Λογισμος ΙΙ, Μεγιστα ελάχιστα και σαγματικά

Α. Ν. Γιαννακόπουλος
Τμήμα Στατιστικής

Ο.Π.Α

Εαρινό Εξάμηνο 2018

Ακρότατα συναρτήσεων πολλών μεταβλητών

Ορισμός

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

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

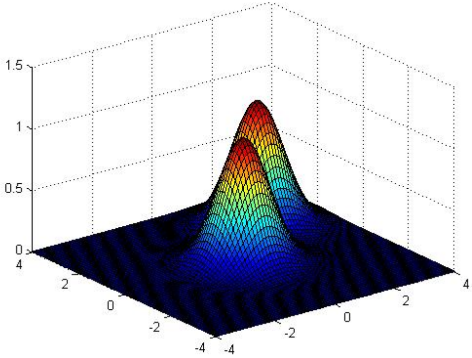
$$f(x_*) \leq f(x), \;\; \forall\, x \in V,$$

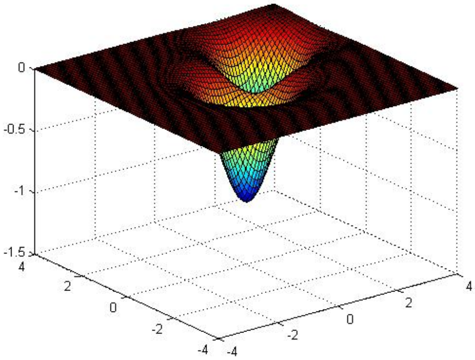
όπου $V \subset U$ είναι μια γειτονιά που περιέχει το σημείο x_*.

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

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

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


1.5
1
0.5
0
4
2
0
-2
-4
-4
-2
0
2
4


0
-0.5
-1
-1.5
4
2
0
-2
-4
-4
-2
0
2
4

Ποιά είναι τα κριτήρια για ένα τοπκό μέγιστο ή ελάχιστο;

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

Ας θεωρήσουμε την συνάρτηση $\phi : \mathbb{R} \to \mathbb{R}$ 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 ϕ.

Όμως τότε ισχύει

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

και

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

Με ακριβώς τον ίδιο τρόπο μπορούμε να δούμε οτι 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 $f : \mathbb{R}^d \to \mathbb{R}$ *μια* C^1 *συνάρτηση.*

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

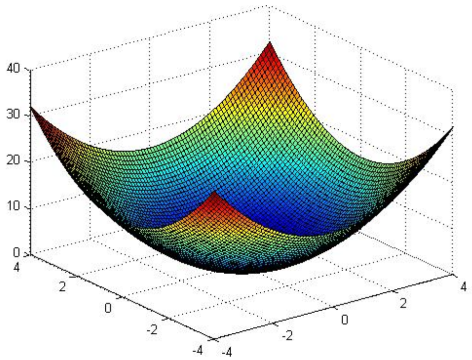
Πρόταση

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

Τα κρίσιμα σημεία όμως δεν είναι απαραίτητα μέγιστα ή ελάχιστα

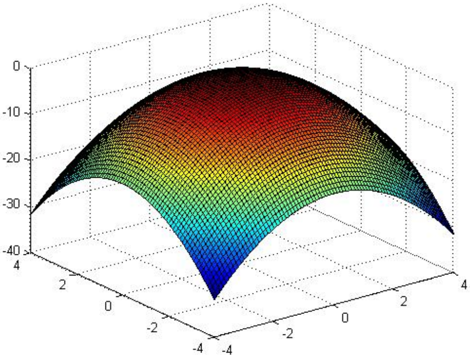
Παράδειγμα

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



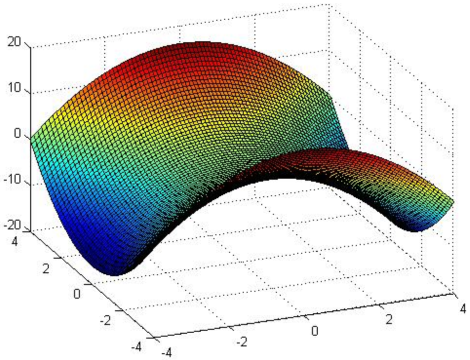
Παράδειγμα

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



Παράδειγμα

Η $f(x_{1,2}) = x_1^2 - x_2^2$ *έχει κρισιμο σημείο το* $(0,0)$ *το οποίο δεν ειναι 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 απο το σημείο x_0,

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

που επειδή x_0 είναι κρίσιμο σημείο γίνεται

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

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

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

τότε

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

και το x_0 είναι τοπικό ελάχιστο.

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

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

τότε

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

και το x_0 είναι τοπικό μέγιστο.

3. Αν δεν ισχύει ούτε το ενα ούτε το άλλο το x_0 είναι σαγματικό.

Ορισμός

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

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

Αν η ανισότητα είναι αυστηρή τότε ονομάζεται θετικα ορισμένος.

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

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

Πρόταση

Ένας συμμετρικός πίνακας A είναι θετικα ημιορισμένος αν και μόνο αν οι ιδιοτιμές του είναι μη αρνητικές.

Είναι επίσης χρησιμο να γνωρίζουμε οτι αν ο A είναι συμμετρικός και θετικα 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 $D^2 f(x_0)$ είναι θετικά ημιορισμένος.

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

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

Παράδειγμα

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

1. *Τοπικό μέγιστο*
2. *Τοπικό ελάχιστο*
3. *Σαγματικο.*

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 $X_1, \cdots, X_n$ *το οποίο πιστεύετε ότι ερχεται απο μια τυχαία μεταβλητή* $X \sim N(\mu, \sigma^2)$ *αλλά δεν γνωρίζετε τις ακριβείς τιμές των παραμετρων* μ *και* σ.

Υπολογίστε τις χρησιμοποιώντας την μέθοδο μέγιστης πιθανοφάνειας.