Βελτιστοποίηση

Μ. Ζαζάνης
Άνοιξη 2019

Κεφάλαιο 1

Τετραγωνικές Μορφές και Κυρτότητα

1.1 Τετραγωνικές μορφές στον $\mathbb{R}^n$ και το θεώρημα του Taylor

Ορισμός 1. *Έστω*

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

συμμετρικός πίνακας $n \times n$ με στοιχεία στους πραγματικούς αριθμούς. Η συνάρτηση $\psi : \mathbb{R}^n \to \mathbb{R}$ που ορίζεται ως

$$\psi(x) = x^T A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{1.1}$$

ονομάζεται τετραγωνική μορφή.

Για παράδειγμα, η τετραγωνική μορφή που αντιστοιχεί στον πίνακα $\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$ είναι η $\psi(x_1, x_2) = x_1^2 - 2x_1x_2 + 2x_2^2$.

Ορισμός 2. *Ένας συμμετρικός πίνακας A, $n \times n$ ονομάζεται*

Θετικά ορισμένος *αν $x^T A x > 0$ για κάθε $x \in \mathbb{R}^n$, $x \neq 0$,*

Θετικά ημιορισμένος *αν $x^T A x \geq 0$ για κάθε $x \in \mathbb{R}^n$ με την ισότητα να ισχύει για κάποιο $x \neq 0$,*

αρνητικά ορισμένος *αν $x^T A x < 0$ για κάθε $x \in \mathbb{R}^n$, $x \neq 0$,*

αρνητικά ημιορισμένος *αν $x^T A x \leq 0$ για κάθε $x \in \mathbb{R}^n$ με την ισότητα να ισχύει για κάποιο $x \neq 0$.*

Θεώρημα 1. *Έστω συμμετρικός πίνακας A, $n \times n$ με πραγματικά στοιχεία. Οι εξής τέσσερεις προτάσεις είναι ισοδύναμες:*

(1) Ο πίνακας A είναι θετικά ορισμένος.

(2) Όλες οι ιδιοτιμές του A είναι θετικές.

(3) Οι ορίζουσες όλων των άνω αριστερά υποπινάκων του A είναι θετικές, δηλαδή

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0 \quad \text{για } k = 1, 2, \ldots, n.$$

(4) Όλοι οι οδηγοί d_i, $i = 1, \ldots, n$ στην απαλοιφή κατά Gauss χωρίς εναλλαγές γραμμών είναι θετικοί

Από το παραπάνω θεώρημα προκύπτει ότι η τετραγωνική μορφή (1.1) παίρνει αυστηρά θετικές τιμές για κάθε $x \neq 0$ αν και μόνο αν ο αντίστοιχος πίνακας A είναι θετικά ορισμένος. Αντίστοιχα η τετραγωνική μορφή παίρνει αυστηρά αρνητικές τιμές για κάθε $x \neq 0$ αν και μόνο αν ο πίνακας A είναι αρνητικά ορισμένος.

Ένα αντίστοιχο θεώρημα ισχύει για αρνητικά ορισμένους πίνακες. Συγκεκριμένα, όλες οι ιδιοτιμές του A στην περίπτωση αυτή είναι αρνητικές και το ίδιο ισχύει για τους οδηγούς d_i. Σε ότι αφορά τις ορίζουσες των άνω αριστερά υποπινάκων τα πρόσημα εναλλάσσονται ως εξής:

$$(-1)^k \begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0 \quad \text{για } k = 1, 2, \ldots, n.$$

1.2 Το Θεώρημα του Taylor για πραγαματικές συναρτήσεις πολλών μεταβλητών

Θεώρημα 2. *(Θεώρημα Taylor) Έστω $f : \mathbb{R}^n \to \mathbb{R}$ μια συνάρτηση με συνεχείς μερικές παραγώγους δεύτερης τάξης. Τότε*

$$f(x+h) = f(x) + \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} h_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} h_i h_j + o(\|h\|^2)$$

Χρησιμοποιούμε τον συμβολισμό $x = (x_1, \ldots, x_n)^T$ και $\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}$. Ο όρος $o(\|h\|^2)$, το λεγόμενο υπόλοιπο του Taylor, τείνει στο μηδέν γρηγορότερα από το $\|h\|^2$ όταν $h \to 0$. Το θεώρημα του Taylor γράφεται σε διανυσματική μορφή ως

$$f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} h^T H(x) h + o(\|h\|^2)$$

όπου ο *Εσσιανός πίνακας (Hessian)*

$$H(x) := \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

είναι συμμετρικός δεδομένου ότι, όταν οι δεύτερες παράγωγοι είναι συνεχείς (σε ένα ανοικτό χωρίο), $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$.

Ορισμός 3. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ δύο φορές συνεχώς παραγωγίσιμη. Το σημείο x^0 θα ονομάζεται* κρίσιμο *αν*

$$\nabla f(x^0) = 0.$$

Θεώρημα 3. *Έστω x^0 κρίσιμο σημείο της f. Αν ο πίνακας $H(x^0)$ είναι θετικά ορισμένος τότε το σημείο αυτό θα είναι σημείο τοπικού ελαχίστου ενώ αν είναι αρνητικά ορισμένος θα είναι σημείο τοπικού μεγίστου.*

(Η απόδειξη του θεωρήματος αυτού είναι άμεση συνέπεια του θεωρήματος του Taylor και μας επιτρέπει να βρούμε τα σημεία τοπικού μεγίστου και τοπικού ελαχίστου.)

Πρόβλημα 1. Να βρείτε τα τοπικά ακρότατα (μέγιστα ή ελάχιστα) της συνάρτησης $f(x, y) = x^2y^2e^{-(x+y)^2}$ για $(x, y) \in \mathbb{R}^2$.

1.3 Κυρτότητα

Ορισμός 4. *Ένα υποσύνολο C του $\mathbb{R}^n$ ονομάζεται κυρτό αν για κάθε $x, y \in C$ και $\theta \in [0, 1]$, $\theta x+(1-\theta)y \in C$.*

Ορισμός 5. *Έστω $B \subset \mathbb{R}^n$. Η* κυρτή θήκη *του B είναι το μικρότερο κυρτό σύνολο το οποίο περιέχει το B, $conv(B) := \{\sum_{i=1}^k \theta_k x^k : k \in \mathbb{N}, x^i \in B, \theta_i \geq 0, \sum_{i=1}^k \theta_i = 1\}$*

Ορισμός 6. *Μια συνάρτηση $f : C \to \mathbb{R}$ ορισμένη πάνω σ' ένα κυρτό σύνολο $C \subset \mathbb{R}^n$ ονομάζεται* κυρτή *(convex) αν, για κάθε $x, y \in C$ και κάθε $\theta \in [0, 1]$*

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y). \tag{1.2}$$

Θα ονομάζεται κοίλη *(concave) αν*

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta)f(y). \tag{1.3}$$

Αν οι (1.2), (1.3) ισχύουν ως ανισότητες τότε η f ονομάζεται αυστηρώς κυρτή ή κοίλη *αντίστοιχως.*

Από τον παραπάνω ορισμό είναι σαφές ότι αν η συνάρτηση f είναι κυρτή στο C τότε η $-f$ είναι κοίλη. Παρατηρήστε επίσης ότι είναι σκόπιμο να ορίσουμε μια κυρτή (ή κοίλη) συνάρτηση είτε σε ολόκληρο το $\mathbb{R}^n$ ή σε ένα κυρτό υποσύνολό του, μια και στον ορισμό, αν τα x και y είναι στο πεδίο ορισμού της συνάρτησης, τότε θα πρέπει και το $\theta x + (1-\theta)y$ να ανήκει επίσης στο πεδίο ορισμού της συνάρτησης. Μπορεί να αποδειχθεί ότι αν η f είναι κυρτή (ή κοίλη) συνάρτηση τότε δεν μπορεί να είναι ασυνεχής στο εσωτερικό του πεδίου ορισμού της. Τυχόν ασυνέχειες, αν υπάρχουν, πρέπει να βρίσκονται στο σύνορο του πεδίου ορισμού.

Το ακόλουθο θεώρημα δείχνει την σχέση ανάμεσα στην κυρτότητα και τις μερικές παραγώγους των κυρτών και κοίλων συναρτήσεων.

Θεώρημα 4. *Έστω $f : C \to \mathbb{R}^n$ κυρτή συνάρτηση με συνεχείς μερικές παραγώγους στο C. Η f είναι κυρτή αν και μόνο αν, για κάθε x_0 στο εσωτερικό του C*

$$f(x) - f(x^0) \ \geq \ \nabla f(x^0) \cdot (x - x^0) \ = \ \sum_{i=1}^n \frac{\partial f(x^0)}{\partial x_i}(x_i - x_i^0). \tag{1.4}$$

Ομοίως, η f είναι κοίλη αν και μόνο αν,

$$f(x) - f(x^0) \;\leq\; \nabla f(x^0) \cdot (x - x^0) \;=\; \sum_{i=1}^{n} \frac{\partial f(x^0)}{\partial x_i}(x_i - x_i^0). \tag{1.5}$$

Η f είναι αυστηρώς κυρτή αν και μόνο η ανισότητα (1.4) είναι αυστηρή. Παρομοίως, η f είναι αυστηρώς κοίλη αν και μόνο αν η ανισότητα (1.5) είναι αυστηρή.

Απόδειξη. Αποδεικνύουμε ότι αν η f είναι κυρτή τότε ισχύει η (1.4). Αφού η f είναι κυρτή, $f(\lambda x + (1-\lambda)x^0) \leq \lambda f(x) + (1-\lambda) f(x^0)$ ή, ισοδύναμα,

$$f(x) - f(x^0) \geq \frac{f(x^0 + \lambda(x - x^0)) - f(x^0)}{\lambda}.$$

Αντίστροφα, υποθέτουμε ότι η (1.4) ισχύει. Θα αποδείξουμε τότε ότι η f είναι κυρτή. Αν $\bar{x} = \lambda x + (1-\lambda)x^0$ τότε

$$\begin{aligned} f(x) - f(\bar{x}) &\geq \nabla f(\bar{x}) \cdot (x - \bar{x}), \\ f(x^0) - f(\bar{x}) &\geq \nabla f(\bar{x}) \cdot (x^0 - \bar{x}). \end{aligned}$$

Πολλαπλασιάζοντας την πρώτη σχέση με λ και την δεύτερη με $1-\lambda$ παίρνουμε

$$\lambda f(x) - \lambda f(\bar{x}) + (1-\lambda) f(x^0) - (1-\lambda) f(\bar{x}) \geq \nabla f(\bar{x}) \cdot (\lambda x - \lambda \bar{x} + (1-\lambda)x^0 - (1-\lambda)\bar{x})$$

ή

$$\lambda f(x) + (1-\lambda) f(x^0) - f(\bar{x}) \geq 0.$$

Η τελευταία σχέση αυτή συνεπάγεται ότι η f είναι κυρτή. □

Αν η f είναι ορισμένη σε ένα κυρτό σύνολο $S \subset \mathbb{R}^n$ και $M_f := \{x \in S$

Στην περίπτωση που η $f : \mathbb{R}^n \to \mathbb{R}$ είναι δύο φορές συνεχώς παραγωγίσιμη, αν $H(x)$ είναι ο Εσσιανός πίνακας των δευτέρων παραγώγων στο σημείο $x \in \mathbb{R}^n$,

$$H(x) := \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix},$$

ισχύει το ακόλουθο

Θεώρημα 5. *Έστω C ένα ανοικτό κυρτό υποσύνολο του $\mathbb{R}^n$ και $f : C \to \mathbb{R}$ δύο φορές συνεχώς παραγωγίσιμη στο C. Τότε*

1. *Η f είναι κυρτή στο C αν και μόνο αν ο Εσσιανός πίνακας $H(x)$ είναι θετικά ημιορισμένος για κάθε $x \in C$.*

2. *Η f είναι κοίλη στο C αν και μόνο αν ο Εσσιανός πίνακας $H(x)$ είναι αρνητικά ημιορισμένος για κάθε $x \in C$.*

Επίσης η f είναι αυστηρώς κυρτή αν και μόνο αν ο Εσσιανός πίνακας είναι θετικά ορισμένος. *Παρομοίως, η f είναι αυστηρώς κοίλη αν και μόνο αν ο Εσσιανός πίνακας είναι* αρνητικά ορισμένος.

Κεφάλαιο 2

Βελτιστοποίηση υπό περιορισμούς ισότητας

Θέλουμε να μεγιστοποιήσουμε την συνάρτηση $f : \mathbb{R}^n \to \mathbb{R}$ κάτω από τους περιορισμούς $g(\mathbf{x}) = 0$ όπου $g : \mathbb{R}^n \to \mathbb{R}$ με $m < n$. Έστω $m, n \in \mathbb{N}$ όπου $m > n$ και f, g_i, $i = 1, 2, \ldots, m$ συναρτήσεις $\mathbb{R}^n \to \mathbb{R}$. Θεωρούμε το πρόβλημα βελτιστοποίησης

$$\begin{aligned} &\min f(\mathbf{x}) \\ &\text{υπό τους περιορισμούς} \\ &g_i(\mathbf{x}) = 0, \quad i = 1, 2, \ldots, m. \end{aligned} \tag{2.1}$$

Έστω $L : \mathbb{R}^{n+m} \to \mathbb{R}$ η συνάρτηση που ορίζεται ως

$$L(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}).$$

Η συνάρτηση αυτή ονομάζεται Λαγκρανζιανή. Βρίσκουμε τα κρίσημα σημεία υπό περιοριμούς ως λύσεις των εξισώσεων

$$\begin{aligned} &\frac{\partial L}{\partial x_j} = 0, \qquad j = 1, 2, \ldots, n \\ &\frac{\partial L}{\partial \lambda_i} = g_i(\mathbf{x}) = 0, \qquad i = 1, 2, \ldots, m. \end{aligned}$$

Τα σημεία μεγίστου βρίσκονται ανάμεσα σ′ αυτά τα σημεία.

Παράδειγμα.

$$\begin{aligned} &\min x_1^2 + x_2^2 + x_3^2 \\ &\text{υπό τους περιορισμούς} \\ &x_1 + x_2 + x_3 = 0 \\ &x_1 + 2x_2 + 3x_3 = 1. \end{aligned}$$

Η Λαγκρανζιανή είναι

$$L = x_1^2 + x_2^2 + x_3^2 - \lambda_1(x_1 + x_2 + x_3) - \lambda_2(x_1 + 2x_2 + 3x_3 - 1)$$

και τα στάσιμα σημεία δίνονται από το σύστημα $\partial L/\partial x_j = 0,\ \partial L/\partial \lambda_i = 0,\ j = 1,2,3,\ i = 1,2.$

$$\begin{array}{rrrrrcl}
-2x_1 & & & -\lambda_1 & -\lambda_2 & = & 0 \\
& -2x_2 & & -\lambda_1 & -2\lambda_2 & = & 0 \\
& & -2x_3 & -\lambda_1 & -3\lambda_2 & = & 0 \\
x_1 & +x_2 & +x_3 & & & = & 0 \\
x_1 & +2x_2 & +3x_3 & & & = & 1
\end{array}$$

Το σύστημα αυτό έχει τη μοναδική λύση $x_1 = -1/2, x_2 = 0, x_3 = 1/2$ και $\lambda_1 = 2, \lambda_2 = -1$.

Η γεωμετρική σημασία του θεωρήματος του Lagrange είναι ότι, προκειμένου να είναι το σημείο $x^* = (x_1^*, \ldots, x_n^*)$ κρίσημο σημείο θα πρέπει η κλίση της f στο σημείο αυτό, $\nabla f(x^*)$ να είναι είναι κάθετη στην υπερεπιφάνεια που ορίζουν οι περιορισμοί g στο ίδιο σημείο. Ο κάθετος αυτός υπόχωρος είναι είναι ο χώρος που δημιουργείται από τα διανύσματα $\nabla g_i(x^*),\ i = 1, \ldots, m$. Οι γραμμικοί αυτοί συνδιασμοί είναι ο υπόχωρος $\{\sum_{i=1}^m \lambda_i \nabla g_i(x^*) : \lambda_i \in \mathbb{R}\}$.

2.1 Ένα παράδειγμα που εξηγεί τη φυσική σημασία των συνθηκών

Έστω $f, g : \mathbb{R}^2 \to \mathbb{R}$ συναρτήσεις συνεχώς παραγωγίσιμες. Ας υποθέσουμε ότι θέλουμε να μεγιστοποιήσουμε την $f(x,y)$ κάτω από τον περιορισμό $g(x,y) = c$. Μια απλή σκέψη είναι η εξής: η σχέση $g(x,y) = c$ επιτρέπει, τουλάχιστον 'τοπικά', να λύσουμε ως προς y και να εκφράσουμε το y ως συνάρτηση του x και της σταθεράς c. Αυτό δικαιολογείται από το θεώρημα της 'πεπλεγμένης συναρτήσεως'. Υποθέτοντας ότι υπάρχει συνάρτηση ϕ (η οποία εξαρτάται από την τιμή του c) τέτοια ώστε $g(x, \phi(x)) = c$ και παραγωγίζοντας ως προς x έχουμε

$$\frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\phi'(x) = 0.$$

Η αναγκαία συνθήκη για να έχει το f τοπικό ακρότατο στο σημείο x είναι

$$\frac{df(x, \phi(x))}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\phi'(x) = 0.$$

Από τις δύο αυτές σχέσεις, υπό την προϋπόθεση ότι $\frac{\partial f(x,y)}{\partial y} \neq 0$ και $\frac{\partial g}{\partial y} \neq 0$, έχουμε

$$\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}} = \frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}} = \phi'(x)$$

ή, ισοδύναμα, υπάρχει $\lambda \in \mathbb{R}$ τέτοιο ώστε

$$\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) = \lambda\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right), \quad \text{ή ισοδύναμα,} \quad \nabla g(x) = \lambda \nabla f(x).$$

Η παραπάνω σχέση εννοείται ότι ισχύει για τα σημεία εκείνα που είναι δεσμευμένα ακρότατα.

2.2 Η οικονομική σημασία των συντελεστών Lagrange

Ας γράψουμε το πρόβλημα μεγιστοποίησης (2.1) στη μορφή

$$\begin{array}{l}
\min f(\mathbf{x}) \\
\text{υπό τους περιορισμούς} \\
g_i(\mathbf{x}) = b_i, \quad i = 1, 2, \ldots, m.
\end{array} \tag{2.2}$$

Η Λαγκρανζιανή είναι

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i (g_i(\mathbf{x}) - b_i)$$

και οι εξισώσεις για τα κρίσιμα σημεία είναι

$$\frac{\partial L}{\partial x_j} = \frac{\partial f}{\partial x_j} - \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j} = 0, \qquad j = 1, \ldots, n, \tag{2.3}$$

$$g_i(\mathbf{x}) = b_i, \qquad i = 1, \ldots, m. \tag{2.4}$$

Έστω $\mathbf{x}^*$, $\boldsymbol{\lambda}^*$, οι τιμές που μεγιστοποιούν το κριτήριο f και $f(\mathbf{x}^*)$ η βέλτιστη τιμή του κριτηρίου. Όλες αυτές οι ποσήτητες εξαρτώνται βέβαια από τα b_i, $i = 1, \ldots, m$. Το ερώτημα που θέτουμε εδώ είναι πώς αλλάζει η βέλτιστη τιμή του κριτηρίου όταν μεταβληθούν λίγο τα b_i. Για το σκοπό αυτό αυτό υπολογίζουμε τις παραγώγους

$$\frac{\partial f(\mathbf{x}^*)}{\partial b_i} = \sum_{k=1}^{n} \frac{\partial f(\mathbf{x}^*)}{\partial x_k} \frac{\partial x_k^*}{\partial b_i}, \qquad i = 1, \ldots, m, \tag{2.5}$$

$$\frac{\partial g_l(\mathbf{x}^*)}{\partial b_i} = \sum_{j=1}^{n} \frac{\partial g_l(\mathbf{x}^*)}{\partial x_j} \frac{\partial x_j^*}{\partial b_i} = \delta_{li}, \qquad i, l = 1, \ldots, m. \tag{2.6}$$

όπου

$$\delta_{ki} = \begin{cases} 1 & \text{αν } k = i, \\ 0 & \text{διαφορετικά.} \end{cases}$$

Τέλος, από την (2.3), στις βέλτιστες τιμές $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$,

$$\frac{\partial f(\mathbf{x}^*)}{\partial x_k} = \sum_{l=1}^{m} \lambda_l^* \frac{\partial g_l(\mathbf{x}^*)}{\partial x_k}. \tag{2.7}$$

Από τις (2.5), (2.7), έχουμε

$$\frac{\partial f(\mathbf{x}^*)}{\partial b_i} = \sum_{k=1}^{n} \left(\sum_{l=1}^{m} \lambda_l^* \frac{\partial g_l(\mathbf{x}^*)}{\partial x_k} \right) \frac{\partial x_k^*}{\partial b_i} = \sum_{l=1}^{m} \lambda_l^* \sum_{k=1}^{n} \frac{\partial g_l(\mathbf{x}^*)}{\partial x_k} \frac{\partial x_k^*}{\partial b_i} = \sum_{l=1}^{m} \lambda_l^* \delta_{li} = \lambda_i^*. \tag{2.8}$$

Συνεπώς, οι τιμές των πολλαπλασιαστών του Lagrange στο βέλτιστο σημείο δείχνουν πόσο θα μεταβαλλόταν οριακά η βέλτιστη τιμή του κριτηρίου αν μεταβαλλόταν οριακά το δεξί μέλος του αντίστοιχου περιορισμού. Αν υποθέσουμε ότι αυξάνουμε το b_i σε $b_i + h$ όπου h μια μικρή ποσότητα, η νέα βέλτιση τιμή του κριτηρίου θα είναι $f(\mathbf{x}^*) + \frac{\partial f(\mathbf{x}^*)}{\partial b_i} h = f(\mathbf{x}^*) + \lambda_i^* h$. Για το λόγο αυτό οι συντελεστές Lagrange υπολογισμένοι στο βέλτιστο ονομάζονται και *σκιώδεις τιμές* (shadow prices).

2.3 Ικανές συνθήκες για ελάχιστο ή μέγιστο σε τετραγωνικές μορφές με περιορισμούς ισότητας

Έστω $Q(x) := x^T A x$ όπου $x \in \mathbb{R}^n$ και ο A είναι συμμετρικός πίνακας $n \times n$, δηλαδή $Q(x) := \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$. Θεωρούμε το πρόβλημα της μεγιστοποίησης ή ελαχιστοποίησης της $Q(x)$ υπό τους γραμμικούς περιορισμούς $Bx = 0$ όπου B είναι ένας πίνακας $k \times n$ (με $k < n$, δηλαδή ο αριθμός των περιορισμών είναι μικρότερος από τον αριθμό των μεταβλητών).

Θεώρημα 6. *Έστω A συμμετρικός πίνακας $n \times n$ και B πίνακας $k \times n$, τέτοιος ώστε* $\text{rank}(B) = k$ *και* $\det B_k \neq 0$.

$$B = \left[\begin{array}{cccc|ccc} b_{11} & b_{12} & \cdots & b_{1k} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2k} & \cdots & b_{2n} \\ \vdots & & & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} & \cdots & b_{kn} \end{array}\right]$$

με B_k το αριστερό $k \times k$ μπλοκ $\underbrace{\begin{bmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & & \vdots \\ b_{k1} & \cdots & b_{kk}\end{bmatrix}}_{B_k}$

Ορίζουμε τον πίνακα $(n + k) \times (n + k)$

$$C = \left[\begin{array}{ccc|ccccc} 0 & \cdots & 0 & b_{11} & \cdots & b_{1k} & \cdots & b_{1n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{k1} & \cdots & b_{kk} & \cdots & b_{kn} \\ \hline b_{11} & \cdots & b_{k1} & a_{11} & \cdots & a_{1k} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ b_{1k} & \cdots & b_{kk} & a_{k1} & \cdots & a_{kk} & \cdots & a_{kn} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ b_{1n} & \cdots & b_{kn} & a_{n1} & \cdots & a_{nk} & \cdots & a_{nn} \end{array}\right].$$

Αν C_r είναι ο άνω αριστερά τετραγωνικός υποπίνακας του C του οποίου το κάτω δεξιά στοιχείο είναι το a_{rr} (και επομένως έχει διάσταση $k + r$) δηλαδή

$$C_r = \begin{bmatrix} 0 & \cdots & 0 & b_{11} & \cdots & b_{1r} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & b_{k1} & \cdots & b_{kr} \\ b_{11} & \cdots & b_{k1} & a_{11} & \cdots & a_{1r} \\ \vdots & & \vdots & \vdots & & \vdots \\ b_{1r} & \cdots & b_{kr} & a_{r1} & \cdots & a_{rr} \end{bmatrix}$$

τότε

1. $x^\top Ax > 0 \forall x \neq 0 : Bx = 0$ *αν και μόνο αν για* $r = k + 1, k + 2, \ldots, n$ $(-1)^k \det C_r > 0$.
2. $x^\top Ax < 0 \forall x \neq 0 : Bx = 0$ *αν και μόνο αν για* $r = k + 1, k + 2, \ldots, n$ $(-1)^r \det C_r > 0$.

Από το παραπάνω Θεώρημα προκύπτει το συμπέρασμα ότι

1. Αν ισχύουν οι συνθήκες της περίπτωσης 1 του Θεωρήματος τότε το σημείο $x = 0$ είναι σημείο ολικού ελαχίστου της $Q(x)$ υπό τους περιορισμούς $Bx = 0$.
2. Αν ισχύουν οι συνθήκες της περίπτωσης 2 του Θεωρήματος τότε το σημείο $x = 0$ είναι σημείο ολικού μεγίστου της $Q(x)$ υπό τους περιορισμούς $Bx = 0$.
3. Σε κάθε άλλη περίπτωση το $x = 0$ δεν είναι ακρότατο.

Παράδειγμα 1: Έστω ότι μας ενδιαφέρει να δούμε κατά πόσον η τετραγωνική μορφή

$$[x_1, x_2]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

υπό τον γραμμικό περιορισμό $x_1 b_1 + x_2 b_2 = 0$ είναι θετικά ή αρνητικά ορισμένη. (Ο πίνακας μπορεί να θεωρηθεί συμμετρικός με $a_{12} = a_{21}$.) Ο γραμμικός περιορισμός μπορεί να εκφραστεί παραμετρικά ως $x_1 = tb_2$, $x_2 = -tb_1$, $t \in \mathbb{R}$ και επομένως η τετραγωνική μορφή γράφεται ως

$$\begin{aligned} t^2[b_2, -b_1]\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} b_2 \\ -b_1 \end{bmatrix} = t^2(b_2^2 a_{11} + b_1^2 a_{22} - 2b_1 b_2 a_{12}) &= -t^2 \begin{vmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{vmatrix} \\ &= (-1)t^2 \det C_2 \end{aligned}$$

αφού στην περίπτωση αυτή

$$C = \begin{bmatrix} 0 & b_1 & b_2 \\ b_1 & a_{11} & a_{12} \\ b_2 & a_{21} & a_{22} \end{bmatrix}.$$

Δεδομένου ότι $k = 1$ και $r = 2$ βλέπουμε ότι ανάλογα με το αν η $\det C_2$ είναι θετική ή αρνητική η τετραγωνική μορφή υπό τον γραμμικό περιορισμό είναι αντίστοιχα αρνητικά ή θετικά ορισμένη.

Παράδειγμα 2: Παρομοίως, έστω ότι θέλουμε να εξετάσουμε αν η τετραγωνική μορφή

$$[x_1, x_2, x_3, x_4]\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

υπό τους γραμμικούς περιορισμούς

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Έχουμε $k = 2$, $n = 4$, και

$$C_3 = \begin{bmatrix} 0 & 0 & b_{11} & b_{12} & b_{13} \\ 0 & 0 & b_{21} & b_{22} & b_{23} \\ b_{11} & b_{21} & a_{11} & a_{12} & a_{13} \\ b_{12} & b_{22} & a_{21} & a_{22} & a_{23} \\ b_{13} & b_{23} & a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad C_4 = \begin{bmatrix} 0 & 0 & b_{11} & b_{12} & b_{13} & b_{14} \\ 0 & 0 & b_{21} & b_{22} & b_{23} & b_{24} \\ b_{11} & b_{21} & a_{11} & a_{12} & a_{13} & a_{14} \\ b_{12} & b_{22} & a_{21} & a_{22} & a_{23} & a_{24} \\ b_{13} & b_{23} & a_{31} & a_{32} & a_{33} & a_{34} \\ b_{14} & b_{24} & a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

Αν $\det C_3 < 0$, και $\det C_4 < 0$ τότε η τετραγωνική μορφή υπό τους δεδομένους περιορισμούς είναι *θετικά ορισμένη.* Αν $\det C_3 < 0$, και $\det C_4 > 0$ τότε η τετραγωνική μορφή υπό τους δεδομένους περιορισμούς είναι *αρνητικά ορισμένη.*

Απόδειξη στην περίπτωση που $k = 1$

Εξετάζουμε εδώ κατά πόσον $x^T Ax > 0$ για κάθε $x \neq 0$ τέτοιο ώστε $b^T x = 0$, όπου $b^T = [b_1, b_2, \ldots, b_n]$. Αυτό σημαίνει ότι για κάθε $\theta > 0$ αρκετά μεγάλο, $x^T Ax + \theta\left(b^T x\right)^2 > 0$ για κάθε $x \neq 0$. Με άλλα λόγια ο πίνακας

$$\begin{bmatrix} a_{11}+\theta b_1^2 & a_{12}+\theta b_1 b_2 & \cdots & a_{1n}+\theta b_1 b_n \\ a_{21}+\theta b_2 b_1 & a_{22}+\theta b_2^2 & \cdots & a_{2n}+\theta b_2 b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+\theta b_n b_1 & a_{n2}+\theta b_n b_2 & \cdots & a_{nn}+\theta b_n^2 \end{bmatrix}$$

είναι θετικά ορισμένος. Αυτό είναι ισοδύναμο με την συνθήκη

$$\Delta_k := \begin{vmatrix} a_{11}+\theta b_1^2 & a_{12}+\theta b_1 b_2 & \cdots & a_{1k}+\theta b_1 b_k \\ a_{21}+\theta b_2 b_1 & a_{22}+\theta b_2^2 & \cdots & a_{2k}+\theta b_2 b_k \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1}+\theta b_k b_1 & a_{k2}+\theta b_k b_2 & \cdots & a_{kk}+\theta b_k^2 \end{vmatrix} > 0, \qquad k = 1, 2, \ldots, n.$$

Δεδομένου ότι ισχύει η σχέση

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ \theta b_1 & a_{11}+\theta b_1^2 & a_{12}+\theta b_1 b_2 & \cdots & a_{1k}+\theta b_1 b_k \\ \theta b_2 & a_{21}+\theta b_2 b_1 & a_{22}+\theta b_2^2 & \cdots & a_{2k}+\theta b_2 b_k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \theta b_k & a_{k1}+\theta b_k b_1 & a_{k2}+\theta b_k b_2 & \cdots & a_{kk}+\theta b_k^2 \end{bmatrix} = \begin{bmatrix} -1 & b_1 & b_2 & \cdots & b_k \\ \theta b_1 & a_{11} & a_{12} & \cdots & a_{1k} \\ \theta b_2 & a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \theta b_k & a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix} \begin{bmatrix} 1 & b_1 & b_2 & \cdots & b_k \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

παίρνοντας ορίζουσες έχουμε

$$\Delta_k := -\theta \begin{vmatrix} 0 & b_1 & b_2 & \cdots & b_k \\ b_1 & a_{11} & a_{12} & \cdots & a_{1k} \\ b_2 & a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_k & a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}.$$

Άσκηση

Πρόβλημα 1

Να εξετάσετε την τετραγωνική μορφή $x_1^2 + x_2^2 + x_3^2 + x_4^2$ υπό τους περιορισμούς $x_1 + x_2 = 0$, $x_1 + x_3 = 0$. (Βέβαια, εδώ μπορούμε να λύσουμε πολύ εύκολα: $x_2 = -x_1$, $x_3 = -x_1$ και επομένως η τετραγωνική μορφή γίνεται $3x_1^2 + x_4^2$ που βέβαια είναι θετικά ορισμένη και επομένως το σημείο $(0, 0, 0, 0)$ είναι σημείο τοπικού (και ολικού) ελαχίστου.)

Η πλαισιωμένη Εσσιανή είναι

$$C = \left[\begin{array}{ccccc|c} 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

Ισχύει ότι $\det C_2 = 3$ και $\det C_3 = 3$ και επομένως (με $k = 2$) $(-1)^k \det C_r > 0$ για $r = 2, 3$. Συνεπώς η τετραγωνική μορφή είναι θετικά ορισμένη και το σημείο $(0, 0, 0, 0)$ είναι σημείο τοπικού ελαχίστου.

2.4 Ικανές συνθήκες για την ύπαρξη τοπικού μεγίστου ή ελαχίστου για προβλήματα με περιορισμούς ισότητας

Θεώρημα 7. *Έστω $f : \mathbb{R}^n \to \mathbb{R}$ και $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$ δύο φορές συνεχώς παραγωγίσιμες συναρτήσεις, με $m < n$. Έστω $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$, και $L(\mathbf{x}, \boldsymbol{\lambda}) := f(\mathbf{x}) - \sum_{i=1}^m \lambda_i (g_i(\mathbf{x}) - b_i)$. Έστω $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ ένα σημείο για το οποίο*

$$\frac{\partial}{\partial x_j} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0, \qquad j = 1, \ldots, n, \tag{2.9}$$

$$\frac{\partial}{\partial \lambda_i} L(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0, \qquad i = 1, \ldots, m. \tag{2.10}$$

Ορίζουμε τον πίνακα $(n+m) \times (n+m)$

$$C(\mathbf{x}^*, \boldsymbol{\lambda}^*) = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_m} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_m} & \cdots & \frac{\partial g_m}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_m} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_m} & \cdots & \frac{\partial g_m}{\partial x_m} & \frac{\partial^2 L}{\partial x_m \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_m^2} & \cdots & \frac{\partial^2 L}{\partial x_m \partial x_n} \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_n} & \cdots & \frac{\partial g_m}{\partial x_n} & \frac{\partial^2 L}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_n \partial x_m} & \cdots & \frac{\partial^2 L}{\partial x_n^2} \end{bmatrix}$$

όπου

$$\frac{\partial^2 L}{\partial x_j \partial x_k} = \frac{\partial^2 f(\mathbf{x}^*)}{\partial x_j \partial x_k} - \sum_{i=1}^m \lambda_i^* \frac{\partial^2 g_i(\mathbf{x}^*)}{\partial x_j \partial x_k}, \qquad \frac{\partial g_i}{\partial x_j} = \frac{\partial g_i(\mathbf{x}^*)}{\partial x_j}.$$

Αν C_r είναι ο άνω αριστερά τετραγωνικός υποπίνακας του C του οποίου το κάτω δεξιά στοιχείο είναι το $\frac{\partial^2 L}{\partial x_r^2}$ (και επομένως έχει διάσταση $m + r$) δηλαδή

$$C_r = \begin{bmatrix} 0 & \cdots & 0 & \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_r} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \frac{\partial g_m}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_r} \\ \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_m}{\partial x_1} & \frac{\partial^2 L}{\partial x_1^2} & \cdots & \frac{\partial^2 L}{\partial x_1 \partial x_r} \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial g_1}{\partial x_r} & \cdots & \frac{\partial g_m}{\partial x_r} & \frac{\partial^2 L}{\partial x_r \partial x_1} & \cdots & \frac{\partial^2 L}{\partial x_r^2} \end{bmatrix}$$

τότε, για το πρόβλημα $\max(\min) f(x_1, \ldots, x_n)$ υπό τους περιορισμούς $g_i(x_1, \ldots, x_n) = b_i$, $i = 1, \ldots, m$:

1. *Το $\mathbf{x}^*$ είναι σημείο τοπικού ελαχίστου (υπό τους δεδομένους περιορισμούς) αν και μόνο αν για $r = m+1, m+2, \ldots, n$, $(-1)^m \det C_r > 0$.*

2. *Το $\mathbf{x}^*$ είναι σημείο τοπικού μεγίστου (υπό τους δεδομένους περιορισμούς) αν και μόνο αν για $r = m+1, m+2, \ldots, n$ $(-1)^r \det C_r > 0$.*

Ασκήσεις

Να βρείτε τα σημεία τοπικού μεγίστου και τοπικού ελαχίστου της συνάρτησης $f(x,y) = x^2 + y^2$ υπό τον περιορισμό $g(x,y) = x^2 + xy + y^2 = 3$. Η Λαγκρανζιανή είναι $L = f(x,y) - \lambda(g(x,y) - 3)$. Οι συνθήκες για τα κρίσημα σημεία είναι

$$\frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - \lambda\frac{\partial g}{\partial x} = 2x - \lambda(2x+y) = 0 \tag{2.11}$$

$$\frac{\partial L}{\partial y} = \frac{\partial f}{\partial y} - \lambda\frac{\partial g}{\partial y} = 2y - \lambda(x+2y) = 0 \tag{2.12}$$

$$\frac{\partial L}{\partial \lambda} = x^2 + xy + y^2 - 3 = 0 \tag{2.13}$$

Οι εξισώσεις (2.11), (2.12) γράφονται ως

$$\begin{bmatrix} 2(1-\lambda) & -\lambda \\ -\lambda & 2(\lambda-1) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Αν η ορίζουσα του παραπάνω πίνακα δεν είναι μηδέν, η μόνη λύση που έχει το σύστημα είναι $x = 0$, $y = 0$, που δεν ικανοποιεί την (2.13). Επομένως εξατάζουμε την περίπτωση $4(1-\lambda)^2 - \lambda^2 = 0$ που δίνει τις λύσεις $\lambda = \frac{2}{3}$ και $\lambda = 2$.

α)$\lambda = \frac{2}{3}$. Η τιμή αυτή με την (2.11) δίνει $x\frac{2}{3} - \frac{2}{3}u = 0$ ή $x = y$. Αυτή η σχέση μαζί με την (2.13) δίνει $3x^2 = 3$ και επομένως έχουμε $x = \pm 1$. Έτσι έχουμε τα σημεία $(x,y,\lambda) = (1,1,\frac{2}{3})$ και $(-1,-1,\frac{2}{3})$.

β) $\lambda = 2$. Η τιμή αυτή με την (2.11) δίνει $-2x - 2y = 0$ ή $y = -x$. Αυτή η σχέση μαζί με την (2.13) δίνει $x^2 = 3$ και επομένως έχουμε $x = \pm\sqrt{3}$. Έτσι έχουμε τα σημεία $(x,y,\lambda) = (\sqrt{3}, -\sqrt{3}, \frac{2}{3})$ και $(-\sqrt{3}, \sqrt{3}, \frac{2}{3})$.

Η πλαισιωμένη Εσσιανή είναι

$$\begin{bmatrix} 0 & g_x & g_y \\ g_x & L_{xx} & L_{xy} \\ g_y & L_{yx} & L_{yy} \end{bmatrix} = \begin{bmatrix} 0 & 2x+y & x+2y \\ 2x+y & 2(1-\lambda) & -\lambda \\ x+2y & -\lambda & 2(1-\lambda) \end{bmatrix}.$$

Ισχύει $k = 1$ και επομένως $k+1 = 2$, άρα εξετάζουμε σε κάθε περίπτωση μια μόνο ορίζουσα.

Το σημείο $(x,y,\lambda) = (1,1,\frac{2}{3})$. Έχουμε

$$\Delta_2 = \begin{vmatrix} 0 & 3 & 3 \\ 3 & \frac{2}{3} & -\frac{2}{3} \\ 3 & -\frac{2}{3} & \frac{2}{3} \end{vmatrix} = -24. \quad \text{Συνεπώς, } (-1)^1\Delta_2 > 0 \text{ και το σημείο αυτό είναι τοπικό ελάχιστο.}$$

Το σημείο $(x,y,\lambda) = (-1,-1,\frac{2}{3})$. Έχουμε

$$\Delta_2 = \begin{vmatrix} 0 & -3 & -3 \\ -3 & \frac{2}{3} & -\frac{2}{3} \\ -3 & -\frac{2}{3} & \frac{2}{3} \end{vmatrix} = -24. \quad \text{Συνεπώς, } (-1)^1\Delta_2 > 0 \text{ και το σημείο αυτό είναι τοπικό ελάχιστο.}$$

Το σημείο $(x, y, \lambda) = (\sqrt{3}, -\sqrt{3}, 2)$. Έχουμε

$\Delta_2 = \begin{vmatrix} 0 & \sqrt{3} & -\sqrt{3} \\ \sqrt{3} & -2 & -2 \\ -\sqrt{3} & -2 & -2 \end{vmatrix} = 24.$ Συνεπώς, $(-1)^2\Delta_2 > 0$ και το σημείο αυτό είναι τοπικό ελάχιστο.

Το σημείο $(x, y, \lambda) = (-\sqrt{3}, \sqrt{3}, 2)$. Έχουμε

$\Delta_2 = \begin{vmatrix} 0 & -\sqrt{3} & \sqrt{3} \\ -\sqrt{3} & -2 & -2 \\ \sqrt{3} & -2 & -2 \end{vmatrix} = 24.$ Συνεπώς, $(-1)^2\Delta_2 > 0$ και το σημείο αυτό είναι τοπικό ελάχιστο.