Μαθηματική Ανάλυση
Διάλεξη 5

Κωνσταντίνος Γιαννουτάκης
Επίκουρος Καθηγητής

Σπύρος Χαλκίδης
Ε.ΔΙ.Π.

Δημήτρης Χρήστου-Βαρσακέλης
Καθηγητής

Οκτώβριος 2023

Θέματα 5ης διάλεξης

- Μονοτονία συναρτήσεων, στάσιμα και κρίσιμα σημεία
- Συναρτήσεις πολλών μεταβλητών
- Μερικές παράγωγοι πρώτης και δεύτερης τάξης
- Εσσιανή μήτρα (ή Εσσιανός πίνακας)
- Σειρές Taylor πολυμεταβλητών συναρτήσεων

Μονοτονία συνάρτησης

Έστω μια συνάρτηση $f(x)$. Αυτή καλείται:

Γνησίως αύξουσα εάν για οποιαδήποτε x_1, x_2 στο πεδίο ορισμού της ισχύει: $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$.

Γνησίως φθίνουσα εάν για οποιαδήποτε x_1, x_2 στο πεδίο ορισμού της ισχύει: $x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$.

Αύξουσα εάν για οποιαδήποτε x_1, x_2 στο πεδίο ορισμού της ισχύει: $x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)$.

Φθίνουσα εάν για οποιαδήποτε x_1, x_2 στο πεδίο ορισμού της ισχύει: $x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2)$.

Μονοτονία συνάρτησης

Εάν η συνάρτηση $f(x)$ είναι συνεχής, τότε:

- Στα διαστήματα του πεδίου ορισμού όπου $f'(x) > 0$, η συνάρτηση είναι γνησίως αύξουσα
- Στα διαστήματα του πεδίου ορισμού όπου $f'(x) < 0$, η συνάρτηση είναι γνησίως φθίνουσα

Βρείτε τα διαστήματα μονοτονίας της συνάρτησης $f(x) = 2x^3 + 3x^2 - 12x + 1$.

Η συνάρτηση είναι συνεχής (γιατί;) οπότε υπολογίζουμε την πρώτη της παράγωγο:
$f'(x) = 2 \cdot 3x^2 + 3 \cdot 2x - 12 = 6x^2 + 6x - 12 = 6(x+2)(x-1)$

x	$-\infty$		-2		1		$+\infty$
$f'(x)$		$+$	0	$-$	0	$+$	
$f(x)$		$\nearrow$		$\searrow$		$\nearrow$	

Μέγιστο συνάρτησης

Σε ένα σημείο $\hat{x}$ **ολικού μεγίστου** ισχύει:

$$f(\hat{x}) \geq f(x) \text{ για όλα τα } x \text{ στο πεδίο ορισμού της,}$$

ενώ σε ένα σημείο x^* **τοπικού μεγίστου**, ισχύει:

$$f(x^*) \geq f(x), \quad x^* - \epsilon \leq x \leq x^* + \epsilon, \epsilon > 0$$

για x σε ένα διάστημα, ενδεχομένως πολύ μικρό, γύρω από το x^*.

Αν η συνάρτηση έχει τοπικό μέγιστο στο σημείο x^* τότε για f παραγωγίσιμη, θα πρέπει η πρώτη παράγωγος της συνάρτησης να μηδενίζεται στο $x = x^*$, δηλαδή

$$f^{'}(x^*) = 0$$

Γιατί $f^{'}(x^{*}) = 0$ αναγκαίο για βέλτιστο στο x^{*}

Σειρά Taylor γύρω από το x^*:

$$f(x^* + \epsilon) = f(x^*) + \underbrace{\left.\frac{df}{dx}\right|_{x*}}_{f'(x^*)} \epsilon + \frac{1}{2} \underbrace{\left.\frac{d^2 f}{dx^2}\right|_{x*}}_{f''(x^*)} \epsilon^2 + O(\epsilon^3)$$

όπου $\left.\frac{df}{dx}\right|_{x*} = f'(x^*)$ και $O(\epsilon^3)$ όροι 3ης ή μεγαλύτερης τάξης ως προς ϵ.
Για να έχουμε **(έστω) τοπικό μέγιστο** στο x^* θα πρέπει για $|\epsilon|$ αρκετά μικρό,

$$f(x^* + \epsilon) - f(x^*) = \left.\frac{df}{dx}\right|_{x*} \epsilon + \frac{1}{2} \left.\frac{d^2 f}{dx^2}\right|_{x*} \epsilon^2 + O(\epsilon^3) < 0.$$

Έστω ότι $f'(x^*) \neq 0$, π.χ. $f'(x^*) > 0$. Θα υπάρχει $\epsilon > 0$ αρκετά μικρό τέτοιο ώστε το δεξί μέλος της ισότητας να είναι **θετικό**, γιατί όταν $\epsilon \to 0$, οι όροι μεγαλύτερων τάξεων, ϵ^2, ϵ^3 κλπ, τείνουν στο μηδέν ¨γρηγορότερα' από το γραμμικό όρο $\left.\frac{df}{dx}\right|_{x*} \epsilon$.
Οδηγούμαστε σε άτοπο.

Γιατί $f'(x^*) = 0$ αναγκαίο για βέλτιστο στο x^*

Τοπικό μέγιστο στο x^*: Για ϵ αρκετά μικρό,

$$f(x^* + \epsilon) - f(x^*) = \left.\frac{df}{dx}\right|_{x*} \epsilon + \frac{1}{2} \left.\frac{d^2 f}{dx^2}\right|_{x*} \epsilon^2 + O(\epsilon^3) < 0.$$

Όμοια, αν $f'(x^*) < 0$ τότε θα υπάρχει $\epsilon < 0$ με $|\epsilon|$ αρκετά μικρό, τέτοιο ώστε το δεξί μέλος της ισότητας να είναι και πάλι **θετικό**, ενώ το αριστερό πρέπει να είναι **αρνητικό**! (άτοπο).

Τελικά λοιπόν, θα πρέπει **οπωσδήποτε** $f'(x^*) = 0$ για μέγιστο στο x^*. Όμοια αποδεικνύεται ότι η **ίδια συνθήκη πρέπει να ισχύει και για τοπικό ελάχιστο** στο x^*.

Αναγκαία συνθήκη

Θεώρημα Fermat

Αν η παραγωγίσιμη συνάρτηση f παίρνει μία ακρότατη τιμή σε ένα σημείο x^*, τότε $f'(x^*) = 0$.

Η συνθήκη πρώτης τάξης $f'(x^*) = 0$ είναι μία **αναγκαία συνθήκη** για να δώσει το x^* μία ακρότατη τιμή στη συνάρτηση. Η συνθήκη αυτή δεν είναι ικανή για μία ακρότατη τιμή, καθότι υπάρχει μία άλλη ομάδα σημείων, τα λεγόμενα **σημεία καμπής**, όπου η παράγωγος είναι δυνατόν να μηδενίζεται, δηλαδή $f'(x^*) = 0$.
Για παράδειγμα για τη συνάρτηση

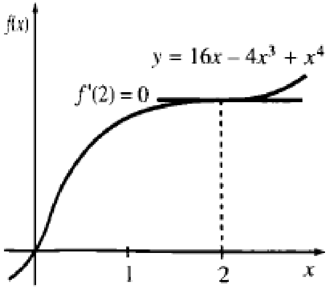
$$y = 16x - 4x^3 + x^4$$

Έχουμε:

$$\frac{dy}{dx} = 16 - 12x^2 + 4x^3 = f'(x)$$

και για $x = 2$ παίρνουμε $f'(2) = 0$.

Γραφική παράσταση



Σχήμα: Συνάρτηση με σημείο καμπής

Το $x = 2$ δεν οδηγεί σε μία ακρότατη τιμή της συνάρτησης. Συμβαίνει η εφαπτομένη της συνάρτησης σε αυτό το σημείο να είναι οριζόντια ($y = 16$).

Στάσιμα και κρίσιμα σημεία

Για μία παραγωγίσιμη συνάρτηση f, το σημείο x^* όπου $f^{'}(x^*) = 0$, χαρακτηρίζεται ως **στάσιμη τιμή** της συνάρτησης. Σε τέτοιες στάσιμες τιμές η συνάρτηση μπορεί να παρουσιάζει ακρότατες τιμές ή σημεία καμπής. Κάθε ακρότατη τιμή μίας συνάρτησης παρουσιάζεται σε μία στάσιμη τιμή, αλλά δεν έχουμε υποχρεωτικά ακρότατη τιμή σε κάθε στάσιμη τιμή.

Κρίσιμα σημεία μιας συνάρτησης ονομάζονται τα στάσιμα σημεία της συνάρτησης μαζί με τα σημεία για τα οποία δεν ορίζεται η παράγωγός τους.

Συνθήκες Δεύτερης Τάξης

Εστω x^* τοπικό βέλτιστο της f. Ας εξετάσουμε πάλι τη σειρά Taylor

$$f(x^* + \epsilon) = f(x^*) + \underbrace{\left.\frac{df}{dx}\right|_{x*}}_{f'(x^*)=0} \epsilon + \frac{1}{2} \underbrace{\left.\frac{d^2 f}{dx^2}\right|_{x*}}_{f''(x^*)} \epsilon^2 + \frac{1}{6} \underbrace{\left.\frac{d^3 f}{dx^3}\right|_{x*}}_{f'''(x^*)} \epsilon^3 + \cdots$$

ή

$$f(x^* + \epsilon) - f(x^*) = \frac{1}{2} \underbrace{\left.\frac{d^2 f}{dx^2}\right|_{x*}}_{f''(x^*)} \epsilon^2 + + \frac{1}{6} \underbrace{\left.\frac{d^3 f}{dx^3}\right|_{x*}}_{f'''(x^*)} \epsilon^3 + \cdots$$

Συνθήκες Δεύτερης Τάξης

Το αν στο x^* έχουμε *min* (δηλ. $f(x^* + \epsilon) - f(x^*) > 0 \ \forall |\epsilon|$ αρκετά μικρό) ή *max* (δηλ. $f(x^* + \epsilon) - f(x^*) < 0 \ \forall |\epsilon|$ αρκετά μικρό) θα κριθεί από το πρόσημο του ¨επόμενου' μη-μηδενικού όρου στη σειρά Taylor:

$$\underbrace{\frac{1}{2}}_{>0} \left.\frac{d^2 f}{dx^2}\right|_{x*} \underbrace{\epsilon^2}_{>0},$$

δηλαδή από το πρόσημο της 2ης παραγώγου:

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

Προφανώς, αν $f''(x^*) > 0$ τότε $f(x^* + \epsilon) - f(x^*) > 0 \ \forall |\epsilon|$ αρκετά μικρό, και το σημείο x^* δίνει **τοπικό ελάχιστο** της f.
Αντίστοιχα, αν $f''(x^*) < 0$ τότε $f(x^* + \epsilon) - f(x^*) < 0 \ \forall |\epsilon|$ αρκετά μικρό, και το σημείο x^* δίνει **τοπικό μέγιστο** της f.

Συνθήκες Δεύτερης Τάξης

Αν η $f(x)$ είναι αυστηρά κοίλη στην περιοχή του x^* και δύο φορές παραγωγίσιμη, τότε η καμπυλότητα της συνάρτησης είναι αρνητική ή διαφορετικά $f''(x^*) < 0$ καθώς στρέφει τα κοίλα κάτω.

> Αν $f'(x^*) = 0$ και $f''(x^*) < 0$, τότε η f παρουσιάζει ένα τοπικό μέγιστο στο x^*.

Αν η $f(x)$ είναι αυστηρά κυρτή στην περιοχή του x^* και δύο φορές παραγωγίσιμη, τότε η καμπυλότητα της συνάρτησης είναι θετική ή διαφορετικά $f''(x^*) > 0$ καθώς στρέφει τα κοίλα άνω.

> Αν $f'(x^*) = 0$ και $f''(x^*) > 0$, τότε η f παρουσιάζει ένα τοπικό ελάχιστο στο x^*.

Στην περίπτωση ενός σημείου καμπής η καμπυλότα της συνάρτησης αλλάζει από κοίλη σε κυρτή ή αντίστροφα. Σε αυτήν την περίπτωση $f''(x^*) = 0$.

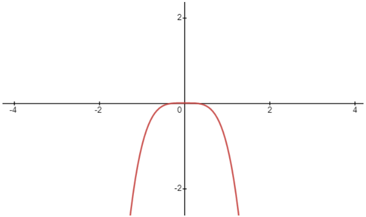
Συνθήκες Δεύτερης Τάξης

Οι συνθήκες δεύτερης τάξης είναι ικανές για να δώσουν ένα τοπικό ακρότατο αλλά όχι αναγκαίες. Δηλαδή, μπορεί να έχουμε ένα τοπικό ακρότατο και παρόλα αυτά να ισχύει $f''(x^*) = 0$.

Για παράδειγμα, η συνάρτηση $f(x) = -x^4$ παρουσιάζει ένα μέγιστο στο $x^* = 0$, αλλά

$$f''(0) = -12(0)^2 = 0$$

.

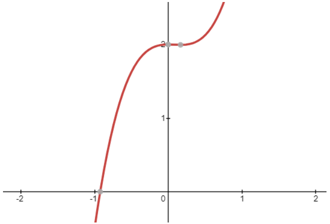


Παράδειγμα

Να μελετηθεί ως προς τα ακρότατα η συνάρτηση:

$$f(x) = y = 2x^3 - 0.5x^2 + 2$$

Έχουμε $\frac{dy}{dx} = 6x^2 - x$, συνεπώς τα στάσιμα σημεία είναι το $x = 0$ και το $x = 1/6$.
$\frac{d^2y}{dx^2} = 12x - 1$.
Στο $x = 0$, $f''(x) = -1 < 0$, συνεπώς έχουμε ένα τοπικό μέγιστο της συνάρτησης.
Στο $x = 1/6$, $f''(x) = 1 > 0$ συνεπώς έχουμε ένα τοπικό ελάχιστο της συνάρτησης.



Άσκηση

Σχεδιάστε τη γραφική παράσταση της συνάρτησης $f(x) = x^4 - 2x^2$, βρίσκοντας τα διαστήματα μονοτονίας, κυρτότητας, τα ακρότατα και τα σημεία καμπής της.

Λύση

Αρχικά βρίσκουμε που τέμνει η συνάρτηση τον άξονα των x:
$f(x) = 0 \Rightarrow x^4 - 2x^2 = 0 \Rightarrow x^2(x^2 - 2) = 0 \Rightarrow x = 0, x = \sqrt{2}, x = -\sqrt{2}$.
Επιπλέον, παρατηρούμε ότι η συνάρτηση είναι άρτια, δηλαδή ισχύει $f(-x) = f(x)$, $\forall x \in \mathbb{R}$, που σημαίνει ότι το γράφημά της είναι συμμετρικό ως προς τον άξονα των y.
Υπολογίζουμε τις ρίζες της πρώτης παραγώγου:
$f'(x) = 4x^3 - 4x = 0 \Rightarrow 4x(x^2 - 1) = 0 \Rightarrow x = 0, x = 1, x = -1$. Για τα διαστήματα μονοτονίας κατασκευάζουμε τον πίνακα:

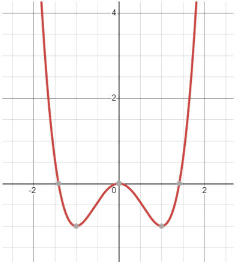
x	$-\infty$		-1		0		1		$+\infty$
$f'(x)$		$-$	0	$+$	0	$-$	0	$+$	
$f(x)$		$\searrow$		$\nearrow$		$\searrow$		$\nearrow$	

Άρα η f έχει τοπικό μέγιστο στο σημείο $x = 0$ (με $f(0) = 0$), και τοπικά ελάχιστα στα σημεία $x = -1$ και $x = 1$ (με $f(1) = f(-1) = -1$).

Λύση

Βρίσκουμε που μηδενίζεται η δεύτερη παράγωγος:
$f''(x) = 0 \Rightarrow 12x^2 - 4 = 0 \Rightarrow x = \frac{1}{\sqrt{3}}$ και $x = -\frac{1}{\sqrt{3}}$.
Τα σημεία αυτά είναι σημεία καμπής, επομένως επεκτείνουμε τον πίνακα μονοτονίας:

x	$-\infty$		-1		$-\frac{1}{\sqrt{3}}$		0		$\frac{1}{\sqrt{3}}$		1		$+\infty$
$f''(x)$		$+$		$+$	0	$-$		$-$	0	$+$		$+$	
$f'(x)$		$-$	0	$+$		$+$	0	$-$		$-$	0	$+$	
$f(x)$													



Τι συμβαίνει αν $f'(x^*) = f''(x^*) = 0$

Αν και η 2η παράγωγος μηδενίζεται στο στάσιμο σημείο x^* τότε η σειρά Taylor δίνει

$$f(x^* + \epsilon) - f(x^*) = \underbrace{\left.\frac{df}{dx}\right|_{x*}}_{f'(x^*)=0} \epsilon + \frac{1}{2} \underbrace{\left.\frac{d^2 f}{dx^2}\right|_{x*}}_{f''(x^*)} \epsilon^2 + \frac{1}{6} \underbrace{\left.\frac{d^3 f}{dx^3}\right|_{x*}}_{f'''(x^*)} \epsilon^3 + \frac{1}{24} \underbrace{\left.\frac{d^4 f}{dx^4}\right|_{x*}}_{f''''(x^*)} \epsilon^4 + \cdots$$

Επειδή ϵ^3 μπορεί να είναι είτε θετικό είτε αρνητικό, θα πρέπει

$$f'''(x^*) = 0$$

διαφορετικά δεν μπορεί να είναι τοπικό βέλτιστο το x^*.
Για παράδειγμα, αν x^* είναι τοπικό μέγιστο, θα πρέπει το αριστερό μέλος να είναι αρνητικό για αρκετά μικρό $|\epsilon|$:

$$f(x^* + \epsilon) - f(x^*) < 0$$

Τι συμβαίνει αν $f'(x^*) = f''(x^*) = 0$

Έστω $f'''(x^*) \neq 0$, π.χ. $f'''(x^*) > 0$. Θα υπάρχει $\epsilon > 0$ αρκετά μικρό τέτοιο ώστε το δεξί μέλος να είναι **θετικό** Οδηγούμαστε σε άτοπο.

Όμοια, αν $f'''(x^*) < 0$ Θα υπάρχει $\epsilon < 0$ με $|\epsilon|$ αρκετά μικρό, τέτοιο ώστε το δεξί μέλος να είναι και πάλι **θετικό** (άτοπο).

Τελικά λοιπόν, θα πρέπει οπωσδήποτε $f'''(x^*) = 0$ για μέγιστο στο x^*. Όμοια αποδεικνύεται ότι η ίδια συνθήκη πρέπει να ισχύει και για ελάχιστο στο x^*.

Συνεπώς: αν $f'(x^*) = f''(x^*) = 0$ τότε αναγκαία συνθήκη για βέλτιστο είναι $f'''(x^*) = 0$.

Το αν το x^* είναι μέγιστο ή ελάχιστο, θα κριθεί από το πρόσημο του επόμενου μη-μηδενικού όρου στη σειρά Taylor,

$$f''''(x^*).$$

Έτσι, ικανή συνθήκη για μέγιστο (ελάχιστο) είναι $f''''(x^*) < 0$ ($f''''(x^*) > 0$).

Παρατηρήστε ότι η αναγκαία και ικανή συνθήκη ¨μεταφέρθηκε' από την 1η και 2η παράγωγο στην 3η και 4η. Αν και αυτές είναι μηδενικές, εφαρμόζουμε τις ίδιες συνθήκες στην 5η και 6η, κ.ο.κ.

Παράδειγμα

Θεωρούμε τη συνάρτηση $f(x) = 10 - x^4$. Έχουμε:

$$f'(x) = -4x^3$$

$$f''(x) = -12x^2$$

$$f'''(x) = -24x$$

$$f^{(4)}(x) = -24$$

Όλες οι πάραγωγοι μέχρι την τρίτη παράγωγο μηδενίζονται στο $x^* = 0$. Αυτό σημαίνει ότι πρέπει να χρησιμοποιήσουμε την τέταρτη παράγωγο ως τελευταίο όρο στο ανάπτυγμα της σειράς Taylor για να διερευνήσουμε το χαρακτήρα της συνάρτησης γύρω από το $x^* = 0$:

$$f(\hat{x}) = f(0) + \frac{f^{(4)}(\zeta)(\hat{x})^4}{24}$$

Παράδειγμα

Δεδομένου ότι $f^{(4)}(x) = -24$ για όλα τα x, αυτό μπορούμε να το διατυπώσουμε ως εξής:
$f(\hat{x}) = f(0) - \hat{x}^4$ ή $f(\hat{x}) - f(0) < -\hat{x}^4$
Για οποιαδήποτε τιμή του $\hat{x} \neq 0$ έχουμε $-\hat{x}^4 < 0$ και επομένως $f(\hat{x}) - f(0) < 0$ ή $f(0) > f(\hat{x})$. Δηλαδή το σημείο $x^* = 0$ δίνει ένα τοπικό μέγιστο αυτής της συνάρτησης.

- Γενικά αν όλες οι παράγωγοι σε ένα σημείο μέχρι και μία παράγωγο περιττής τάξης είναι μηδέν, ενώ η επόμενη παράγωγος άρτιας τάξης είναι αρνητική, τότε έχουμε ένα τοπικό μέγιστο σε αυτό το σημείο, ενώ αν είναι θετική τότε έχουμε ένα τοπικό ελάχιστο σε αυτό το σημείο.

Παράδειγμα 2

Θεωρούμε τη συνάρτηση $f(x) = x^3$. Έχουμε:

$$f'(x) = 3x^2$$

$$f''(x) = 6x$$

$$f'''(x) = 6$$

Με την $f'''(x)$ να είναι η πρώτη από τις παραγώγους ανώτερης τάξης που δε μηδενίζεται στο x^* πρέπει να χρησιμοποιήσουμε την τρίτη παράγωγο ως τον τελευταίο όρο του αναπτύγματος της σειράς Taylor για να διερευνήσουμε το χαρακτήρα της συνάρτησης γύρω από το $x^* = 0$. Δηλαδή,

$$f(\hat{x}) = f(0) + \frac{f'''(\zeta)(\hat{x})^3}{6}$$

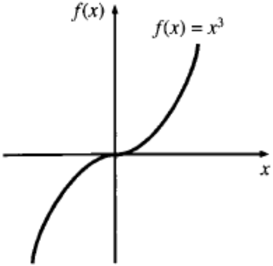
για ζ ανάμεσα στο x^* και το $\hat{x}$.

Παράδειγμα 2

Επειδή $f'''(x) = 6$ για όλα τα x μπορούμε να γράψουμε αυτήν την εξίσωση ως:

$$f(\hat{x}) = f(0) + \hat{x}^3$$

Για $\hat{x} > 0$ παίρνουμε $f(\hat{x}) > f(0)$ ενώ για $\hat{x} < 0$ παίρνουμε $f(\hat{x}) < f(0)$. Επομένως το $x^* = 0$ δε δίνει ούτε μέγιστο ούτε ελάχιστο. Πρόκειται για ένα σημείο καμπής, όπως φαίνεται και στο διάγραμμα.



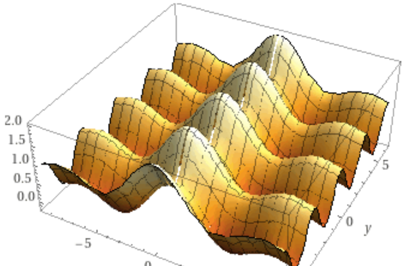
Σχήμα: Γραφική παράσταση της $f(x) = x^3$

Συναρτήσεις πολλών μεταβλητών

Συνάρτηση (βαθμωτή) n μεταβλητών ονομάζεται μια αντιστοιχία που απεικονίζει κάθε n-άδα $(x_1, x_2, \ldots x_n)$ του $\mathbb{R}^n$ (ή κάθε σημείο του n-διάστατου χώρου) σε έναν πραγματικό αριθμό. Το πεδίο ορισμού μιας τέτοιας συνάρτησης είναι ένα υποσύνολο του $\mathbb{R}^n$.

$$f : X \to \mathbb{R}, \text{ όπου } X \subseteq \mathbb{R}^n$$

Παράδειγμα: Η συνάρτηση $f : \mathbb{R}^* \times \mathbb{R} \to \mathbb{R}$ με $f(x, y) = \frac{\sin(x)}{x} + \cos^2(y)$



Ισοσταθμικά σύνολα

Ισοσταθμικό σύνολο (level set) της συνάρτησης $y = f(x_1, x_2, \cdots, x_n)$ είναι το σύνολο

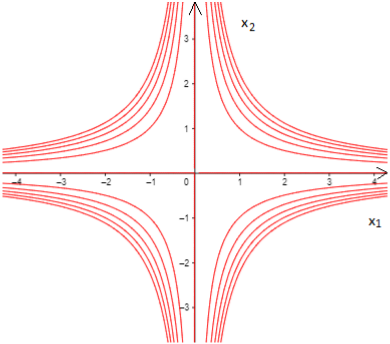
$$L = \{(x_1, \cdots, x_n)\} \in \mathbb{R}^n : f(x_1, x_2, \cdots, x_n) = c\}$$

για ένα δεδομένο αριθμό $c \in \mathbb{R}$.

Παράδειγμα Ισοσταθμικών συνόλων

Έστω ότι θέλουμε να βρούμε τα ισοσταθμικά σύνολα της συνάρτησης $y = f(x_1, x_2) = x_1^2 x_2^2$.

Θέτουμε $x_1^2 x_2^2 = c \iff x_2 = \pm\sqrt{\frac{c}{x_1^2}}$. Τα ισοσταθμικά σύνολα της συνάρτησης απεικονίζονται στο παρακάτω διάγραμμα:



Σχήμα: Ισοσταθμικά σύνολα της συνάρτησης $y = x_1^2 x_2^2$

Εφαρμογές Ισοσταθμικών Καμπυλών

Στα Οικονομικά τα ισοσταθμικά σύνολα τα συναντάμε:

- στη θεωρία καταναλωτή (εκεί όπου αναφέρονται ως **καμπύλες αδιαφορίας**), όπου οποιοδήποτε σημείο πάνω στην ίδια καμπύλη έχει την ίδια χρησιμότητα.
- στη θεωρία παραγωγού (εκεί όπου αναφέρονται ως **καμπύλες ισοπαραγωγής**), όπου οποιοδήποτε σημείο πάνω στην ίδια καμπύλη αντιστοιχεί στο ίδιο επίπεδο παραγωγής.

Παραγώγιση πολυμεταβλητών συναρτήσεων

Η **μερική παράγωγος** μίας συνάρτησης $y = f(x_1, x_2, \cdots, x_n)$ (μπορεί να γραφεί και ως $y = f(\mathbf{x})$ όπου $\mathbf{x} \in \mathbb{R}^n$) ως προς τη μεταβλητή x_i είναι:

$$\frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, \cdots, x_i + \Delta x_i, \cdots x_n) - f(x_1, \cdots, x_i, \cdots, x_n)}{\Delta x_i}$$

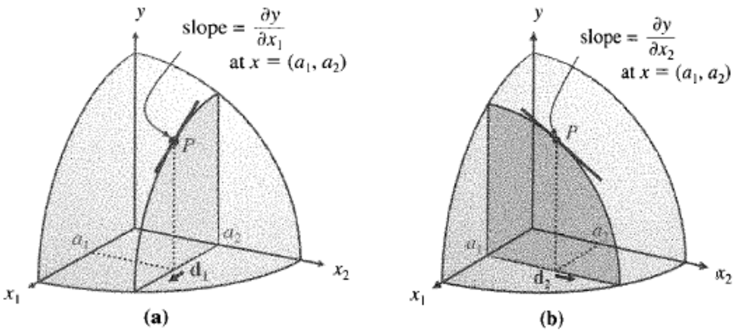
Χρησιμοποιούνται εναλλακτικά οι συμβολισμοί $\frac{\partial f}{\partial x_i}$ ή $f_i(\mathbf{x})$ ή απλά f_i.

Για μία συνάρτηση $z = f(x, y)$, οι μερικές παράγωγοι ως προς x και y είναι:

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}, \quad \frac{\partial z}{\partial y} = \lim_{\Delta x \to 0} \frac{f(x, y + \Delta x) - f(x, y)}{\Delta x}.$$

Γεωμετρική ερμηνεία

Η μερική παράγωγος εκφράζει το ρυθμό μεταβολής της συνάρτησης ως προς μία μεταβλητή, όταν όλες οι υπόλοιπες μεταβλητές παραμένουν σταθερές.



Σχήμα: Μερικές Παράγωγοι

Παραγώγιση πολυμεταβλητών συναρτήσεων

Αντί να υπολογίζουμε τις μερικές παραγώγους με βάση τις αρχές που εισάγει ο ορισμός τους, μπορούμε να τις βρούμε χρησιμοποιώντας τους κανόνες παραγώγισης, όπως κάναμε για τις μονομεταβλητές συναρτήσεις. Επειδή όταν υπολογίζουμε την $\frac{\partial f}{\partial x_i}$ κρατούμε σταθερές όλες τις μεταβλητές εκτός από την x_i μπορούμε να θεωρήσουμε όλους τους όρους της συνάρτησης $f(\mathbf{x})$ που δεν εξαρτώνται από το x_i ως μία σταθερά c και στη συνέχεια να χρησιμοποιήσουμε τους κανόνες παραγώγισης για συναρτήσεις μίας μεταβλητής.

Ιδιότητες

Εάν υπάρχουν οι μερικές παράγωγοι ως προς τη μεταβλητή x των συναρτήσεων f, g, τότε ισχύουν τα ακόλουθα:

- $$\frac{\partial(f \pm g)}{\partial x} = \frac{\partial f}{\partial x} \pm \frac{\partial g}{\partial x}$$
- $$\frac{\partial(\lambda f)}{\partial x} = \lambda \frac{\partial f}{\partial x}, \quad \lambda \in \mathbb{R}$$
- $$\frac{\partial(fg)}{\partial x} = g\frac{\partial f}{\partial x} + f\frac{\partial g}{\partial x}$$
- $$\frac{\partial(f/g)}{\partial x} = \frac{g\frac{\partial f}{\partial x} - f\frac{\partial g}{\partial x}}{g^2}$$

Παράδειγμα

Έστω $y = f(x_1, x_2) = x_1^2 x_2$. Η μεταβλητή x_2 κρατιέται σταθερή όταν υπολογίζουμε την $\frac{\partial f}{\partial x_1}$. Αν θέσουμε $x_2 = c$, όπου c μία σταθερά, τότε η συνάρτηση παίρνει τη μορφή:

$$y = cx_1^2$$

και επομένως:

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = \frac{d[cx_1^2]}{dx_1} = 2cx_1$$

Αντικαθιστούμε το c με x_2 και έχουμε

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 2x_1 x_2$$

Άσκηση

Εάν $f(x, y, z) = 3x^2y^3 - 2xy^2 + 4y^4 + z^2$, υπολογίστε τις μερικές παραγώγους της f ως προς x, y και z.

Λύση

Αφού $f(x,y,z) = 3x^2y^3 - 2xy^2 + 4y^4 + z^2$, τότε έχουμε:

$$\frac{\partial f}{\partial x} = 6xy^3 - 2y^2$$

$$\frac{\partial f}{\partial y} = 9x^2y^2 - 4xy + 16y^3$$

$$\frac{\partial f}{\partial z} = 2z$$

Μερικές παράγωγοι δεύτερης τάξης

Κάθε μία παράγωγος δεύτερης τάξης μίας συνάρτησης $f : \mathbb{R}^n \to \mathbb{R}$ συμβολίζεται ως:

$$f_{ij} = \frac{\partial f_i(x)}{\partial x_j}, i, j = 1, 2, \cdots, n$$

όπου η f_{ij} βρίσκεται παραγωγίζοντας πρώτα τη συνάρτηση $f(\mathbf{x})$ ως προς την μεταβλητή x_i και κατόπιν παραγωγίζοντας το αποτέλεσμα $f_i(\mathbf{x})$ ως προς τη μεταβλητή x_j.

Μερικές παράγωγοι δεύτερης τάξης

Για μία συνάρτηση $f(x, y)$, οι μερικές παράγωγοι δεύτερης τάξης είναι οι ακόλουθες:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = f_{x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$$

$$\frac{\partial^2 f}{\partial y^2} = f_{yy} = f_{y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$$

$$\frac{\partial^2 f}{\partial y \partial x} = f_{xy} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Οι δύο τελευταίες ονομάζονται μεικτές ή σταυροειδείς παράγωγοι.

Παράδειγμα

Έστω $f(x,y) = x^2y^3 + x^4y + xe^y$. Τότε οι μερικές παράγωγοι 2ης τάξης είναι:

1. $$\frac{\partial^2 f}{\partial x^2} = f_{xx} = \frac{\partial}{\partial x}(2xy^3 + 4x^3y + e^y) = 2y^3 + 12x^2y$$
2. $$\frac{\partial^2 f}{\partial y^2} = f_{yy} = \frac{\partial}{\partial y}(3x^2y^2 + x^4 + xe^y) = 6x^2y + xe^y$$
3. $$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = \frac{\partial}{\partial x}(3x^2y^2 + x^4 + xe^y) = 6xy^2 + 4x^3 + e^y$$
4. $$\frac{\partial^2 f}{\partial y \partial x} = f_{xy} = \frac{\partial}{\partial y}(2xy^3 + 4x^3y + e^y) = 6xy^2 + 4x^3 + e^y$$

Μεικτές παράγωγοι

Δύο μεικτές παράγωγοι τάξης $k \geq 2$ μίας συνάρτησης $f(x, y)$ με πεδίο ορισμού ένα ανοικτό σύνολο είναι ίσες εάν:

- Όλες οι μερικές παράγωγοι μέχρι τάξης k είναι συνεχείς
- Εάν ο συνολικός αριθμός παραγωγίσεων ως προς κάθε μεταβλητή είναι ο ίδιος και στις δύο μεικτές παραγώγους

Διάνυσμα κλίσης

Τις μερικές παραγώγους πρώτης τάξης είναι χρήσιμο να τις ¨αντιμετωπίζουμε' ως **στοιχεία ενός διανύσματος**-στήλης το οποίο ονομάζουμε **διάνυσμα κλίσης** το οποίο συμβολίζουμε ως ∇f ή $\frac{\partial f}{\partial x}$:

$$\frac{\partial f}{\partial x} \ \text{ή} \ \nabla f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

Γενικά, κάθε μερικη παράγωγος θα είναι συνάρτηση των στοιχείων του x. Όταν γράφουμε

$$\left.\frac{\partial f}{\partial x}\right|_{x^*} \ \text{ή} \ \nabla f(x^*)$$

εννοούμε το διάνυσμα κλίσης της $f : \mathbb{R}^n \to \mathbb{R}$ **υπολογισμένο στο σημείο** με συντεταγμένες $x^* = [x_1^*, x_2^*, \cdots x_n^*]^T$.

Παράδειγμα

Βρείτε το διάνυσμα κλίσης για τη συνάρτηση $f(x) = 5 - 2x_1 + 3x_2$.

Οι πρώτες μερικές παράγωγοι της συνάρτησης είναι $f_1 = -2$ και $f_2 = 3$. Συνεπώς το διάνυσμα κλίσης είναι:

$$\nabla f = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

Σε αυτη την περίπτωση το διάνυσμα κλίσης είναι σταθερό, άρα θα έχει την ίδια τιμή σε όλα τα σημεία του $\mathbb{R}^2$.

Εσσιανή μήτρα

Θα είναι ιδιαίτερα χρήσιμο να συγκεντρώσουμε όλες τις παραγώγους δεύτερης τάξης της $f : \mathbb{R}^n \to \mathbb{R}$ μία $n \times n$ μήτρα.
Για συνάρτηση δύο μεταβλητών $y = f(x_1, x_2)$ υπάρχουν 4 μερικές παράγωγοι δεύτερης τάξης:

$$f_{11} \equiv \frac{\partial^2 f}{\partial x_1 \partial x_1}, \quad f_{12} \equiv \frac{\partial^2 f}{\partial x_1 \partial x_2}, \quad f_{21} \equiv \frac{\partial^2 f}{\partial x_2 \partial x_1}, \quad f_{22} \equiv \frac{\partial f}{\partial x_2 \partial x_2}$$

Αυτές σχηματίζουν τη μήτρα

$$\nabla_2 F \equiv \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}$$

Για συνάρτηση τριών μεταβλητών $y = f(x_1, x_2, x_3)$ υπάρχουν 9 παράγωγοι δεύτερης τάξης,

$$\nabla_2 F \equiv \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

και ούτω καθεξής...

Εσσιανή μήτρα

Γενικά, για $f : \mathbb{R}^n \to \mathbb{R}$,

$$\nabla_2 F \equiv \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}$$

ή

$$\nabla_2 F \equiv [f_{ij}], \quad i, j = 1, 2, ..., n.$$

με το στοιχείο στη σειρά i και στήλη j της $\nabla_2 F$ να αντιπροσωπεύει το αποτέλεσμα της παραγώγισης της f πρώτα ως προς x_i και κατόπιν ως προς x_j.

Η μήτρα $\nabla_2 f$ ονομάζεται **Εσσιανή μήτρα**. Συμβολίζεται επίσης ως $\frac{\partial^2 f}{\partial x \partial x}$ ή και με το γράμμα H.

Παράδειγμα

Να βρεθούν οι μερικές παράγωγοι πρώτης και δεύτερης τάξης της συνάρτησης $f(x_1, x_2) = x_1^2 x_2$ και να παρουσιαστούν υπό μορφή διανύσματος/μήτρας αντίστοιχα.

Οι μερικές παράγωγοι πρώτης τάξης είναι:

$$f_1 = 2x_1x_2,\ f_2 = x_1^2$$

Ενώ οι μερικές παράγωγοι δεύτερης τάξης είναι:

$$f_{11} = 2x_2,\ f_{12} = 2x_1,\ f_{21} = 2x_1,\ f_{22} = 0$$

Τοποθετώντας τις παραγώγους αυτές σε διάνυσμα και μήτρα αντίστοιχα έχουμε:

$$\nabla f = \begin{bmatrix} 2x_1x_2 \\ x_1^2 \end{bmatrix} \text{ και } \nabla_2 F = \begin{bmatrix} 2x_2 & 2x_1 \\ 2x_1 & 0 \end{bmatrix}$$

Σύνοψη: συμβολισμοί

Μερικές παράγωγοι:

$$\frac{\partial f}{\partial x_i} = f_i, \quad i = 1, ..., n.$$

Διάνυσμα κλίσης:

$$\frac{\partial f}{\partial x} = \nabla f, \quad n \times 1.$$

Δεύτερες μερικές παράγωγοι:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = f_{ij}, \quad i, j = 1, ..., n.$$

Εσσιανός πίνακας

$$\frac{\partial^2 f}{\partial x \partial x} = \nabla_2 f, \quad n \times n.$$

Σειρά Taylor πολυμεταβλητής συνάρτησης

Έστω $f : \mathbb{R}^n \to \mathbb{R}$ και $x^*, \epsilon \in \mathbb{R}^n$. Η σειρά Taylor της f στο σημείο x^* είναι

$$f(x^* + \epsilon) = f(x^*) + \underbrace{\frac{\partial f}{\partial x_1}\epsilon_1 + \frac{\partial f}{\partial x_2}\epsilon_2 + \cdots + \frac{\partial f}{\partial x_n}\epsilon_n}_{\text{όροι 1ης τάξης}} +$$

$$\underbrace{\begin{aligned} &+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x_1 \partial x_1}\epsilon_1^2 + \frac{\partial^2 f}{\partial x_1 \partial x_2}\epsilon_1\epsilon_2 + \cdots \frac{\partial^2 f}{\partial x_1 \partial x_n}\epsilon_1\epsilon_n\right. \\ &+ \frac{\partial^2 f}{\partial x_2 \partial x_1}\epsilon_2\epsilon_1 + \frac{\partial^2 f}{\partial x_2 \partial x_2}\epsilon_2^2 + \cdots \frac{\partial^2 f}{\partial x_2 \partial x_n}\epsilon_2\epsilon_n \\ &\left.+ \frac{\partial^2 f}{\partial x_n \partial x_1}\epsilon_n\epsilon_1 + \frac{\partial^2 f}{\partial x_n \partial x_2}\epsilon_n\epsilon_2 + \cdots \frac{\partial^2 f}{\partial x_n \partial x_n}\epsilon_n^2\right)\end{aligned}}_{\text{όροι 2ης τάξης}} + \cdots$$

$$= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\epsilon_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\epsilon_i\epsilon_j + \ldots$$

Σειρά Taylor πολυμεταβλητής συνάρτησης

ή

$$f(x^* + \epsilon) = f(x^*) + \underbrace{\left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \cdots \frac{\partial f}{\partial x_n}\right]}_{\nabla f^T} \cdot \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} +$$

$$+\frac{1}{2}\left[\epsilon_1, \epsilon_2, \cdots \epsilon_n\right] \underbrace{\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}}_{\nabla_2 f} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} + \cdots$$

Σε συμπαγή μορφή:

$$f(x^* + \epsilon) - f(x^*) = \left(\left.\frac{\partial f}{\partial x}\right|_{x^*} \right)^T \epsilon + \frac{1}{2}\epsilon^T \underbrace{\left(\left.\frac{\partial^2 f}{\partial x \partial x}\right|_{x^*} \right)}_{H} \epsilon + \cdots$$

Συνθήκες για βέλτιστο στο x^*

$$f(x^*+\epsilon)-f(x^*)=\left(\frac{\partial f}{\partial x}\bigg|_{x^*}\right)^T\epsilon+\frac{1}{2}\epsilon^T\underbrace{\left(\frac{\partial^2 f}{\partial x\partial x}\bigg|_{x^*}\right)}_{H}\epsilon+\cdots$$

Για βέλτιστο, είναι αναγκαίο

$$\partial f/\partial x=0\Rightarrow\partial f/\partial x_i=0\quad\forall i=1,2,...,n.$$

(αν οποιαδήποτε $\partial f/\partial x_i\neq 0$, μπορούμε να επιλέξουμε το πρόσημο του αντίστοιχου στοιχείου ϵ_i, ώστε να οδηγηθούμε σε άτοπο - όπως κάναμε για $f:\mathbb{R}\rightarrow\mathbb{R}$).
Στο στάσιμο x^*,

$$f(x^*+\epsilon)-f(x^*)=\left(\frac{\partial f}{\partial x}\bigg|_{x^*}\right)^T\epsilon+\frac{1}{2}\epsilon^T\underbrace{\left(\frac{\partial^2 f}{\partial x\partial x}\bigg|_{x^*}\right)}_{H}\epsilon+\cdots$$

το **πρόσημο της διαφοράς** $f(x^*+\epsilon)-f(x^*)$ καθορίζεται από τον όρο $\epsilon^T H\epsilon$

Συνθήκες για βέλτιστο στο x^*

Αν $\epsilon^T H\epsilon > 0, \forall \epsilon \neq 0$, τότε για $||\epsilon||$ αρκετά μικρό, θα είναι $f(x^* + \epsilon) - f(x^*) > 0$, δηλαδή έχουμε **ελάχιστο** στο x^*.

Αν $\epsilon^T H\epsilon < 0, \forall \epsilon \neq 0$, τότε για $||\epsilon||$ αρκετά μικρό, θα είναι $f(x^* + \epsilon) - f(x^*) < 0$, δηλαδή έχουμε **μέγιστο** στο x^*.

Αναγκαία συνθήκη: $\frac{\partial f}{\partial x} = 0$

Ίκανή συνθήκη: $\epsilon^T H\epsilon > 0$ για ελάχιστο, $\epsilon^T H\epsilon < 0$ για μέγιστο.

Συγκρίνετε με τις αντίστοιχες συνθήκες για $f : \mathbb{R} \to \mathbb{R}$!

Αναγκαίες και Ικανές συνθήκες για βέλτιστο πολυμεταβλητής συνάρτησης

Αναγκαία συνθήκη για να είναι το $\mathbf{x}^*$ τοπικό βέλτιστο της (δύο φορές παραγωγίσιμης) $y = f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ είναι:

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

ή

$$f_i(\mathbf{x}^*) = 0, \quad i = 1, \cdots, n.$$

Ικανές συνθήκες: $\forall \epsilon \in \mathbb{R}^n, \epsilon \neq 0, \epsilon^T H \epsilon > 0$ για τοπικό ελάχιστο ή < 0 για τοπικό μέγιστο. Δηλαδή η Εσσιανή μήτρα H είναι **θετικά ορισμένη** για τοπικό ελάχιστο ή **αρνητικά ορισμένη** για τοπικό μέγιστο.

Προσοχή: τα στοιχεία του H είναι γενικά συναρτήσεις των στοιχείων του x^*, οπότε $H = H(x^*)$. **Συνήθως όμως θα γράφουμε απλά** H και θα εννοούμε ότι τα στοιχεία του H υπολογίζονται στο στάσιμο σημείο (η άλλο σημείο ενδιαφέροντος).

Παράδειγμα

Να δείξετε ότι η $f(x) = (x_1 - 2)^2 + 4(x_2 + 1)^2$ έχει τοπικό ελάχιστο στο σημείο $x^* = [2, -1]^T$.

1. Βρίσκω τα σημεία που μηδενίζουν το διάνυσμα κλίσης

$$\frac{\partial f}{\partial x} = 0 \Rightarrow \left[\begin{array}{c} f_1 \\ f_2 \end{array}\right] = 0 \Rightarrow \left[\begin{array}{c} 2(x_1 - 2) \\ 8(x_2 + 1) \end{array}\right] = 0 \Rightarrow \begin{array}{rl} x_1 & = 2 \\ x_2 & = -1 \end{array}$$

Άρα το $x^* = [2, -1]^T$ είναι όντως στάσιμο.

2. Για να είναι τ.ελάχιστο θα πρέπει

$$\epsilon^T H \epsilon > 0 \quad \forall \epsilon \neq 0, \epsilon = [\epsilon_1, \epsilon_2]^T \in \mathbb{R}^2.$$

όπου

$$H = \frac{\partial^2 f}{\partial x \partial x} = \left[\begin{array}{cc} f_{11} & f_{12} \\ f_{21} & f_{22} \end{array}\right]$$

Παράδειγμα

Υπολογίζω:

$$f_{11} = \partial f_1/\partial x_1 = \frac{\partial}{\partial x_1}(2(x_1 - 2)) = 2, \quad f_{12} = \partial f_1/\partial x_2 = \frac{\partial}{\partial x_2}(2(x_1 - 2)) = 0$$

$$f_{21} = \partial f_2/\partial x_1 = \frac{\partial}{\partial x_1}(8(x_2 + 1)) = 0, \quad f_{22} = \partial f_2/\partial x_2 = \frac{\partial}{\partial x_2}(8(x_2 + 1)) = 8.$$

Συνεπώς, η ικανή συνθήκη για ελάχιστο είναι

$$[\epsilon_1 \epsilon_2] \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} > 0 \quad \forall \epsilon \neq 0.$$

$$\Leftrightarrow 2\epsilon_1^2 + 8\epsilon_2^2 > 0$$

το οποίο προφανώς ισχύει μιας και τα στοιχεία του ϵ δε μπορούν να είναι και τα δύο μηδενικά.