ΠΡΟΧΕΙΡΕΣ ΣΗΜΕΙΩΣΕΙΣ
ΜΗ ΓΡΑΜΜΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ
ΓΙΑ ΤΟ ΜΑΘΗΜΑ
"ΓΡΑΜΜΙΚΟΣ ΚΑΙ ΜΗ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ"

Ν. Γ. ΧΡΗΣΤΑΚΗΣ

ΗΡΑΚΛΕΙΟ
ΙΟΥΝΙΟΣ 20O5

ΜΗ ΓΡΑΜΜΙΚΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ

– Προβλήματα όπου οι αντικειμενικές συναρτήσεις ή/και οι περιορισμοί δεν είναι γραμμικές συναρτήσεις, αποτελούν προβλήματα μη γραμμικού προγραμματισμού (ΜΓΠ)

Παράδειγμα προβλήματος ΜΓΠ

Μεγιστοποίηση ρυθμού απόδοσης σε μια διαδικασία παραγωγής:
Έστω το κλασσικό πρόβλημα παραγωγής n τύπων προϊόντων με μείξη m τύπων πρώτων υλών. Συμβολίζουμε με b_i τη διαθέσιμη ποσότητα της πρώτης ύλης i και a_{ij} την ποσότητα πρώτης ύλης i που απαιτείται για παρασκευή μίας μονάδας του προϊόντος j. Έστω x_j ο αριθμός προϊόντων j που θα παραχθούν. Κάθε j-προϊόν έχει οικονομική απόδοση r_j και απαιτεί d_j εργατοώρες για να παραχθεί. Η διαδικασία παραγωγής έχει ένα πάγιο κόστος c και απαιτεί d_0 εργατοώρες για να ξεκινήσει. Συνεπώς, η συνολική απόδοση της διαδικασίας είναι $\sum_{j=1}^{n} r_j x_j - c$, ενώ οι απαιτούμενες εργατοώρες είναι $\sum_{j=1}^{n} d_j x_j + d_0$. Ο ρυθμός απόδοσης της διαδικασίας ορίζεται ως η απόδοση ανά εργατοώρα. Για τη μεγιστοποίηση του ρυθμού απόδοσης, κατασκευάζουμε το πρόβλημα ΜΓΠ

$$\max \frac{\sum_{j=1}^{n} r_j x_j - c}{\sum_{j=1}^{n} d_j x_j + d_0}$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i$$

$$i = 1, 2, \ldots, m$$

$$x_j \ge 0 \quad j = 1, 2, \ldots, n$$

Το πρόβλημα αυτό ανήκει στην κατηγορία προβλημάτων ΚΛΑΣΜΑΤΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ

Άλλες κατηγορίες προβλημάτων ΜΓΠ

- προβλήματα ΔΙΑΧΩΡΙΣΙΜΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (π.χ. πρόβλημα μεταφοράς προϊόντων όπου το κόστος μεταφοράς ανά μονάδα προϊόντος δεν είναι σταθερό αλλά μειώνεται όσο αυξάνεται η ποσότητα του μεταφερόμενου προϊόντος)
- προβλήματα ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (όπου οι περιορισμοί είναι γραμμικοί και η αντικειμενική συνάρτηση είναι ένα πολυώνυμο 2ου βαθμού)

Η μαθηματική διατύπωση του γενικού προβλήματος ΜΓΠ είναι η εξής:

Δίνεται πραγματική συνάρτηση $f: \mathbb{R}^n \rightarrow \mathbb{R}$ και ένα σύνολο $F \subseteq \mathbb{R}^n$. Σκοπός είναι να βρεθεί $\underline{x}^* \in F$:

$$f(\underline{x}^*) = \max_{x \in F} f(x)$$

όπου τόσο η f όσο και το F δεν εκφράζονται αποκλειστικά από γραμμικές σχέσεις ως προς τις μεταβλητές $\underline{x} = (x_1, x_2, \ldots, x_n)^T$

- Εάν $F = \mathbb{R}^n$ το πρόβλημα ΜΓΠ ονομάζεται πρόβλημα βελτιστοποίησης ΧΩΡΙΣ ΠΕΡΙΟΡΙΣΜΟΥΣ (unconstrained optimization)
- Όταν το F είναι γνήσιο υποσύνολο του $\mathbb{R}^n$, τότε το πρόβλημα ΜΓΠ ονομάζεται πρόβλημα βελτιστοποίησης ΜΕ ΠΕΡΙΟΡΙΣΜΟΥΣ (constrained optimization).

Η εφικτή περιοχή F εκφράζεται / καθορίζεται μέσω περιορισμών και συναρτήσεων όχι απαραίτητα γραμμικών.

Συνήθως, στην γενική του μορφή το πρόβλημα ΜΓΠ εκφράζεται ως:

$$\max f(\underline{x})$$

με περιορισμούς

$$h_1(\underline{x}) = 0$$
$$\vdots$$
$$h_m(\underline{x}) = 0$$
$$g_1(\underline{x}) \leq 0$$
$$\vdots$$
$$g_p(\underline{x}) \leq 0$$

όπου: $m \leq n$ και οι συναρτήσεις f, h_i $(i=1,\dots,m)$ και g_j $(j=1,\dots,p)$ είναι 2-διαφορίσιμες συναρτήσεις επί του $\mathbb{R}^n$, και συνεχείς.

Η διαφορισιμότητα είναι πολύ σημαντική μιας και αφού γάχνω να ελαχιστοποιήσω ή να μεγιστοποιήσω κάποιες συναρτήσεις, γνωρίζω από τον Διαφορικό Λογισμό ότι θα πρέπει να εξετάσω τις πρώτες και δεύτερες παραγώγους τους.

Στο σημείο αυτό ακολουθούν κάποιοι βασικοί ορισμοί και θεωρήματα που είναι απαραίτητα για τη συνέχεια.

Ορισμός

Έστω $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$ διαφορίσιμη συνάρτηση. Το ανάδελτα της f σε σημείο $\underline{x} \in F$ είναι το διάνυσμα

$$\nabla f(\underline{x}) = \left[\frac{\partial f(\underline{x})}{\partial x_1}, \dots, \frac{\partial f(\underline{x})}{\partial x_n} \right]$$

Ο πίνακας Hesse της f στο $\underline{x} \in F$ είναι ο:

$$Hf(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(\underline{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\underline{x})}{\partial x_n^2} \end{bmatrix}$$

Το ανάδελτα και ο πίνακας Hesse στον λογισμό πολλών μεταβλητών έχουν τη σημασία της πρώτης και δεύτερης παραγώγου αντίστοιχα

Έχουμε ήδη αναφερθεί σε τετραγωνικές μορφές και τους συμμετρικούς πίνακές τους και έχουμε δώσει τα κριτήρια για το πότε αυτοί είναι θετικά ή αρνητικά (ημιο-ρισμένοι. Εδώ θα επαναλάβουμε κάποιες ικανές και αναγκαίες συνθήκες για να είναι ένας πραγματικός συμμετρικός πίνακας A θετικά ορισμένος:

i) Οι ιδιοτιμές του A να είναι θετικές

ii) Οι οδηγοί κύριες ελάσσονες ορίζουσες να είναι θετικές

(για λεπτομέρειες αναφερθείτε στα εισαγωγικά μαθήματα)

Θα δούμε τώρα κάποια βασικά θεωρήματα ύπαρξης τοπικού και ολικού μέγιστου τα οποία θα βοηθήσουν στην επίλυση προβλημάτων είτε με είτε χωρίς περιορισμούς

Θεώρημα (Αναγκαία συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ 2-διαφορίσιμη συνεχής συνάρτηση με τοπικό μέγιστο $\underline{x}^* \in \mathbb{R}^n$. Τότε

i) $\nabla f(\underline{x}^*) = 0$

ii) Ο πίνακας $Hf(\underline{x}^*)$ είναι αρνητικά ημιορισμένος

Θεώρημα (Ικανή συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ διαφορίσιμη συνάρτηση και $x^* \in \mathbb{R}^n$:

i) $\nabla f(x^*) = 0$

ii) Ο πίνακας $Hf(\underline{x}^*)$ είναι αρνητικά ορισμένος

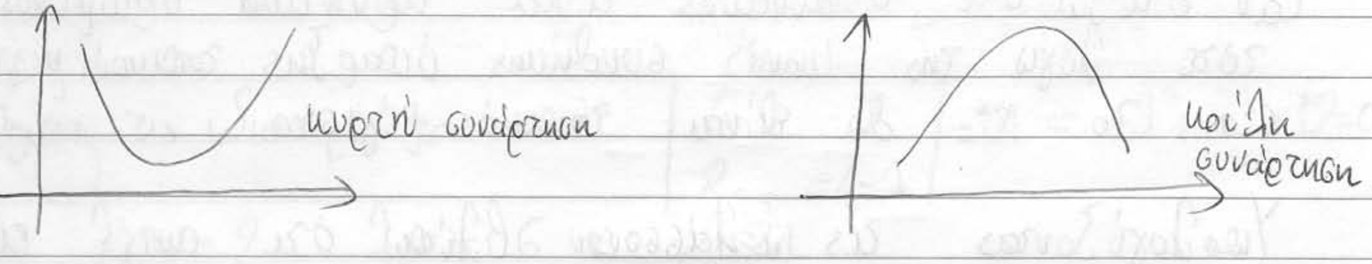
τότε το $\underline{x}^*$ είναι τοπικό μέγιστο.

Θεώρημα ολικού μεγίστου

Έστω $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$ κοίλη συνάρτηση και F κυρτό σύνολο. Τότε κάθε τοπικό μέγιστο $\underline{x}^*$ της f είναι και ολικό μέγιστο, και συνεπώς λύση του προβλήματος ΜΓΠ χωρίς περιορισμούς $\max f(\underline{x})$, $\underline{x} \in \mathbb{R}^n$

Θεώρημα καθορισμού κυρτότητας (κοιλότητας)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$ 2-διαφορίσιμη συνεχής συνάρτηση. Η f είναι κυρτή (κοίλη) αν και μόνο αν ο πίνακας $Hf(\underline{x})$ είναι θετικά (αρνητικά) ημιορισμένος $\forall\, \underline{x} \in \mathbb{R}^n$.



Παράδειγμα επίλυσης προβλήματος ΜΓΠ χωρίς περιορισμούς

$$\max (x_1 + 2x_3 + x_2 x_3 - x_1^2 - x_2^2 - x_3^2)$$
$$x_1, x_2, x_3 \in \mathbb{R}.$$

Η αναγκαία συνθήκη τοπικού μεγίστου $\nabla f(\underline{x}) = 0$ είναι:

$$\frac{\partial f}{\partial x_1}(\underline{x}) = 1 - 2x_1 = 0$$

$$\frac{\partial f}{\partial x_2}(\underline{x}) = x_3 - 2x_2 = 0$$

$$\frac{\partial f}{\partial x_3}(\underline{x}) = 2 + x_2 - 2x_3 = 0$$

$$\Downarrow$$

Το $\underline{x}^* = [1/2, 2/3, 4/3]^T$ είναι υποψήφιο μέγιστο. Εξετάζουμε τώρα τον πίνακα Hesse σε τυχόν σημείο $\underline{x}$:

$$Hf(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{bmatrix} \longrightarrow \text{συμμετρικός πίνακας}$$

Εάν δείξω ότι ο πίνακας είναι αρνητικά ορισμένος, τότε λόγω της ικανής συνθήκης ύπαρξης τοπικού μεγίστου το x^* θα είναι τοπικό μέγιστο

Υπολογίζοντας τις ιδιοτιμές του βλέπω ότι αυτές είναι -2, -3 και -1, όλες αρνητικές $\Rightarrow$ ο $Hf(x)$ είναι αρνητικά ορισμένος. Ταυτόχρονα, λόγω του ιδίου γεγονότος η f είναι κοίλη συνάρτηση και αφού η εφικτή περιοχή $F = \mathbb{R}^n$ είναι κυρτό σύνολο, το σημείο $x^* = (1/2, 2/3, 4/3)$ είναι και ολικό μέγιστο, άρα και λύση του προβλήματος ΜΓΠ.

Παράδειγμα προβλήματος ΜΓΠ. χωρίς περιορισμούς

$$\max\left(4x_1 + 6x_2 - 2x_1^2 - 2x_1x_2 - 2x_2^2\right)$$
$$x_1, x_2 \in \mathbb{R}$$

Αναγκαία συνθήκη τοπικού μεγίστου $\nabla f(\underline{x}) = 0$:

$$\left.\begin{aligned} \frac{\partial f}{\partial x_1} &= 4 - 4x_1 - 2x_2 = 0 \\ \frac{\partial f}{\partial x_2} &= 6 - 2x_1 - 4x_2 = 0 \end{aligned}\right\} \Rightarrow \overset{\text{πιθανό μέγιστο}}{x^* = [1/3,\ 4/3]^T}$$

Εξετάζω τον $Hf(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} -4 & -2 \\ -2 & -4 \end{bmatrix}$

Βρίσκω τις ιδιοτιμές του από: $\begin{vmatrix} -4-\lambda & -2 \\ -2 & -4-\lambda \end{vmatrix} = 0 \Rightarrow +\lambda^2 + 8\lambda + 12 = 0$

$\Rightarrow \quad \lambda_1 = -2 \qquad \lambda_2 = -6$

$\lambda_1, \lambda_2 < 0 \Rightarrow$ το x^* είναι τοπικό μέγιστο

Επειδή όμως ο $Hf(x)$ είναι αρνητικά ορισμένος η f είναι κοίλη $\Rightarrow$ το x^* είναι ολικό μέγιστο επειδή η εφικτή περιοχή $F = \mathbb{R}^2$

Επίλυση προβλημάτων ΜΓΠ. με εξισωτικούς περιορισμούς

Η γενική μορφή προβλημάτων αυτού του τύπου ανάγονται σε:

$$\max f(\underline{x})$$

με περιορισμούς

$$h_1(\underline{x}) = 0$$
$$\vdots$$
$$h_m(\underline{x}) = 0$$

όπου $\underline{x} \in \mathbb{R}^n$, $m \leq n$ και $f \in C^2(\mathbb{R}^n)$, $h_i \in C^2(\mathbb{R}^n)$ $i=1,2,\ldots,m$, δηλαδή οι f και h_i είναι συνεχώς 2-διαφορίσιμες συναρτήσεις επί του $\mathbb{R}^n$.

Με διανυσματικό συμβολισμό $\underline{h} = [h_1, h_2, \ldots, h_m]^T$, το πρόβλημα ΜΓΠ. γράφεται ως:

$$\max f(\underline{x})$$

με περιορισμούς

$$\underline{h}(\underline{x}) = \underline{0}$$
$$\underline{x} \in \mathbb{R}^n$$

Για την επίλυση προβλημάτων τέτοιου τύπου θα χρησιμοποιήσουμε στοιχεία από τη γεωμετρία καμπυλών και επιφανειών.

Η εφικτή περιοχή F είναι τώρα κάποια υπερεπιφάνεια εμφυτευμένη στον $\mathbb{R}^n$. Εάν οι περιορισμοί που την ορίζουν:

$$h_1(\underline{x}) = 0$$
$$h_2(\underline{x}) = 0$$
$$\vdots$$
$$h_m(\underline{x}) = 0$$

είναι ανεξάρτητοι, η υπερεπιφάνεια που ορίζουν έχει

διάσταση $n-m$. Εάν επιπλέον οι h_i είναι C^1 συναρτήσεις, τότε η υπερεπιφάνεια ονομάζεται λεία.
Στα προβλήματα ΜΓΠ οι υπερεπιφάνειες των εφικτών περιοχών είναι λείες, μας και η αρχική θεώρηση είναι ότι $h_i \in C^2(\mathbb{R}^n)$

Ορισμοί

i) $\forall\ \underline{x} \in F$, το ανάδελτα της $\underline{h}$ στο $\underline{x}$ ορίζεται ως:

$$\nabla \underline{h}(\underline{x}) = \begin{bmatrix} \frac{\partial h_1(\underline{x})}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \vdots & & & \\ \frac{\partial h_m(x)}{\partial x_1} & \frac{\partial h_m(\underline{x})}{\partial x_2} & \cdots & \frac{\partial h_m(x)}{\partial x_n} \end{bmatrix} =$$

$$= \begin{bmatrix} \nabla h_1(\underline{x}) \\ \vdots \\ \nabla h_m(\underline{x}) \end{bmatrix} \quad (m \times n) \text{ πίνακας}$$

ii) $\forall\ \underline{x} \in F$ και $\underline{\lambda} \in \mathbb{R}^m$ ορίζουμε τον πίνακα

$$H(\underline{\lambda}^T \underline{h}) = \sum_{i=1}^{m} \lambda_i\, H h_i(\underline{x}) \quad \text{όπου}$$

$\underline{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_m]^T$ διάνυσμα και ο πίνακας $H h_i(\underline{x})$ ο πίνακας Hesse της συνάρτησης h_i στο $\underline{x}$

iii) Εφαπτόμενο επίπεδο $M_{\underline{x}^*}^F$ της F στο $\underline{x}^* \in F$ ορίζεται το σύνολο των παραγώγων στο $\underline{x}^*$ όλων των διαφορίσιμων καμπυλών που περνάνε από το $\underline{x}^*$

iv) Σημείο $x^* \in F$ λέγεται κανονικό σημείο των περιορισμών $\underline{h}(\underline{x}) = \underline{0}$ εάν τα διανύσματα $\nabla h_1(\underline{x}^*)$, $\nabla h_2(\underline{x}^*)$, ..., $\nabla h_m(\underline{x}^*)$ είναι γραμμικά ανεξάρτητα.

Θεώρημα (Ορισμός εφαπτόμενου επιπέδου)

Έστω $x^* \in F$ ένα κανονικό σημείο των περιορισμών $\underline{h}(\underline{x}) = \underline{0}$ που ορίζουν την επιφάνεια F. Τότε το εφαπτόμενο επίπεδο $M^F_{x^*}$ της F στο $\underline{x}^*$ είναι ίσο με:

$$\boxed{M^{\underline{h}}_{\underline{x}^*} = \{ \underline{y} \in \mathbb{R}^n : \nabla \underline{h}(\underline{x}) \underline{y} = \underline{0} \}}$$

Με βάση τους ορισμούς και το θεώρημα αυτό θα αναπτύξουμε τις συνθήκες για τοπικά μέγιστα κάτω από εξισωτικούς περιορισμούς.

Θεώρημα (Ασθενής αναγκαία συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$, $\underline{h}: \mathbb{R}^n \to \mathbb{R}^m$ $(m \leq n)$ C^1-συναρτήσεις και F η επιφάνεια που ορίζεται από τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$. Έστω x^* κανονικό σημείο των περιορισμών $\underline{h}(\underline{x}) = \underline{0}$, το οποίο είναι τοπικό μέγιστο της f υπό αυτούς τους περιορισμούς. Τότε

$$\nabla f(x^*) \underline{y} = 0$$

$\forall \ \underline{y} \in M^F_{x^*}$

Θεώρημα (Αναγκαία συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$, $\underline{h}: \mathbb{R}^n \to \mathbb{R}^m$ $(m \leq n)$ C^2-συναρτήσεις και F η επιφάνεια που ορίζεται από τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$. Αν το $\underline{x}^*$ είναι τοπικό μέγιστο της f υπό τους περιορισμούς $\underline{h}(x) = \underline{0}$ και επιπλέον είναι κανονικό σημείο της F, τότε

i) $\exists \ \underline{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_m]^T \in \mathbb{R}^m$:

$$\nabla f(\underline{x}^*) - \underline{\lambda}^T \nabla \underline{h}(\underline{x}^*) = \underline{0}$$

ii) Ο πίνακας
$$L(\underline{x}^*) = Hf(\underline{x}^*) - H(\underline{\lambda}^T \underline{h}(\underline{x}^*))$$
είναι αρνητικά ημιορισμένος στο εφαπτόμενο επίπεδο $M^F_{\underline{x}^*}$, δηλαδή ισχύει ότι $\underline{y}^T L(\underline{x}^*) \underline{y} \leq 0 \quad \forall\, \underline{y} \in M^F_{\underline{x}^*}$

Θεώρημα (Ικανή συνθήκη τοπικού μεγίστου)
Έστω $f: \mathbb{R}^n \to \mathbb{R}$, $\underline{h}: \mathbb{R}^n \to \mathbb{R}^m$ $(m \leq n)$ C^2-συναρτήσεις και F η επιφάνεια που ορίζεται από τους περιορισμούς $\underline{h}(x) = \underline{0}$.
Έστω $\underline{x}^* \in F$ και $\underline{\lambda} \in \mathbb{R}^m$:
i) $\nabla f(\underline{x}^*) - \underline{\lambda}^T \nabla \underline{h}(\underline{x}^*) = \underline{0}$
ii) Ο πίνακας
$$L(\underline{x}^*) = Hf(\underline{x}^*) - H(\underline{\lambda}^T \underline{h}(\underline{x}^*))$$
είναι αρνητικά ορισμένος στο επίπεδο $M^{\underline{h}}_{\underline{x}^*} = \{\underline{y} \in \mathbb{R}^n : \nabla \underline{h}(\underline{x}^*)\underline{y} = \underline{0}\}$, δηλαδή: $\underline{y}^T L(\underline{x}^*) \underline{y} < 0$ για $\underline{y} \in M^{\underline{h}}_{\underline{x}^*}$ και $\underline{y} \neq \underline{0}$.

Τότε, το $\underline{x}^*$ είναι τοπικό μέγιστο της f υπό τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$.

Από τις συνθήκες που αναπτύξαμε γίνεται φανερό ότι για την εύρεση τοπικών ακροτάτων σημαντικό ρόλο παίζει η ΣΥΝΑΡΤΗΣΗ LAGRANGE

$$\ell(\underline{x}, \underline{\lambda}) = f(\underline{x}) - \underline{\lambda}^T \underline{h}(\underline{x})$$

όπου οι περιορισμοί $\underline{h}(\underline{x}) = \underline{0}$ και οι αναγκαίες συνθήκες για την ύπαρξη τοπικών ακροτάτων γράφονται ενιαία ως:
$$\nabla \ell(\underline{x}^*, \underline{\lambda}^*) = \underline{0}$$

ή πιο αναλυτικά:

$$\frac{\partial f(\underline{x}^*)}{\partial x_j} - \sum_{i=1}^{m} \lambda_i^* \frac{\partial h_i(\underline{x}^*)}{\partial x_j} = 0, \quad (j=1,2,\dots,n)$$

$$h_i(\underline{x}^*) = 0, \quad (i=1,2,\dots,m)$$

Επίσης, ο πίνακας $L(\underline{x}^*)$ μπορεί να θεωρηθεί ως ο πίνακας Hesse της $\ell(\underline{x}, \underline{\lambda})$, όπου τα λ θα λογίζονται ως παράμετροι και όχι ως μεταβλητές, κι έτσι θα πάρουμε παραγώγους μόνο ως προς τα $\underline{x}$.

Μεθοδολογία επίλυσης για εύρεση τοπικού μεγίστου

- Κατασκευάζουμε τη συνάρτηση Lagrange $\ell(\underline{x}, \underline{\lambda})$ και επιλύουμε την $\nabla \ell(\underline{x}, \underline{\lambda}) = \underline{0}$
- Ελέγχουμε για κάθε λύση ξεχωριστά εάν ο $L(\underline{x}^*)$ είναι αρνητικά ορισμένος.
- Εάν η εφικτή περιοχή F είναι φραγμένη, τότε το μέγιστο από τα τοπικά μέγιστα θα είναι και ολικό μέγιστο, δηλ. λύση του προβλήματος ΜΓΠ.

Η μέθοδος αυτή είναι γνωστή ως ΜΕΘΟΔΟΣ ΠΟΛΛΑΠΛΑΣΙΑΣΤΩΝ LAGRANGE.

Παράδειγμα 1 επίλυσης προβλήματος ΜΓΠ με εξισωτικούς περιορισμούς

Να ευρεθούν τα τοπικά ακρότατα (x_1^*, x_2^*, x_3^*) της:

$$f(x_1, x_2, x_3) = x_1 \cdot x_2 \cdot x_3$$

όταν

$$h(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1 = 0$$

με $\underline{x}^* > \underline{0}$.

Λύση: Η συνάρτηση Lagrange είναι:

$$\ell(x_1, x_2, x_3, \lambda) = x_1 \cdot x_2 \cdot x_3 - \lambda(x_1^2 + x_2^2 + x_3^2 - 1)$$

$$\nabla \ell(\underline{x}, \underline{\lambda}) = 0 \Rightarrow$$

$$\frac{\partial \ell}{\partial x_1} = x_2 x_3 - 2\lambda x_1 = 0$$

$$\frac{\partial \ell}{\partial x_2} = x_1 x_3 - 2\lambda x_2 = 0$$

$$\frac{\partial \ell}{\partial x_3} = x_1 x_2 - 2\lambda x_3 = 0$$

$$\frac{\partial \ell}{\partial \lambda} = -h(x_1, x_2, x_3) = 0 \Rightarrow h(x_1, x_2, x_3) = 0 \Rightarrow x_1^2 + x_2^2 + x_3^2 - 1 = 0$$

Πολλ/ζοντας την 1η εξίσωση με x_1, τη 2η με x_2 και την 3η με x_3 βλέπουμε ότι:

$$2\lambda x_1^2 = 2\lambda x_2^2 = 2\lambda x_3^2 = x_1 x_2 x_3$$

$$\Rightarrow x_1^2 = x_2^2 = x_3^2 = \frac{x_1 x_2 x_3}{2\lambda}$$

Αντικαθιστώντας στην 4η εξίσωση για να βρω το λ^* και κρατώντας από τις εξισώσεις και τον περιορισμό $\underline{x}^* > \underline{0}$ ότι $x_1^* = x_2^* = x_3^* = 2\lambda$ βρίσκω ότι

$$\lambda^* = \frac{1}{\sqrt{12}} = \frac{\sqrt{3}}{6} \Rightarrow$$

$$[x_1^*, x_2^*, x_3^*, \lambda^*] = [\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/6]$$

Υπολογίζω τώρα το επίπεδο $M^{1}_{\underline{x}^*}$

$$\nabla h(x_1^*, x_2^*, x_3^*) = \left[\frac{\partial h(\underline{x}^*)}{\partial x_1}, \frac{\partial h(\underline{x}^*)}{\partial x_2}, \frac{\partial h}{\partial x_3}(\underline{x}^*)\right] =$$

$$= [2x_1^*, 2x_2^*, 2x_3^*]$$

$$= [2\sqrt{3}/3, 2\sqrt{3}/3, 2\sqrt{3}/3]$$

$$M^{1}_{\underline{x}^*} = \{\underline{y} = [y_1, y_2, y_3]^T \in \mathbb{R}^3 : \nabla h(\underline{x}^*)\,\underline{y} = 0\}$$

$$= \left\{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : \frac{2\sqrt{3}}{3}y_1 + \frac{2\sqrt{3}}{3}y_2 + \frac{2\sqrt{3}}{3}y_3 = 0\right\}$$

$$= \{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : y_1 + y_2 + y_3 = 0\}$$

Πρέπει να υπολογίσω τώρα και την $L(x^*)$, η οποία όπως είπαμε θα είναι ο πίνακας των 2ων παραγώγων του $\ell(\underline{x}, \underline{\lambda}^*)$ μόνο ως προς $\underline{x}$.

Έτσι: $\ell(x, \underline{\lambda}^*) = x_1x_2x_3 - \frac{\sqrt{3}}{6}(x_1^2 + x_2^2 + x_3^2 - 1)$

$$L(\underline{x}^*) = \begin{bmatrix} \frac{\partial^2 \ell(x^*)}{\partial x_1^2} & \frac{\partial^2 \ell(\underline{x}^*)}{\partial x_1 \partial x_2} & \frac{\partial^2 \ell}{\partial x_1 \partial x_3}(\underline{x}^*) \\ \frac{\partial^2 \ell}{\partial x_2 \partial x_1}(x^*) & \frac{\partial^2 \ell}{\partial x_2^2}(x^*) & \frac{\partial^2 \ell}{\partial x_2 \partial x_3}(\underline{x}^*) \\ \frac{\partial^2 \ell}{\partial x_3 \partial x_1}(x^*) & \frac{\partial^2 \ell}{\partial x_3 \partial x_2}(x^*) & \frac{\partial^2 \ell}{\partial x_3^2}(\underline{x}^*) \end{bmatrix}$$

$$= \begin{bmatrix} -2\lambda^* & x_3^* & x_2^* \\ x_3^* & -2\lambda^* & x_1^* \\ x_2^* & x_1^* & -2\lambda^* \end{bmatrix} = \frac{\sqrt{3}}{3}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$$

Για να αποδείξουμε ότι ο $L(\underline{x}^*)$ είναι αρνητικά ορισμένος στο $M^h_{\underline{x}^*}$ πρέπει να πάρουμε μια βάση του $M^h_{\underline{x}^*}$:

$$M^h_{\underline{x}^*} = \{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : y_1 + y_2 + y_3 = 0\}$$

$$= \{[y_1, y_2, -y_1 - y_2]^T : y_1, y_2 \in \mathbb{R}\}$$

$$= \{y_1[1, 0, -1]^T + y_2[0, 1, -1]^T : y_1, y_2 \in \mathbb{R}\}$$

$$\Downarrow$$

ελέγχω την τετραγωνική μορφή $\underline{y}^T L(\underline{x}^*)\, \underline{y}$:

$$[y_1, y_2, -y_1 - y_2]\ \frac{\sqrt{3}}{3}\begin{bmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \\ -y_1 - y_2 \end{bmatrix} =$$

$$= \frac{\sqrt{3}}{3}\left(-2y_1^2 - 2y_2^2 - 2(y_1 + y_2)^2\right)$$

Η παράσταση αυτή ΠΡΟΦΑΝΩΣ είναι < 0 $\forall$ $y_1, y_2 \in \mathbb{R}$ και $y_1, y_2 \neq 0$

Άρα το σημείο $[x_1^*, x_2^*, x_3^*] = [\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3]$ είναι τοπικό μέγιστο και μάλιστα είναι και ολικό μέγιστο επειδή η εφικτή περιοχή είναι φραγμένη.

Παράδειγμα 2 επίλυσης προβλήματος ΜΓΠ με εξισωτικούς περιορισμούς

Να βρεθούν τα τοπικά μέγιστα της

$$f(x, y, z) = -(x+y+z)$$

όταν

$$x^2 + y = 3$$
$$x + 3y + 2z = 7$$

Λύση:

$$h_1(x,y,z) = x^2 + y - 3 = 0$$
$$h_2(x,y,z) = x + 3y + 2z - 7 = 0$$

Η συνάρτηση Lagrange:

$$\ell(x,y,z,\lambda_1,\lambda_2) = -x-y-z-\lambda_1(x^2+y-3)-\lambda_2(x+3y+2z-7)$$

$$\nabla \ell(\underline{x}, \underline{\lambda}) = 0$$

$$\Downarrow$$

$$\frac{\partial \ell}{\partial x} = -1 - 2\lambda_1 x - \lambda_2 = 0$$

$$\frac{\partial \ell}{\partial y} = -1 - \lambda_1 - 3\lambda_2 = 0$$

$$\frac{\partial \ell}{\partial z} = -1 - 2\lambda_2 = 0$$

$$\frac{\partial \ell}{\partial \lambda_1} = 0 \Rightarrow h_1(\underline{x}) = 0 \Rightarrow x^2 + y - 3 = 0$$

$$\frac{\partial \ell}{\partial \lambda_2} = 0 \Rightarrow h_2(\underline{x}) = 0 \Rightarrow x + 3y + 2z - 7 = 0$$

$$\Downarrow$$

Πιθανό ακρότατο το:

$$[x^*, y^*, z^*, \lambda_1^*, \lambda_2^*] = [-1/2, 11/4, -3/8, 1/2, -1/2]$$

Εύρεση των $\nabla h_1(\underline{x})$ και $\nabla h_2(\underline{x})$ για τον καθορισμό του επιπέδου $M^h_{\underline{x}^*}$:

$$\nabla h_1(\underline{x}^*) = \left[\frac{\partial h_1(\underline{x}^*)}{\partial x}, \frac{\partial h_1(\underline{x}^*)}{\partial y}, \frac{\partial h_1(\underline{x}^*)}{\partial z}\right] =$$

$$= [2x^*, 1, 0] = [-1, 1, 0]$$

και αναλόγως

$$\nabla h_2(\underline{x}^*) = [1, 3, 2]$$

$$\Downarrow$$

$$M^h_{\underline{x}^*} = \{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : \nabla h_1(\underline{x}^*)\underline{y} = 0,\ \nabla h_2(\underline{x}^*)\underline{y} = 0\}$$

$$= \{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : -y_1 + y_2 = 0,\ y_1 + 3y_2 + 2y_3 = 0\}$$

Για να βρούμε τον πίνακα $L(\underline{x}^*)$ παίρνω τη συνάρτηση Lagrange ως:

$$\ell(\underline{x}, \lambda_1^*, \lambda_2^*) = -x-y-z - \frac{1}{2}(x^2+y-3) + \frac{1}{2}(x+3y+2z-7)$$

$$\Downarrow$$

$$L(\underline{x}^*) = \begin{bmatrix} \frac{\partial^2 \ell}{\partial x^2}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial x \partial y}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial x \partial z}(\underline{x}^*, \underline{\lambda}^*) \\ \frac{\partial^2 \ell}{\partial y \partial x}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial y^2}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial y \partial z}(\underline{x}^*, \underline{\lambda}^*) \\ \frac{\partial^2 \ell}{\partial z \partial x}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial z \partial y}(\underline{x}^*, \underline{\lambda}^*) & \frac{\partial^2 \ell}{\partial z^2}(\underline{x}^*, \underline{\lambda}^*) \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Για να εξετάσω αν ο $L(\underline{x}^*)$ είναι αρνητικά ορισμένος στο $M^h_{\underline{x}^*}$ πρέπει πρώτα να βρω μια βάση του $M^h_{\underline{x}^*}$:

$$M^h_{\underline{x}^*} = \{[y_1, y_2, y_3]^T \in \mathbb{R}^3 : -y_1 + y_2 = 0,\ y_1 + 3y_2 + 2y_3 = 0\}$$

$$= \{[y_1, y_1, -2y_1]^T \in \mathbb{R}^3 : y_1 \in \mathbb{R}\}$$

$$= \{y_1 [1, 1, -2]^T : y_1 \in \mathbb{R}\}$$

Εξετάζω τώρα την τετραγωνική μορφή $\underline{y}^T L(\underline{x}^*)\, \underline{y}$

$$[y_1, y_1, -2y_1] \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_1 \\ -2y_1 \end{bmatrix}$$

$$= -y_1^2 < 0 \quad \text{ΠΡΟΦΑΝΩΣ} \quad \forall\, y_1 \neq 0$$

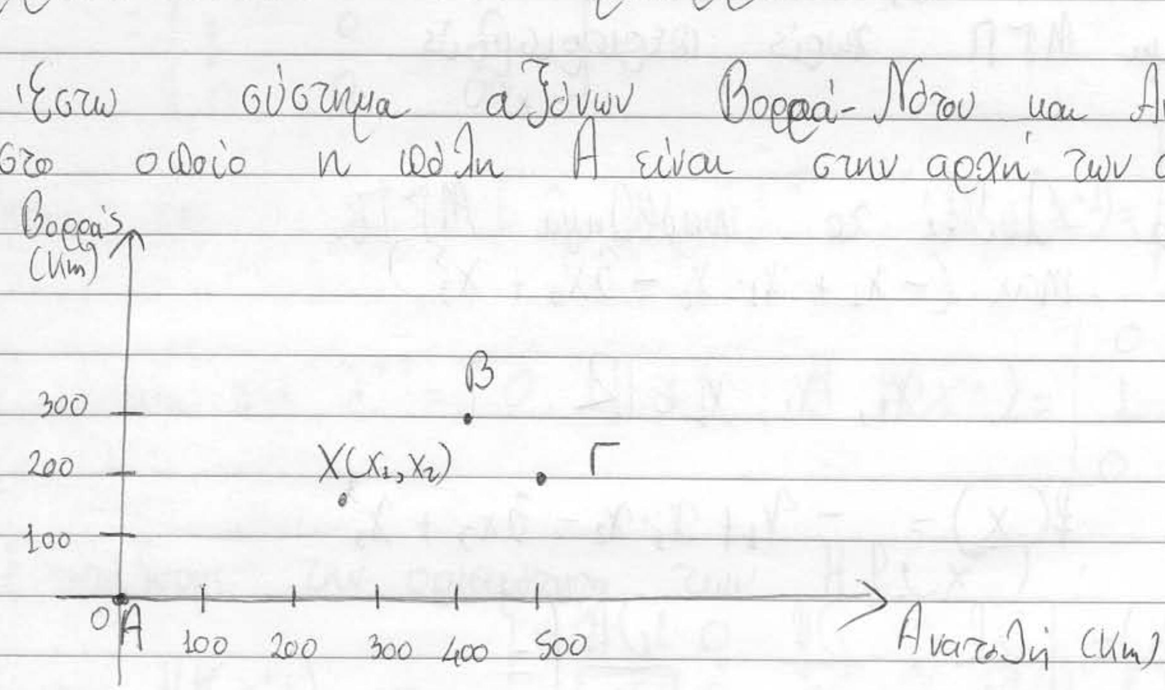
$$\Downarrow$$

το σημείο $[x^*, y^*, z^*] = [-1/2, 11/4, -3/8]$ είναι το μοναδικό τοπικό μέγιστο της f με τους δοσμένους περιορισμούς.

Ασκήσεις

1. Μια εταιρεία θέλει να εγκαταστήσει ένα εργοστάσιο που παράγει 100 τόνους τσιμέντου ημερησίως και το οποίο διανέμεται στις πόλεις Α, Β, Γ σε ποσότητες 50, 30 και 20 τόνων αντίστοιχα. Η πόλη Β βρίσκεται 300 Km βόρεια και 400 Km ανατολικά της Α, ενώ η Γ βρίσκεται 200 Km βόρεια και 500 Km ανατολικά της Α. Αν η μεταφορά ενός τόνου τσιμέντου προς οποιαδήποτε πόλη κοστίζει 0,01 ευρώ/Km, να διατυπωθεί ένα πρόβλημα μη γραμμικού προγραμματισμού για την εύρεση της θέσης του εργοστασίου που ελαχιστοποιεί το ημερήσιο συνολικό κόστος μεταφοράς.

Έστω σύστημα αξόνων Βορρά-Νότου και Ανατολής-Δύσης στο οποίο η πόλη Α είναι στην αρχή των αξόνων:



Έστω Χ με συντεταγμένες (x_1, x_2) η θέση του εργοστασίου. Οι αποστάσεις του Χ από τις πόλεις Α, Β, Γ θα είναι

$$d_A = \sqrt{x_1^2 + x_2^2},$$
$$d_B = \sqrt{(x_1 - 400)^2 + (x_2 - 300)^2},$$
$$d_\Gamma = \sqrt{(x_1 - 500)^2 + (x_2 - 200)^2},$$

Το συνολικό κόστος μεταφοράς θα είναι

$$0,01(50 d_A + 30 d_B + 20 d_\Gamma)$$

$$\Downarrow$$

Το πρόβλημα μη γραμμικού προγραμματισμού για την εύρεση της θέσης του εργοστασίου θα είναι

$$\min \left[0,01\left(50\sqrt{x_1^2+x_2^2} + 30\sqrt{(x_1-400)^2+(x_2-300)^2} + 20\sqrt{(x_1-500)^2+(x_2-200)^2}\right)\right]$$

$$x_1, x_2 \in \mathbb{R}$$

Πρόβλημα ΜΓΠ χωρίς περιορισμούς

2. Να λυθεί το πρόβλημα ΜΓΠ.

$$\max\ (-x_1 + x_1 \cdot x_2 - 3x_3 + x_3^3)$$

$$x_1, x_2, x_3 \in \mathbb{R}$$

Λύση: $f(\underline{x}) = -x_1 + x_1 \cdot x_2 - 3x_3 + x_3^3$

$$\nabla f(\underline{x}) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}\right] = \left[-1+x_2,\ x_1,\ -3+3x_3^2\right]$$

$$\nabla f(\underline{x}) = 0 \Rightarrow \begin{cases} -1+x_2 = 0 \\ x_1 = 0 \\ -3+3x_3^2 = 0 \end{cases} \Rightarrow \begin{cases} x_1 = 0 \\ x_2 = 1 \\ x_3 = \pm 1 \end{cases}$$

Άρα 2 είναι τα πιθανά ακρότατα $[x_1^*, x_2^*, x_3^*]^T$, τα $[0, 1, 1]^T$ και $[0, 1, -1]^T$

Πρέπει τώρα να καθορίσουμε τον πίνακα $Hf(\underline{x})$ στα $\underline{x}^*$:

$$Hf(\underline{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^3} \end{bmatrix} =$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 6x_3 \end{bmatrix}$$

Άρα για $\underline{x}^{*1} = [0, 1, 1]^T$ $Hf(\underline{x}^{*1}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 6 \end{bmatrix}$

και για $\underline{x}^{*2} = [0, 1, -1]^T$ $Hf(\underline{x}^{*2}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -6 \end{bmatrix}$

Εξετάζουμε την ορισμότητα των $Hf(\underline{x}^*)$:

Για τον $Hf(\underline{x}^{*1}) \Rightarrow \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 6-\lambda \end{vmatrix} = (6-\lambda)(\lambda^2-1) = 0$

$$\Rightarrow \begin{cases} \lambda = 6 \\ \lambda = 1 \\ \lambda = -1 \end{cases}$$

$\Rightarrow$ μη-ορισμένος

Για τον $Hf(\underline{x}^{*2}) \Rightarrow \begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & -6-\lambda \end{vmatrix} = (-6-\lambda)(\lambda^2-1) = 0 \Rightarrow \begin{cases} \lambda = -6 \\ \lambda = 1 \\ \lambda = -1 \end{cases}$

$\Rightarrow$ μη-ορισμένος

εφόσον δεν ορίζεται ορισμότητα, αυτό σημαίνει ότι η αντικειμενική συνάρτηση μπορεί να πάρει οσοδήποτε μεγάλες τιμές $\Rightarrow$ το πρόβλημα είναι μη φραγμένο και δεν μπορούμε λοιπόν να υπολογίσουμε βέλτιστη λύση.

3. Να ευρεθεί η λύση στο πρόβλημα ΜΤΠ

$$\max\left[(x_1-3)^2+(x_2-4)^2\right]$$

αν

$$x_1^2+x_2^2=25$$

$$x_1, x_2 \in \mathbb{R}$$

Λύση: $f(\underline{x}) = (x_1-3)^2+(x_2-4)^2$

$$h(\underline{x}) = x_1^2+x_2^2-25=0$$

Βλέπουμε ότι η συνάρτηση για μεγιστοποίηση ορίζει έναν κύκλο με κέντρο το σημείο $(3,4)$ στον $\mathbb{R}^2$. Ο περιορισμός ορίζει έναν κύκλο στον $\mathbb{R}^2$ με κέντρο την αρχή των αξόνων και ακτίνα 5.

Κατασκευάζουμε τη συνάρτηση Lagrange:

$$\ell(\underline{x},\underline{\lambda}) = f(\underline{x}) - \lambda h(\underline{x}) =$$
$$= (x_1-3)^2+(x_2-4)^2 - \lambda(x_1^2+x_2^2-25)$$

Για την ύπαρξη τοπικού ακρότατου απαιτείται

$$\nabla \ell(\underline{x},\underline{\lambda}) = \underline{0}$$

$$\Rightarrow \frac{\partial \ell}{\partial x_1} = 2(x_1-3) - 2\lambda x_1 = 0$$

$$\frac{\partial \mathcal{L}}{\partial x_2} = 2(x_2-4) - 2\lambda x_2 = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = 0 \Rightarrow x_1^2 + x_2^2 - 25 = 0$$

$$\Rightarrow \left.\begin{array}{l} x_1(1-\lambda) = 3 \\ x_2(1-\lambda) = 4 \\ x_1^2 + x_2^2 = 25 \end{array}\right\} \Rightarrow$$ έχει 2 λύσεις επειδή από την $3^η$ εξίσωση $\lambda = \begin{cases} 2 \\ 0 \end{cases}$

Πιθανά ακρότατα τα: $[x_1^*, x_2^*, \lambda^*] = \begin{cases} [-3, -4, 2] \\ [3, 4, 0] \end{cases}$

Για να υπολογίσω την ορισμότητα της $2^{ης}$ παραγώγου πρέπει πρώτα να υπολογίσω για τα σημεία αυτά το επίπεδο $M_{\underline{x}^*}^h$. Για το σκοπό αυτό πρέπει να υπολογίσω το $\nabla h(\underline{x}^*)$

Για το $\nabla h(\underline{x}^*)$ έχω ότι:

$$\nabla h(\underline{x}^*) = \left[\frac{\partial h(\underline{x}^*)}{\partial x_1}, \frac{\partial h(\underline{x}^*)}{\partial x_2}\right] = [2x_1^*, 2x_2^*] =$$

$$= \begin{cases} [6, 8] & \text{για } \lambda^* = 0 \\ [-6, -8] & \text{για } \lambda^* = 2 \end{cases}$$

Άρα $M_{\underline{x}^*}^h = \left\{ \underline{y} = [y_1, y_2]^T \in \mathbb{R}^2 : \nabla h(\underline{x}) \underline{y} = 0 \right\}$

$$= \begin{cases} \{[y_1, y_2]^T \in \mathbb{R}^2 : 6y_1 + 8y_2 = 0\} \\ \{[y_1, y_2]^T \in \mathbb{R}^2 : -6y_1 - 8y_2 = 0\} \end{cases} =$$

$$= \{[y_1, y_2]^T \in \mathbb{R}^2 : 3y_1 + 4y_2 = 0\} = \{[y_1, -\tfrac{3}{4}y_1]^T : y_1 \in \mathbb{R}\}$$

$$= \{y_1 [1, -3/4]^T : y_1 \in \mathbb{R}\}$$

Όπως βλέπω, για τα 2 πιθανά ακρότατα τα επίπεδα τους $M^{\mathbf{h}}_{x^*}$ ταυτίζονται:

$$L(x^*) = \begin{bmatrix} \frac{\partial^2 \ell}{\partial x_1^2} & \frac{\partial^2 \ell}{\partial x_1 \partial x_2} \\ \frac{\partial^2 \ell}{\partial x_2 \partial x_1} & \frac{\partial^2 \ell}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2-2\lambda^* & 0 \\ 0 & 2-2\lambda^* \end{bmatrix}$$

$$\Rightarrow \begin{cases} \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \text{ για } \lambda^* = 0 \\ \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \text{ για } \lambda^* = 2 \end{cases}$$

Άρα για $\lambda^* = 0$:

$$[y_1, -3/4\, y_1] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} y_1 \\ -3/4\, y_1 \end{bmatrix} = \frac{25}{8} y_1^2 > 0 \quad \forall y_1 \neq 0$$

$\Rightarrow$ Ο πίνακας είναι θετικά ορισμένος για $[3, 4, 0]$

Για $\lambda^* = 2$:

$$[y_1, -3/4\, y_1] \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ -3/4\, y_1 \end{bmatrix} = -\frac{25}{8} y_1^2 < 0 \quad \forall y_1 \neq 0$$

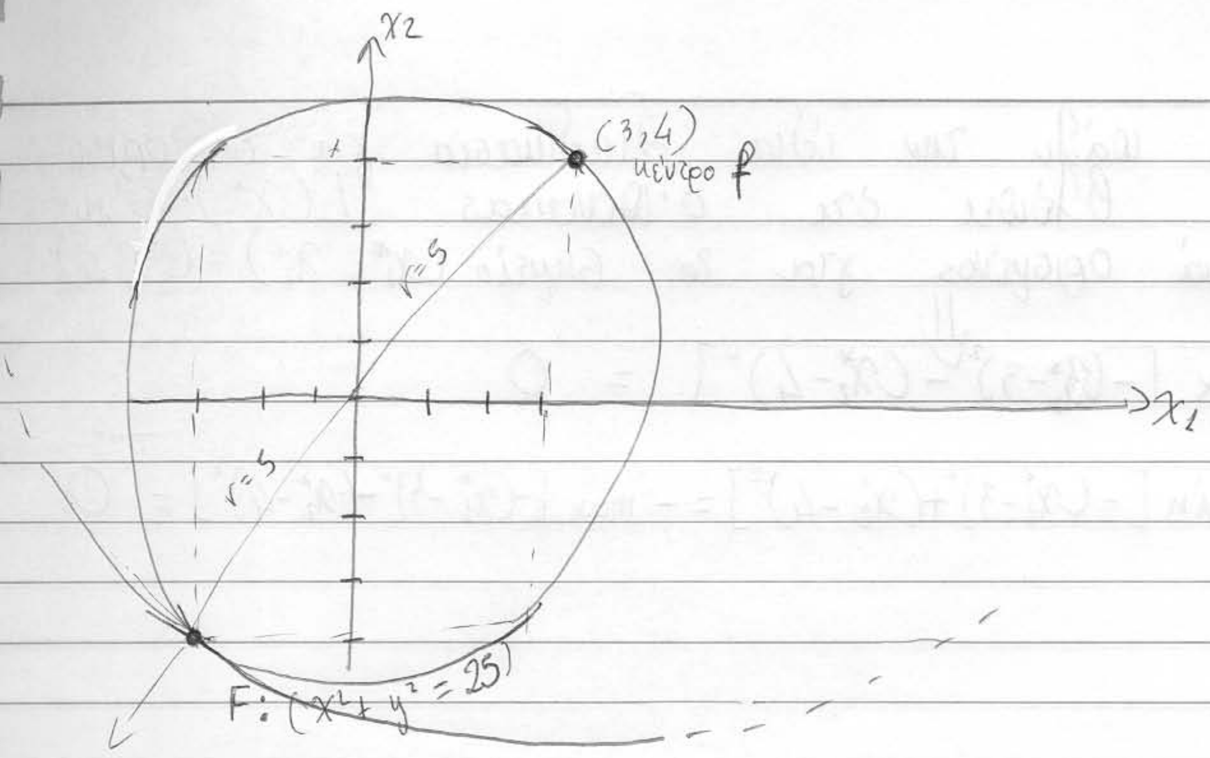
$\Rightarrow$ Ο πίνακας είναι αρνητικά ορισμένος για $[-3, -4, 2]$

Συνεπώς: αυτό το σημείο είναι το τοπικό μέγιστο το οποίο είναι προφανώς και ολικό μέγιστο γιατί η εφικτή περιοχή είναι φραγμένη (είναι η περιφέρεια του κύκλου $x^2 + y^2 = 25$).

$$f_{max} = (-3-3)^2 + (-4-4)^2 = 100 = 10^2$$

$$\Downarrow$$

σημείο στην περιφέρεια κύκλου με κέντρο $[3, 4]$ και ακτίνα 10.



max f

Προφανώς επειδή η τιμή της αντικειμενικής συνάρτησης λόγω των τετραγώνων είναι πάντα μη αρνητική, εάν είχα πρόβλημα ελαχιστοποίησης, αυτό θα έδινταν από το σημείο [3, 4] όπου η f παίρνει την μικρότερη επιτρεπόμενη τιμή της, δηλαδή 0.

Αυτό γίνεται επίσης προφανές από τη θετική ορισμένότητα του πίνακα $L(\underline{x}^*)$ στο σημείο αυτό στο επίπεδο $M^{\pm}_{\underline{x}^*}$, κατ' αντιστοιχία με τις συναρτήσεις μιας μεταβλητής.

Επίσης, εάν ήταν πρόβλημα ελαχιστοποίησης, θα μπορούσε κάποιος να το μετατρέψει σε πρόβλημα μεγιστοποίησης ως: $\min f(x) = -\max(-f(x)) \Rightarrow$
$\Rightarrow \min[(x_1-3)^2+(x_2-4)^2] = -\max[-(x_1-3)^2-(x_2-4)^2]$
και να επιλύσει το πρόβλημα ως μεγιστοποίησης
⇓
τότε πάλι θα κατέληγε σε 2 σημεία ως πιθανά αποτελέσματα:

$$[x_1^*, x_2^*, \lambda^*] = \begin{cases} [3, 4, 0] \\ [-3, -4, -2] \end{cases}$$

Κάνοντας πάλι την ίδια διαδικασία για την ορισμότητα, βλέπουμε ότι ο πίνακας $L(x^*)$ είναι αρνητικά ορισμένος για το σημείο $(x_1^*, x_2^*) = (3, 4)$

$$\max\left[-(x_1^*-3)^2 - (x_2^*-4)^2\right] = 0$$

$$\Rightarrow \min\left[(x_1^*-3)^2 + (x_2^*-4)^2\right] = -\max\left[-(x_1^*-3)^2 - (x_2^*-4)^2\right] = 0$$

Συνθήκες Kuhn-Tucker για επίλυση προβλημάτων ΜΓΠ με περιορισμούς

Η γενική μορφή προβλήματος ΜΓΠ όπως έχουμε ήδη αναφέρει είναι:

$$\begin{aligned}
&\max f(\underline{x})\\
&h_1(\underline{x}) = 0\\
&\quad\vdots\\
&h_m(\underline{x}) = 0\\
&g_1(\underline{x}) \leq 0\\
&\quad\vdots\\
&g_p(\underline{x}) \leq 0\\
&\underline{x} \in \mathbb{R}^n
\end{aligned}$$

<u>Ορισμοί</u>

i) Ο ανισωτικός περιορισμός $g_k(\underline{x}) \leq 0$ του προβλήματος ΜΓΠ ονομάζεται ΕΝΕΡΓΟΣ για ένα σημείο $\underline{x}^*$ της εφικτής περιοχής F αν ισχύει σαν ισότητα, δηλαδή $g_k(\underline{x}^*) = 0$. Διαφορετικά λέγεται <u>ανενεργός</u>. Ένας εξισωτικός περιορισμός θεωρείται πάντα ενεργός.

ii) Έστω σημείο $\underline{x}^*$ που ικανοποιεί τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$, $\underline{g}(\underline{x}) \leq 0$ και έστω J το σύνολο των ενεργών ανισωτικών περιορισμών για το $\underline{x}^*$, δηλαδή

$$J = \{ j : g_j(\underline{x}^*) = 0 \}$$

Τότε, το $\underline{x}^*$ ονομάζεται <u>κανονικό σημείο των περιορισμών</u> αν τα διανύσματα $\nabla h_i(\underline{x}^*)$ $(i = 1, 2, \ldots, m)$, $\nabla g_k(\underline{x}^*)$ $(k \in J)$ είναι γραμμικά ανεξάρτητα

Σημειώνουμε ότι τα τοπικά μέγιστα (εάν υπάρχουν) επηρεάζονται <u>ΜΟΝΟ από τους ενεργούς περιορισμούς</u>. Έτσι, εάν ξέραμε εκ των προτέρων ποιοι περιορισμοί σε μια λύση $\underline{X}^*$ είναι ενεργοί, θα μπορούσαμε να αγνοήσουμε τους ανενεργούς περιορισμούς, τους ενεργούς να τους μετατρέψουμε σε εξισώσεις και να λύσουμε ένα πρόβλημα ΜΓΠ με εξισωτικούς περιορισμούς με τη μέθοδο πολλ/στών Lagrange.

Ένας τρόπος επίλυσης θα ήταν να επιλέγαμε ένα υποσύνολο ανισωτικών περιορισμών, να τους μετατρέψαμε σε εξισωτικούς και, αγνοώντας τους υπόλοιπους, να λύναμε το πρόβλημα ΜΓΠ με αυτούς. Αυτό πρέπει να γίνει για όλους τους δυνατούς συνδυασμούς περιορισμών. Οι λύσεις θα ήταν λύσεις του αρχικού προβλήματος, εφόσον ικανοποιούσαν τις συνθήκες του αρχικού προβλήματος. Η διαδικασία αυτή είναι όμως αναποτελεσματική μιας και ο δυνατός συνδυασμός περιορισμών για την δημιουργία υποσυνόλων εξισωτικών περιορισμών είναι 2^p (όπου υπολογίζεται και η περίπτωση όλοι οι περιορισμοί να είναι ανενεργοί $\forall \, \underline{x} \in F$).

Η ακολουθούμενη διαδικασία αποτελεί γενίκευση της θεωρίας που έχουμε αναπτύξει ήδη με εξισωτικούς περιορισμούς. Οι συνθήκες (αναγκαίες και ικανές) που διέπουν τη μεθοδολογία επίλυσης προβλημάτων ΜΓΠ με ανισωτικούς περιορισμούς ονομάζονται ΣΥΝΘΗΚΕΣ KUHN-TUCKER.

Θεώρημα (Αναγκαία συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$, $\underline{h}: \mathbb{R}^n \to \mathbb{R}^m$ $(m \leq n)$, $\underline{g}: \mathbb{R}^n \to \mathbb{R}^p$ C^2-συναρτήσεις και F η εφικτή περιοχή που ορίζεται από τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$, $\underline{g}(\underline{x}) \leq \underline{0}$. Αν το $\underline{x}^* \in F$ είναι τοπικό μέγιστο της f υπό τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$, $\underline{g}(\underline{x}) \leq 0$ και είναι επιπλέον και κανονικό σημείο της F, τότε

i) Υπάρχουν $\underline{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_m]^T \in \mathbb{R}^m$ και $\underline{\mu} = [\mu_1, \mu_2, \ldots, \mu_p]^T \in \mathbb{R}^p$:

$$\underline{\mu} \geq \underline{0},$$

$$\nabla f(\underline{x}^*) - \underline{\lambda}^T \nabla \underline{h}(\underline{x}^*) - \underline{\mu}^T \nabla \underline{g}(\underline{x}^*) = \underline{0},$$

$$\mu_k g_k(\underline{x}^*) = 0 \quad (k = 1, 2, \ldots, p)$$

ii) Ο πίνακας

$$L(\underline{x}^*) = Hf(\underline{x}^*) - H(\underline{\lambda}^T \underline{h}(\underline{x}^*)) - H(\underline{\mu}^T \underline{g}(\underline{x}^*))$$

είναι αρνητικά ημιορισμένος στο εφαπτόμενο επίπεδο των ενεργών περιορισμών στο $\underline{x}^*$

$$M^F_{\underline{x}^*} = \{\underline{y} \in \mathbb{R}^n : \nabla \underline{h}(\underline{x}^*)\underline{y} = \underline{0}, \ \nabla g_j(\underline{x}^*)\underline{y} = 0, \ j \in J\}$$

όπου $J = \{j : g_j(\underline{x}^*) = 0\}$

δηλαδή ισχύει ότι $\underline{y}^T L(\underline{x}^*)\underline{y} \leq 0$, $\underline{y} \in M^F_{\underline{x}^*}$

Οι συνθήκες της i) μαζί με τους περιορισμούς ονομάζονται ΣΥΝΘΗΚΕΣ KUHN-TUCKER.

Ενισχύοντας τη συνθήκη ii) παίρνουμε τις ικανές συνθήκες τοπικού μεγίστου

Θεώρημα (Ικανή συνθήκη τοπικού μεγίστου)

Έστω $f: \mathbb{R}^n \to \mathbb{R}$, $\underline{h}: \mathbb{R}^n \to \mathbb{R}^m$, $\underline{g}: \mathbb{R}^n \to \mathbb{R}^p$, C^2-συναρτήσεις και F η εφικτή περιοχή που ορίζεται από τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$, $\underline{g}(\underline{x}) \leq \underline{0}$. Έστω $\underline{x}^* \in \mathbb{R}^n$, $\underline{\lambda} \in \mathbb{R}^m$ και $\underline{\mu} \in \mathbb{R}^p$:

i) $\underline{\mu} \geq 0$

$$\nabla f(\underline{x}^*) - \underline{\lambda}^T \nabla \underline{h}(\underline{x}^*) - \underline{\mu}^T \nabla \underline{g}(\underline{x}^*) = \underline{0}$$

$$\mu_k g_k(\underline{x}^*) = 0, \ (k = 1, 2, \ldots, p)$$

ii) Ο πίνακας

$$L(\underline{x}^*) = H f(\underline{x}^*) - H(\underline{\lambda}^T \underline{h}(\underline{x}^*)) - H(\underline{\mu}^T \underline{g}(\underline{x}^*))$$

είναι αρνητικά ορισμένος στον υπόχωρο

$$M' = \{ \underline{y} \in \mathbb{R}^n : \nabla \underline{h}(\underline{x}^*) \underline{y} = 0, \ \nabla g_j(\underline{x}^*) \underline{y} = 0, \ j \in J' \}$$

όπου $J' = \{ j : g_j(\underline{x}^*) = 0, \ \mu_j > 0 \}$ το σύνολο των ενεργών ανισωτικών περιορισμών του $\underline{x}^*$ για τις οποίες ο συντελεστής μ_j είναι και μεγαλύτερος από το 0. Θυμίζουμε ότι για την αρνητική ορισμότητα ισχύει ότι $\underline{y}^T L(\underline{x}^*) \underline{y} < 0$, $\underline{y} \in M'$ και $\underline{y} \neq \underline{0}$.

Εάν ισχύουν τα i) και ii) τότε το $\underline{x}^*$ είναι τοπικό μέγιστο της f υπό τους περιορισμούς $\underline{h}(\underline{x}) = \underline{0}$, $\underline{g}(\underline{x}) \leq \underline{0}$.

Οι συνθήκες Kuhn-Tucker μαζί με τους περιορισμούς μπορούν να γραφούν αναλυτικά ως:

$$\mu_k \geq 0 \quad (k=1,2,\ldots,p)$$

$$\frac{\partial f}{\partial x_j}(\underline{x}^*) - \sum_{i=1}^{m} \lambda_i \frac{\partial h_i}{\partial x_j}(\underline{x}^*) - \sum_{k=1}^{p} \mu_k \frac{\partial g_k}{\partial x_j}(\underline{x}^*) = 0 \quad (j=1,2,\ldots,n)$$

$$\mu_k \, g_k(\underline{x}^*) = 0 \quad (k=1,2,\ldots,p)$$

$$g_k(\underline{x}^*) \leq 0 \quad (k=1,2,\ldots,p)$$

$$h_i(\underline{x}^*) = 0 \quad (i=1,2,\ldots,m)$$

Οι συνθήκες αυτές βοηθούν στην επίλυση προβλημάτων ΜΓΠ μέσω του εντοπισμού τοπικών μεγίστων. Για μια σημαντική όμως κατηγορία προβλημάτων με ανισωτικούς περιορισμούς εξασφαλίζουν ολικά μέγιστα. Η κατηγορία αυτή ονομάζεται ΚΥΡΤΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΣ και αφορά προβλήματα χωρίς εξισωτικούς περιορισμούς της μορφής:

$$\max f(\underline{x})$$

με περιορισμούς

$$g_1(\underline{x}) \leq 0$$
$$g_2(\underline{x}) \leq 0$$
$$\vdots$$
$$g_p(\underline{x}) \leq 0$$

όπου η f είναι κοίλη συνάρτηση και $g_k(\underline{x})$ $(k=1,2,\ldots,p)$ είναι κυρτές συναρτήσεις (και φυσικά, κατ' επέκταση και η εφικτή περιοχή F είναι κυρτό σύνολο). Τα προβλήματα αυτά ονομάζονται προβλήματα κυρτού προγραμματισμού (Κ.Π)

Πόρισμα 1

Έστω το πρόβλημα ΚΠ

$$\max f(\underline{x})$$
$$g_1(\underline{x}) \le 0$$
$$\vdots$$
$$g_p(\underline{x}) \le 0$$

με εφικτή περιοχή F. Αν υπάρχουν $\underline{x}^* \in \mathbb{R}^n$, $\underline{\mu} \in \mathbb{R}^p$:

$$\mu_k \ge 0 \quad (k=1,\dots,p)$$
$$\frac{\partial f}{\partial x_j}(\underline{x}^*) - \sum_{k=1}^{p} \mu_k \frac{\partial g_k}{\partial x_j}(\underline{x}^*) = 0 \quad (k=1,2,\dots,p)$$
$$\mu_k g_k(\underline{x}^*) = 0 \quad (k=1,2,\dots,p)$$
$$g_k(\underline{x}^*) \le 0 \quad (k=1,2,\dots,p)$$

τότε το $\underline{x}^*$ είναι λύση του προβλήματος ΚΠ, είναι δηλαδή ολικό μέγιστο της f υπό τους περιορισμούς $\underline{g}(\underline{x}) \le 0$.

Σημείωση: Η συνθήκη αρνητικής ορισμότητας του $L(\underline{x}^*)$ και μάλιστα εδώ στον $\mathbb{R}^n$ είναι δεδομένο ότι ισχύει αφού η f είναι κοίλη (αρνητικές ιδιοτιμές $Hf(\underline{x}^*)$), οι $\underline{g}$ κυρτές (θετικές ιδιοτιμές $Hg_k(\underline{x}^*)$) και $\underline{\mu} \ge 0$, άρα και οι ιδιοτιμές του $H(f - \underline{\mu}^T \underline{g})$ στο $\underline{x}^*$ θα είναι αρνητικές, όπου $L(\underline{x}^*) = Hf(\underline{x}^*) - H(\underline{\mu}^T \underline{g}(\underline{x}^*)) = H(f(\underline{x}^*) - \underline{\mu}^T \underline{g}(\underline{x}^*))$

Πόρισμα 2

Έστω το πρόβλημα ΜΓΠ

$$\max f(\underline{x})$$
$$\underline{x} \geq \underline{0}$$

με f κοίλη. Αν υπάρχει $\underline{x}^* \geq 0$:

$$\frac{\partial f}{\partial x_j}(\underline{x}^*) \leq 0 \quad (j=1,2,\ldots,n)$$

$$x_j^* \frac{\partial f}{\partial x_j}(\underline{x}^*) = 0 \quad (j=1,2,\ldots,n)$$

τότε το $\underline{x}^*$ είναι λύση του προβλήματος

[**Σημείωση:** Η απόδειξή του είναι προφανής δεδομένου ότι $p=n$, $g_j(\underline{x}) = -x_j$ και $\frac{\partial f}{\partial x_j}(x^*) + \mu_j = 0$, όπου $\mu_j \geq 0$, $j=1,2,\ldots,p$]

Πόρισμα 3

Έστω το πρόβλημα ΜΓΠ

$$\max f(\underline{x})$$
$$g_1(\underline{x}) \leq 0$$
$$\vdots$$
$$g_p(\underline{x}) \leq 0$$
$$\underline{x} \geq \underline{0}$$

με f κοίλη και g_K $(K=1,2,\ldots,p)$ κυρτές. Αν υπάρχουν $\underline{x}^* \in \mathbb{R}^n$ και $\underline{\mu} \in \mathbb{R}^p$:

$$\mu_K \geq 0 \quad (K=1,\ldots,p)$$

$$\frac{\partial f}{\partial x_j}(\underline{x}^*) - \sum_{K=1}^{p} \mu_K \frac{\partial g_K}{\partial x_j}(\underline{x}^*) \leq 0 \quad (j=1,2,\ldots,n)$$

$$x_j^*\left(\frac{\partial f(\underline{x}^*)}{\partial x_j} - \sum_{k=1}^{p} \mu_k \frac{\partial g_k}{\partial x_j}(\underline{x}^*)\right) = 0 \quad (j=1,2,\dots,n)$$

$$\mu_k\, g_k(\underline{x}^*) = 0 \quad (k=1,2,\dots p)$$

$$g_k(\underline{x}^*) \le 0 \quad (k=1,2,\dots p)$$

$$x_j^* \ge 0 \quad (j=1,2,\dots,n)$$

τότε το $\underline{x}^*$ είναι λύση του προβλήματος

[Σημείωση: Η απόδειξή του βασίζεται στην ίδια ιδέα που γίνεται η απόδειξη του θεωρήματος 2]

Παράδειγμα 1 επίλυσης με συνθήκες KT

Να λυθεί το πρόβλημα ΜΓΠ

$$\max\left[\ln(1+x_1+x_2)\right]$$

αν

$$x_1 + 2x_2 \le 5$$

$$x_1, x_2 \ge 0$$

Λύση: Στην περίπτωση αυτή $p=1$

$$f(\underline{x}) = \ln(1+x_1+x_2)$$

$$g_1(\underline{x}) = x_1 + 2x_2 - 5 \le 0$$

Η συνάρτηση $g_1(\underline{x})$ ως γραμμική είναι κυρτή ενώ η $f(\underline{x})$ είναι κοίλη. Παίρνοντας την 1η παράγωγο ως προς x_1 ή x_2 αυτή δίνει $\frac{1}{1+x_1+x_2}$ και ο πίνακας Hesse τότε της $f(\underline{x})$ είναι:

$$Hf(\underline{x}) = \begin{bmatrix} -\frac{1}{(1+x_1+x_2)^2} & -\frac{1}{(1+x_1+x_2)^2} \\ -\frac{1}{(1+x_1+x_2)^2} & -\frac{1}{(1+x_1+x_2)^2} \end{bmatrix} = \begin{bmatrix} -A & -A \\ -A & -A \end{bmatrix}$$

όπου $A = -\frac{1}{(1+x_1+x_2)^2}$ και $0 < A \leq 1$ αφού $x_1, x_2 \geq 0$

Παίρνοντας τις κύριες ελάσσονες ορίζουσες του πίνακα αυτού βλέπουμε ότι αυτές είναι $-A$, 0, $-A$, όλες ≤ 0, άρα ο $Hf(\underline{x})$ είναι αρνητικά ημιορισμένος στον $\mathbb{R}^n$ $\Rightarrow$ η συνάρτηση $f(\underline{x})$ είναι κοίλη.

Βάσει του θεωρήματος 3 τότε, έχουμε:

$$\mu_1 \geq 0 \qquad (KT1)$$

$$\frac{1}{1+x_1+x_2} - \mu_1 \cdot 1 \leq 0 \qquad (KT2)$$

$$\frac{1}{1+x_1+x_2} - \mu_1 \cdot 2 \leq 0 \qquad (KT3)$$

$$x_1\left(\frac{1}{1+x_1+x_2} - \mu_1 \cdot 1\right) = 0 \qquad (KT4)$$

$$x_2\left(\frac{1}{1+x_1+x_2} - \mu_1 \cdot 2\right) = 0 \qquad (KT5)$$

$$\mu_1(x_1 + 2x_2 - 5) = 0 \qquad (KT6)$$

$$x_1 + 2x_2 - 5 \leq 0 \qquad (KT7)$$

$$x_1 \geq 0 \qquad (KT8)$$

$$x_2 \geq 0 \qquad (KT9)$$

Αυτές είναι και οι συνθήκες Kuhn-Tucker του προβλήματος.

Η επίλυση του συστήματος αυτού δεν είναι τετριμμέvn. Ο ευκολότερος τρόπος είναι να ξεχωρίσουμε περιπτώσεις ανάλογα με το αν $\mu_1=0$ ή >0, $x_1=0$ ή >0 και $x_2=0$ ή >0. Οι δυνατοί συνδυασμοί είναι 8:

- 1 με όλα θετικά
- 1 με όλα μηδενικά
- $\frac{3!}{1!(3-1)!}$ με ένα μη μηδενικό και τα άλλα μηδέν
- $\frac{3}{2!(3-2)!}$ με δύο μη μηδενικά και τα άλλα μηδέν

Περίπτωση 1 : $\mu_1=0$

Τότε $(KT2) \Rightarrow 1+x_1+x_2 \leq 0$ που δεν είναι συμβατό με τις $(KT8)$, $(KT9)$ $\Rightarrow$ δεν δίνει λύση

Περίπτωση 2 $\mu_1>0$, $x_1=0$, $x_2=0$

Δεν ικανοποιείται η $(KT6)$ $\Rightarrow$ δεν δίνει λύση

Περίπτωση 3 $\mu_1>0$, $x_1>0$, $x_2=0$

$(KT6) \Rightarrow x_1=5$

$(KT4) \Rightarrow 5\left(\frac{1}{1+5+0}-\mu_1 \cdot 1\right)=0 \Rightarrow \mu_1=\frac{1}{6}$

Η $(x_1^*, x_2^*, \mu_1^*)=(5, 0, 1/6)$ ικανοποιεί όλες τις συνθήκες και άρα είναι λύση του προβλήματος

Περίπτωση 4 $\mu_1>0$, $x_1=0$, $x_2>0$

$(KT6) \Rightarrow x_2=5/2$

$(KT5) \Rightarrow \frac{5}{2}\left(\frac{1}{1+0+5/2}-2\mu_1\right)=0 \Rightarrow \mu_1=\frac{1}{7}$

Όμως για $\mu_1 = 1/7$ δεν ικανοποιείται η (ΚΤ 2) $\Rightarrow$ η περίπτωση αυτή δε δίνει λύση

<u>Περίπτωση 5</u> $\mu_1 > 0$, $x_1 > 0$, $x_2 > 0$ τότε από (ΚΤ4) και (ΚΤ5) έχουμε ότι:

$$\frac{1}{1+x_1+x_2} - \mu_1 = \frac{1}{1+x_1+x_2} - 2\mu_1 = 0$$

$\Rightarrow \mu_1 = 2\mu_1 \Rightarrow$ αδύνατο για $\mu_1 \neq 0 \Rightarrow$
$\Rightarrow$ δεν δίνει λύση

$\Downarrow$

Η μοναδική λύση του προβλήματος είναι αυτή της περίπτωσης 3 με $(x_1^*, x_2^*) = (5, 0)$ και $f_{max} = \ln 6$

<u>Παράδειγμα 2 επίλυσης με συνθήκες ΚΤ</u>

Να επιλυθεί το πρόβλημα ΜΓΠ.

$$\min(x_1^2 + x_2^2 - 2x_1 - 3x_2)$$

όταν
$$2 - x_1 - x_2 \geq 0$$
$$x_1, x_2 \geq 0$$

<u>Λύση:</u> Το πρόβλημα γράφεται στη μορφή

$$-\max(-x_1^2 - x_2^2 + 2x_1 + 3x_2)$$
$$x_1 + x_2 - 2 \leq 0$$
$$x_1, x_2 \geq 0$$

δηλαδή $p = 1$ και
$$f(x_1, x_2) = -x_1^2 - x_2^2 + 2x_1 + 3x_2$$

$$g_1(x_1, x_2) = x_1 + x_2 - 2 \leq 0$$

Η g_1 είναι γραμμική και συνεπώς κυρτή. Για την $f(\underline{x})$ έχουμε:

$$Hf(\underline{x}) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

που είναι αρνητικά ορισμένος $\forall\ \underline{x} \in \mathbb{R}^2$. Άρα, η $f(\underline{x})$ είναι κοίλη και συνεπώς, λόγω του θεωρήματος 3, έχουμε:

$$\mu_1 \geq 0 \quad (KT1)$$
$$-2x_1 + 2 - \mu_1 \leq 0 \quad (KT2)$$
$$-2x_2 + 3 - \mu_1 \leq 0 \quad (KT3)$$
$$x_1(-2x_1 + 2 - \mu_1) = 0 \quad (KT4)$$
$$x_2(-2x_2 + 3 - \mu_1) = 0 \quad (KT5)$$
$$\mu_1(x_1 + x_2 - 2) = 0 \quad (KT6)$$
$$x_1 + x_2 - 2 \leq 0 \quad (KT7)$$
$$x_1 \geq 0 \quad (KT8)$$
$$x_2 \geq 0 \quad (KT9)$$

Πάλι κανονικά θα έπρεπε να εξετάσουμε 8 διαφορετικές περιπτώσεις, λόγω όμως των συνθηκών, κάποιες θα τις εξετάσω ομαδικά.

Περίπτωση 1 $\quad \mu_1 = 0$

Τότε από (KT2) και (KT3) έχω ότι

$$x_1 \geq 1$$
$$x_2 \geq 3/2$$

Τότε, το αριστερό μέλος της (KT7) δίνει:

$AM(KT7) \Rightarrow \quad x_1 + x_2 - 2 \geq 1 + 3/2 - 2 = 1/2 > 0$

το οποίο έρχεται σε αντίθεση με τη συνθήκη (ΚΤ7). Συνεπώς, η περίπτωση αυτή δεν δίνει λύση

Περίπτωση 2: $\mu_1 > 0, \quad x_1 > 0$

Τότε από (ΚΤ4) και (ΚΤ6) παίρνουμε το σύστημα:

$$-2x_1 + 2 - \mu_1 = 0$$
$$x_1 + x_2 - 2 = 0$$

οπότε τα x_1 και μ_1 ως συνάρτηση του x_2 είναι:

$$x_1 = 2 - x_2$$
$$\mu_1 = 2x_2 - 2$$

Άρα, η (ΚΤ5) γίνεται:

$$(KT5) \Rightarrow \quad x_2(-2x_2 + 3 - 2x_2 + 2) = 0$$

$$\Rightarrow \begin{cases} x_2 = 0 \\ -4x_2 + 5 = 0 \Rightarrow x_2 = 5/4 \end{cases}$$

Αντικαθιστώντας το x_2 στις εξισώσεις για x_1, μ_1 παίρνουμε ότι:

$(x_1^*, x_2^*, \mu_1) = (2, 0, -2) \Rightarrow$ απορρίπτεται λόγω (ΚΤ1)

ή $(x_1^*, x_2^*, \mu_1) = (3/4, 5/4, 1/2)$ η οποία ικανοποιεί όλες τις συνθήκες (ΚΤ1)-(ΚΤ9)

Περίπτωση 3 $\quad \mu_1 > 0, \quad x_1 = 0$

Τότε, από (ΚΤ6) έχουμε:

$(KT6) \Rightarrow \overset{>0}{x_1} + x_2 - 2 = 0 \Rightarrow$

$\Rightarrow x_2 = 2$

Τότε, αφού $x_2 \neq 0$ από (ΚΤ5) έχουμε:

$(KT5) \Rightarrow -2x_2 + 3 - \mu_1 = 0 \xRightarrow{x_2 = 2} \mu_1 = -1$

το οποίο έρχεται σε αντίθεση με την (ΚΤ1)

Συνεπώς η λύση του προβλήματος είναι αυτή που προέκυψε από την περίπτωση 2, δηλαδή

$(x_1^*, x_2^*) = (3/4, 5/4)$ με ελάχιστη τιμή της αρχικής συνάρτησης την

$$\left(\frac{3}{4}\right)^2 + \left(\frac{5}{4}\right)^2 - 2 \cdot \frac{3}{4} - 3 \cdot \frac{5}{4} = -\frac{25}{8}$$

Τετραγωνικός Προγραμματισμός (ΤΠ)

Τα προβλήματα ΤΠ είναι προβλήματα ΜΓΠ στα οποία οι περιορισμοί είναι γραμμικοί και η αντικειμενική συνάρτηση είναι ένα πολυώνυμο $2^{ου}$ βαθμού. Στη γενική τους μορφή τα προβλήματα αυτά γράφονται ως:

$$\max\left(\underline{c}^T \underline{x} + \underline{x}^T D \underline{x}\right)$$
$$A\underline{x} \leq \underline{b}$$
$$\underline{x} \geq 0$$

όπου $\underline{c}, \underline{x} \in M_{n\times 1}$, $\underline{b} \in M_{m\times 1}$, $A \in M_{m\times n}$ και $D \in M_{n\times n}$ συμμετρικός $n\times n$ πίνακας με $d_{ij} = d_{ji}$ για $i \neq j$

Στην περίπτωση που η αντικειμενική συνάρτηση είναι κοίλη, τότε έχουμε ένα πρόβλημα ΚΥΡΤΟΥ ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ (ΚΤΠ). Αυτό συμβαίνει αν και μόνο αν ο πίνακας D (που συμπίπτει με τον πίνακα Hesse της αντικειμενικής συνάρτησης) είναι αρνητικά ημιορισμένος. Στην περίπτωση αυτή μπορούμε να ακολουθήσουμε τη μέθοδο που περιγράψαμε για τα προβλήματα ΚΠ με τις συνθήκες Kuhn-Tucker. Υπάρχει ένας αλγόριθμος ο οποίος έχει αναπτυχθεί για τέτοια προβλήματα και ονομάζεται ΤΡΟΠΟΠΟΙΗΜΕΝΗ ΜΕΘΟΔΟΣ SIMPLEX. Η μέθοδος αυτή λειτουργεί όπως η μέθοδος Simplex με μια μικρή τροποποίηση ώστε να ικανοποιούνται οι συνθήκες ΚΤ του Πορίσματος 3 που έχουμε ήδη δει

Για να καταλάβουμε πώς εφαρμόζονται οι συνθήκες ΚΤ γράφουμε το πρόβλημα ΚΤΠ στη γενική του μορφή αναλυτικά:

$$\max(c_1x_1 + c_2x_2 + \dots + c_nx_n + d_{11}x_1^2 + d_{22}x_2^2 + \dots +$$
$$+ d_{nn}x_n^2 + 2d_{12}x_1 \cdot x_2 + 2d_{13}x_1 \cdot x_3 + \dots +$$
$$+ 2d_{1n}x_1 \cdot x_n + 2d_{23}x_2 \cdot x_3 + 2d_{24}x_2 \cdot x_4 + \dots +$$
$$+ 2d_{2n}x_2 \cdot x_n + \dots + 2d_{n-1\,n}x_{n-1} \cdot x_n)$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1j}x_j + \dots + a_{1n}x_n \leq b_1$$
$$\vdots$$
$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{ij}x_j + \dots + a_{in}x_n \leq b_i$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mj}x_j + \dots + a_{mn}x_n \leq b_m$$

$$x_1, x_2, \dots, x_n \geq 0$$

Για να εφαρμόσουμε το Πόρισμα 3 πρέπει να υπολογίσουμε τις παραγώγους στο αριστερό μέλος της 2ης και 3ης συνθήκης του Πορίσματος 3:

$$\frac{\partial f}{\partial x_j} - \sum_{k=1}^{p} \mu_k \frac{\partial g_k}{\partial x_j} =$$

$$= c_j + 2d_{jj}x_j + 2d_{1j}x_1 + 2d_{2j}x_2 + \dots + 2d_{j-1\,j}x_{j-1} +$$
$$+ 2d_{j\,j+1}x_{j+1} + 2d_{j\,j+2}x_{j+2} + \dots + 2d_{jn}x_n -$$
$$- \mu_1 a_{1j} - \mu_2 a_{2j} - \dots - \mu_m a_{mj} =$$

$$= c_j + 2\sum_{l=1}^{n} d_{lj}x_l - \sum_{i=1}^{m} \mu_i a_{ij}$$

Έτσι, οι συνθήκες ΚΤ του προβλήματος 3 θα πάρουν τη μορφή:

$$\mu_i \geq 0 \quad (i=1,2,\dots,m)$$

$$c_j + 2\sum_{l=1}^{n} d_{lj} x_l - \sum_{i=1}^{m} \mu_i a_{ij} \leq 0 \quad (j=1,2,\dots,n)$$

$$x_j \left(c_j + 2\sum_{l=1}^{n} d_{lj} x_l - \sum_{i=1}^{m} \mu_i a_{ij} \right) = 0 \quad (j=1,2,\dots,n)$$

$$\mu_i \left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right) = 0 \quad (i=1,2,\dots,m)$$

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad (i=1,2,\dots,m)$$

$$x_j \geq 0 \quad (j=1,2,\dots,n)$$

Θέτοντας

$$y_j = -\left(c_j + 2\sum_{l=1}^{n} d_{lj} x_l - \sum_{i=1}^{m} \mu_i a_{ij} \right) \quad (j=1,\dots,n)$$

$$\lambda_i = -\left(\sum_{j=1}^{n} a_{ij} x_j - b_i \right)$$

ή, σε διανυσματική μορφή

$$\underline{y} = -(\underline{c} + 2D\underline{x} - A^T \underline{\mu}) =$$

$$= -2D\underline{x} + A^T\underline{\mu} - \underline{c} \quad (E1)$$

$$\lambda = -(A\underline{x} - \underline{b}) = -A\underline{x} + \underline{b} \quad (E2)$$

οι παραπάνω συνθήκες γίνονται:

$$\underline{\mu} \geq \underline{0}$$

$$\underline{y} \geq \underline{0}$$

$$x_j y_j = 0 \quad (j = 1, 2, \ldots, n)$$
$$\mu_i \lambda_i = 0 \quad (i = 1, 2, \ldots, m)$$

$$\underline{\lambda} \geq \underline{0}$$
$$\underline{x} \geq \underline{0}$$

Έτσι, ουσιαστικά το πρόβλημα ανάγεται στο να βρεθούν $\underline{x}, \underline{y}, \underline{\mu}, \underline{\lambda}$:

$$\boxed{\begin{array}{l} (E1) \Rightarrow \ -2D\underline{x} + A^T\underline{\mu} - \underline{y} = \underline{c} \\ (E2) \Rightarrow \ A\underline{x} + \underline{\lambda} = \underline{b} \\ \underline{x} \geq \underline{0}, \ \underline{\mu} \geq \underline{0}, \ \underline{y} \geq \underline{0}, \ \underline{\lambda} \geq \underline{0} \\ x_j y_j = 0 \quad (j = 1, 2, \ldots n) \\ \mu_i \lambda_i = 0 \quad (i = 1, 2, \ldots, m) \end{array}}$$

Εάν δεν υπήρχαν οι 2 τελευταίες συνθήκες το πρόβλημα θα ήταν ένα πρόβλημα ΓΠ. και θα μπορούσε να χρησιμοποιηθεί η μέθοδος Simplex όπως την έχουμε ήδη δει για να το επιλύσουμε. Οι 2 τελευταίες συνθήκες ονομάζονται ΠΕΡΙΟΡΙΣΜΟΙ ΣΥΜΠΛΗΡΩΜΑΤΙΚΟΤΗΤΑΣ και οι μεταβλητές y_j και λ_i $(i=1, \ldots, m, \ j = 1, \ldots n)$ ονομάζονται

ΣΥΜΠΛΗΡΩΜΑΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ ΤΩΝ x_j και μ_i, αντίστοιχα. Αυτό που μας λένε οι συμπληρωματικοί περιορισμοί είναι ότι:

- εάν $x_j \neq 0 \Rightarrow y_j = 0$ και αντίστροφα
- εάν $\mu_i \neq 0 \Rightarrow \lambda_i = 0$ και αντίστροφα

Η τροποποιημένη μέθοδος Simplex που χρησιμοποιείται για την επίλυση αυτού του προβλήματος είναι ουσιαστικά η μέθοδος Simplex, μόνο που λόγω των συμπληρωματικών περιορισμών, όταν η x_j είναι βασική (και συνεπώς $\neq 0$) στον πίνακα Simplex, δεν επιτρέπεται στην y_j να γίνει από μη-βασική, βασική γιατί τότε δεν θα ικανοποιείται ο περιορισμός $x_j y_j = 0$. Φυσικά ισχύει και το αντίστροφο (δηλ. για την y_j στη βάση και τη x_j μη βασική). Και οι δυο μεταβλητές επιτρέπεται να είναι ταυτόχρονα μη βασικές.

Παρόμοιες συνθήκες ισχύουν και για τα μ_i, λ_i λόγω του περιορισμού $\mu_i \lambda_i = 0$.

Για να γίνει κατανοητό αυτό, δίνουμε για παράδειγμα ένα πρόβλημα με 2 μεταβλητές x_1 και x_2 και 1 ανισωτικό περιορισμό (εκτός από το $\underline{x} \geq 0$). Τότε, τα ζευγάρια των συμπληρωματικών θα είναι 3, δηλαδή:

$$\left.\begin{array}{l}(x_1, y_1)\\(x_2, y_2)\\(\mu_1, \lambda_1)\end{array}\right\} \Rightarrow \text{δεν επιτρέπεται να είναι δηλαδή μαζί στη βάση: } \left\{\begin{array}{l}\text{το } x_1 \text{ με το } y_1\\\text{το } x_2 \text{ με το } y_2\\\text{το } \mu_1 \text{ με το } \lambda_1\end{array}\right.$$

Το πρόβλημα ΤΠ που λύνουμε έχει αντικειμενική συνάρτηση $Z = 0$, μιας και το αρχικό πρόβλημα μεγιστοποίησης έχει ληφθεί υπόψη μέσω των περιορισμών και των συνθηκών Kuhn-Tucker. Η αντικειμενική συνάρτηση "κατασκευάζεται" με την προσθήκη τεχνητών μεταβλητών για να μπορέσουμε να πάρουμε μια αρχική βασική εφικτή λύση και να ξεκινήσουμε τη μέθοδο Simplex. Επιλέγουμε επίλυση με τη μέθοδο του μεγάλου-Μ, έτσι ώστε η αντικειμενική συνάρτηση να γίνει

$$Z = -M(t_1 + t_2 + \ldots + t_k) \; , 0 \leq k \leq m$$

και το πρόβλημα ΤΠ γίνεται πρόβλημα μεγιστοποίησης ως:

$$\max \left(Z = -M(t_1 + \ldots + t_k) \right)$$

Παράδειγμα επίλυσης προβλήματος με τροποποιημένη μέθοδο Simplex

Να λυθεί το πρόβλημα ΜΤΠ.

$$\max \left(4x_1 + 6x_2 - 2x_1^2 - 2x_1 \cdot x_2 - 2x_2^2 \right)$$

όταν
$$x_1 + 2x_2 \leq 2$$
$$x_1, x_2 \geq 0$$

Λύση: Οι γραμμικοί περιορισμοί είναι κυρτές συναρτήσεις. Για να δούμε εάν έχουμε πρόβλημα ΚΤΠ ώστε να εφαρμόσουμε την τροποποιημένη μέθοδο Simplex, πρέπει να εξετάσουμε εάν η αντικειμενική συνάρτηση

είναι κοίλη. Το ανάδελτά της είναι:
$\nabla f = [4 - 4x_1 - 2x_2 \, , \; 6 - 2x_1 - 4x_2]$

και ο πίνακας Hesse:

$$Hf(\underline{x}) = \begin{bmatrix} -4 & -2 \\ -2 & -4 \end{bmatrix}$$

Για την ορισμότητά του έχουμε ότι:

$$\begin{vmatrix} -4-\lambda & -2 \\ -2 & -4-\lambda \end{vmatrix} = (-4-\lambda)^2 - 4 = 0 \Rightarrow (4+\lambda) = \pm 2$$
$$\Rightarrow \lambda = \begin{cases} -2 \\ -6 \end{cases} < 0$$

$\Rightarrow Hf(\underline{x})$ αρνητικά ορισμένος $\forall \underline{x} \in \mathbb{R}^2 \Rightarrow$
$\Rightarrow f(\underline{x})$ είναι κοίλη

Συνεπώς, έχουμε πρόβλημα ΚΤΠ και έτσι μπορούμε να εφαρμόσουμε την τροποποιημένη μέθοδο Simplex. Εδώ:

$$\underline{c} = \begin{bmatrix} 4 \\ 6 \end{bmatrix} \quad D = \begin{bmatrix} -2 & -1 \\ -1 & -2 \end{bmatrix} \quad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
$$A = [1, 2] \, , \quad \underline{b} = [2]$$

Οι σχέσεις λοιπόν:

$$-2D\underline{x} + A^T\underline{\mu} - \underline{y} = \underline{c}$$

$$A\underline{x} \qquad + \underline{\lambda} = \underline{b}$$

$$\underline{x} \geq \underline{0}, \; \underline{y} \geq \underline{0}, \; \underline{\mu} \geq \underline{0}, \; \underline{\lambda} \geq \underline{0}$$

$$x_j y_j = 0 \qquad j = 1, \dots, n$$

$$\mu_i \lambda_i = 0 \qquad i = 1, \dots, m$$

γίνονται

$$4x_1 + 2x_2 + \mu_1 - y_1 = 4$$
$$2x_1 + 4x_2 + 2\mu_1 - y_2 = 6$$
$$x_1 + 2x_2 + \lambda_1 = 2$$
$$x_1, x_2, \mu_1, y_1, y_2, \lambda_1 \geq 0$$
$$x_1 y_1 = x_2 y_2 = 0 = \mu_1 \lambda_1$$

Συνεπώς, οι μεταβλητές που δεν μπορούν να είναι μαζί στη βάση είναι οι: x_1 και y_1

x_2 και y_2

μ_1 και λ_1

Για την επίλυση του προβλήματος αυτού εισάγω 2 τεχνητές μεταβλητές t_1 και t_2 στις 2 πρώτες εξισώσεις ώστε να ξεκινήσω από μια βασική εφικτή λύση. Το πρόβλημα λοιπόν γράφεται ως:

$$4x_1 + 2x_2 + \mu_1 - y_1 + t_1 = 4$$
$$2x_1 + 4x_2 + 2\mu_1 - y_2 + t_2 = 6$$
$$x_1 + 2x_2 + \lambda_1 = 2$$
$$x_1, x_2, \mu_1, y_1, y_2, \lambda_1, t_1, t_2 \geq 0$$

ώστε

$$\max(-Mt_1 - Mt_2 = -10M - (-6M)x_1 - (-6M)x_2 - (-3M)\mu_1 - My_1 - My_2)$$

όπου $M >> 0$

Έτσι ο 1ος πίνακας Simplex του προβλήματος γίνεται:

1ος Simplex

BM	x_1	x_2	μ_1	y_1	y_2	λ_1	t_1	t_2	θ
← t_1	4	2	1	-1	0	0	1	0	4
t_2	2	4	2	0	-1	0	0	1	6
λ_1	1	2	0	0	0	1	0	0	2
$\bar{c}_j$	-6M	-6M	$\cancel{-3M}$	M	M	0	0	0	-10M

Αφού η λ_1 είναι ΒΜ, η μ_1 δεν μπορεί να είναι εισερχόμενη μεταβλητή, οπότε η στήλη του μ_1 εξαιρείται κατά την εφαρμογή του κανόνα ελαχίστου πηλίκου

$\bar{c}_{x_1}$ και $\bar{c}_{x_2} < 0 \Rightarrow$ είτε x_1 είτε x_2 εισερχόμενες μεταβλητές. Παρατηρούμε όμως ότι στη στήλη του x_1 το ελάχιστο πηλίκο είναι στη γραμμή της τεχνητής μεταβλητής t_1, οπότε έτσι τη βγάζουμε από τη βάση $\Rightarrow$ x_1: εισερχόμενη

t_1: εξερχόμενη

$\Downarrow$

2ος Simplex

BM	x_1	x_2	μ_1	y_1	y_2	λ_1	t_1	t_2	θ
x_1	1	1/2	1/4	-1/4	0	0	1/4	0	1
t_2	0	3	3/2	1/2	-1	0	-1/2	1	4
← λ_1	0	3/2	-1/4	1/4	0	1	-1/4	0	1
$\bar{c}_j$	0	-3M	$\cancel{-\frac{3}{2}M}$	$\cancel{-1/2M}$	M	0	3/2M	0	-4M

Αφού η λ_1 και η x_1 είναι στη βάση, λόγω των συμπληρωματικών περιορισμών εξαιρούνται από τον καθορισμό της νέας εισερχόμενης μεταβλητής οι y_1 και μ_1

Αφού $\bar{c}_{x_2} < 0 \Rightarrow$ από κανόνα ελάχιστου πηλίκου:

x_2: εισερχόμενη

λ_1: εξερχόμενη

$\Downarrow$

3ος Simplex

BM	x_1	x_2	μ_1	y_1	y_2	λ_1	t_1	t_2	θ
x_1	1	0	1/3	-1/3	0	-1/3	1/3	0	2/3
$\leftarrow t_2$	0	0	2	0	-1	-2	0	1	2
x_2	0	1	-1/6	1/6	0	2/3	-1/6	0	2/3
$\bar{c}_j$	0	0	-2M	$\cancel{0}$	$\cancel{M}$	2M	M	0	-2M

Επειδή οι x_1 και x_2 είναι ΒΜ, οι συμπληρωματικές τους y_1 και y_2 εξαιρούνται από τη διαδικασία εύρεσης του ελάχιστου πηλίκου

$\bar{c}_{\mu_1} < 0 \Rightarrow$ κανόνας ελάχιστου πηλίκου στη στήλη του $\mu_1 \Rightarrow$

$\Rightarrow$ μ_1: εισερχόμενη

t_2: εξερχόμενη

4ος Simplex

BM	x_1	x_2	μ_1	y_1	y_2	λ_1	t_1	t_2	θ
x_1	1	0	0	-1/3	1/6	0	1/3	-1/6	1/3
μ_1	0	0	1	0	-1/2	-1	0	1/2	1
x_2	0	1	0	1/6	-1/12	1/2	-1/6	1/12	5/6
$\bar{c}_j$	0	0	0	0	0	0	M	M	0

Βλέπουμε ότι $\bar{c}_j \geq 0 \quad \forall j$ => άρα αυτόσδα είναι και ο τελευταίος πίνακας Simplex, με βέλτιστη λύση τη

$(x_1^*, x_2^*, \mu_1^*, y_1^*, y_2^*, \lambda_1^*, t_1^*, t_2^*) = (1/3, 5/6, 1, 0, 0, 0, 0, 0)$
και η βέλτιστη τιμή της αντικειμενικής συνάρτησης θα είναι η:

$$4x_1^* + 6x_2^* - 2x_1^* - 2x_1^* \cdot x_2^* - 2x_2^{*2} = 150/36.$$

Ασκήσεις

1. Να λυθεί το πρόβλημα ΜΓΠ

$$\max \left[\ln(1+x_1) + x_1 - x_2 \right]$$

αν

$$x_1 \le 1$$
$$x_1 + x_2 \le 2$$
$$x_1, x_2 \ge 0$$

Λύση:

$$f(\underline{x}) = \ln(1+x_1) + x_1 - x_2$$
$$g_1(\underline{x}) = x_1 - 1 \le 0$$
$$g_2(\underline{x}) = x_1 + x_2 - 2 \le 0$$

Πρώτα πρέπει να εξετάσουμε εάν έχουμε πρόβλημα ΚΠ ώστε να εφαρμόσουμε τις συνθήκες Kuhn-Tucker όπως αυτές ορίζονται από το Πόρισμα 3.

Οι $g_1(\underline{x})$, $g_2(\underline{x})$ ως γραμμικές είναι κυρτές. Για την $f(\underline{x})$ έχουμε:

$$\nabla f(\underline{x}) = \left[\frac{1}{1+x_1} + 1, \ -1 \right]$$

και

$$Hf(\underline{x}) = \begin{bmatrix} -\frac{1}{(1+x_1)^2} & 0 \\ 0 & 0 \end{bmatrix}$$

Ο $Hf(\underline{x})$ είναι αρνητικά ημιορισμένος στο $\mathbb{R}$ $\forall x_1, x_2 \in \mathbb{R}$ (είναι προφανές ότι οι ιδιοτιμές του είναι ≤ 0 $\forall x_1 \ge 0$)

$$\Downarrow$$

έχω πρόβλημα ΚΠ και εφαρμόζω τις συνθήκες ΚΤ του Πορίσματος 3

$$\mu_1 \geq 0 \qquad (1)$$
$$\mu_2 \geq 0 \qquad (2)$$
$$\frac{1}{1+x_1} + 1 - \mu_1 - \mu_2 \leq 0 \qquad (3)$$
$$-1 - \mu_1 \cdot 0 - \mu_2 \leq 0 \qquad (4)$$
$$x_1 \left(\frac{1}{1+x_1} + 1 - \mu_1 - \mu_2 \right) = 0 \qquad (5)$$
$$x_2 \left(-1 \quad - \mu_2 \right) = 0 \qquad (6)$$
$$\mu_1 (x_1 - 1) = 0 \qquad (7)$$
$$\mu_2 (x_1 + x_2 - 2) = 0 \qquad (8)$$
$$x_1 - 1 \leq 0 \qquad (9)$$
$$x_1 + x_2 - 2 \leq 0 \qquad (10)$$
$$x_1 \geq 0 \qquad (11)$$
$$x_2 \geq 0 \qquad (12)$$

Για την εύρεση της λύσης θα πρέπει να εξετάσω περιπτώσεις για $\mu_1 \geq 0$, $\mu_2 \geq 0$, $x_1 \geq 0$, $x_2 \geq 0$.

Έχω όμως αναγκαστικά ότι $x_2 = 0$, γιατί αλλιώς η (6) θα οδηγούσε σε $\mu_2 = -1$ $\Rightarrow$ άτοπο λόγω της (2).

Περίπτωση 1 $\quad \mu_1 = 0 \quad \mu_2 = 0$

Σ' αυτή την περίπτωση η (3) δεν ικανοποιείται αφού $x_1 \geq 0$ κι έτσι η παράσταση του αριστερού μέλους δεν μπορεί ποτέ να γίνει ≤ 0

Περίπτωση 2 $\quad \mu_1 = 0 \quad \mu_2 > 0$

Τότε από την (8) σε συνδυασμό με το ότι $x_2 = 0$

έχουμε ότι $x_1 = 2$. Γι' αυτή όμως την τιμή του x_1 δεν ικανοποιείται η συνθήκη (9), οπότε η περίπτωση αυτή δεν δίνει λύση

Περίπτωση 3 $\quad \mu_1 > 0, \ \mu_2 = 0$

Τότε $(7) \Rightarrow x_1 = 1$

$$\Downarrow$$

$$(5) \Rightarrow \mu_1 = 3/2$$

Η λύση $[x_1^*, x_2^*, \mu_1^*, \mu_2^*] = [1, 0, 3/2, 0]$ ικανοποιεί όλες τις συνθήκες, οπότε αποτελεί λύση του προβλήματος

Περίπτωση 4 $\quad \mu_1 > 0 \quad \mu_2 > 0$

Τότε οι συνθήκες (7) και (8) για $x_2 = 0$ δίνουν 2 τιμές για το $x_1 \Rightarrow$ άτοπο $\Rightarrow$ η περίπτωση αυτή δεν δίνει λύση

Έτσι, η μοναδική λύση του προβλήματος είναι η:
$(x_1^*, x_2^*) = (1, 0)$ με μέγιστη τιμή της αντικειμενικής συνάρτησης την: $[\ln(2) + 1]$

2. Να λυθεί το πρόβλημα ΜΓΠ.

$$\max(9 - 2x_1 - 2x_1^2 + 2x_1 \cdot x_2 - x_2^2 - 2x_3)$$

όταν $x_1, x_2, x_3 \geq 0$

Λύση:

Πρώτα πρέπει να δούμε εάν η

$f(\underline{x}) = 9 - 2x_1 - 2x_1^2 + 2x_1 \cdot x_2 - x_2^2 - 2x_3$

είναι κοίλη ώστε να μπορέσουμε να εφαρμόσουμε το Πόρισμα 2 για προβλήματα τέτοιας μορφής

$$\nabla f(\underline{x}) = [-2 - 4x_1 + 2x_2,\ 2x_1 - 2x_2,\ -2]$$

$$Hf(\underline{x}) = \begin{bmatrix} -4 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Ο πίνακας $Hf(\underline{x})$ είναι αρνητικά ημιορισμένος γιατί όλες οι ιδιοτιμές του πίνακα είναι ≤ 0 (οι ιδιοτιμές του πίνακα είναι 0 και $\frac{-6 \pm \sqrt{20}}{2}$)

Συνεπώς το Πόρισμα 2 είναι εφαρμόσιμο και έτσι θα έχω ότι:

$$\frac{\partial f}{\partial x_j} \leq 0, \quad j = 1, 2, 3$$

$$x_j \frac{\partial f}{\partial x_j} = 0, \quad j = 1, 2, 3$$

Άρα:

$$-2-4x_1+2x_2 \le 0 \quad (1)$$
$$2x_1-2x_2 \le 0 \quad (2)$$
$$-2 \le 0 \quad (3)$$
$$x_1(-2-4x_1+2x_2) = 0 \quad (4)$$
$$x_2(2x_1-2x_2) = 0 \quad (5)$$
$$x_3(-2) = 0 \quad (6)$$

Από την (6) έχω κατ'ανάγκη ότι $x_3 = 0$.
Θα διακρίνω περιπτώσεις για τα $x_1 \ge 0$ και $x_2 \ge 0$

Περίπτωση 1 $\quad x_1 = 0 \;,\; x_2 = 0$
Βλέπουμε ότι ικανοποιούνται όλες οι συνθήκες (1) - (6) και η $(x_1^*, x_2^*, x_3^*) = (0,0,0)$ είναι λύση του προβλήματος

Περίπτωση 2 $\quad x_1 = 0 \;,\; x_2 > 0$

Τότε $(5) \Rightarrow x_1 = x_2 = 0 \Rightarrow$ άτοπο αφού υποθέσαμε ότι $x_2 > 0 \Rightarrow$ η περίπτωση αυτή δε δίνει λύση

Περίπτωση 3 $\quad x_1 > 0 \;,\; x_2 = 0$
Τότε $(4) \Rightarrow x_1 = -1/2 < 0 \Rightarrow$ άτοπο
$\Rightarrow$ Η περίπτωση αυτή δε δίνει λύση

Περίπτωση 4 $\quad x_1 > 0 \;,\; x_2 > 0$

Τότε
$$\left.\begin{aligned} (4) \Rightarrow -2-4x_1+2x_2 &= 0 \\ 2x_1-2x_2 &= 0 \end{aligned}\right\} \Rightarrow x_1 = x_2 = -1$$

$\Rightarrow$ άτοπο $\Rightarrow$ η περίπτωση αυτή δε δίνει λύση

Άρα, η μοναδική λύση είναι αυτή της Περίπτωσης 1
με $(x_1^*, x_2^*, x_3^*) = (0, 0, 0)$

3. Να λυθεί το πρόβλημα ΚΤΠ

$$\max(2x_1 - 2x_2 - 2x_3 - x_1^2 - x_2^2 - x_3^2)$$
$$x_1 + x_2 + x_3 \leq 1$$
$$4x_1 + 2x_2 \leq 2$$
$$x_1, x_2, x_3 \geq 0$$

Λύση: Το πρόβλημα γράφεται στη μορφή

$$\max(\underline{c}^T \underline{x} + \underline{x}^T D \underline{x})$$
$$A\underline{x} \leq \underline{b}$$
$$\underline{x} \geq 0$$

όπου $\underline{c} = [2, -2, -2]^T$

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \quad A = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 0 \end{bmatrix} \quad \underline{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$\underline{x} = [x_1, x_2, x_3]^T$

Το πρόβλημα μετασχηματίζεται σε πρόβλημα ΓΠ μέσω των:

$$-2D\underline{x} + A^T\underline{\mu} - \underline{y} = \underline{c}$$
$$A\underline{x} + \underline{\lambda} = \underline{b}$$
$$\underline{x} \geq 0, \ \underline{\mu} \geq 0, \ \underline{y} \geq 0, \ \underline{\lambda} \geq 0$$

όπου $\underline{y} = [y_1, y_2, y_3]^T$

$\underline{\mu} = [\mu_1, \mu_2]^T$ και $\underline{\lambda} = [\lambda_1, \lambda_2]^T$

Άρα:

$$\begin{aligned}
2x_1 \quad\quad + \mu_1 + 4\mu_2 - y_1 \quad\quad &= 2 \\
2x_2 \quad + \mu_1 + 2\mu_2 \quad - y_2 \quad &= -2 \\
2x_3 + \mu_1 \quad\quad - y_3 \quad &= -2 \\
x_1 + x_2 + x_3 \quad\quad + \lambda_1 \quad &= 1 \\
4x_1 + 2x_2 \quad\quad + \lambda_2 &= 2
\end{aligned}$$

$$x_1, x_2, x_3, \mu_1, \mu_2, y_1, y_2, y_3, \lambda_1, \lambda_2 \geq 0$$

$$x_1 y_1 = x_2 y_2 = x_3 y_3 = 0 = \mu_1 \lambda_1 = \mu_2 \lambda_2$$

Πολλ/ζοντας την 2η και 3η εξίσωση με (-1) και προσθέτοντας στην 1η εξίσωση μια τεχνητή μεταβλητή t_1 έχουμε:

$$\begin{aligned}
2x_1 \quad\quad + \mu_1 + 4\mu_2 - y_1 \quad\quad + t_1 &= 2 \\
-2x_2 \quad - \mu_1 - 2\mu_2 \quad + y_2 \quad &= 2 \\
-2x_3 - \mu_1 \quad\quad + y_3 \quad &= 2 \\
x_1 + x_2 + x_3 \quad\quad + \lambda_1 \quad &= 1 \\
4x_1 + 2x_2 \quad\quad + \lambda_2 &= 2
\end{aligned}$$

για $\max\ (-Mt_1 =$

$$= -2M - (-2M)x_1 - (-M)\mu_1 - (-4M)\mu_2 - My_1)$$

$$x_1, x_2, x_3, \mu_1, \mu_2, y_1, y_2, y_3, \lambda_1, \lambda_2 \geq 0$$

και $M >> 0$

Λόγω συμπληρωματικότητας στη βάση δεν μπορούν να είναι μαζί οι:

$$\left\{\begin{matrix} x_1 \text{ και } y_1 \\ x_2 \text{ και } y_2 \\ x_3 \text{ και } y_3 \end{matrix}\right\}, \quad \left\{\begin{matrix} \mu_1 \text{ και } \lambda_1 \\ \mu_2 \text{ και } \lambda_2 \end{matrix}\right\}$$

1ος Simplex

BM	x_1	x_2	x_3	μ_1	μ_2	y_1	y_2	y_3	λ_1	λ_2	t_1	θ
t_1	2	0	0	1	4	-1	0	0	0	0	1	2
y_2	0	-2	0	-1	-2	0	1	0	0	0	0	2
y_3	0	0	-2	-1	0	0	0	1	0	0	0	2
λ_1	1	1	1	0	0	0	0	0	1	0	0	1
← λ_2	4	2	0	0	0	0	0	0	0	1	0	2
$\bar{c}_j$	$-2M$	$\cancel{0}$	$\cancel{0}$	$\cancel{-M}$	$\cancel{-4M}$	M	0	0	0	0	0	$-2M$

Λόγω συμπληρωματικότητας στη βάση δεν μπορούν να μπουν οι x_2, x_3, μ_1, μ_2.

$\bar{c}_{x_1} < 0$ και από ελάχιστο πηλίκο $\Rightarrow$ x_1 : εισερχόμενη
λ_2 : εξερχόμενη

$\Downarrow$

2ος Simplex

BM	x_1	x_2	x_3	μ_1	μ_2	y_1	y_2	y_3	λ_1	λ_2	t_1	θ
← t_1	0	-1	0	1	4	-1	0	0	0	-1/2	1	1
y_2	0	-2	0	-1	-2	0	1	0	0	0	0	2
y_3	0	0	-2	-1	0	0	0	1	0	0	0	2
λ_1	0	1/2	1	0	0	0	0	0	1	-1/4	0	1/2
x_1	1	1/2	0	0	0	0	0	0	0	1/4	0	1/2
$\bar{c}_j$	0	$\cancel{M}$	$\cancel{0}$	$\cancel{-M}$	$-4M$	$\cancel{M}$	0	0	0	$\frac{1}{2}M$	0	$-M$

Λόγω συμπληρωματικότητας στη βάση δεν μπορούν να μπουν οι: x_2, x_3, μ_1, y_1

$\bar{c}_{\mu_2} < 0$ και από ελάχιστο πηλίκο $\Rightarrow$ μ_2 : εισερχόμενη
t_1 : εξερχόμενη

3os Simplex

BM	x_1	x_2	x_3	μ_1	μ_2	y_1	y_2	y_3	λ_1	λ_2	t_1	θ
μ_2	0	$-1/4$	0	$1/4$	1	$-1/4$	0	0	0	$-1/8$	$1/4$	$1/4$
y_2												$5/2$
y_3												2
λ_1												$1/2$
x_1												$1/2$
$\bar{c}_j$	0	0	0	0	0	0	0	0	0	0	M	0

$\bar{c}_j \geq 0 \quad \forall j \implies$ βέλτιστη λύση του αρχικού προβλήματος η:

$$\underline{X}_{opt} = [x_1^*, x_2^*, x_3^*]^T = [1/2, 0, 0]^T$$

με τιμή αντικειμενικής συνάρτησης $3/4$