Θεωρία Βελτιστοποίησης

Βασίλειος Μαχαιράς
Πολιτικός Μηχανικός Ph.D.

Κλασικές Τεχνικές Βελτιστοποίησης

Πανεπιστήμιο Θεσσαλίας
Σχολή Θετικών Επιστημών
Τμήμα Πληροφορικής

Διάλεξη 2η/2017

Μαθηματική Βελτιστοποίηση

- Η «Μαθηματική Βελτιστοποίηση» (Mathematical Optimization) ή απλά «βελτιστοποίηση» είναι μια ιδιαίτερη κατηγορία προβλημάτων.
- Βελτιστοποίηση είναι η διαδικασία εύρεσης του «βέλτιστου» αποτελέσματος για δεδομένες συνθήκες.
- Υπάρχουν προβλήματα βελτιστοποίησης παντού!

Ιστορία

- **Θεωρία (convex analysis): 1900–1970**

Algorithms

- 1947: simplex algorithm for linear programming (Dantzig)
- 1960s: early interior-point methods (Fiacco & McCormick, Dikin, ...)
- 1970s: ellipsoid method and other subgradient methods
- 1980s: polynomial-time interior-point methods for linear programming (Karmarkar 1984)
- late 1980s–now: polynomial-time interior-point methods for nonlinear convex optimization (Nesterov & Nemirovski 1994)

Applications

- before 1990: mostly in operations research; few in engineering
- since 1990: many new applications in engineering (control, signal processing, communications, circuit design, . . .); new problem classes (semidefinite and second-order cone programming, robust optimization)

Βελτιστοποίηση προβλημάτων

- Κάθε πρόβλημα βελτιστοποίησης έχει κάποια χαρακτηριστικά.
- Το πρόβλημα κατηγοριοποιείται ανάλογα με τα χαρακτηριστικά του.
- Υπάρχουν τεχνικές επίλυσης των προβλημάτων βελτιστοποίησης… αλλά ΟΧΙ όλων!

Βελτιστοποίηση προβλημάτων

- Η μαθηματική έκφραση ενός προβλήματος βελτιστοποίησης είναι:

$$\min_{x} f(x)$$

$$\acute{o}\pi o \upsilon \quad x \in R^n$$

$$\upsilon\pi\acute{o}\ \tau o\upsilon\varsigma\ \pi\varepsilon\rho\iota o\rho\iota\sigma\mu o\acute{\upsilon}\varsigma:$$

$$\mathrm{g}_{\mathrm{i}}(x) \leq 0 \ , \quad i = 1,\dots,m$$

$$\mathrm{h}_{\mathrm{j}}(x) = 0 \ , \quad \mathrm{j} = 1,\dots,n$$

Βελτιστοποίηση προβλημάτων

- Η μαθηματική έκφραση ενός προβλήματος βελτιστοποίησης είναι:

$$\min_x f(x)$$

$$ή \ldots \underset{x}{\text{minimize}}\, f_0(x)$$

$$\textit{όπου} \quad x \in R^n$$

$$ή \ldots f_0(x) \rightarrow \min$$

$$\textit{υπό τους περιορισμούς}:$$

$$\mathrm{g_i}(x) \le 0 \;, \quad i = 1, \ldots, m$$

$$\mathrm{h_j}(x) = 0 \;, \quad \mathrm{j} = 1, \ldots, n$$

Βελτιστοποίηση προβλημάτων

$\min_{x} f(x)$

όπου $x \in R^n$

υπό τους περιορισμούς:

$\mathrm{g_i}(x) \le 0 \ , \qquad i = 1,\ldots,m$

$\mathrm{h_j}(x) = 0 \ , \qquad \mathrm{j} = 1,\ldots,l$

$x = (x_1,\ldots,x_n)$ μεταβλητές βελτιστοποίησης

$f : R^n \to R$ *αντικειμενική συνάρτηση*

(ή συνάρτηση κόστους)

$\mathrm{g_i} : R^n \to R \quad i = 1,\ldots,m :$ συναρτήσεις περιορισμών

Επιχειρησιακή Έρευνα

- Η επιχειρησιακή Έρευνα (Operational ή Operations Research) είναι ο κλάδος της επιστήμης με αντικείμενο την εύρεση βέλτιστων λύσεων για τη λήψη αποφάσεων σε προβλήματα, στα οποία απαιτείται αποτελεσματική κατανομή των διαθέσιμων – αλλά πάντα περιορισμένων – πόρων.
- Το μάθημα της «Θεωρίας βελτιστοποίησης» αφορά ειδικότερα στις μεθόδους βελτιστοποίησης, όμως η βιβλιογραφία του κλάδου της Επιχειρησιακής Έρευνας προσφέρει αρκετά παραδείγματα.

Κλασικές τεχνικές βελτιστοποίησης

Κλασικές τεχνικές βελτιστοποίησης

- Οι κλασικές μέθοδοι βελτιστοποίησης είναι χρήσιμες στην επίλυση συναρτήσεων, οι οποίες είναι *συνεχείς* και *παραγωγίσιμες*.
- Οι μέθοδοι είναι αναλυτικοί και εφαρμόζουν τεχνικές του διαφορικού λογισμού για να βρουν τα βέλτιστα σημεία.
- Συχνά τα πρακτικά προβλήματα έχουν συναρτήσεις που δεν είναι συνεχείς ή παραγωγίσιμες, οπότε η αναλυτική επίλυση έχει περιορισμένη πρακτική εφαρμογή. Οι κλασικές όμως μέθοδοι αποτελούν τη βάση για την ανάπτυξη των αριθμητικών τεχνικών που θα αναπτυχθούν σε επόμενες διαλέξεις.

Βελτιστοποίηση συναρτήσεων μιας μεταβλητής

Βελτιστοποίηση μιας μεταβλητής

A function of one variable $f(x)$ is said to have a *relative* or *local minimum* at $x = x^*$ if $f(x^*) \le f(x^* + h)$ for all sufficiently small positive and negative values of h. Similarly, a point x^* is called a *relative* or *local maximum* if $f(x^*) \ge f(x^* + h)$ for all values of h sufficiently close to zero. A function $f(x)$ is said to have a *global* or *absolute minimum* at x^* if $f(x^*) \le f(x)$ for all x, and not just for all x close to x^*, in the domain over which $f(x)$ is defined. Similarly, a point x^* will be a global maximum of $f(x)$ if $f(x^*) \ge f(x)$ for all x in the domain. Figure 2.1 shows the difference between the local and global optimum points.

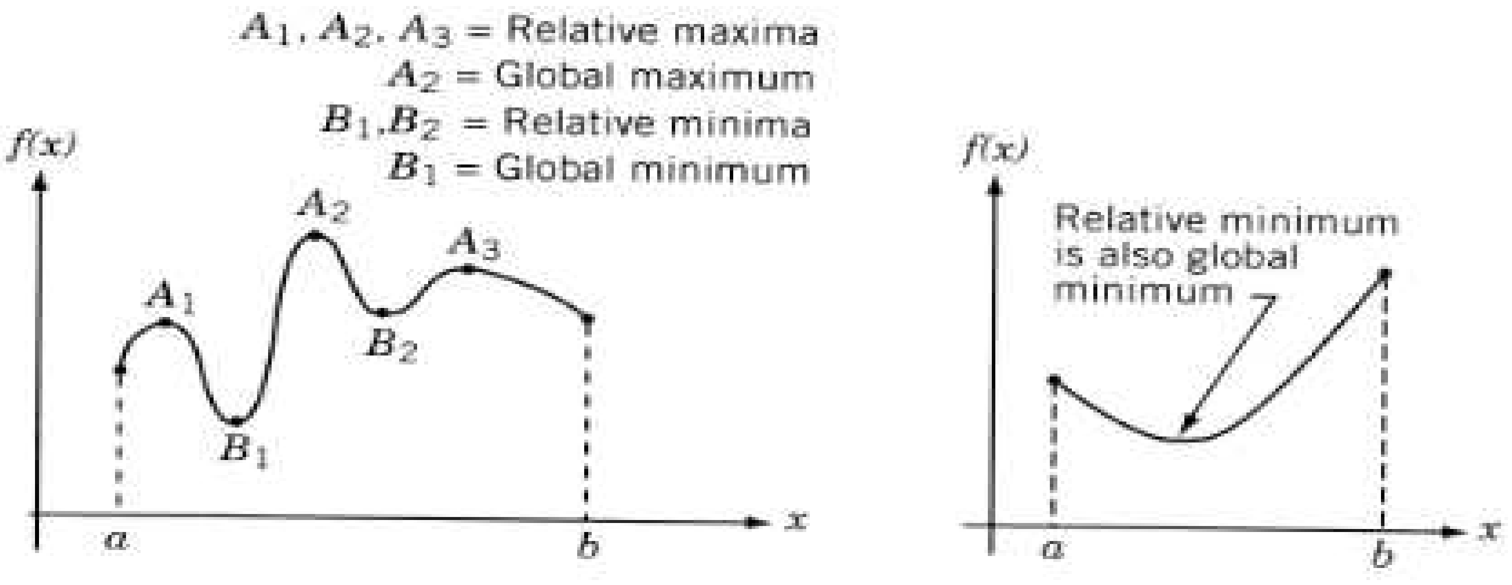


Figure 2.1 Relative and global minima.

Αναγκαία συνθήκη ύπαρξης ακρότατου Θεώρημα 1ο

Theorem 2.1 Necessary Condition If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $df(x)/dx = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$.

- Αν μια συνάρτηση f(x) έχει ακρότατο στη θέση x^* (ελάχιστο ή μέγιστο) μεταξύ δύο τιμών της μεταβλητής x **και υπάρχει η 1η παράγωγος της συνάρτησης στο σημείο x^*** τότε $f`(x^*)=0$

- Το 1ο θεώρημα ΔΕΝ ισχύει όταν η παράγωγος της συνάρτησης δεν υπάρχει.
- Το 1ο θεώρημα δεν αναφέρει τι συμβαίνει όταν τα ακρότατα υπάρχουν στα συνοριακά σημεία της μεταβλητής.

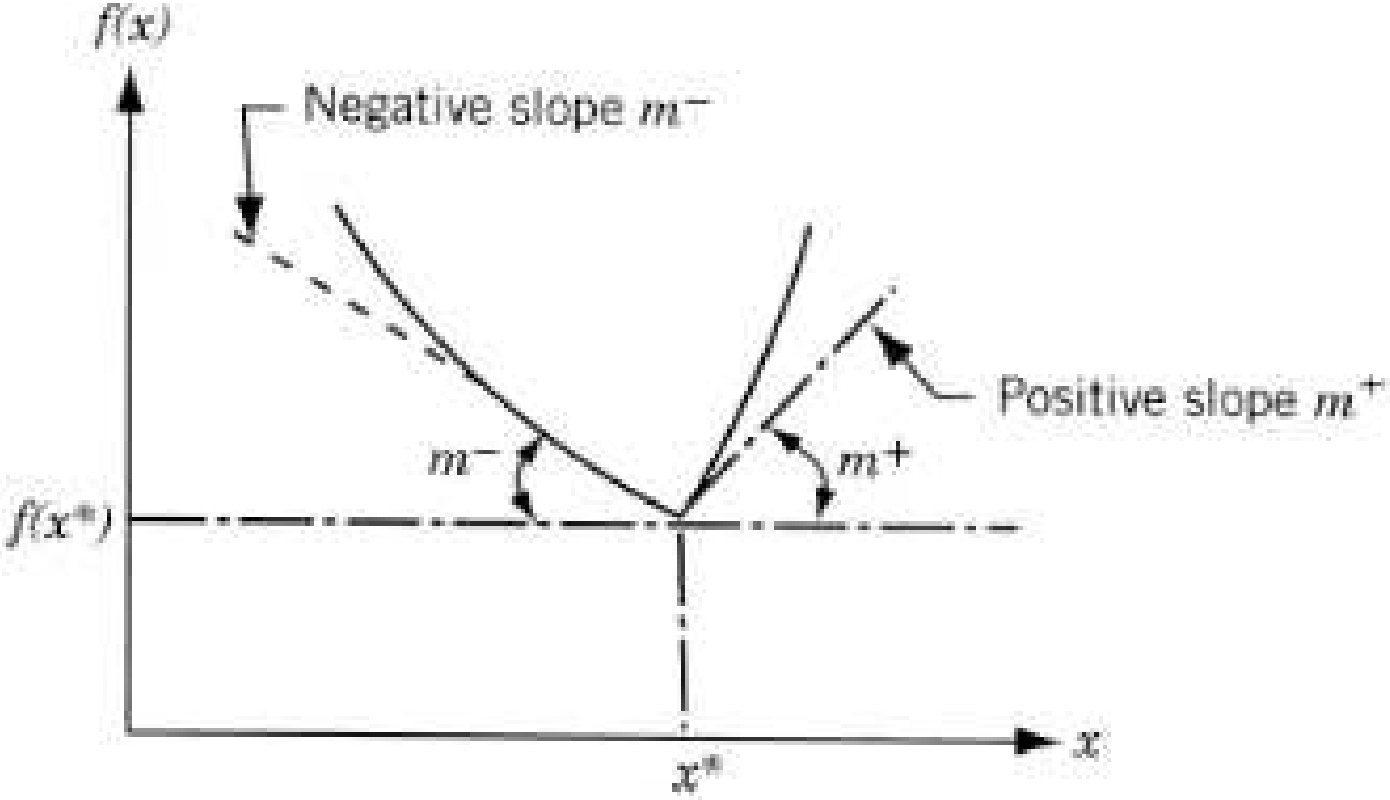
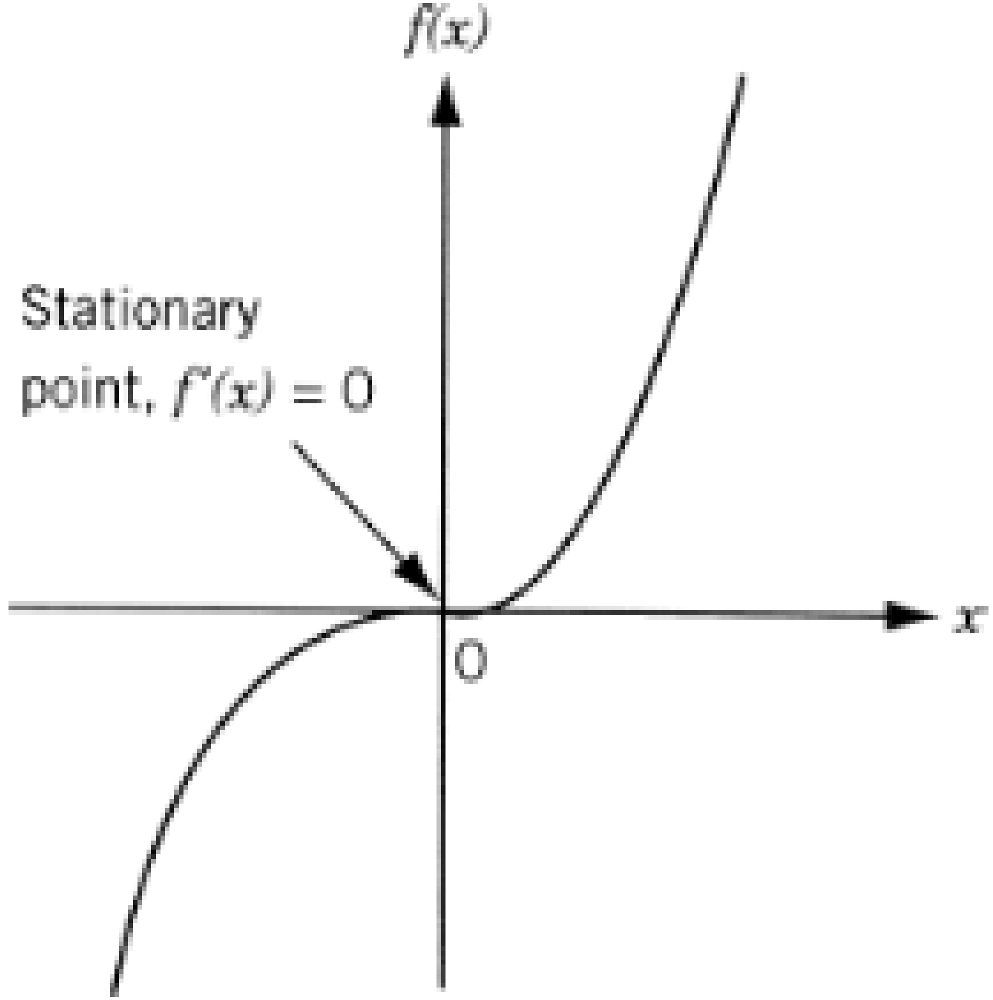


Figure 2.2 Derivative undefined at x^*.

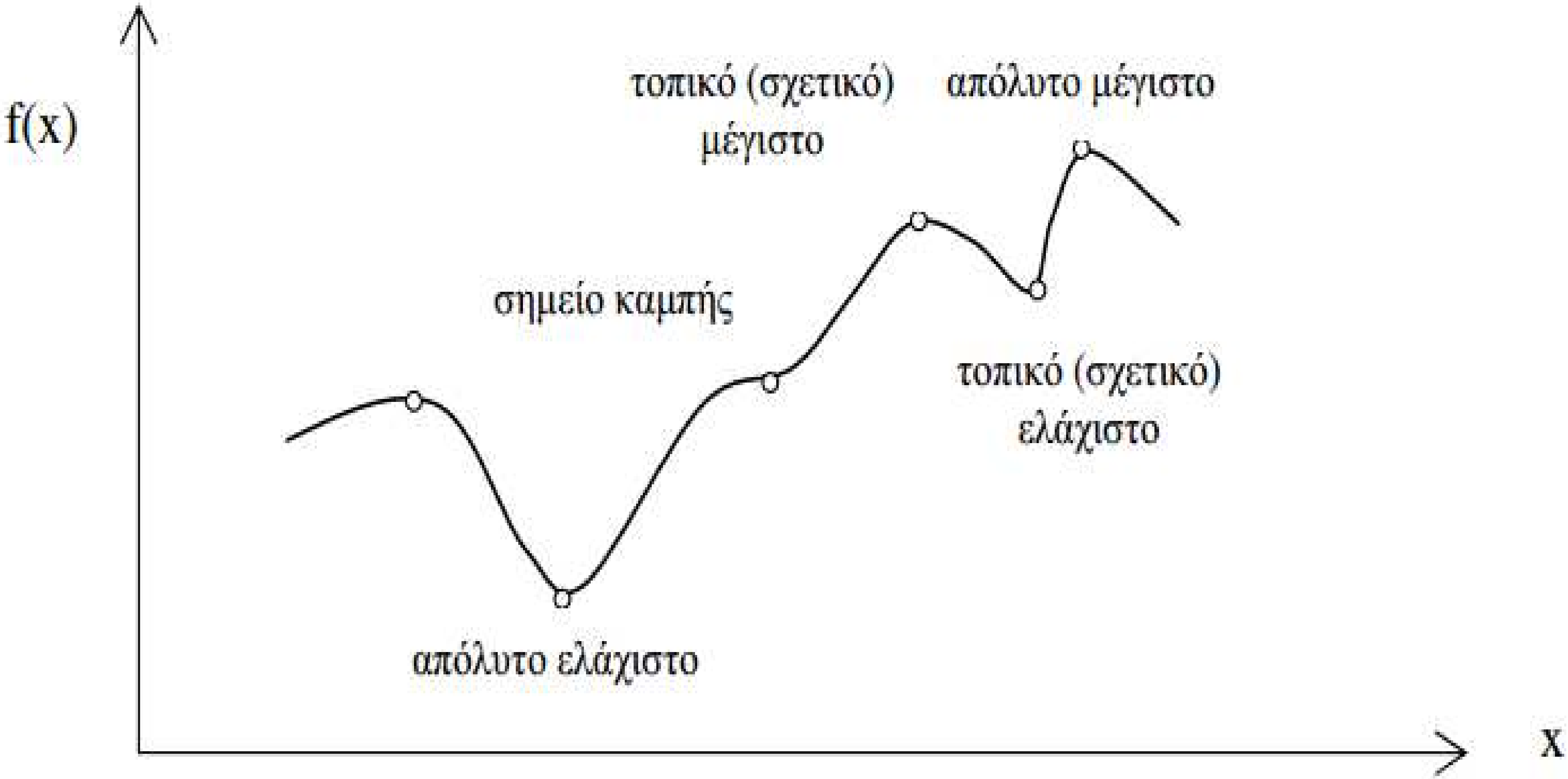
- Το 1ο θεώρημα ΔΕΝ εξασφαλίζει ότι θα υπάρχει ακρότατο σε κάθε σημείο που μηδενίζεται η παράγωγος.
- Εκεί που μηδενίζεται

είτε είναι ακρότατο είτε

είναι «στάσιμο σημείο»



Ακρότατα συνάρτησης μιας μεταβλητής



Θεώρημα 2ο – ικανή συνθήκη ακρότατου

- Για να προσδιοριστεί εάν το ακρότατο (εκεί που μηδενίζεται η 1η παράγωγος) είναι μέγιστο ή ελάχιστο χρειάζεται να ελεγχθεί η τιμή της 2ης ή ανώτερης παραγώγου στα κρίσιμα σημεία.
- Υπολογίζουμε τη n-οστή παράγωγο της συνάρτησης έως ότου βρούμε τιμή διάφορη του μηδενός. Δηλαδή:
- $f`(x^*)=f``(x^*)=\ldots=f^{(n-1)}(x^*)=0, \quad$ ενώ $f^{(n)}(x^*)\neq 0$
- Αν το n είναι περιττός αριθμός τότε στο x^* δεν είναι ούτε μέγιστο ούτε ελάχιστο.
- Αν το n είναι άρτιος αριθμός τότε:
- στο x^* είναι τοπικό μέγιστο αν $f^{(n)}(x^*)<0$,
- στο x^* είναι τοπικό ελάχιστο αν $f^{(n)}(x^*)>0$.

Example 2.1 Determine the maximum and minimum values of the function

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5$$

SOLUTION Since $f'(x) = 60(x^4 - 3x^3 + 2x^2) = 60x^2(x-1)(x-2)$, $f'(x) = 0$ at $x = 0$, $x = 1$, and $x = 2$. The second derivative is

$$f''(x) = 60(4x^3 - 9x^2 + 4x)$$

At $x = 1$, $f''(x) = -60$ and hence $x = 1$ is a relative maximum. Therefore,

$$f_{\text{max}} = f(x = 1) = 12$$

At $x = 2$, $f''(x) = 240$ and hence $x = 2$ is a relative minimum. Therefore,

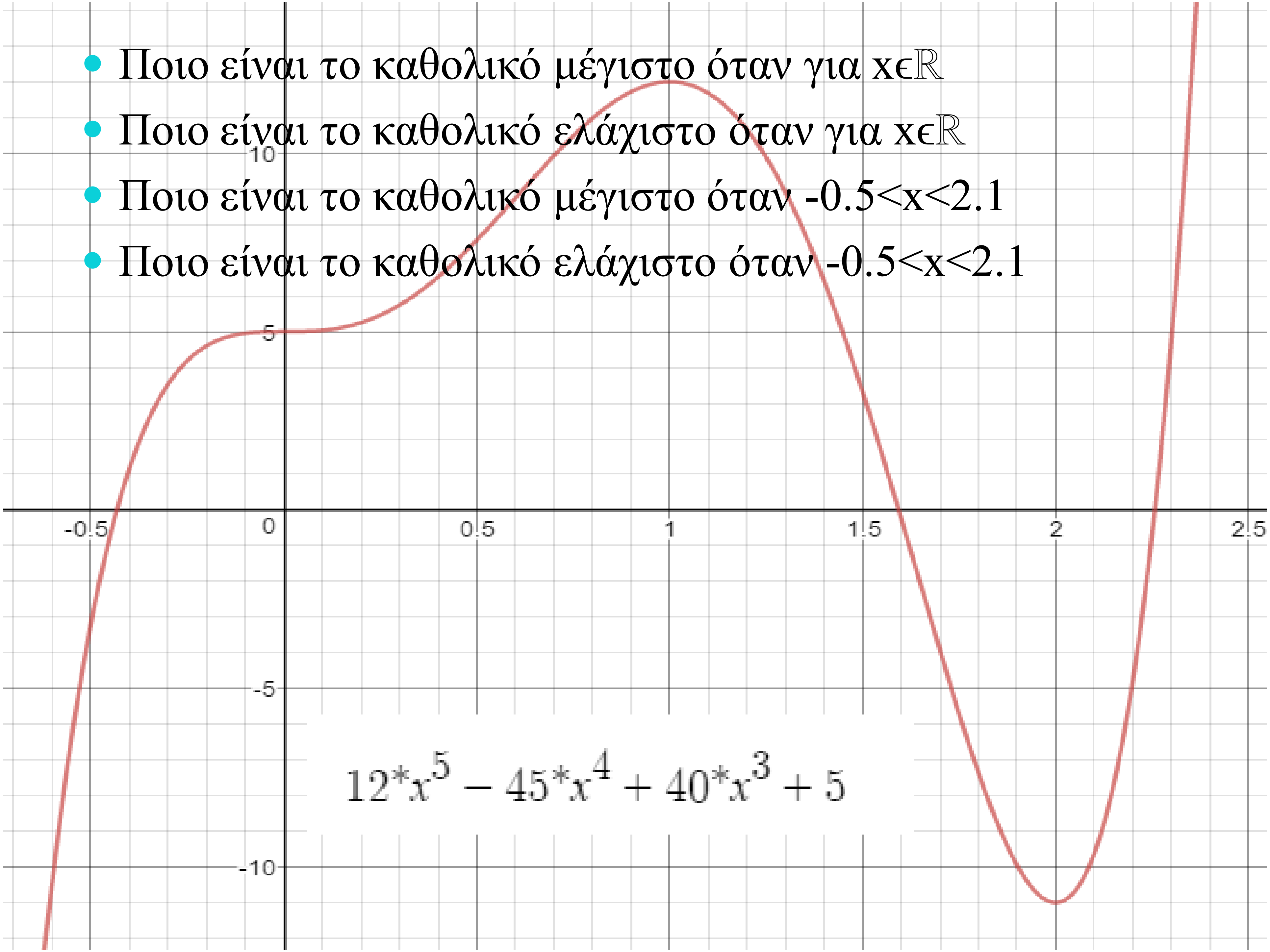
$$f_{\text{min}} = f(x = 2) = -11$$

At $x = 0$, $f''(x) = 0$ and hence we must investigate the next derivative:

$$f'''(x) = 60(12x^2 - 18x + 4) = 240 \quad \text{at} \quad x = 0$$

Since $f'''(x) \neq 0$ at $x = 0$, $x = 0$ is neither a maximum nor a minimum, and it is an inflection point.

- Ποιο είναι το καθολικό μέγιστο όταν για x$\epsilon\mathbb{R}$
- Ποιο είναι το καθολικό ελάχιστο όταν για x$\epsilon\mathbb{R}$
- Ποιο είναι το καθολικό μέγιστο όταν -0.5<x<2.1
- Ποιο είναι το καθολικό ελάχιστο όταν -0.5<x<2.1



$$12{*}x^5 - 45{*}x^4 + 40{*}x^3 + 5$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, χωρίς περιορισμούς

Συνάρτηση πολλών μεταβλητών

- Έστω f(**X**) συνάρτηση πολλών μεταβλητών με $\mathbf{X} \in \mathbb{R}^n$ (είναι διάνυσμα $[x_1, x_2, \ldots, x_n]$)
- **Αναγκαία** συνθήκη ύπαρξης ακρότατου:
- Αν η f(**X**) έχει ακρότατο (ελάχιστο ή μέγιστο) στο σημείο $\mathbf{X}=\mathbf{X}^*$ και οι πρώτες μερικές παράγωγοι της f(**X**) υπάρχουν στο $\mathbf{X}^*$ τότε όλες είναι ίσες με μηδέν:

$$\frac{\partial f}{\partial x_1}(\mathbf{X}^*) = \frac{\partial f}{\partial x_2}(\mathbf{X}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{X}^*) = 0$$

Συνάρτηση πολλών μεταβλητών

- Έστω f(**X**) συνάρτηση πολλών μεταβλητών με $\mathbf{X} \in \mathbb{R}^n$ (είναι διάνυσμα $[x_1, x_2, \ldots, x_n]$)
- **Ικανή** συνθήκη για τοπικό ακρότατο:
- Το σημείο $\mathbf{X}^*$ είναι τοπικό ελάχιστο όταν ο πίνακας των 2ων μερικών παραγώγων (Εσσιανός πίνακας – Hessian) υπολογισμένος στο σημείο $\mathbf{X}^*$ είναι *θετικά ορισμένος*.
- Το σημείο $\mathbf{X}^*$ είναι τοπικό μέγιστο όταν ο πίνακας των 2ων μερικών παραγώγων (Εσσιανός πίνακας – Hessian) υπολογισμένος στο σημείο $\mathbf{X}^*$ είναι *αρνητικά ορισμένος*.

Εσσιανός πίνακας – Hessian

- Εσσιανός πίνακας – Hessian είναι ο συμμετρικός πίνακας των $2^{\omega\nu}$ μερικών παραγώγων:

$$\mathbf{J}|_{\mathbf{X}=\mathbf{X}^*} = \left[\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{\mathbf{X}=\mathbf{X}^*} \right]$$

Έστω $\underline{u}$ το διάνυσμα των μεταβλητών απόφασης $\quad \underline{u} = [u_1\ u_2 \ldots u_m]^T \in R^m$

Ζητούνται οι τιμές που $\quad \min\limits_{\underline{u}} J = J(\underline{u})$

Διάνυσμα Κλίσης: $\quad \nabla J = \left[\frac{\partial J}{\partial u_1} \cdots \frac{\partial J}{\partial u_m} \right]^T$

Εσσιανός Πίνακας Hessian: Είναι το συμμετρικό μητρώο (m,m)

$$H(J) = \begin{bmatrix} \frac{\partial^2 J}{\partial u_1^2} & \cdots & \frac{\partial^2 J}{\partial u_1 \partial u_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 J}{\partial u_m \partial u_1} & \cdots & \frac{\partial^2 J}{\partial u_m^2} \end{bmatrix}$$

Θετικά/αρνητικά ορισμένος πίνακας

Ορισμός: Ο συμμετρικός πίνακας A(n,n) είναι θετικά ορισμένος $A > 0$,

αν $\forall x \in R^n \neq 0$: $x^T A x > 0$

Ο έλεγχος της συνθήκης $H(J) > 0$ (τοπικό ελάχιστο) γίνεται με τους εξής δύο τρόπους:

(1) όλες οι ιδιοτιμές του H είναι θετικές

Οι ιδιοτιμές υπολογίζονται από την εξίσωση $\det(H - \lambda I) = 0$ και ικανοποιούν τη σχέση $H x = \lambda x$.

Προφανώς επειδή $H > 0$: $x^T H x = \lambda x^T x > 0$ και επειδή $x^T x = \Sigma x^2 > 0 \Rightarrow \lambda > 0$

(2) το πρόσημο όλων των (m το πλήθος) κύριων μερικών οριζουσών του H είναι θετικό

Κύριες μερικές ορίζουσες είναι οι ορίζουσες των τετραγωνικών πινάκων που διαμορφώνονται αρχίζοντας από το στοιχείο $\dfrac{\partial^2 J}{\partial x_1^2} = J_{x_1x_1}$

δηλαδή $\Delta_1 = J_{x_1x_1}$, $\Delta_2 = J_{x_1x_1} J_{x_2x_2} - J^2_{x_1x_2}$, κ.λπ.

Θετικά/αρνητικά ορισμένος πίνακας

Ο έλεγχος της συνθήκης H(J) < 0 (τοπικό μέγιστο) γίνεται με τους εξής δύο τρόπους:

(1) όλες οι ιδιοτιμές του H είναι αρνητικές

(2) το πρόσημο των κύριων μερικών οριζουσών του H είναι $(-1)^j$, $j = 1, 2, \ldots m$, δηλαδή εναλλάσσεται αρχίζοντας από –1.

Παρατηρήσεις

(1) αν ο πίνακας Hessian είναι ημιορισμένος (≥) χρειάζεται να ελεγχθούν οι παράγωγοι ανώτερης τάξης όπως και στην περίπτωση μιας μεταβλητής
(2) στην περίπτωση δύο μεταβλητών, αν ο πίνακας Hessian δεν είναι ούτε θετικά ούτε αρνητικά ορισμένος στο στάσιμο σημείο, το σημείο είναι μικτό (saddle point ή σημείο σάγματος) και η συνάρτηση έχει μέγιστο ως προς τη μια μεταβλητή και ελάχιστο ως προς την άλλη. Μικτά σημεία μπορεί να υπάρχουν και για περισσότερες από δύο μεταβλητές.

Υπενθύμιση μερικών στοιχείων άλγεβρας…

Ιδιοτιμές και Ιδιοδιανύσματα

Αν ο A είναι ένας $n \times n$ πίνακας και το $\mathbf{x}$ είναι ένα διάνυσμα στον $\mathbb{R}^n$, τότε το $A\mathbf{x}$ είναι και αυτό ένα διάνυσμα στον $\mathbb{R}^n$. Συνήθως δεν υπάρχει καμία σχέση ανάμεσα στο διάνυσμα $\mathbf{x}$ και στο διάνυσμα $A\mathbf{x}$. Αν όμως το $\mathbf{x}$ είναι ένα μη μηδενικό διάνυσμα τέτοιο ώστε το $A\mathbf{x}$ να είναι βαθμωτό πολλαπλάσιο του $\mathbf{x}$, τότε υπάρχει μία γεωμετρική σχέση ανάμεσα στα $\mathbf{x}$ και $A\mathbf{x}$. Για παράδειγμα, αν ο A είναι ένας 2×2 πίνακας και το $\mathbf{x}$ είναι ένα μη μηδενικό διάνυσμα στον $\mathbb{R}^2$ τέτοιο ώστε το $A\mathbf{x}$ να είναι βαθμωτό πολλαπλάσιο του $\mathbf{x}$, τότε για κάθε διάνυσμα $\mathbf{y}$ που ανήκει στην ευθεία που διέρχεται από την αρχή των αξόνων την οποία ορίζει το $\mathbf{x}$, το $A\mathbf{y}$ ανήκει στην ευθεία αυτή.

Ορισμός. Αν ο A είναι ένας $n \times n$ πίνακας, τότε ένα μη μηδενικό διάνυσμα $\mathbf{x}$ στον $\mathbb{R}^n$ ονομάζεται ***ιδιοδιάνυσμα*** του A αν το $A\mathbf{x}$ είναι βαθμωτό πολλαπλάσιο του $\mathbf{x}$, δηλαδή αν

$$A\mathbf{x} = \lambda\mathbf{x}$$

για κάποιο βαθμωτό λ. Το βαθμωτό λ ονομάζεται ***ιδιοτιμή*** του A και λέμε ότι το x είναι ένα ιδιοδιάνυσμα του A που ***αντιστοιχεί*** στην λ.

Ιδιοτιμές και Ιδιοδιανύσματα

Παράδειγμα 1 Το διάνυσμα $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ είναι ένα ιδιοδιάνυσμα του

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

το οποίο αντιστοιχεί στην ιδιοτιμή $\lambda = 3$, εφόσον

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}. \quad \triangle$$

Ιδιοτιμές και Ιδιοδιανύσματα

Για να βρούμε τις ιδιοτιμές του $n \times n$ πίνακα A ξαναγράφουμε την

$$A\mathbf{x} = \lambda \mathbf{x}$$

στη μορφή

$$A\mathbf{x} = \lambda I \mathbf{x}$$

ή ισοδύναμα

$$(\lambda I - A)\mathbf{x} = \mathbf{0}. \tag{1}$$

Για να είναι το λ μία ιδιοτιμή του A, πρέπει να υπάρχει μια μη τετριμμένη λύση αυτού του ομογενούς συστήματος. Γνωρίζουμε ότι το ομογενές σύστημα (1) έχει μη τετριμμένες λύσεις αν και μόνο αν

$$\det(\lambda I - A) = 0.$$

Ιδιοτιμές και Ιδιοδιανύσματα

Η εξίσωση $\det(\lambda I - A) = 0$ ονομάζεται ***χαρακτηριστική εξίσωση*** του A. Τα βαθμωτά τα οποία ικανοποιούν αυτή την εξίσωση είναι οι ιδιοτιμές του A. Αν την αναπτύξουμε η $\det(\lambda I - A)$ είναι ένα πολυώνυμο του λ το οποίο ονομάζεται ***χαρακτηριστικό πολυώνυμο*** του A.

Μπορούμε να δείξουμε ότι αν ο A είναι ένας $n \times n$ πίνακας, τότε το χαρακτηριστικό πολυώνυμο του A είναι βαθμού n με πραγματικούς συντελεστές και ο συντελεστής του μεγιστοβάθμιου όρου λ^n είναι 1. Άρα το χαρακτηριστικό πολυώνυμο ενός $n \times n$ πίνακα έχει τη μορφή

$$p(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1\lambda + \alpha_0,$$

Ορίζουσες πίνακα-Determinants

In the case of a 2 × 2 matrix, the specific formula for the determinant:

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Similarly, suppose we have a 3 × 3 matrix *A*, and we want the specific formula for its determinant |*A*|:

$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= aei + bfg + cdh - ceg - bdi - afh.$$

Επιστροφή στη ροή…

Θετικά/αρνητικά ορισμένος πίνακας

Ο έλεγχος της συνθήκης H(J) < 0 (τοπικό μέγιστο) γίνεται με τους εξής δύο τρόπους:

(1) όλες οι ιδιοτιμές του H είναι αρνητικές

(2) το πρόσημο των κύριων μερικών οριζουσών του H είναι $(-1)^j$, j = 1, 2, …m, δηλαδή εναλλάσσεται αρχίζοντας από –1.

Παρατηρήσεις

(1) αν ο πίνακας Hessian είναι ημιορισμένος (≥) χρειάζεται να ελεγχθούν οι παράγωγοι ανώτερης τάξης όπως και στην περίπτωση μιας μεταβλητής
(2) στην περίπτωση δύο μεταβλητών, αν ο πίνακας Hessian δεν είναι ούτε θετικά ούτε αρνητικά ορισμένος στο στάσιμο σημείο, το σημείο είναι μικτό (saddle point ή σημείο σάγματος) και η συνάρτηση έχει μέγιστο ως προς τη μια μεταβλητή και ελάχιστο ως προς την άλλη. Μικτά σημεία μπορεί να υπάρχουν και για περισσότερες από δύο μεταβλητές.

Σημείο Σάγματος – Saddle Point

In the case of a function of two variables, $f(x, y)$, the Hessian matrix may be neither positive nor negative definite at a point (x^*, y^*) at which

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$$

In such a case, the point (x^*, y^*) is called a *saddle point*. The characteristic of a saddle point is that it corresponds to a relative minimum or maximum of $f(x, y)$ with respect to one variable, say, x (the other variable being fixed at $y = y^*$) and a relative maximum or minimum of $f(x, y)$ with respect to the second variable y (the other variable being fixed at x^*).

Σημείο Σάγματος – Saddle Point

As an example, consider the function $f(x, y) = x^2 - y^2$. For this function,

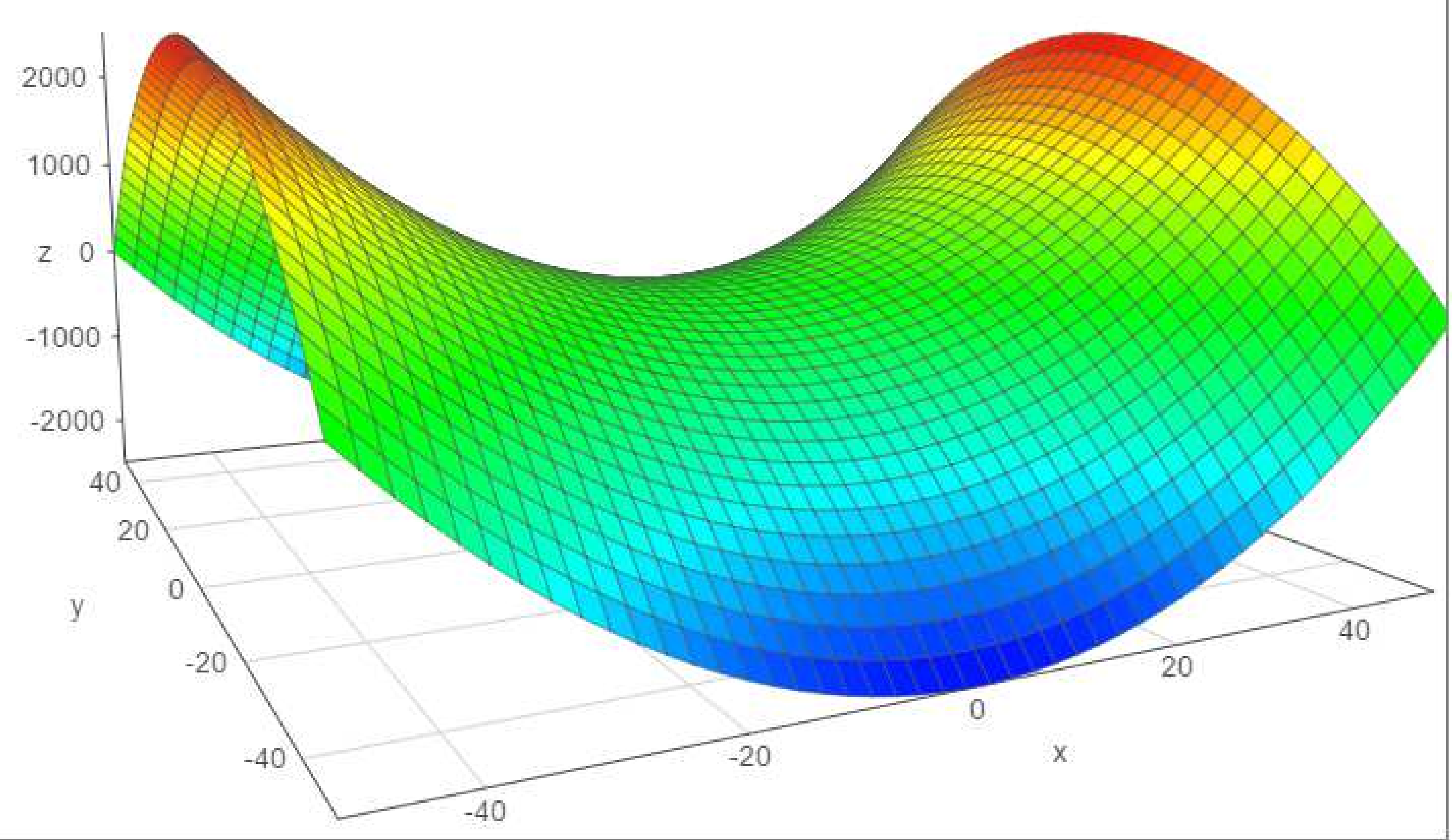
$$\frac{\partial f}{\partial x} = 2x \quad \text{and} \quad \frac{\partial f}{\partial y} = -2y$$

These first derivatives are zero at $x^* = 0$ and $y^* = 0$. The Hessian matrix of f at (x^*, y^*) is given by

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

Since this matrix is neither positive definite nor negative definite, the point $(x^* = 0, y^* = 0)$ is a saddle point. The function is shown graphically in Fig. 2.5. It can be seen that $f(x, y^*) = f(x, 0)$ has a relative minimum and $f(x^*, y) = f(0, y)$ has a relative maximum at the saddle point (x^*, y^*). Saddle points may exist for functions of more than two variables also. The characteristic of the saddle point stated above still holds provided that x and y are interpreted as vectors in multidimensional cases.

Σημείο Σάγματος – Saddle Point



Σημείο Σάγματος – Saddle Point

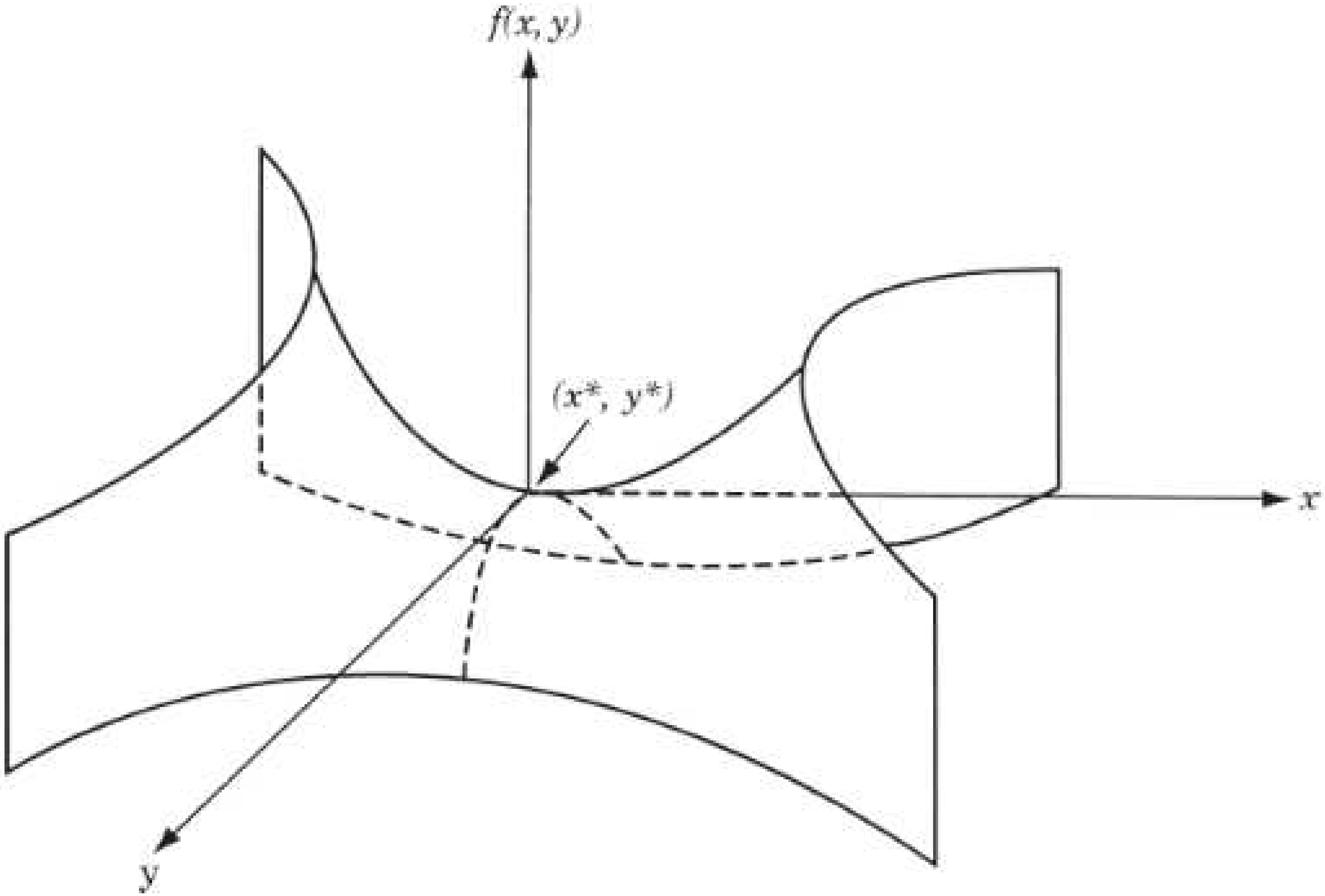


Figure 2.5 Saddle point of the function $f(x, y) = x^2 - y^2$.

Παράδειγμα συνάρτησης με 2 μεταβλητές, χωρίς περιορισμούς

Example 2.5 Find the extreme points of the function

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

SOLUTION The necessary conditions for the existence of an extreme point are

$$\frac{\partial f}{\partial x_1} = 3x_1^2 + 4x_1 = x_1(3x_1 + 4) = 0$$

$$\frac{\partial f}{\partial x_2} = 3x_2^2 + 8x_2 = x_2(3x_2 + 8) = 0$$

These equations are satisfied at the points

$$(0, 0), \quad (0, -\tfrac{8}{3}), \quad (-\tfrac{4}{3}, 0), \quad \text{and} \quad (-\tfrac{4}{3}, -\tfrac{8}{3})$$

Παράδειγμα συνάρτησης με 2 μεταβλητές, χωρίς περιορισμούς

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

To find the nature of these extreme points, we have to use the sufficiency conditions. The second-order partial derivatives of f are given by

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 + 4$$

$$\frac{\partial^2 f}{\partial x_2^2} = 6x_2 + 8$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 0$$

The Hessian matrix of f is given by

$$\mathbf{J} = \begin{bmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{bmatrix}$$

If $J_1 = |6x_1 + 4|$ and $J_2 = \begin{vmatrix} 6x_1 + 4 & 0 \\ 0 & 6x_2 + 8 \end{vmatrix}$, the values of J_1 and J_2 and the nature of the extreme point are as given below:

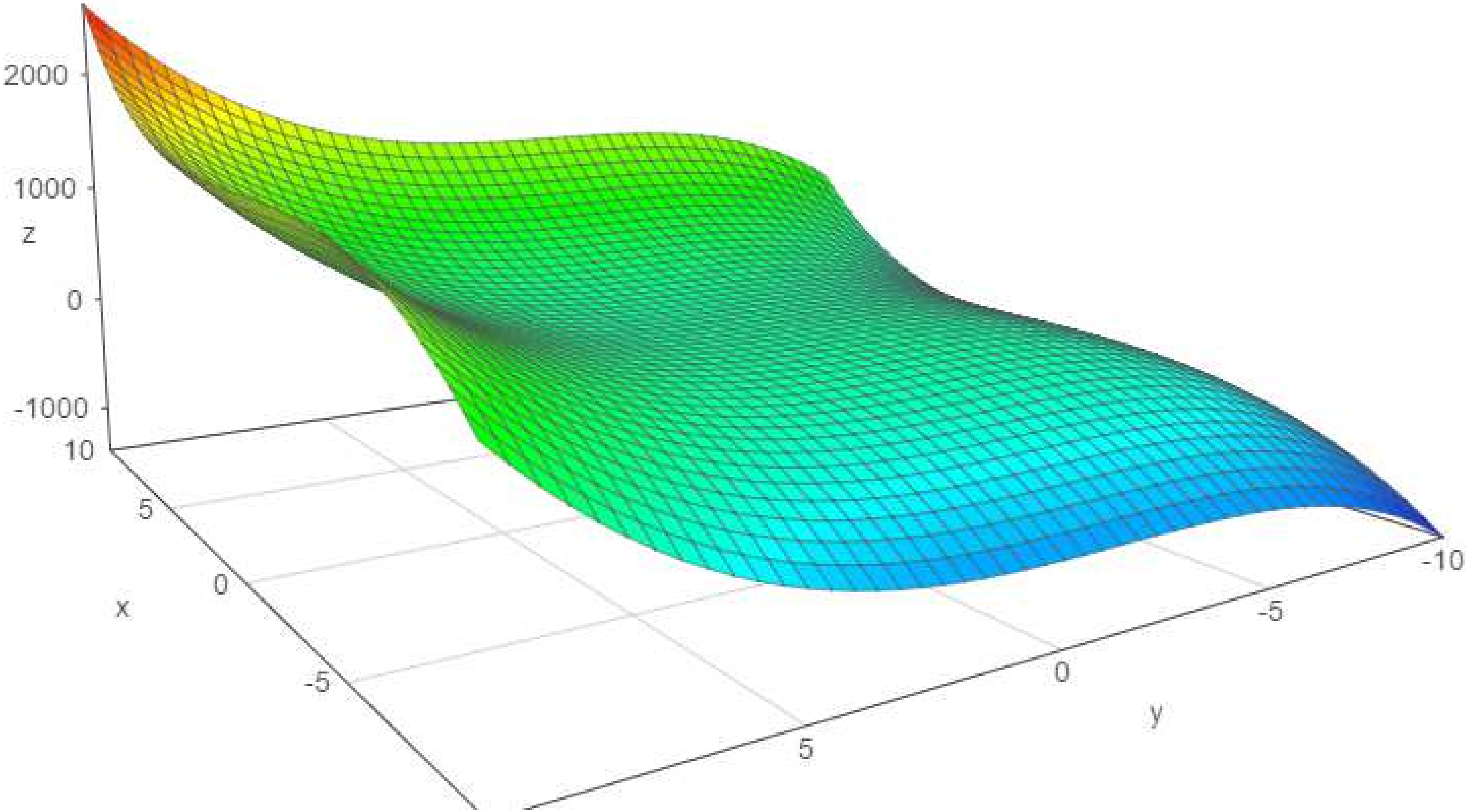
Παράδειγμα συνάρτησης με 2 μεταβλητές, χωρίς περιορισμούς

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$

Point $\mathbf{X}$	Value of J_1	Value of J_2	Nature of $\mathbf{J}$	Nature of $\mathbf{X}$	$f(\mathbf{X})$
$(0, 0)$	+4	+32	Positive definite	Relative minimum	6
$(0, -\frac{8}{3})$	+4	−32	Indefinite	Saddle point	418/27
$(-\frac{4}{3}, 0)$	−4	−32	Indefinite	Saddle point	194/27
$(-\frac{4}{3}, -\frac{8}{3})$	−4	+32	Negative definite	Relative maximum	50/3

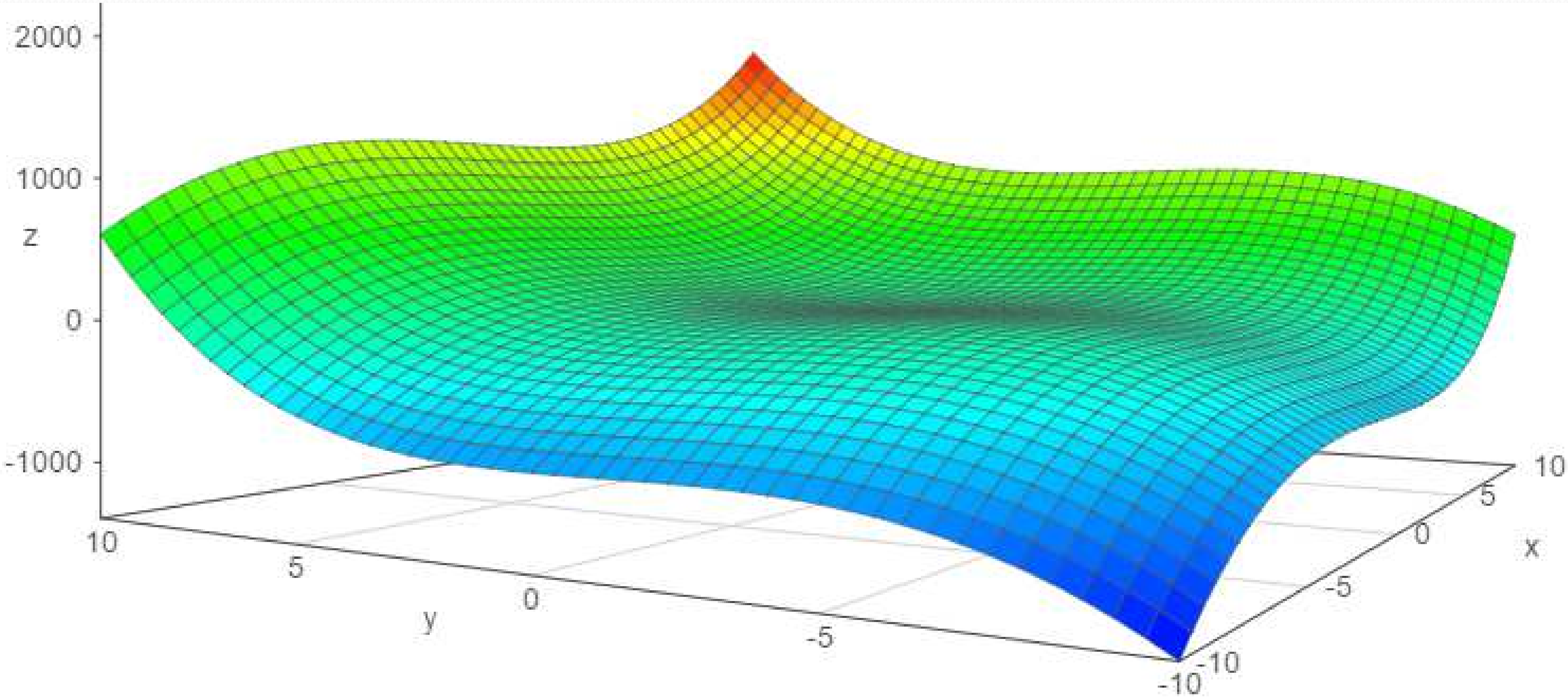
Παράδειγμα συνάρτησης με 2 μεταβλητές, χωρίς περιορισμούς

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$



Παράδειγμα συνάρτησης με 2 μεταβλητές, χωρίς περιορισμούς

$$f(x_1, x_2) = x_1^3 + x_2^3 + 2x_1^2 + 4x_2^2 + 6$$



Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας

$$\text{Minimize } f = f(\mathbf{X})$$

subject to

$$g_j(\mathbf{X}) = 0, \quad j = 1, 2, \ldots, m$$

where

$$\mathbf{X} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix}$$

Here m is less than or equal to n; otherwise (if $m > n$), the problem becomes overdefined and, in general, there will be no solution. There are several methods available for the solution of this problem. The methods of direct substitution, constrained variation, and Lagrange multipliers are discussed in the following sections.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας Επίλυση με απευθείας αντικατάσταση.

Solution by Direct Substitution

For a problem with n variables and m equality constraints, it is theoretically possible to solve simultaneously the m equality constraints and express any set of m variables in terms of the remaining $n - m$ variables. When these expressions are substituted into the original objective function, there results a new objective function involving only $n - m$ variables. The new objective function is not subjected to any constraint, and hence its optimum can be found by using the unconstrained optimization techniques discussed in Section 2.3.

This method of direct substitution, although it appears to be simple in theory, is not convenient from a practical point of view. The reason for this is that the constraint equations will be nonlinear for most of practical problems, and often it becomes impossible to solve them and express any m variables in terms of the remaining $n - m$ variables. However, the method of direct substitution might prove to be very simple and direct for solving simpler problems, as shown by the following example.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με απευθείας αντικατάσταση.

Example 2.6 Find the dimensions of a box of largest volume that can be inscribed in a sphere of unit radius.

SOLUTION Let the origin of the Cartesian coordinate system x_1, x_2, x_3 be at the center of the sphere and the sides of the box be $2x_1$, $2x_2$, and $2x_3$. The volume of the box is given by

$$f(x_1, x_2, x_3) = 8x_1x_2x_3 \tag{E$_1$}$$

Since the corners of the box lie on the surface of the sphere of unit radius, x_1, x_2, and x_3 have to satisfy the constraint

$$x_1^2 + x_2^2 + x_3^2 = 1 \tag{E$_2$}$$

This problem has three design variables and one equality constraint. Hence the equality constraint can be used to eliminate any one of the design variables from the objective function. If we choose to eliminate x_3, Eq. (E$_2$) gives

$$x_3 = (1 - x_1^2 - x_2^2)^{1/2} \tag{E$_3$}$$

Thus the objective function becomes

$$f(x_1, x_2) = 8x_1x_2(1 - x_1^2 - x_2^2)^{1/2} \tag{E$_4$}$$

which can be maximized as an unconstrained function in two variables.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με απευθείας αντικατάσταση.

The necessary conditions for the maximum of f give

$$\frac{\partial f}{\partial x_1} = 8x_2 \left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_1^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right] = 0 \qquad (E_5)$$

$$\frac{\partial f}{\partial x_2} = 8x_1 \left[(1 - x_1^2 - x_2^2)^{1/2} - \frac{x_2^2}{(1 - x_1^2 - x_2^2)^{1/2}} \right] = 0 \qquad (E_6)$$

Equations (E_5) and (E_6) can be simplified to obtain

$$1 - 2x_1^2 - x_2^2 = 0$$

$$1 - x_1^2 - 2x_2^2 = 0$$

from which it follows that $x_1^* = x_2^* = 1/\sqrt{3}$ and hence $x_3^* = 1/\sqrt{3}$. This solution gives the maximum volume of the box as

$$f_{\max} = \frac{8}{3\sqrt{3}}$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με απευθείας αντικατάσταση.

To find whether the solution found corresponds to a maximum or a minimum, we apply the sufficiency conditions to $f(x_1, x_2)$ of Eq. (E_4). The second-order partial derivatives of f at (x_1^*, x_2^*) are given by

$$\frac{\partial^2 f}{\partial x_1^2} = -\frac{32}{\sqrt{3}} \text{ at } (x_1^*, x_2^*)$$

$$\frac{\partial^2 f}{\partial x_2^2} = -\frac{32}{\sqrt{3}} \text{ at } (x_1^*, x_2^*)$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -\frac{16}{\sqrt{3}} \text{ at } (x_1^*, x_2^*)$$

Since

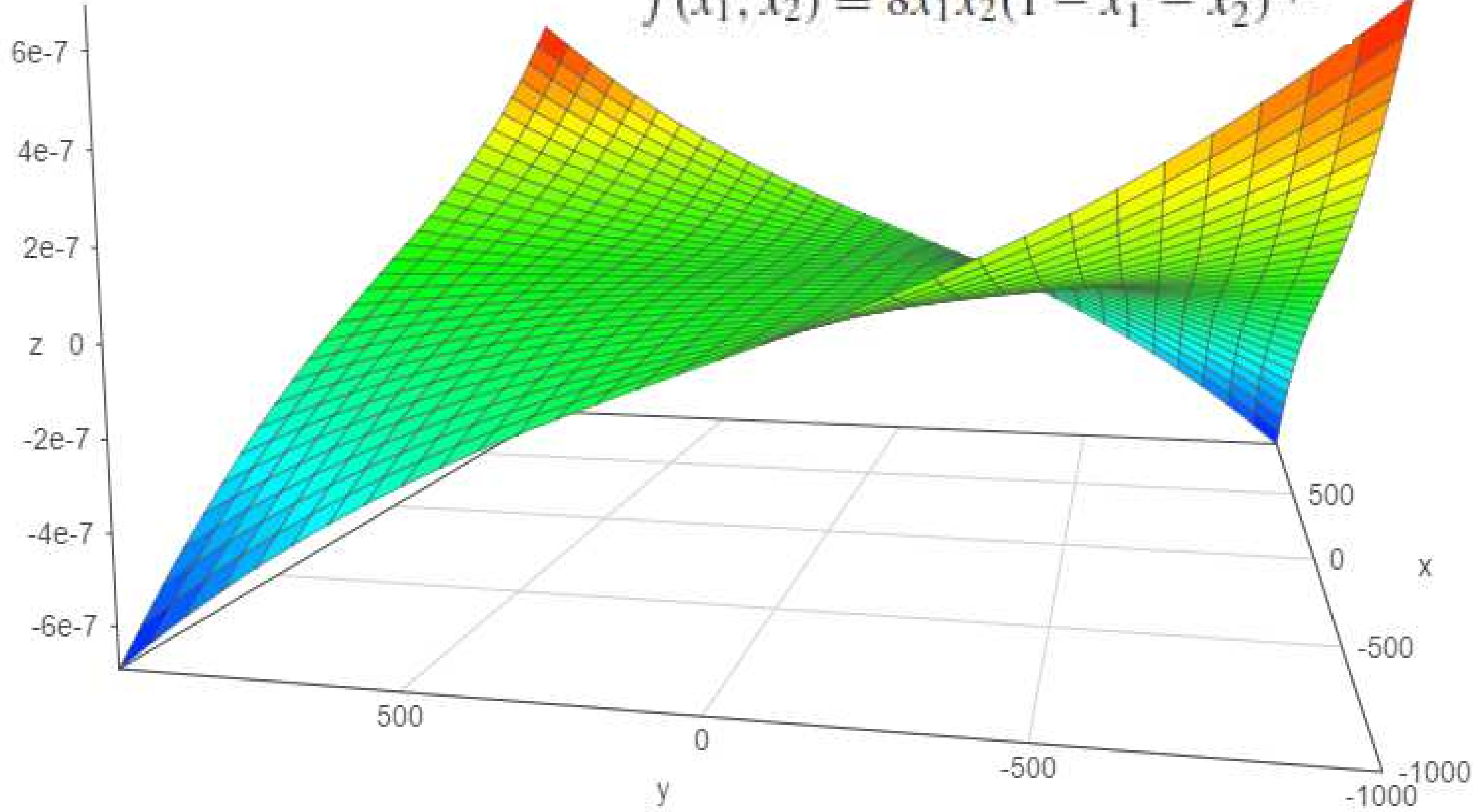
$$\frac{\partial^2 f}{\partial x_1^2} < 0 \quad \text{and} \quad \frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 > 0$$

the Hessian matrix of f is negative definite at (x_1^*, x_2^*). Hence the point (x_1^*, x_2^*) corresponds to the maximum of f.

…Τοπικό μέγιστο!

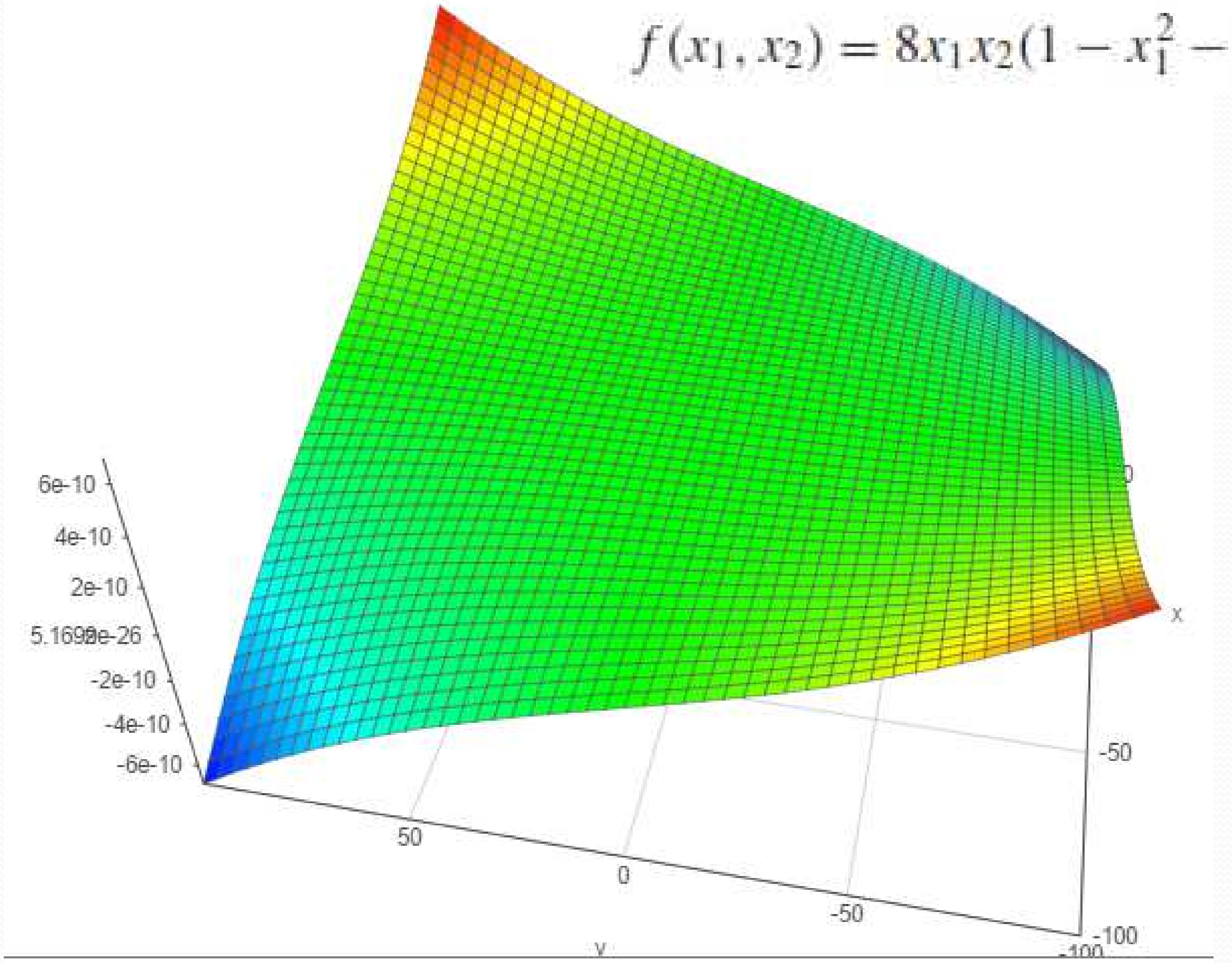
Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με απευθείας αντικατάσταση.

$$f(x_1, x_2) = 8x_1x_2(1 - x_1^2 - x_2^2)^{1/2}$$



Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με απευθείας αντικατάσταση.

$$f(x_1, x_2) = 8x_1x_2(1 - x_1^2 - x_2^2)^{1/2}$$



Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο της περιορισμένης μεταβολής (constrained variation).

Solution by the Method of Constrained Variation

The basic idea used in the method of constrained variation is to find a closed-form expression for the first-order differential of f (df) at all points at which the constraints $g_j(\mathbf{X}) = 0$, $j = 1, 2, \ldots, m$, are satisfied. The desired optimum points are then obtained by setting the differential df equal to zero. Before presenting the general method,

Για την περίπτωση 2 μεταβλητών η απαραίτητη συνθήκη ύπαρξης ακρότατου είναι:

$$\left(\frac{\partial f}{\partial x_1} \frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2} \frac{\partial g}{\partial x_1} \right) \Bigg|_{(x_1^*, x_2^*)} = 0 \tag{2.25}$$

Equation (2.25) represents a necessary condition in order to have (x_1^*, x_2^*) as an extreme point (minimum or maximum).

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο της περιορισμένης μεταβολής (constrained variation).

Example 2.7 A beam of uniform rectangular cross section is to be cut from a log having a circular cross section of diameter $2a$. The beam has to be used as a cantilever beam (the length is fixed) to carry a concentrated load at the free end. Find the dimensions of the beam that correspond to the maximum tensile (bending) stress carrying capacity.

SOLUTION From elementary strength of materials, we know that the tensile stress induced in a rectangular beam (σ) at any fiber located a distance y from the neutral axis is given by

$$\frac{\sigma}{y} = \frac{M}{I}$$

where M is the bending moment acting and I is the moment of inertia of the cross section about the x axis. If the width and depth of the rectangular beam shown in Fig. 2.7 are $2x$ and $2y$, respectively, the maximum tensile stress induced is given by

$$\sigma_{\max} = \frac{M}{I}y = \frac{My}{\frac{1}{12}(2x)(2y)^3} = \frac{3}{4}\frac{M}{xy^2}$$

Thus for any specified bending moment, the beam is said to have maximum tensile stress carrying capacity if the maximum induced stress ($\sigma_{\max}$) is a minimum. Hence we need to minimize k/xy^2 or maximize Kxy^2, where $k = 3M/4$ and $K = 1/k$, subject to the constraint

$$x^2 + y^2 = a^2$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο της περιορισμένης μεταβολής (constrained variation).

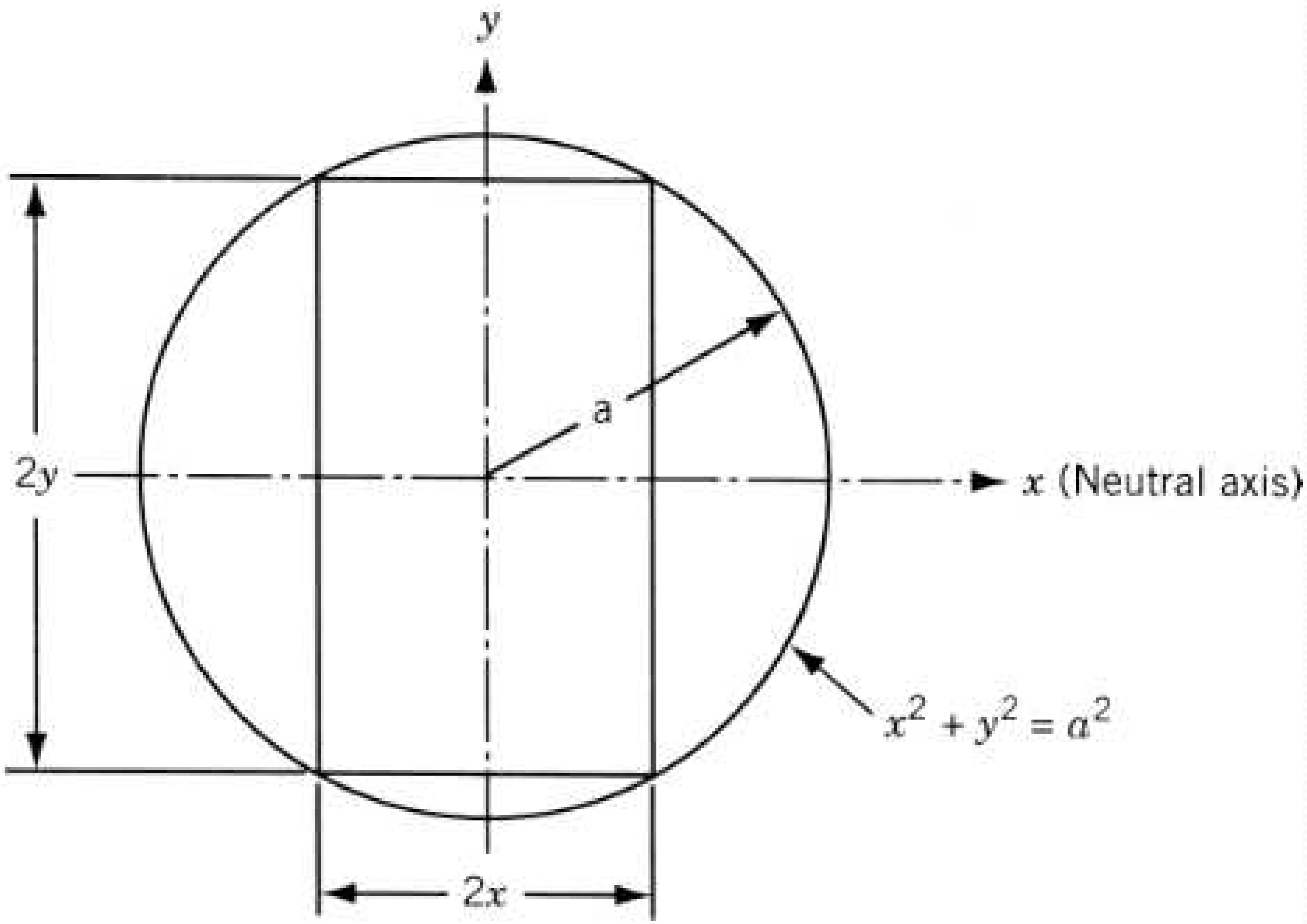


Figure 2.7 Cross section of the log.

This problem has two variables and one constraint; hence Eq. (2.25) can be applied for finding the optimum solution. Since

$$f = kx^{-1}y^{-2} \tag{E$_1$}$$

$$g = x^2 + y^2 - a^2 \tag{E$_2$}$$

we have

$$\frac{\partial f}{\partial x} = -kx^{-2}y^{-2}$$

$$\frac{\partial f}{\partial y} = -2kx^{-1}y^{-3}$$

$$\frac{\partial g}{\partial x} = 2x$$

$$\left.\left(\frac{\partial f}{\partial x_1}\frac{\partial g}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial g}{\partial x_1}\right)\right|_{(x_1^*, x_2^*)} = 0 \qquad (2.25) \qquad \frac{\partial g}{\partial y} = 2y$$

Equation (2.25) gives

$$-kx^{-2}y^{-2}(2y) + 2kx^{-1}y^{-3}(2x) = 0 \quad \text{at} \quad (x^*, y^*)$$

that is,

$$y^* = \sqrt{2}x^* \tag{E$_3$}$$

Thus the beam of maximum tensile stress carrying capacity has a depth of $\sqrt{2}$ times its breadth. The optimum values of x and y can be obtained from Eqs. (E$_3$) and (E$_2$) as

$$x^* = \frac{a}{\sqrt{3}} \quad \text{and} \quad y^* = \sqrt{2}\frac{a}{\sqrt{3}}$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Problem with Two Variables and One Constraint. Consider the problem

$$\text{Minimize } f(x_1, x_2) \tag{2.31}$$

subject to

$$g(x_1, x_2) = 0$$

For this problem, the necessary condition for the existence of an extreme point at $\mathbf{X} = \mathbf{X}^*$ was found in Section 2.4.2 to be

$$\left(\frac{\partial f}{\partial x_1} - \frac{\partial f/\partial x_2}{\partial g/\partial x_2}\frac{\partial g}{\partial x_1}\right)\bigg|_{(x_1^*, x_2^*)} = 0 \tag{2.32}$$

By defining a quantity λ, called the *Lagrange multiplier*, as

$$\lambda = -\left(\frac{\partial f/\partial x_2}{\partial g/\partial x_2}\right)\bigg|_{(x_1^*, x_2^*)} \tag{2.33}$$

Πρέπει να είναι μη μηδενικό

Equation (2.32) can be expressed as

$$\left(\frac{\partial f}{\partial x_1} + \lambda\frac{\partial g}{\partial x_1}\right)\bigg|_{(x_1^*, x_2^*)} = 0 \tag{2.34}$$

and Eq. (2.33) can be written as

$$\left(\frac{\partial f}{\partial x_2} + \lambda\frac{\partial g}{\partial x_2}\right)\bigg|_{(x_1^*, x_2^*)} = 0 \tag{2.35}$$

In addition, the constraint equation has to be satisfied at the extreme point, that is,

$$g(x_1, x_2)|_{(x_1^*, x_2^*)} = 0 \tag{2.36}$$

Thus Eqs. (2.34) to (2.36) represent the necessary conditions for the point (x_1^*, x_2^*) to be an extreme point.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Example 2.9 Find the solution of Example 2.7 using the Lagrange multiplier method:

$$\text{Minimize } f(x, y) = kx^{-1}y^{-2}$$

subject to

$$g(x, y) = x^2 + y^2 - a^2 = 0$$

SOLUTION The Lagrange function is

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y) = kx^{-1}y^{-2} + \lambda(x^2 + y^2 - a^2)$$

The necessary conditions for the minimum of $f(x, y)$ [Eqs. (2.38)] give

$$\frac{\partial L}{\partial x} = -kx^{-2}y^{-2} + 2x\lambda = 0 \qquad (\mathrm{E}_1)$$

$$\frac{\partial L}{\partial y} = -2kx^{-1}y^{-3} + 2y\lambda = 0 \qquad (\mathrm{E}_2)$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - a^2 = 0 \qquad (\mathrm{E}_3)$$

Equations (E_1) and (E_2) yield

$$2\lambda = \frac{k}{x^3y^2} = \frac{2k}{xy^4}$$

from which the relation $x^* = (1/\sqrt{2})y^*$ can be obtained. This relation, along with Eq. (E_3), gives the optimum solution as

$$x^* = \frac{a}{\sqrt{3}} \quad \text{and} \quad y^* = \sqrt{2}\frac{a}{\sqrt{3}}$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Necessary Conditions for a General Problem. The equations derived above can be extended to the case of a general problem with n variables and m equality constraints:

$$\text{Minimize } f(\mathbf{X}) \tag{2.39}$$

subject to

$$g_j(\mathbf{X}) = 0, \quad j = 1, 2, \ldots, m$$

The Lagrange function, L, in this case is defined by introducing one Lagrange multiplier λ_j for each constraint $g_j(\mathbf{X})$ as

$$L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$$
$$= f(\mathbf{X}) + \lambda_1 g_1(\mathbf{X}) + \lambda_2 g_2(\mathbf{X}) + \cdots + \lambda_m g_m(\mathbf{X}) \tag{2.40}$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

By treating L as a function of the $n+m$ unknowns, $x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m$, the necessary conditions for the extremum of L, which also correspond to the solution of the original problem stated in Eq. (2.39), are given by

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \lambda_j \frac{\partial g_j}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n \tag{2.41}$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\mathbf{X}) = 0, \qquad j = 1, 2, \ldots, m \tag{2.42}$$

Equations (2.41) and (2.42) represent $n+m$ equations in terms of the $n+m$ unknowns, x_i and λ_j. The solution of Eqs. (2.41) and (2.42) gives

$$\mathbf{X}^* = \begin{Bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} \quad \text{and} \quad \lambda^* = \begin{Bmatrix} \lambda_1^* \\ \lambda_2^* \\ \vdots \\ \lambda_m^* \end{Bmatrix}$$

The vector $\mathbf{X}^*$ corresponds to the relative constrained minimum of $f(\mathbf{X})$ (sufficient conditions are to be verified) while the vector λ^* provides the sensitivity information, as discussed in the next subsection.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Sufficiency Conditions for a General Problem. A sufficient condition for $f(\mathbf{X})$ to have a constrained relative minimum at $\mathbf{X}^*$ is given by the following theorem.

Theorem 2.6 Sufficient Condition A sufficient condition for $f(\mathbf{X})$ to have a relative minimum at $\mathbf{X}^*$ is that the quadratic, Q, defined by

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \, \partial x_j} dx_i \, dx_j \tag{2.43}$$

evaluated at $\mathbf{X} = \mathbf{X}^*$ must be positive definite for all values of $d\mathbf{X}$ for which the constraints are satisfied.

1. If

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \, \partial x_j} (\mathbf{X}^*, \boldsymbol{\lambda}^*) dx_i \, dx_j$$

is negative for all choices of the admissible variations dx_i, $\mathbf{X}^*$ will be a constrained maximum of $f(\mathbf{X})$.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

2. It has been shown by Hancock [2.1] that a necessary condition for the quadratic form Q, defined by Eq. (2.43), to be positive (negative) definite for all admissible variations $d\mathbf{X}$ is that each root of the polynomial z_i, defined by the following determinantal equation, be positive (negative):

$$\begin{vmatrix} L_{11}-z & L_{12} & L_{13} & \cdots & L_{1n} & g_{11} & g_{21} & \cdots & g_{m1} \\ L_{21} & L_{22}-z & L_{23} & \cdots & L_{2n} & g_{12} & g_{22} & \cdots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \cdots & L_{nn}-z & g_{1n} & g_{2n} & \cdots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \cdots & g_{1n} & 0 & 0 & \cdots & 0 \\ g_{21} & g_{22} & g_{23} & \cdots & g_{2n} & 0 & 0 & \cdots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \cdots & g_{mn} & 0 & 0 & \cdots & 0 \end{vmatrix} = 0 \tag{2.44}$$

where

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \, \partial x_j}(\mathbf{X}^*, \boldsymbol{\lambda}^*) \tag{2.45}$$

$$g_{ij} = \frac{\partial g_i}{\partial x_j}(\mathbf{X}^*) \tag{2.46}$$

3. Equation (2.44), on expansion, leads to an $(n-m)$th-order polynomial in z. If some of the roots of this polynomial are positive while the others are negative, the point $\mathbf{X}^*$ is not an extreme point.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Example 2.10 Find the dimensions of a cylindrical tin (with top and bottom) made up of sheet metal to maximize its volume such that the total surface area is equal to $A_0 = 24\pi$.

SOLUTION If x_1 and x_2 denote the radius of the base and length of the tin, respectively, the problem can be stated as

$$\text{Maximize } f(x_1, x_2) = \pi x_1^2 x_2$$

subject to

$$2\pi x_1^2 + 2\pi x_1 x_2 = A_0 = 24\pi$$

The Lagrange function is

$$L(x_1, x_2, \lambda) = \pi x_1^2 x_2 + \lambda(2\pi x_1^2 + 2\pi x_1 x_2 - A_0)$$

Υπενθύμιση:

The Lagrange function, L, in this case is defined by introducing one Lagrange multiplier λ_j for each constraint $g_j(\mathbf{X})$ as

$$L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m)$$
$$= f(\mathbf{X}) + \lambda_1 g_1(\mathbf{X}) + \lambda_2 g_2(\mathbf{X}) + \cdots + \lambda_m g_m(\mathbf{X}) \qquad (2.40)$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

and the necessary conditions for the maximum of f give

$$\frac{\partial L}{\partial x_1} = 2\pi x_1 x_2 + 4\pi \lambda x_1 + 2\pi \lambda x_2 = 0 \tag{E$_1$}$$

$$\frac{\partial L}{\partial x_2} = \pi x_1^2 + 2\pi \lambda x_1 = 0 \tag{E$_2$}$$

$$\frac{\partial L}{\partial \lambda} = 2\pi x_1^2 + 2\pi x_1 x_2 - A_0 = 0 \tag{E$_3$}$$

Equations (E_1) and (E_2) lead to

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

that is,

$$x_1 = \tfrac{1}{2} x_2 \tag{E$_4$}$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

and Eqs. (E_3) and (E_4) give the desired solution as

$$x_1^* = \left(\frac{A_0}{6\pi}\right)^{1/2}, \; x_2^* = \left(\frac{2A_0}{3\pi}\right)^{1/2}, \; \text{and } \lambda^* = -\left(\frac{A_0}{24\pi}\right)^{1/2}$$

This gives the maximum value of f as

$$f^* = \left(\frac{A_0^3}{54\pi}\right)^{1/2}$$

If $A_0 = 24\pi$, the optimum solution becomes

$$x_1^* = 2, \quad x_2^* = 4, \quad \lambda^* = -1, \quad \text{and} \quad f^* = 16\pi$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

To see that this solution really corresponds to the maximum of f, we apply the sufficiency condition of Eq. (2.44). In this case

$$L_{11} = \left.\frac{\partial^2 L}{\partial x_1^2}\right|_{(\mathbf{X}^*,\lambda^*)} = 2\pi x_2^* + 4\pi\lambda^* = 4\pi$$

$$L_{12} = \left.\frac{\partial^2 L}{\partial x_1 \partial x_2}\right|_{(\mathbf{X}^*,\lambda^*)} = L_{21} = 2\pi x_1^* + 2\pi\lambda^* = 2\pi$$

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

$$L_{22} = \left.\frac{\partial^2 L}{\partial x_2^2}\right|_{(\mathbf{X}^*,\lambda^*)} = 0$$

$$g_{11} = \left.\frac{\partial g_1}{\partial x_1}\right|_{(\mathbf{X}^*,\lambda^*)} = 4\pi x_1^* + 2\pi x_2^* = 16\pi$$

$$g_{12} = \left.\frac{\partial g_1}{\partial x_2}\right|_{(\mathbf{X}^*,\lambda^*)} = 2\pi x_1^* = 4\pi$$

Thus Eq. (2.44) becomes

$$\begin{vmatrix} 4\pi - z & 2\pi & 16\pi \\ 2\pi & 0 - z & 4\pi \\ 16\pi & 4\pi & 0 \end{vmatrix} = 0$$

$$\begin{vmatrix} L_{11} - z & L_{12} & L_{13} & \dots & L_{1n} & g_{11} & g_{21} & \dots & g_{m1} \\ L_{21} & L_{22} - z & L_{23} & \dots & L_{2n} & g_{12} & g_{22} & \dots & g_{m2} \\ \vdots & & & & & & & & \\ L_{n1} & L_{n2} & L_{n3} & \dots & L_{nn} - z & g_{1n} & g_{2n} & \dots & g_{mn} \\ g_{11} & g_{12} & g_{13} & \dots & g_{1n} & 0 & 0 & \dots & 0 \\ g_{21} & g_{22} & g_{23} & \dots & g_{2n} & 0 & 0 & \dots & 0 \\ \vdots & & & & & & & & \\ g_{m1} & g_{m2} & g_{m3} & \dots & g_{mn} & 0 & 0 & \dots & 0 \end{vmatrix} = 0 \qquad (2.44)$$

that is,

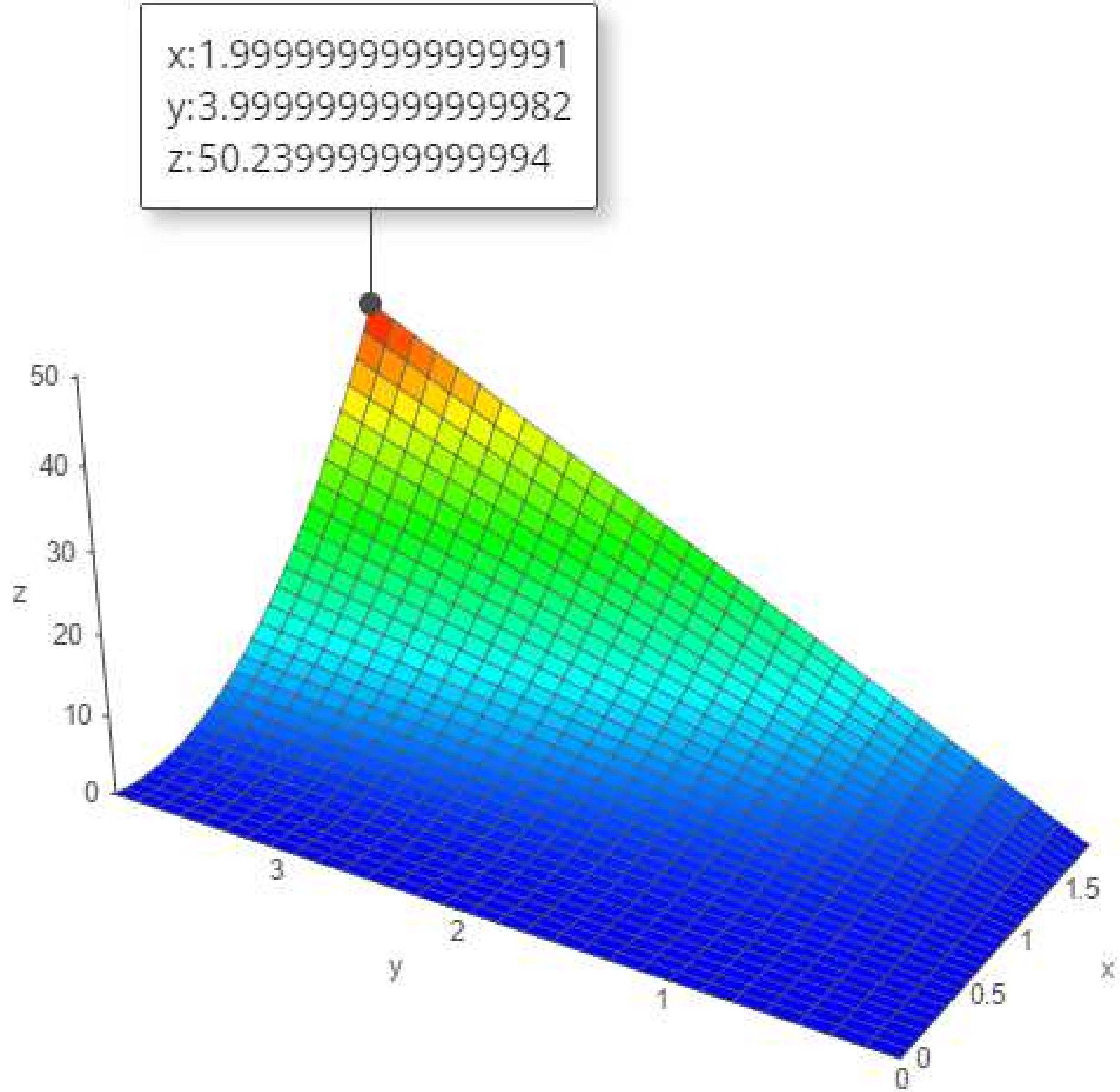
$$272\pi^2 z + 192\pi^3 = 0$$

This gives

$$z = -\tfrac{12}{17}\pi$$

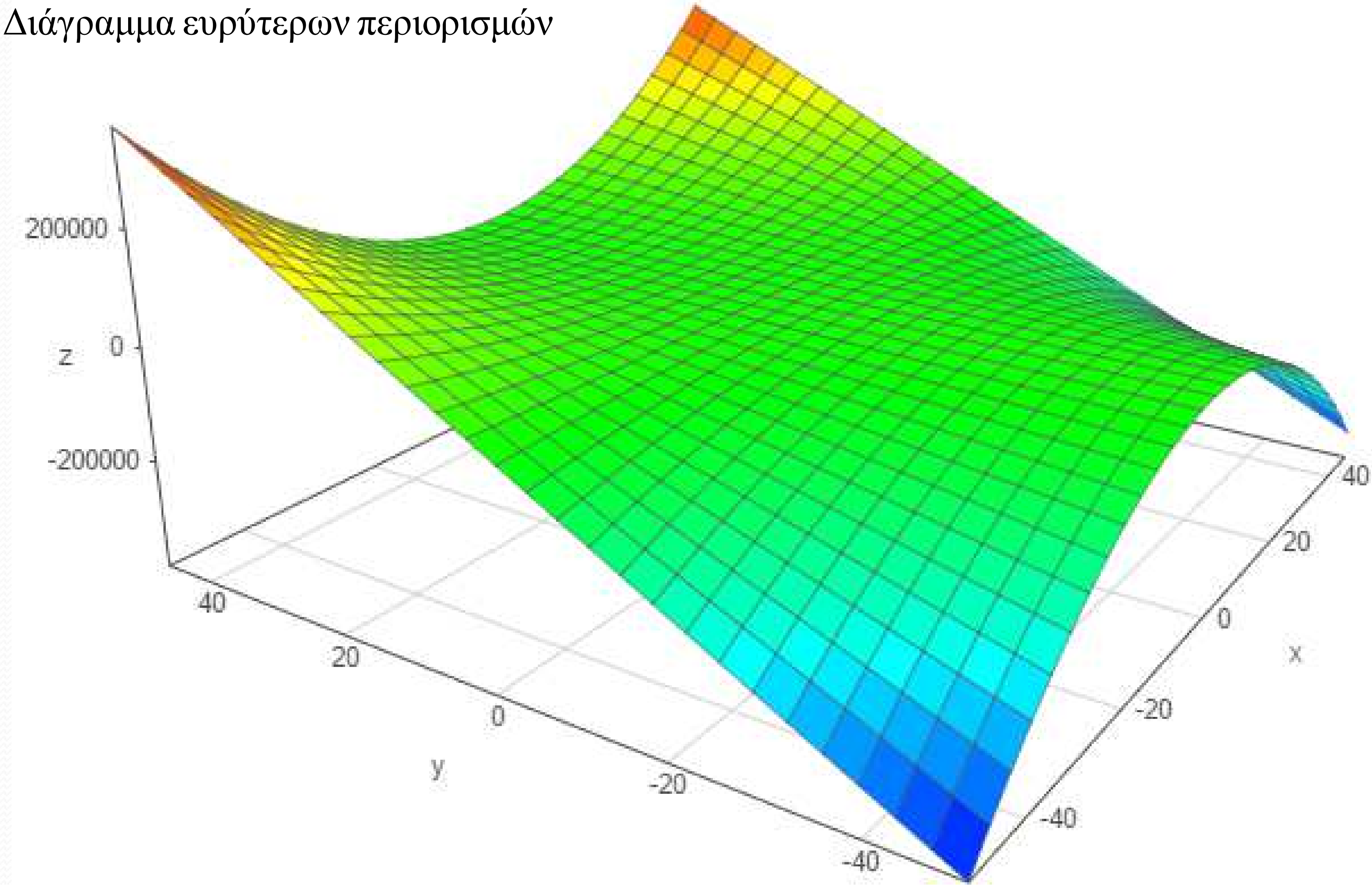
Since the value of z is negative, the point (x_1^*, x_2^*) corresponds to the maximum of f.

Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).



Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ισότητας. Επίλυση με τη μέθοδο των πολλαπλασιαστών Lagrange (Lagrange Multipliers).

Διάγραμμα ευρύτερων περιορισμών



Βελτιστοποίηση συναρτήσεων με πολλές μεταβλητές, με περιορισμούς ανισότητας.

$$\text{Minimize } f(\mathbf{X})$$

subject to

$$g_j(\mathbf{X}) \le 0, \quad j = 1, 2, \ldots, m \tag{2.58}$$

The inequality constraints in Eq. (2.58) can be transformed to equality constraints by adding nonnegative slack variables, y_j^2, as

$$g_j(\mathbf{X}) + y_j^2 = 0, \quad j = 1, 2, \ldots, m \tag{2.59}$$

where the values of the slack variables are yet unknown. The problem now becomes

$$\text{Minimize } f(\mathbf{X})$$

subject to

$$G_j(\mathbf{X}, \mathbf{Y}) = g_j(\mathbf{X}) + y_j^2 = 0, \qquad j = 1, 2, \ldots, m \tag{2.60}$$

where $\mathbf{Y} = \{y_1, y_2, \ldots, y_m\}^{\mathrm{T}}$ is the vector of slack variables.

This problem can be solved conveniently by the method of Lagrange multipliers. For this, we construct the Lagrange function L as

$$L(\mathbf{X}, \mathbf{Y}, \boldsymbol{\lambda}) = f(\mathbf{X}) + \sum_{j=1}^{m} \lambda_j G_j(\mathbf{X}, \mathbf{Y}) \tag{2.61}$$

- Οι αναλυτικές μέθοδοι υπολογίζουν τη θέση των ακρότατων, τα οποία μπορεί να είναι καθολικά ή τοπικά. Όμως το καθολικό ακρότατο μπορεί να βρίσκεται στα συνοριακά σημεία ή σε σημεία που δεν υπάρχει η παράγωγος, άρα πρέπει να γίνεται έλεγχος και εκεί.
- Ο υπολογισμός των ακρότατων βασίζεται στον υπολογισμό της κλίσης της αντικειμενικής συνάρτησης. Όμως η μηδενική κλίση δεν εξασφαλίζει την ύπαρξη ακρότατου.

- Ποιες είναι οι αναγκαίες και ικανές συνθήκες για την ύπαρξη ακρότατου μιας συνάρτησης f(x)?
- Απάντηση:
- Αν μια συνάρτηση f(x) έχει ακρότατο στη θέση x^* (ελάχιστο ή μέγιστο) μεταξύ δύο τιμών της μεταβλητής x **και υπάρχει η $1^η$ παράγωγος της συνάρτησης στο σημείο x^*** τότε $f`(x^*)=0$
- Για να προσδιοριστεί εάν το ακρότατο (εκεί που μηδενίζεται η $1^η$ παράγωγος) είναι μέγιστο ή ελάχιστο χρειάζεται να ελεγχθεί η τιμή της $2^{ης}$ ή ανώτερης παραγώγου στα κρίσιμα σημεία.