΟΣ

Το παραπάνω πρόβλημα αποτελεί πρόβλημα διαχωρίσιμου προγραμματισμού με στόχο τη μεγιστοποίηση του κέρδους. Τροποποιώντας την μοντελοποίηση του αρχικού προβλήματος γραμμικού προγραμματισμού ώστε να προσαρμοστεί στα νέα δεδομένα παίρνουμε την παρακάτω μοντελοποίηση.

Όπως έχουμε ήδη ορίσει στο κεφάλαιο του διαχωρίσιμου προγραμματισμού[95], έστω n_j το πλήθος των γραμμικών τμημάτων στην $f_j(x_j)$, έτσι ώστε τα

$$x_j = \sum_{k=1}^{n_j} x_{jk}.$$

Επομένως, έχουμε $n = 2,\ n_1 = 2\ , n_2 = 2$ και έστω οι μεταβλητές απόφασης

$$x_{1K} = \textit{ο αριθμός των παρτίδων του προϊόντος } 1 \textit{ σε κανονικό χρόνο/εβδομάδα}$$
$$x_{1Y} = \textit{ο αριθμός των παρτίδων του προϊόντος } 1 \textit{ σε υπερωρίες/εβδομάδα}$$
$$x_{2K} = \textit{ο αριθμός των παρτίδων του προϊόντος } 2 \textit{ σε κανονικό χρόνο/εβδομάδα}$$
$$x_{2Y} = \textit{ο αριθμός των παρτίδων του προϊόντος } 2 \textit{ σε υπερωρίες/εβδομάδα}$$

επομένως,

$$x_1 = x_{1K} + x_{1Y}$$

και

$$x_2 = x_{2K} + x_{2Y}$$

και η αντικειμενική συνάρτηση

$$\text{max}imize \quad Z = 3x_{1K} + 2x_{1Y} + 5x_{2K} + x_{2Y}$$

υπό τους περιορισμούς

<u>Περιορισμοί διαθέσιμου χρόνου λειτουργίας</u>

$$\begin{aligned} x_{1K} + x_{1Y} \qquad\qquad\qquad\quad &\leq 4 \\ 2(x_{2K} + x_{2Y}) &\leq 12 \\ 3(x_{1K} + x_{1Y}) + 2(x_{2K} + x_{2Y}) &\leq 18 \end{aligned}$$

[95] Βλέπε στο παρόν σελ. 59.

Περιορισμοί για την παραγωγή του καινούργιου προϊόντος

(χρήση του 25% της παραγωγικής δυνατότητας του μηχανήματος 1 για την παραγωγή του προϊόντος 1)

$$x_{1K} \leq 3$$

$$x_{1Y} \leq 1$$

(χρήση του 50% της παραγωγικής δυνατότητας του μηχανήματος 2 για την παραγωγή του προϊόντος 2)

$$x_{2K} \leq 3$$

$$x_{2Y} \leq 3$$

Περιορισμοί μη αρνητικότητας

$$x_{1K} \geq 0$$

$$x_{1Y} \geq 0$$

$$x_{2K} \geq 0$$

$$x_{2Y} \geq 0$$

Παρατηρούμε ότι οι περιορισμοί για την παραγωγή του καινούργιου προϊόντος με τους περιορισμούς μη αρνητικότητας, καθιστούν τους πρώτους δύο συναρτησιακούς περιορισμούς μη αναγκαίους και επομένως, μπορούν να διαγραφούν.

Ωστόσο, οι παραπάνω περιορισμοί δεν εξασφαλίζουν ότι θα καλυφθεί όλος ο κανονικός χρόνος εργασίας και στη συνέχεια θα γίνουν υπερωρίες. Με βάση, όμως, την σημαντική ιδιότητα του διαχωρίσιμου προγραμματισμού[96], ο αλγόριθμος μεγιστοποίησης της $f(\vec{x})$ δίνει αυτόματα μέγιστη προτεραιότητα στη χρήση του x_{jK} όταν αυξάνεται το x_j από το μηδέν, χωρίς να συμπεριληφθεί ο ειδικός περιορισμός άμεσα στο μοντέλο.

[96] Βλέπε στο παρόν σελ. 62.

Με βάση την παραπάνω μοντελοποίηση, το πρόβλημα διαχωρίσιμου προγραμματισμού έχει μετατραπεί σε πρόβλημα γραμμικού προγραμματισμού και κατ' επέκταση μπορεί να επιλυθεί με τη μέθοδο simplex.

Λύνοντας, λοιπόν, το παραπάνω πρόβλημα με τη μέθοδο simplex με τη βοήθεια του λογισμικού προγράμματος excel – solver παίρνουμε την βέλτιστη λύση

$$x_{1K} = 3$$

$$x_{1Y} = 1$$

$$x_{2K} = 3$$

$$x_{2Y} = 0$$

με μέγιστο κέρδος 26.000 €.

Βιβλιογραφία

- Μπακόπουλος, Α. & Χρυσοβέργης Ι., *Εισαγωγή στην Αριθμητική Ανάλυση*, Αθήνα 1999
- Αργυρός, Σπύρος, *Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης*, Αθήνα [2]2004
- Monge, Gaspard, *Mémoire sur la théorie des déblais et des remblais, Histoire de l'Académie Royale des Sciences de Paris* 1781
- Karush, William, *Minima of Functions of Several Variables with Inequalities as Side Conditions*, M.S. thesis, Department of Mathematics, University of Chicago 1939
- Kuhn, H.W. & Tucker, A.W., *Nonlinear Programming,* Berkeley 1951
- Dupuy , Trevor N., *Numbers, Predictions and War*, Indianapolis 1979
- Trefethen, Lloyd N. & Bau, David III, *Numerical Linear Algebra*, Philadelphia 1997
- Hillier, Federich S. & Liebermann, Gerald J., *Introduction to Operations Research*, New York [7]2001
- Chiang, Mung, *Geometric Programming for Communication Systems*, MA 2005
- Hubbard, John H. & Hubbard, Barbara Burke, *Διανυσματικός Λογισμός, Γραμμική Άλγεβρα και Διαφορικές Μορφές – Μια ενοποιημένη Προσέγγιση,* μετάφραση των Β. Μεταφτσή και Α. Τσολομύτη, Πάτρα 2006
- Sun, Wenyu & Yuan, Ya-Xiang, *Optimization Theory and Methods — Nonlinear Programming*, NY 2006
- Boyd, Stephen & Kim, Seung-Jean & Vandenberghe, Lieven & Hassibi, Arash, *A Tutorial On Geometric Programming*, Published online 2007
- Jahn, Johannes, *Introduction to the Theory of Nonlinear Optimization*, Berlin [3]2007
- www.phpsimplex.com/en/history.htm, 25/07/2011