

# Λογισμος II, Μεγιστα ελαχιστα και σαγματικά

A. N. Γιαννακόπουλος  
Τμήμα Στατιστικής

Ο.Π.Α

Εαρινό Εξάμηνο 2024

# Ακρότατα συναρτήσεων πολλών μεταβλητών

## Ορισμός

Έστω  $f : U \subset \mathbb{R}^d \rightarrow \mathbb{R}$  μια συνάρτηση.

Το σημείο  $x_* \in U$  θα ονομάζεται **τοπικό ελάχιστο** της  $f$  αν

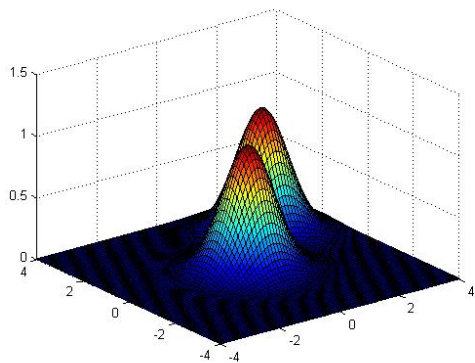
$$f(x_*) \leq f(x), \quad \forall x \in V,$$

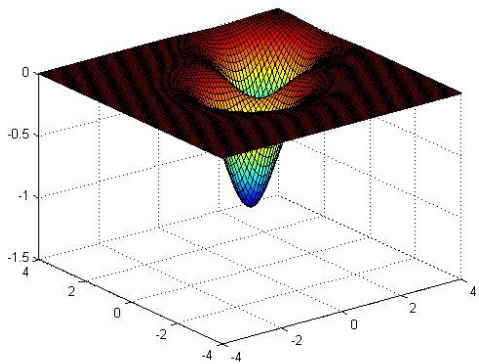
όπου  $V \subset U$  είναι μια γειτονιά που περιέχει το σημείο  $x_*$ .

Το σημείο  $x^* \in U$  θα ονομάζεται **τοπικό μέγιστο** της  $f$  αν

$$f(x^*) \geq f(x), \quad \forall x \in V,$$

όπου  $V \subset U$  είναι μια γειτονιά που περιέχει το σημείο  $x^*$ .





## Ποιά είναι τα κριτήρια για ένα τοπικό μέγιστο ή ελάχιστο;

Ας υποθέσουμε ότι  $x_*$  είναι ένα τοπικό ελάχιστο της  $f$  και ότι η  $f$  είναι  $C^1$  σε μια γειτονιά του σημείου  $x_*$ .

Ας θεωρήσουμε την συνάρτηση  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  που ορίζεται ως

$$\phi(t) = f(x_* + th), \text{ για οποιοδήποτε } h \in \mathbb{R}^d.$$

Αν  $x_*$  τοπικό ελάχιστο της  $f$  τότε το  $t = 0$  είναι τοπικό ελάχιστο της  $\phi$ .

Όμως τότε ισχύει

$$0 = \left. \frac{d\phi}{dt}(t) \right|_{t=0} = \langle Df(x_*), h \rangle$$

και

$$Df(x_*) = 0.$$

Με ακριβώς τον ίδιο τρόπο μπορούμε να δούμε ότι αν  $x^*$  είναι ένα τοπικό μέγιστο της  $f$  και η  $f$  είναι  $C^1$  σε μία γειτονιά του σημείου  $x^*$  τότε

$$Df(x^*) = 0.$$

## Ορισμός

Έστω  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  μια  $C^1$  συνάρτηση.

Τα σημεία  $x \in \mathbb{R}^d$  για τα οποία ισχύει  $Df(x) = 0$  ονομάζονται **κρίσιμα σημεία** της συνάρτησης  $f$ .

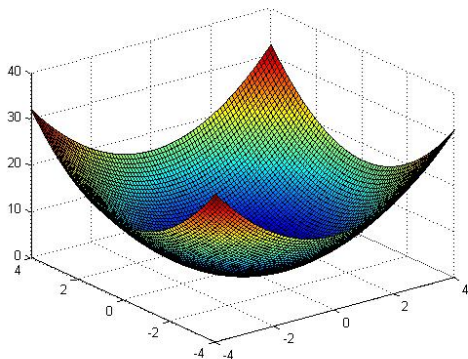
## Πρόταση

Τα τοπικά μέγιστα ή ελάχιστα μιας  $C^1$  συνάρτησης  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  είναι κρίσιμα σημεία της.

Τα κρίσιμα σημεία όμως δεν είναι απαραίτητα μέγιστα ή ελάχιστα

## Παράδειγμα

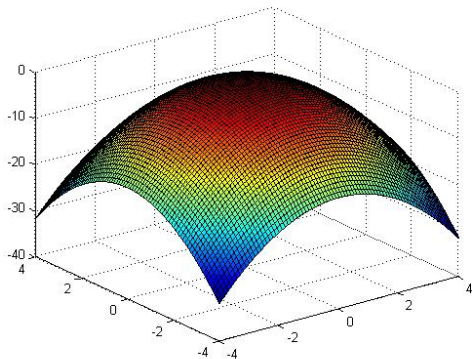
Η  $f(x_{1,2}) = x_1^2 + x_2^2$  έχει κρισιμο σημείο το  $(0, 0)$  το οποίο είναι ελάχιστο.





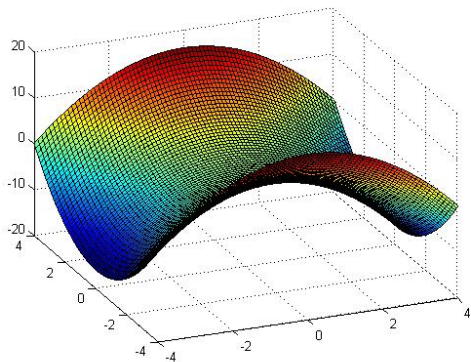
## Παράδειγμα

Η  $f(x_{1,2}) = -x_1^2 - x_2^2$  έχει κρισιμο σημείο το  $(0, 0)$  το οποίο είναι μέγιστο.



## Παράδειγμα

Η  $f(x_{1,2}) = x_1^2 - x_2^2$  έχει κρισιμο σημείο το  $(0, 0)$  το οποίο δεν είναι ούτε μέγιστο ούτε ελάχιστο.



## Ορισμός

Ένα κρίσιμο σημείο το οποίο δεν είναι ούτε τοπικό μέγιστο ούτε τοπικό ελάχιστο ονομάζεται **σαγματικό σημείο** (saddle point).

Σε ένα σαγματικό σημείο υπάρχουν διευθύνσεις κατά τις οποίες η συναρτηση θα **αυξάνεται** και διευθύνσεις κατά τις οποίες η συναρτηση θα **μειώνεται**.

Πώς θα ξέρουμε αν ένα κρίσιμο σημείο είναι τοπικό μέγιστο, ελάχιστο ή σαγματικό;

Ας υποθέσουμε ότι  $x_0 \in \mathbb{R}^d$  είναι ένα κρίσιμο σημείο μιας  $C^2$  συνάρτησης.

Χρησιμοποιούμε το ανάπτυγμα Taylor για την  $f$  γύρω από το σημείο  $x_0$ ,

$$f(x_0 + tz) = f(x_0) + t\langle Df(x_0), z \rangle + \frac{1}{2}t^2\langle z, (D^2f)(x_0)z \rangle + O(\|z\|^3)$$

που επειδή  $x_0$  είναι κρίσιμο σημείο γίνεται

$$f(x_0 + tz) = f(x_0) + \frac{1}{2}t^2\langle z, (D^2f)(x_0)z \rangle + O(\|z\|^3)$$

- 1 Αν ο πίνακας του Hess  $D^2f(x_0)$  έχει την ιδιότητα

$$\langle z, (D^2f)(x_0)z \rangle \geq 0, \quad \forall z \in \mathbb{R}^d,$$

τότε

$$f(x_0 + tz) \geq f(x_0), \quad \forall z \in \mathbb{R}^d,$$

και το  $x_0$  είναι **τοπικό ελάχιστο**.

- 2 Αν ο πίνακας του Hess  $D^2f(x_0)$  έχει την ιδιότητα

$$\langle z, (D^2f)(x_0)z \rangle \leq 0, \quad \forall z \in \mathbb{R}^d,$$

τότε

$$f(x_0 + tz) \leq f(x_0), \quad \forall z \in \mathbb{R}^d,$$

και το  $x_0$  είναι **τοπικό μέγιστο**.

- 3 Αν δεν ισχύει ούτε το ένα ούτε το άλλο το  $x_0$  είναι **σαγματικό**.

## Ορισμός

Ένας πίνακας  $A \in \mathbb{R}^{d \times d}$  ονομάζεται **θετικά ημιορισμένος** αν

$$\langle z, Az \rangle \geq 0, \quad \forall z \in \mathbb{R}^d.$$

Αν η ανισότητα είναι αυστηρή τότε ονομάζεται **θετικά ορισμένος**.

Ένας πίνακας  $A \in \mathbb{R}^{d \times d}$  ονομάζεται **αρνητικά ημιορισμένος** αν ο  $-A$  είναι θετικά ημιορισμένος.

Απο το θεώρημα του Euler γνωρίζουμε επίσης ότι ο πίνακας  $D^2f(x_0)$  είναι συμμετρικός.

### Πρόταση

*Ένας συμμετρικός πίνακας  $A$  είναι θετικά ημιορισμένος αν και μόνο αν οι ιδιοτιμές του είναι μη αρνητικές.*

Είναι επίσης χρησιμο να γνωρίζουμε ότι αν ο  $A$  είναι συμμετρικός και θετικά ημιορισμένος τότε υπάρχει πίνακας  $B$  τέτοιος ώστε

$$A = B^T B$$

Αν ο  $A$  είναι συμμετρικός και θετικά ορισμένος τότε ο  $B$  είναι και αντιστρέψιμος.

Μπορούμε λοιπόν να συνοψίσουμε την παραπάνω ανάλυση μας στην εξής προταση.

### Πρόταση

*Ένα κρίσιμο σημείο είναι τοπικό ελάχιστο αν ο πίνακας  $D^2f(x_0)$  είναι θετικά ημιορισμένος.*

*Ένα κρίσιμο σημείο είναι τοπικό μέγιστο αν ο πίνακας  $D^2f(x_0)$  είναι αρνητικά ημιορισμένος.*

*Ένα κρίσιμο σημείο είναι σαγματικό αν ο πίνακας  $D^2f(x_0)$  δεν είναι ούτε θετικά ούτε αρνητικά ημιορισμένος.*



## Παράδειγμα

Βρείτε ένα κριτήριο για να χαρακτηρίσετε αν ένα κρίσιμο σημείο μιας συνάρτησης  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  είναι

- 1 Τοπικό μέγιστο
- 2 Τοπικό ελάχιστο
- 3 Σαγματικό.

## Παράδειγμα

Βρείτε τα κρίσιμα σημεία της  $f(x_1, x_2) = 12x_1^2 + x_2^3 - 12x_1x_2$  και χαρακτηρίστε τα.

## Παράδειγμα

Έχετε ένα στατιστικό δείγμα  $X_1, \dots, X_n$  το οποίο πιστεύετε ότι ερχεται από μια τυχαία μεταβλητή  $X \sim N(\mu, \sigma^2)$  αλλά δεν γνωρίζετε τις ακριβείς τιμές των παραμετρών  $\mu$  και  $\sigma$ .

Υπολογίστε τις χρησιμοποιώντας την μέθοδο μέγιστης πιθανοφάνειας.