Ακρότατα πραγματικών συναρτήσεων

13.1 Ορισμός. Έστω $U \subseteq R^n$, $x_0 \in U$ και $f : U \subseteq R^n \to R$ συνάρτηση τότε:

1)Το x_0 λέγεται <u>τοπικό ελάχιστο</u> της f αν υπάρχει περιοχή V του x_0 ώστε $f(x_0) \leq f(x)$ για κάθε $x \in V \bigcap U$.

Το x_0 λέγεται <u>τοπικό μέγιστο</u> της f αν υπάρχει περιοχή V του x_0 ώστε $f(x_0) \geq f(x)$ για κάθε $x \in V \bigcap U$.

Το x_0 λέγεται <u>τοπικό ή σχετικό ακρότατο</u> της f αν είναι είτε τοπικό ελάχιστο είτε τοπικό μέγιστο της f.

Υποθέτουμε περαιτέρω ότι U ανοικτό υποσύνολο στον R^n και f διαφορίσιμη συνάρτηση, το σημείο x_0 λέγεται <u>κρίσιμο</u> σημείο της f αν, $\nabla f(x_0) = 0 \Leftrightarrow \frac{\partial f}{\partial x_i}(x_0) = 0$ για κάθε $i = 1, 2, ..., n$. Ένα κρίσιμο σημείο που <u>δεν</u> είναι τοπικό ακρότατο λέγεται <u>σαγματικό</u> σημείο για την f.

Κατ' αναλογία με τον Λογισμό συναρτήσεων μιας μεταβλητής ισχύει η ακόλουθη γενίκευση του θεωρήματος Fermat:

13.2 Θεώρημα Αν $U \subseteq R^n$ ανοικτό, η $f : U \subseteq R^n \to R$ είναι διαφορίσιμη στο U και η f παρουσιάζει τοπικό ακρότατο στο $x_0 \in U$ τότε $\nabla f(x_0) = 0$, δηλαδή το x_0 είναι κρίσιμο σημείο για την f.

Απόδειξη: Υποθέτουμε χωρίς περιορισμό της γενικότητας ότι η f παρουσιάζει τοπικό μέγιστο στο $x_0 \in U$. Αν $h \in R^n$ με $h \neq 0$ περιορίζουμε την f στην ευθεία $\ell(t) = x_0 + th$, δηλαδή θεωρούμε την $g(t) = f(x_0 + th)$ η οποία ορίζεται σε κατάλληλο διάστημα $(-\delta, \delta)$, $\delta > 0$. Παρατηρούμε ότι, $g(0) = f(x_0) \geq f(x_0 + th) = g(t)$ για κάθε $t \in (-\delta_1, \delta_1)$ για κάποιο $0 < \delta_1 \leq \delta$, δηλαδή η g παρουσιάζει τοπικό μέγιστο στο x_0. Επειδή η g ως σύνθεση διαφορίσιμων (της $\ell(t)$ και $f(x)$) είναι διαφορίσιμη στο $(-\delta, \delta)$ έπεται από το θεώρημα Fermat για διαφορίσιμες συναρτήσεις μιας μεταβλητής ότι $g'(0) = 0$. Όμως από τον κανόνα της αλυσίδας, $g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \cdot h_i = \nabla f(x_0) \cdot h$, $h = (h_1, ..., h_n)$.

Το ίδιο συμπέρασμα έπεται από το γεγονός ότι, $g'(0)$ ισούται με την παράγωγο της f στο x_0 στην κατεύθυνση h, δηλαδή $g'(0) = \nabla f(x_0) \cdot h = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \cdot h_i$, όπου $h = (h_1, ..., h_n)$ με $\|h\| = 1$.

Έπεται από τα παραπάνω ότι: $\nabla f(x_0) \cdot h = 0$ για κάθε $h \in R^n \Leftrightarrow \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \cdot h_i = 0$ για κάθε $h = (h_1, ..., h_n) \in R^n$. Έτσι συμπεραίνουμε ότι, $\nabla f(x_0) = 0 \Leftrightarrow \frac{\partial f}{\partial x_i}(x_0) = 0$ για κάθε $i = 1, 2, ..., n$.

Σημείωση. Το προηγούμενο αποτέλεσμα μας λέει ότι αν αναζητούμε τα τοπικά ακρότατα μιας διαφορίσιμης συνάρτησης $f : U \subseteq R^n \to R$, τότε θα πρέπει να ψάξουμε στα κρίσιμα σημεία. Εφόσον:

x_0 **τοπικό ακρότατο της** $f \Rightarrow x_0$ **κρίσιμο σημείο της** f**.**

Παραδείγματα: 1) Να βρεθούν τα κρίσιμα σημεία της $f(x,y) = x^2 y + xy^2$

Έχουμε ότι, $\frac{\partial f}{\partial x} = 2xy + y^2, \frac{\partial f}{\partial y} = 2xy + x^2$.

Εξισώνοντας τις μερικές παραγώγους με μηδέν παίρνουμε το σύστημα :

$\begin{cases} 2xy + y^2 = 0 \\ 2xy + x^2 = 0 \end{cases}$ αφαιρώντας, παίρνουμε $x^2 = y^2 \Leftrightarrow x = \pm y$.

Αντικαθιστώντας το $x = y$ στην πρώτη εξίσωση, βρίσκουμε ότι $2y^2 + y^2 = 3y^2 = 0$ οπότε $y = 0$ και άρα $x = 0$. Αν $x = -y$, τότε $-2y^2 + y^2 = -y^2 = 0$, δηλαδή $y = 0$ και επομένως $x = 0$.

Επομένως το <u>μοναδικό κρίσιμο σημείο</u> της f είναι το $(0,0)$.

Επειδή $f(x,x) = 2x^3$ το οποίο παίρνει και θετικές και αρνητικές για x κοντά στο 0, το $(0,0)$ <u>δεν</u> είναι τοπικό ακρότατο της f, επομένως είναι <u>σαγματικό</u> σημείο της f.

2)Έστω $f(x,y) = x^2 + y^2$.

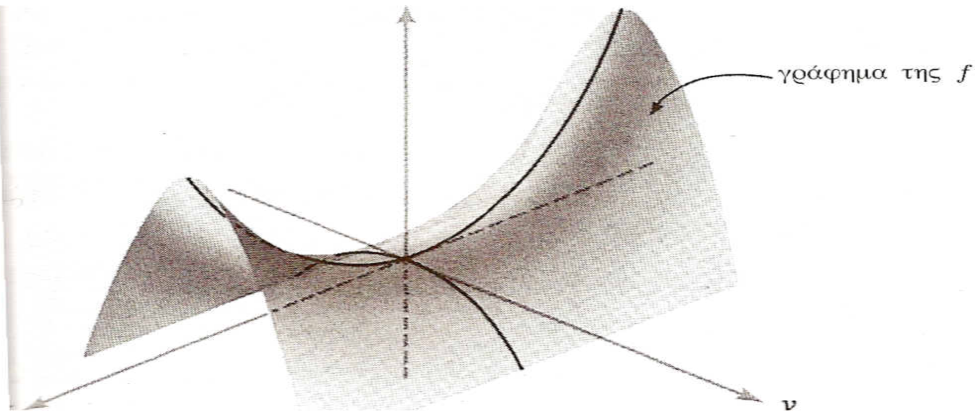
Παρατηρούμε ότι $\frac{\partial f}{\partial x} = 2x$ και $\frac{\partial f}{\partial y} = 2y$.

Επομένως το σύστημα: $\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \Leftrightarrow \begin{cases} 2x = 0 \\ 2y = 0 \end{cases} \Leftrightarrow x = y = 0$.

Έτσι το μοναδικό <u>κρίσιμο</u> σημείο της f είναι το $(0,0)$. Επειδή $f(0,0) = 0 \le f(x,y) = x^2 + y^2$ για κάθε $(x,y) \in R^2$ το σημείο $(0,0)$ είναι τοπικό (μάλιστα ολικο) ελάχιστο για την f.

3) Έστω $f(x,y) = x^2 - y^2$

Παρατηρούμε ότι $\frac{\partial f}{\partial x} = 2x$, $\frac{\partial f}{\partial y} = -2y$. Συνεπώς το σύστημα $\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases}$, έχει ως μόνη λύση το $(0,0)$. Συμπεραίνουμε ότι το μόνο <u>κρίσιμο</u> σημείο της f είναι το $(0,0)$. Εξετάζοντας απ' ευθείας τις τιμές της f σε σημείο κοντά στην αρχή των αξόνων, βλέπουμε ότι $f(x,0) = x^2 \ge f(0,0) = 0$ και $f(0,y) = -y^2 \le f(0,0) = 0$.



Έπεται ότι το $(0,0)$ δεν είναι τοπικό ακρότατο για την f, επομένως είναι σαγματικό σημείο για την f.

4) Έστω $f(x,y)=x\cdot y$. Τότε έχουμε $\frac{\partial f}{\partial x}=y$, $\frac{\partial f}{\partial y}=x$ και το σύστημα $\begin{cases}\frac{\partial f}{\partial x}=0\\ \frac{\partial f}{\partial y}=0\end{cases}$,έχει ως μόνη λύση το σημείο $(0,0)$. Το σημείο $(0,0)$ είναι το μόνο κρίσιμο σημείο της f και είναι σαγματικό σημείο αφού για x,y ομόσημα έχουμε $f(x,y)=xy>0$ και για x,y ετερόσημα έχουμε, $f(x,y)=xy<0$.

Παρατηρούμε ότι, $f(x,y)=\frac{1}{4}\left[(x+y)^2-(x-y)^2\right]$, άρα αν θέσουμε $u=x+y$ και $v=x-y$ η συνάρτηση γίνεται $f\left(x(u,v),y(u,v)\right)=\frac{1}{4}\left(u^2-v^2\right)$, που είναι η περίπτωση του παραδείγματος (3).

5)Θεωρούμε την συνάρτηση, $f(x,y)=\lambda x^2+\mu y^2+ax^3+\beta x^2y+\gamma xy^2+\delta y^3$, όπου $\lambda>0,\mu>0$ και a,β,γ,δ τυχόντες πραγματικοί αριθμοί.

Παρατηρούμε ότι, αν θέσομε $\varphi(x,y)=ax^3+\beta x^2y+\gamma xy^2+\delta y^3$ τότε,

$$\lim_{(x,y)\to(0,0)}\frac{\varphi(x,y)}{\lambda x^2+\mu y^2}=0.$$

Πράγματι, αρκεί να παρατηρήσουμε ότι, $\frac{|x|^3}{\lambda x^2+\mu y^2}=\frac{|x|}{\lambda}\frac{\lambda x^2}{\lambda x^2+\mu y^2}\le\frac{|x|}{\lambda}\xrightarrow[x\to0]{}0$,

$\frac{|xy^2|}{\lambda x^2+\mu y^2}=\frac{|x|}{\lambda}\frac{\lambda y^2}{\lambda x^2+\mu y^2}\le\frac{|x|}{\lambda}\xrightarrow[x\to0]{}0$ και παρομοίως για τις συναρτήσεις, $\frac{y^3}{\lambda x^2+\mu y^2}\xrightarrow[y\to0]{}0$ και $\frac{x^2y}{\lambda x^2+\mu y^2}\xrightarrow[y\to0]{}0$.

Έπεται ότι, υπάρχει $r>0:\left|\frac{\varphi(x,y)}{\lambda x^2+\mu y^2}\right|<\frac{1}{2}$, όταν $0<\sqrt{x^2+y^2}<r$.

Συνεπώς, $f(x,y)\ge\left(\lambda x^2+\mu y^2\right)-|\varphi(x,y)|\ge\frac{1}{2}\left(\lambda x^2+\mu y^2\right)\ge f(0,0)=0$ για $\sqrt{x^2+y^2}<r$, που σημαίνει ότι η f έχει τοπικό ελάχιστο στο $(0,0)$.

Παρατηρούμε ότι το δευτεροβάθμιο μέρος της συνάρτησης κυριαρχεί και καθορίζει την συμπεριφορά της συνάρτησης πλησίον του σημείου $(0,0)$.

Τετραγωνικές μορφές. Μια πολυωνυμική συνάρτηση $f: R^n \to R$ ομογενής και 2ου βαθμού ονομάζεται τετραγωνική μορφή ή και τετραγωνική συνάρτηση (όλα τα μονώνυμα της f είναι της μορφής $x_i x_j, 1 \le i, j \le n$).

Παραδείγματα 1) $f(x,y) = ax^2 + \beta xy + \gamma y^2, (x,y) \in R^2, a, \beta, \gamma \in R$ είναι τετραγωνική μορφή δύο μεταβλητών (γενική μορφή).

2) Η $f(x,y,z) = ax^2 + \beta y^2 + \gamma z^2 + \delta xy + \varepsilon xz + \zeta yz$, $(x,y,z) \in R^3$, $a, \beta, \gamma, \delta, \varepsilon, \zeta \in R$ είναι τετραγωνική μορφή 3 μεταβλητών (γενική μορφή).

3) Οι $f(x,y) = x^2 + y^2 + 2$ και $g(x,y) = x^2 + y^2 + x + 2xy$ δεν είναι τετραγωνικές μορφές αφού στην f εμφανίζεται σταθερός όρος (μονώνυμο μηδενικού βαθμού) και στην g το πρωτοβάθμιο μονώνυμο x.

4)Η $f(x,y) = 2x^2 + y^2$ και $f(x,y,z) = x^2 + y^2 - xz$ είναι τετραγωνικές μορφές 2 και 3 μεταβλητών αντίστοιχα.

Παρατηρήσεις: (1) Μια τετραγωνική μορφή $f: R^n \to R$ γράφεται στην γενική της μορφή ως, $f(x_1,...,x_n) = \sum_{i,j=1}^{n} a_{ij} x_i x_j \left(= \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \le i \ne j \le n} a_{ij} x_i x_j \right) (1)$

Με την βοήθεια του $n \times n$ πίνακα $\mathrm{A} = (a_{ij})$ μπορούμε να γράψουμε την (1) ως εξής,

$$f(x_1,...,x_n) = (x_1,...,x_n)\mathrm{A}\begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \quad (2)$$

Μπορούμε ακόμη να υποθέσουμε ότι ο $\mathrm{A} = (a_{ij})$ είναι συμμετρικός (δηλαδή $a_{ij} = a_{ji}$ για κάθε i, j) αφού η f δεν μεταβάλλεται αν αντικαταστήσουμε το a_{ij} με το $b_{ij} = \dfrac{a_{ij} + a_{ji}}{2}$, αφού $x_i x_j = x_j x_i$. (Άρα, $f(x_1,...,x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + 2 \sum_{1 \le i < j \le n} a_{ij} x_i x_j$.)

2)Ο τετραγωνικός χαρακτήρας της f φαίνεται και από την

$$f(\lambda x_1,...,\lambda x_n) = \lambda^2 f(x_1,...,x_n), \quad \lambda \in R.$$

Μας ενδιαφέρουν κυρίως οι τετραγωνικές μορφές δύο μεταβλητών και δευτερευόντως οι τετραγωνικές μορφές τριών ή περισσοτέρων μεταβλητών.

Παράδειγμα: Έστω $f(x,y) = ax^2 + \beta xy + \gamma y^2$ η γενική τετραγωνική μορφή δύο μεταβλητών. Θέτοντας $\beta' = \dfrac{\beta}{2}$, μπορούμε να γράψουμε:

$f(x,y)=ax^2+2\beta' xy+\gamma y^2$. Η f δίνεται από τον συμμετρικό πίνακα $\begin{pmatrix} a & \beta' \\ \beta' & \gamma \end{pmatrix}=\begin{pmatrix} a & \dfrac{\beta}{2} \\ \dfrac{\beta}{2} & \gamma \end{pmatrix}$, αφού $f(x,y)=ax^2+2\beta' xy+\gamma y^2=(x,y)\begin{pmatrix} a & \beta' \\ \beta' & \gamma \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$.

Έτσι μια τετραγωνική μορφή δύο μεταβλητών θα γράφεται:

$$f(x,y)=ax^2+2\beta xy+\gamma y^2 .$$

Σημειώνουμε ότι αν ο πίνακας της f είναι αντιστρέψιμος, δηλαδή αν $\alpha\gamma-\beta^2\neq 0$, τότε το $(0,0)$ είναι το μοναδικό κρίσιμο σημείο της f (γιατί;)

13.3 Ορισμός Μια τετραγωνική μορφή $f:R^n\to R$ λέγεται:

(α) Θετικά ορισμένη, αν $f(x)>0$ για κάθε $x\in R^n$ με $x\neq 0$ και

(β) Αρνητικά ορισμένη, αν $f(x)<0$ για κάθε $x\in R^n$ με $x\neq 0$.

(γ) Η f λέγεται θετικά ημιορισμένη αν $f(x)\geq 0$ για κάθε $x\in R^n$.

Ανάλογα ορίζεται και η αρνητικά ημιορισμένη τετραγωνική μορφή.

(δ) Η f λέγεται αόριστη αν υπάρχουν $x,y\in R^n$ ώστε $f(x)<0<f(y)$

Προφανώς, $f(0)=0$ για κάθε τετραγωνική μορφή, αφού δεν περιέχει σταθερό όρο (διάφορο του μηδενός).

Παρατηρούμε ότι: 1) Αν η f είναι θετικά (αντιστοίχως αρνητικά) ορισμένη τότε το $0\in R^n$ είναι ολικό ελάχιστο (αντιστοίχως ολικό μέγιστο) για την f. Ανάλογες παρατηρήσεις ισχύουν και για τις ημιορισμένες τετραγωνικές μορφές.

2) Αν η f είναι θετικά ή αρνητικά ορισμένη τότε ο συμμετρικός πίνακας Α που ορίζει την f (υπενθυμίζουμε ότι, $f(x_1,...,x_n)=(x_1,...,x_n)\mathrm{A}\begin{pmatrix} x_1 \\ . \\ . \\ x_n \end{pmatrix}$ για κάποιο συμμετρικό πίνακα Α) είναι αντιστρέψιμος (γιατί;).

Η συμπεριφορά μιας τετραγωνικής μορφής δύο μεταβλητών κοντά στο σημείο $(0,0)$ περιγράφεται στην επόμενη

13.4 Πρόταση. Έστω $f(x,y)=ax^2+2\beta xy+\gamma y^2$ τετραγωνική μορφή δύο μεταβλητών. Τότε, ισχύουν τα ακόλουθα:

(ι) Αν $a>0$ και $a\gamma-\beta^2>0$ τότε η f είναι θετικά ορισμένη άρα το $(0,0)$ είναι ολικό ελάχιστο για την f.

(ιι) Αν $a<0$ και $a\gamma-\beta^2>0$ τότε η f είναι αρνητικά ορισμένη άρα το $(0,0)$ είναι ολικό μέγιστο για την f.

(ιιι) Αν $a\gamma-\beta^2<0$, το σημείο $(0,0)$ είναι σαγματικό σημείο της f.

(ιv) Αν $a\gamma-\beta^2=0$ τότε η f στο $(0,0)$ έχει, (α) μέγιστο ή ελάχιστο ανάλογα αν $a<0$ ή $a>0$ και (β) αν $a=0$, μέγιστο ή ελάχιστο ανάλογα αν $\gamma<0$ ή $\gamma>0$.

Απόδειξη: Αν $a=\gamma=0$ τότε $f(x,y)=2\beta xy$ με $\beta\neq 0$ και εύκολα βλέπουμε ότι το $(0,0)$ είναι σαγματικό σημείο της f (δες και το παράδειγμα (4)). Παρατηρούμε ακόμη ότι τότε $a\gamma-\beta^2=-\beta^2<0$ και η περίπτωση αυτή εντάσσεται στην (ιιι). Υποθέτουμε λοιπόν χωρίς περιορισμό της γενικότητας ότι $a\neq 0$. Αντιμετωπίζοντας την f ως συνάρτηση του x ή του y και ενθυμούμενοι ότι ένα τριώνυμο παραγοντοποιείται «συμπληρώνοντας τα τετράγωνα» έχουμε ,

$$f(x,y)=a\left(x^2+2\frac{\beta}{a}xy+\frac{\gamma}{a}y^2\right)=a\left[x^2+2x\cdot\frac{\beta y}{a}+\left(\frac{\beta y}{a}\right)^2+\frac{\gamma}{a}y^2-\left(\frac{\beta y}{a}\right)^2\right]=$$

$=a\left[\left(x+\frac{\beta}{a}y\right)^2+\frac{a\gamma-\beta^2}{a^2}\cdot y^2\right]$. (Αν $\gamma\neq 0$, τότε εναλλάσσοντας το a με το γ και το x με το y έχομε: $f(x,y)=\gamma\left[\left(y+\frac{\beta}{\gamma}x\right)^2+\frac{a\gamma-\beta^2}{\gamma^2}\cdot x^2\right]$)

Έπεται ότι: (ι) Αν $a>0$ και $a\gamma-\beta^2>0$, στο σημείο $(0,0)$ η f έχει ελάχιστη τιμή.

(ιι) Αν $a<0$ και $a\gamma-\beta^2>0$ στο σημείο $(0,0)$ η f παίρνει μέγιστη τιμή.

(ιιι) Αν $a\gamma-\beta^2<0$ το σημείο $(0,0)$ είναι σαγματικό σημείο της f.

Πράγματι, αν $y>0$ τότε οι ρίζες της εξίσωσης $f(x,y)=0$ ως προς x, είναι οι $\rho_1(y)=y\rho_1$ και $\rho_2(y)=y\rho_2$, όπου, $\rho_i=\dfrac{-\beta\pm\sqrt{\beta^2-a\gamma}}{a}, i=1,2$.

Έπεται εύκολα ότι για $y>0$ με y μικρό αριθμό, υπάρχουν $x_1,x_2\in R$ με $|x_1|,|x_2|$ μικρούς αριθμούς ώστε $f(x_1,y)>0$ και $f(x_2,y)<0$ (ο x_1 θα είναι εκτός του διαστήματος των ριζών $\rho_1(y)$ και $\rho_2(y)$ και ο x_2 θα επιλεγεί εντός αυτού του διαστήματος). Συνεπώς το $(0,0)$ είναι σαγματικό σημείο της f.

Σημειώνουμε ότι το τελευταίο συμπέρασμα ισχύει ακόμη και αν $a=0$ (με την προϋπόθεση ότι $a\gamma-\beta^2<0$). Πράγματι αν $\gamma\neq 0$, εναλλάσσοντας τους ρόλους των a και γ βρίσκουμε $f(x,y)=\gamma\left[\left(y+\frac{\beta}{\gamma}x\right)^2-\frac{\beta^2x^2}{\gamma^2}\right]$.

Από όπου εύκολα συμπεραίνουμε ότι το $(0,0)$ είναι πράγματι σαγματικό σημείο της f.

(ιν) Αν $a\gamma-\beta^2=0$ και $a\neq 0$ τότε $f(x,y)=a\left(x+\frac{\beta y}{a}\right)^2$.

Από όπου έπεται αμέσως ο πρώτος ισχυρισμός.

Αν $a\gamma-\beta^2=0$ και $a=0$, τότε $\beta=0$ και άρα $f(x,y)=\gamma y^2$, από όπου έπεται ο δεύτερος ισχυρισμός.

Παρατήρηση. Έστω $f\left(x_1,...,x_n\right)=\sum_{1\le i,j\le n} a_{ij}x_i\cdot x_j$, μια τετραγωνική μορφή που ορίζεται από τον συμμετρικό πίνακα $\mathrm{A}=\left(a_{ij}\right)_{1\le i,j\le n}$. Στο μάθημα της Γραμμικής Άλγεβρας αποδεικνύεται ότι:

1)Η f είναι θετικά ορισμένη ($f\left(x\right)>0$ για κάθε $x\in R^n$ με $x\neq 0$) αν και μόνο αν όλες οι ιδιοτιμές του πίνακα Α είναι θετικές.

2)Η f είναι αρνητικά ορισμένη ($f\left(x\right)<0$ για κάθε $x\in R^n$ με $x\neq 0$) αν και μόνο αν όλες οι ιδιοτιμές του πίνακα Α είναι αρνητικές.

3) Η f είναι θετικά ημιορισμένη ($f\left(x\right)\ge 0$ για κάθε $x\in R^n$) αν και μόνο αν όλες οι ιδιοτιμές του Α είναι ≥ 0.

Η f είναι αρνητικά ημιορισμένη ($f\left(x\right)\le 0$ για κάθε $x\in R^n$) αν και μόνο αν όλες οι ιδιοτιμές του Α είναι ≤ 0.

4) Η f είναι αόριστη αν και μόνο αν ο Α έχει αρνητικές και θετικές ιδιοτιμές. Δηλαδή η f δεν είναι θετικά αλλά ούτε αρνητικά ημιορισμένη.

Επίσης υπενθυμίζουμε τα ακόλουθα για ένα τετραγωνικό πίνακα A.

5) Οι ιδιοτιμές του Α είναι οι ρίζες της χαρακτηριστικής εξίσωσης (χαρακτηριστικού πολυωνύμου) του Α, δηλαδή της $\det\left(\mathrm{A}-\lambda\mathrm{I}\right)=0$, όπου Ι ο ταυτοτικός πίνακας.

6) Για την ορίζουσα $\det\mathrm{A}$ του Α ισχύει ότι $\det\mathrm{A}=$ το γινόμενο των ιδιοτιμών του Α.

7) Αν ο Α είναι επιπλέον συμμετρικός τότε οι ιδιοτιμές του είναι όλες πραγματικές.

Είναι εύκολο να ελέγξουμε τους ισχυρισμούς (1), (2), (3) και (4) στην περίπτωση μιας τετραγωνικής μορφής δύο μεταβλητών.

Έστω, $f\left(x,y\right)=ax^2+2\beta xy+\gamma y^2=\left(x,y\right)\begin{pmatrix} a & \beta \\ \beta & \gamma \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$.

Το χαρακτηριστικό πολυώνυμο του πίνακα $\mathrm{A}=\begin{pmatrix} a & \beta \\ \beta & \gamma \end{pmatrix}$ είναι το $\varphi\left(\lambda\right)=\det\begin{vmatrix} a-\lambda & \beta \\ \beta & \gamma-\lambda \end{vmatrix}$. Επομένως

$\varphi\left(\lambda\right)=0\Leftrightarrow\varphi\left(\lambda\right)=\left(a-\lambda\right)\left(\gamma-\lambda\right)-\beta^2=0\Leftrightarrow\lambda^2-\left(a+\gamma\right)\lambda+\left(a\gamma-\beta^2\right)=0$.

Η διακρίνουσα του τριωνύμου $\varphi\left(\lambda\right)$ είναι η $\Delta=\left(a-\gamma\right)^2+4\beta^2$.

Οι πραγματικές ρίζες του χαρακτηριστικού πολυωνύμου $\varphi\left(\lambda\right)$ είναι οι, $\lambda_1=\dfrac{a+\gamma+\sqrt{\Delta}}{2}$ και $\lambda_2=\dfrac{a+\gamma-\sqrt{\Delta}}{2}$. Οι ρίζες αυτές είναι οι ιδιοτιμές του πίνακα Α. Το γινόμενο των ριζών του $\varphi\left(\lambda\right)$ είναι το $\mathrm{P}=\det\mathrm{A}=a\gamma-\beta^2$ το δε άθροισμα των ριζών του είναι το $S=a+\gamma$.

Εύκολα τώρα διαπιστώνουμε τους παραπάνω ισχυρισμούς .

Παρατηρούμε ότι αν οι ρίζες είναι ετερόσημες, δηλαδή το $\mathrm{P}=a\gamma-\beta^2<0$ (η f είναι αόριστη τετραγωνική μορφή)τότε το $\left(0,0\right)$ είναι σαγματικό σημείο της f.

13.5 Ορισμός. Έστω $U \subseteq R^n$ ανοικτό, $x_0 \in U$ και $f : U \to R$ συνάρτηση της κλάσης C^2. Η <u>Εσσιανή</u> (Hessian) της f στο x_0 είναι η τετραγωνική μορφή $Hf(x_0) : R^n \to R$ που ορίζεται ως,

$$Hf(x_0)(h) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j \,,\ h = (h_1, ..., h_n) \in R^n \,.$$

Παρατηρούμε ότι: 1) Η $\mathrm{H}f(x_0)$ είναι ακριβώς το <u>διαφορικό δεύτερης τάξης</u> της f στο x_0, δηλαδή $\mathrm{H}f(x_0) = D_2 f(x_0)$.

Υπενθυμίζουμε ότι από την συνέχεια των μερικών παραγώγων δεύτερης τάξης της f, έχουμε ότι αν $h = (h_1, ..., h_n) \in R^n$ τότε: $D_2 f(x_0)(h) = \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2}(x_0) h_i^2 + 2 \sum_{1 \le i < j \le n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j$ και άρα η Εσσιανή είναι πράγματι τετραγωνική μορφή.

Ακόμη ότι το ανάπτυγμα Taylor δεύτερης τάξης της f στο x_0 είναι το $f(x_0 + h) = f(x_0) + D_1 f(x_0)(h) + \frac{1}{2!} D_2 f(x_0)(h) + R_2(h, x_0)$ όπου $R_2(h, x_0)$ είναι το υπόλοιπο Taylor της f στο x_0 για το οποίο ισχύει $\lim_{h \to 0} \frac{R_2(h, x_0)}{\|h\|^2} = 0$.

2) Ο πίνακας (Hesse) που δίνει την τετραγωνική μορφή $\mathrm{H}f(x_0)$, δηλαδή την Εσσιανή, είναι ο συμμετρικός πίνακας

$$\left[\mathrm{H}f(x_0)\right] = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0)..... & \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0)...... & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) \\ .... & .... & .... \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0)... & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix}$$

των δευτέρων μερικών παραγώγων της f στο x_0.

Παρατηρούμε ότι, $\mathrm{H}f(x_0)(h) = (h_1,...,h_n)\cdot\left[\mathrm{H}f(x_0)\right]\cdot\begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ h_n \end{pmatrix} = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j$,

όπου $h = (h_1,...,h_n) \in R^n$ και ακόμη ότι $\left[\mathrm{H}f(x_0)\right] = \begin{pmatrix} \nabla\left(\frac{\partial f}{\partial x_1}\right) \\ \cdot \\ \cdot \\ \nabla\left(\frac{\partial f}{\partial x_n}\right) \end{pmatrix}$.

13.6 Λήμμα Έστω $\mathrm{H}: R^n \to R$ τετραγωνική μορφή. Αν η H είναι θετικά ορισμένη τότε υπάρχει σταθερά $\mathrm{M} > 0$ ώστε, $\mathrm{H}(h) \geq \mathrm{M}\cdot\|h\|^2$ για κάθε $h \in R^n$.

Απόδειξη: Έστω $\mathrm{H}(h) = \sum_{i,j=1}^{n} a_{ij} h_i h_j$ και $S^{n-1} = \left\{h \in R^n : \|h\| = 1\right\}$, ορίζουμε την συνάρτηση $g : S^{n-1} \to R : g(h) = \mathrm{H}(h)$, δηλαδή περιορίζουμε την H στο S^{n-1}.
Επειδή η g είναι συνεχής (η H είναι συνεχής στον R^n ως πολυώνυμο) και η επιφάνεια S^{n-1} της μοναδιαίας σφαίρας του R^n είναι συμπαγές σύνολο η g επιτυγχάνει ελάχιστη τιμή στην S^{n-1}, η οποία είναι βεβαίως θετικός αριθμός, αφού η H είναι θετικά ορισμένη. Έστω $\mathrm{M} = \min\left\{g(h) : \|h\| = 1\right\} > 0$. Επειδή η H είναι τετραγωνική μορφή θα έχουμε, ότι αν $h \neq 0$ τότε, $\mathrm{H}(h) = \mathrm{H}\left(\frac{h}{\|h\|}\cdot\|h\|\right) = \|h\|^2\cdot\mathrm{H}\left(\frac{h}{\|h\|}\right) = \|h\|^2\cdot g\left(\frac{h}{\|h\|}\right) \geq \|h\|^2\cdot\mathrm{M}$. Αν $h = 0$ το αποτέλεσμα είναι προφανώς σωστό.

Πρόκειται να αποδείξουμε ένα αποτέλεσμα για διαφορίσιμες συναρτήσεις πολλών μεταβλητών που γενικεύει το ακόλουθο κλασσικό αποτέλεσμα για μια πραγματική μεταβλητή.
Αν $\mathrm{I} \subseteq R$ είναι ανοικτό διάστημα, $x_0 \in \mathrm{I}$ και $f : \mathrm{I} \to R$ είναι μια συνάρτηση της κλάσης C^2 ώστε $f'(x_0) = 0$ τότε:

(ι) Αν $f''(x_0) > 0$, το x_0 είναι τοπικό ελάχιστο της f

(ιι) Αν $f''(x_0) < 0$, το x_0 είναι τοπικό μέγιστο της f.

Έστω $U \subseteq R^n$ ανοικτό, $x_0 \in U$ και $f : U \to R$ συνάρτηση της κλάσης C^3 ώστε το x_0 είναι κρίσιμο σημείο της f. Επειδή τότε $\nabla f(x_0) = 0$, το ανάπτυγμα Taylor της f στο x_0, δεύτερης τάξης γράφεται ως εξής:

$$f(x_0 + h) = f(x_0) + \frac{1}{2}\mathrm{H}f(x_0)(h) + R_2(h, x_0).$$

Η γενίκευση στην οποία αναφερθήκαμε έχει ως ακολούθως.

13.7 Θεώρημα (Κριτήριο της δεύτερης παραγώγου .) Αν η $f : U \subseteq R^n \to R$ είναι της κλάσης C^3 και το $x_0 \in U$ είναι ένα κρίσιμο σημείο της f τότε:

(ι) Αν η Εσσιανή $\mathrm{H}f(x_0)$ είναι <u>θετικά</u> ορισμένη, δηλαδή αν οι ιδιοτιμές του Εσσιανού πίνακα $[\mathrm{H}f(x_0)]$ της $\mathrm{H}f(x_0)$ είναι όλες θετικές , τότε το x_0 είναι <u>τοπικό ελάχιστο</u> για την f .

(ιι) Αν η Εσσιανή $\mathrm{H}f(x_0)$ είναι <u>αρνητικά</u> ορισμένη, δηλαδή αν οι ιδιοτιμές του $[\mathrm{H}f(x_0)]$ είναι όλες αρνητικές, τότε το x_0 είναι <u>τοπικό</u> μέγιστο για την f .

(ιιι) Αν η Εσσιανή $\mathrm{H}f(x_0)$ δεν είναι θετικά αλλά ούτε αρνητικά ημιορισμένη, δηλαδή ο $[\mathrm{H}f(x_0)]$ έχει αρνητικές και θετικές ιδιοτιμές τότε το x_0 είναι σαγματικό σημείο της f .

Απόδειξη: Όπως παρατηρήσαμε ήδη ο τύπος του Taylor στο x_0, δεύτερης τάξης γράφεται $f(x_0+h) = f(x_0) + \frac{1}{2}\mathrm{H}f(x_0)(h) + R_2(h, x_0)$ όπου $\lim_{h\to 0}\frac{R_2(h,x_0)}{\|h\|^2} = 0$.

(ι) Επειδή η $H(f)(x_0)$ είναι θετικά ορισμένη από το προηγούμενο Λήμμα έπεται ότι υπάρχει $\mathrm{M} > 0$ ώστε $\mathrm{H}f(x_0)(h) \geq \mathrm{M}\|h\|^2$ για κάθε $h \in R^n$.

Αφού $\frac{R_2(h,x_0)}{\|h\|^2} \xrightarrow[h\to 0]{} 0$, υπάρχει $\delta > 0 : 0 < \|h\| < \delta$ τότε $|R_2(h,x_0)| < \frac{\mathrm{M}}{2}\|h\|^2$.

Συνεπώς, αν $0 < \|h\| < \delta$ τότε, $0 < \frac{1}{2}\mathrm{H}f(x_0) + R_2(h,x_0) = f(x_0+h) - f(x_0)$. Άρα $f(x_0) < f(x_0+h)$ για κάθε $h \neq 0$ με $\|h\| < \delta$, δηλαδή το x_0 είναι (γνήσιο !) τοπικό ελάχιστο για την f .

Η απόδειξη στην περίπτωση που η $\mathrm{H}f(x_0)$ είναι <u>αρνητικά</u> ορισμένη είναι παρόμοια (εξάλλου έπεται αν εφαρμόσουμε τα παραπάνω στην $-f$) και έτσι παραλείπεται.

(ιιι) Έστω $h_1, h_2 \in R^n$ ώστε, $\mu_1 = \frac{1}{2}\mathrm{H}f(x_0)(h_1) > 0$ και $\mu_2 = \frac{1}{2}\mathrm{H}f(x_0)(h_2) < 0$. Είναι σαφές ότι υπάρχει $\delta > 0$ ώστε : $0 < t < \delta \Rightarrow \mu_1 - \frac{|R_2(th_1,x_0)|}{\|th_1\|^2}\cdot\|h_1\|^2 > 0$ και $\mu_2 + \frac{|R_2(th_2,x_0)|}{\|th_2\|^2}\cdot\|h_2\|^2 < 0$.

Έπεται ότι αν $0 < t < \delta$ τότε,

$$f(x_0+th_1) = f(x_0) + \frac{1}{2}\mathrm{H}f(x_0)(th_1) + R_2(th_1,x_0) =$$

$$f(x_0) + \frac{1}{2}\mathrm{H}f(x_0)(h_1)t^2 + R_2(th_1,x_0) == f(x_0) + \mu_1 t^2 + R_2(th_1,x_0) > f(x_0),$$

(εφόσον $\mu_1 t^2 - |R_2(th_1,x_0)| > 0 \Rightarrow \mu_1 t^2 + R_2(th_1,x_0) \geq \mu_1 t^2 - |R_2(th_1,x_0)| > 0$).

Ανάλογα βρίσκουμε:

$f\left(x_0+th_2\right)=f\left(x_0\right)+\frac{1}{2}\mathrm{H}f\left(x_0\right)\left(th_2\right)+R_2\left(th_2,x_0\right)=$

$f\left(x_0\right)+\mu_2 t^2+R_2\left(th_2,x_0\right)<f\left(x_0\right)$.

Άρα για $0<t<\delta$ έχουμε, $f\left(x_0+th_2\right)<f\left(x_0\right)<f\left(x_0+th_1\right)$, από όπου συμπεραίνουμε ότι η f δεν έχει τοπικό ακρότατο στο x_0. Δηλαδή το x_0 είναι σαγματικό σημείο της f.

Ακολούθως εφαρμόζουμε το προηγούμενο θεώρημα και την πρόταση 13.4 σε μια συνάρτηση δύο μεταβλητών.

Έστω λοιπόν $z=f\left(x,y\right)$ πραγματική συνάρτηση της κλάσης C^3, ορισμένη στο ανοικτό υποσύνολο U του R^2 και $\vec{x}=\left(x_0,y_0\right)\in U$ κρίσιμο σημείο της f, δηλαδή $\frac{\partial f}{\partial x}\left(\vec{x}\right)=\frac{\partial f}{\partial y}\left(\vec{x}\right)=0$.

Θέτομε, $a=\frac{\partial^2 f}{\partial x^2}\left(\vec{x}\right),\beta=\frac{\partial^2 f}{\partial x\partial y}\left(\vec{x}\right),\gamma=\frac{\partial^2 f}{\partial y^2}\left(\vec{x}\right)$ και $\Delta=a\gamma-\beta^2$.

Η ποσότητα $\Delta=a\gamma-\beta^2$ ονομάζεται διακρίνουσα και ισούται με την ορίζουσα του Εσσιανού πίνακα της f στο $\vec{x}=\left(x_0,y_0\right)$. (Παρατηρούμε ότι ο Εσσιανός πίνακας της f στο ακόλουθο θεώρημα είναι αντιστρέψιμος, εφόσον η ορίζουσά του ισούται με $\Delta=a\gamma-\beta^2\neq 0$.)

13.8 Θεώρημα Ισχύουν τα ακόλουθα:

(ι) Αν $a>0$ και $a\gamma-\beta^2>0$, η f έχει τοπικό ελάχιστο στο $\left(x_0,y_0\right)$

(ιι) Αν $a<0$ και $a\gamma-\beta^2>0$, η f έχει τοπικό μέγιστο στο $\left(x_0,y_0\right)$

(ιιι) Αν $a\gamma-\beta^2<0$, το $\left(x_0,y_0\right)$ είναι σαγματικό σημείο της f.

Απόδειξη: Ας συμβολίσουμε με Q την Εσσιανή $\mathrm{H}f\left(x_0,y_0\right)$ της f στο $\vec{x}=\left(x_0,y_0\right)$. Τότε έχουμε, $Q\left(h_1,h_2\right)=\frac{\partial^2 f}{\partial x^2}\left(\vec{x}\right)h_1^2+2\frac{\partial^2 f}{\partial x\partial y}\left(\vec{x}\right)h_1h_2+\frac{\partial^2 f}{\partial y^2}\left(\vec{x}\right)h_2^2$, όπου $\left(h_1,h_2\right)\in R^2$.

Συνεπώς σύμφωνα με τον συμβολισμό που υιοθετήσαμε

$$Q\left(h_1,h_2\right)=ah_1^2+2\beta h_1h_2+\gamma h_2^2,\left(h_1,h_2\right)\in R^2.$$

Οι ισχυρισμοί (ι) , (ιι) και (ιιι) είναι τώρα εύκολη συνέπεια του προηγούμενου θεωρήματος σε συνδυασμό με την πρόταση 13.4.

Σημείωση. Μια ενδιαφέρουσα απευθείας απόδειξη του ισχυρισμού (ιιι) του θεωρήματος 13.8 έχει ως ακολούθως:

Παρατηρούμε ότι επειδή

$a\gamma-\beta^2<0$ από τον ισχυρισμό (ιιι) της πρότασης 13.4 το $(0,0)$ είναι σαγματικό σημείο της Q, επομένως υπάρχουν σημεία $\vec{u}=(u_0,u_1)$ και $\vec{v}=(v_0,v_1)$ του R^2 ώστε $Q(\vec{u})>0$ και $Q(\vec{v})<0$.

Υποθέτουμε, όπως μπορούμε, ότι $\|\vec{u}\|=\|\vec{v}\|=1$.

Έστω $\delta>0:\mathrm{B}(\vec{x},\delta)\subseteq U$, θεωρούμε τις συναρτήσεις (της κλάσης C^3) που ορίζονται από τους τύπους, $\varphi(t)=f(\vec{x}+t\vec{u})$ και $\psi(t)=f(\vec{x}+t\vec{v})$ για $|t|<\delta$.

Εφαρμόζοντας τον κανόνα αλυσίδας δύο φορές στις συναρτήσεις φ και ψ (δηλαδή στις φ και φ' και ψ και ψ'), υπολογίζουμε ότι

$\varphi'(0)=\nabla f(\vec{x})\cdot\vec{u}=0$ και $\varphi''(0)=Q(\vec{u})>0$

$\psi'(0)=\nabla f(\vec{x})\cdot\vec{v}=0$ και $\psi''(0)=Q(\vec{u})<0$.

Έπεται από την θεωρία της μιας πραγματικής μεταβλητής ότι η φ επιτυγχάνει στο 0 τοπικό ελάχιστο και η ψ τοπικό μέγιστο. Συμπεραίνουμε ότι,

$f(\vec{x}+t\vec{v})=\psi(t)<\psi(0)=f(\vec{x})=\varphi(0)<\varphi(t)=f(\vec{x}+t\vec{u})$ για αρκετά μικρό $|t|$.

Είναι σαφές από αυτές τις ανισότητες ότι η f δεν μπορεί να έχει στο σημείο $\vec{x}=(x_0,y_0)$ τοπικό ακρότατο, έτσι το $\vec{x}=(x_0,y_0)$ είναι πράγματι σαγματικό σημείο της f.

Οι υπολογισμοί, με τον κανόνα της αλυσίδας, των φ' και φ'' στην προηγούμενη απόδειξη έχουν ως εξής: Για $t\in(-\delta,\delta)$ έχουμε,

$$\varphi'(t)=\frac{\partial f}{\partial x}(\vec{x}+t\vec{u})\cdot u_0+\frac{\partial f}{\partial y}(\vec{x}+t\vec{u})\cdot u_1,\ \text{άρα},\ \varphi''(t)=\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial x}(\vec{x}+t\vec{u})\cdot u_0\right]\cdot u_0+$$

$$+\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial x}(\vec{x}+t\vec{u})\cdot u_0\right]\cdot u_1+\frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}(\vec{x}+t\vec{u})\cdot u_1\right]\cdot u_0+\frac{\partial}{\partial y}\left[\frac{\partial f}{\partial y}(\vec{x}+t\vec{u})\cdot u_1\right]\cdot u_1=$$

$$=\frac{\partial^2 f}{\partial x^2}(\vec{x}+t\vec{u})\cdot u_0^2+2\frac{\partial^2 f}{\partial y\partial x}(\vec{x}+t\vec{u})\cdot u_0u_1+\frac{\partial^2 f}{\partial y^2}(\vec{x}+t\vec{u})\cdot u_1^2$$

Συνεπώς, $\varphi'(0)=\nabla f(\vec{x})\cdot\vec{u}=0$ και $\varphi''(0)=\dfrac{\partial^2 f}{\partial x^2}(\vec{x})\cdot u_0^2+2\dfrac{\partial^2 f}{\partial x\partial y}(\vec{x})\cdot u_0u_1+$

$+\dfrac{\partial^2 f}{\partial y^2}(\vec{x})\cdot u_1^2=Q(\vec{u})>0$.

Οι υπολογισμοί για τις ψ' και ψ'' είναι παρόμοιοι.

Παρατηρήσεις: 1) Η συμπεριφορά μιας συνάρτησης f της κλάσης C^3 κοντά στο κρίσιμο σημείο x_0 υπό την προϋπόθεση ότι ο Εσσιανός πίνακας $[\mathrm{H}f(x_0)]$ είναι αντιστρέψιμος (στην περίπτωση που η f είναι δύο μεταβλητών αυτό σημαίνει ότι η διακρίνουσα, $\Delta=a\gamma-\beta^2\neq0$) καθορίζεται από το <u>δευτεροβάθμιο</u> μέρος της, δηλαδή την Εσσιανή $\mathrm{H}f(x_0)$ της f. Τα κρίσιμα σημεία στα οποία ο $[\mathrm{H}f(x_0)]$ είναι αντιστρέψιμος ονομάζονται <u>μη εκφυλισμένα</u>, τα υπόλοιπα κρίσιμα σημεία όπου ο $\left[\mathrm{H}f\left(x_0\right)\right]$ δεν είναι αντιστρέψιμος ονομάζονται <u>εκφυλισμένα</u> .

2)Η μελέτη μιας C^3 συνάρτησης δύο μεταβλητών συνοψίζεται ως εξής:

(α) Βρίσκουμε τα κρίσιμα σημεία της f λύνοντας το σύστημα $\begin{cases} \dfrac{\partial f}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} = 0 \end{cases}$

(β) Υπολογίζουμε τις Εσσιανές των κρίσιμων σημείων της f :

(ι) κάποιες από αυτές μπορεί να είναι θετικά ορισμένες, υποδεικνύοντας τα σχετικά ελάχιστα, κάποιες μπορεί να είναι αρνητικά ορισμένες υποδεικνύοντας τα σχετικά μέγιστα ($\Delta = a\gamma - \beta^2 > 0$).

(ιι) Κάποιες μπορεί να μην είναι <u>ούτε</u> θετικά <u>ούτε</u> αρνητικά ορισμένες υποδεικνύοντας τα <u>σαγματικά σημεία</u> (αν $\Delta < 0$).

(γ) Τα κρίσιμα σημεία για τα οποία $\Delta = a\gamma - \beta^2 \neq 0$ (μη εκφυλισμένα κρίσιμα σημεία) είναι (τοπικά) μέγιστα και ελάχιστα ή σαγματικά σημεία. Τα υπόλοιπα κρίσιμα σημεία όπου $\Delta = 0$ εκφυλισμένα κρίσιμα σημεία), συνήθως ελέγχονται απευθείας.

3) Το ‘‘ Κριτήριο της δεύτερης παραγώγου’’ (θεώρημα 13.7) μπορεί να αποφανθεί υπό την προϋπόθεση ότι το κρίσιμο σημείο είναι μη εκφυλισμένο, διότι τότε όλες οι ιδιοτιμές του Εσσιανού πίνακα είναι μη μηδενικές, από όπου έπεται ότι μια από τις τρείς περιπτώσεις του θεωρήματος ισχύει. Παραδείγματα όπως τα $f\left(x,y\right) = \pm\left(x^4 + y^4\right)$ και $x^4 - y^4$, δείχνουν ότι η f μπορεί να έχει ελάχιστο, μέγιστο ή σαγματικό σημείο αντίστοιχα στο $\left(0,0\right)$, ενώ ο Εσσιανός πίνακας στο σημείο αυτό είναι ο μηδενικός πίνακας

Παραδείγματα: 1) Να μελετηθούν τα κρίσιμα σημεία της συνάρτησης

$$f\left(x,y\right) = x^3 + x^2 y - y^2 - 4y, \quad \left(x,y\right) \in R^2 .$$

Λύση $\dfrac{\partial f}{\partial x} = 3x^2 + 2xy, \dfrac{\partial f}{\partial y} = x^2 - 2y - 4$. Επιλύουμε το σύστημα

$$\begin{cases} \dfrac{\partial f}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} = 0 \end{cases} \Leftrightarrow \begin{cases} 3x^2 + 2xy = 0 \\ x^2 - 2y - 4 = 0 \end{cases} \Leftrightarrow \begin{cases} x\left(3x + 2y\right) = 0 \\ x^2 - 2y = 4 \end{cases} \Leftrightarrow \left.\begin{cases} x = 0 \\ x^2 - 2y = 4 \end{cases}\right\} \vee \left\{\begin{matrix} 3x + 2y = 0 \\ x^2 - 2y = 4 \\ x \neq 0 \end{matrix}\right\} .$$

Το πρώτο σύστημα έχει ως λύσεις τις $x = 0$ και $y = -2$.

Το δεύτερο σύστημα είναι ισοδύναμο με το $\left\{\begin{matrix} x^2 + 3x - 4 = 0 \\ x^2 - 2y = 4 \\ x \neq 0 \end{matrix}\right\}$ και έχει ως λύσεις τις

$x = 1$ και $y = -\dfrac{3}{2}$ και $x = -4$ και $y = 6$

Έτσι οι λύσεις του αρχικού συστήματος είναι: $\left(0,-2\right), \left(1, -\dfrac{3}{2}\right), \left(-4,6\right)$.

Έπεται ότι τα κρίσιμα σημεία της f είναι $\left(0,-2\right), \left(1, -\dfrac{3}{2}\right), \left(-4,6\right)$

Οι δεύτερες μερικές παράγωγοι της f είναι:

$\frac{\partial^2 f}{\partial x^2} = 6x + 2y$, $\frac{\partial^2 f}{\partial y^2} = -2$ και $\frac{\partial^2 f}{\partial x \partial y} = 2x = \frac{\partial^2 f}{\partial y \partial x}$. Στο κρίσιμο σημείο (x_0, y_0) η Εσσιανή είναι: $Q(h_1, h_2) = \frac{\partial^2 f}{\partial x^2}(x_0, y_0)h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)h_1 h_2 + \frac{\partial^2 f}{\partial y^2}(x_0, y_0)h_2^2$.

(I) Στο κρίσιμο σημείο $(0, -2)$ έχουμε: $a = -4, \beta = 0, \gamma = -2$. Επομένως $a < 0$ και $a\gamma - \beta^2 = (-4)\cdot(-2) - 0^2 = 8 > 0$. Άρα η f έχει στο $(0, -2)$ τοπικό μέγιστο.

(II) Στο κρίσιμο σημείο $\left(1, -\frac{3}{2}\right)$ έχουμε: $a = \frac{\partial^2 f}{\partial x^2}\left(1, -\frac{3}{2}\right) = 6 + 2\left(-\frac{3}{2}\right) = 6 - 3 = 3$, $\beta = \frac{\partial^2 f}{\partial x \partial y}\left(1, -\frac{3}{2}\right) = 2$ και $\gamma = \frac{\partial^2 f}{\partial y^2}\left(1, -\frac{3}{2}\right) = -2$.

Συνεπώς $a\gamma - \beta^2 = 3\cdot(-2) - 2^2 = -10 < 0$ και άρα το $\left(1, -\frac{3}{2}\right)$ είναι σαγματικό σημείο για την f.

(III) Στο κρίσιμο σημείο $(-4, 6)$ έχουμε: $a = \frac{\partial^2 f}{\partial x^2}(-4, 6) = 6\cdot(-4) + 2\cdot 6 = -24 + 12 = -12$, $\beta = \frac{\partial^2 f}{\partial x \partial y}(-4, 6) = 2\cdot(-4) = -8$ και $\gamma = \frac{\partial^2 f}{\partial y^2}(-4, 6) = -2$. Συνεπώς, $a\gamma - \beta^2 = (-12)\cdot(-2) - (-8)^2 = 24 - 64 = -40 < 0$ και το $(-4, 6)$ είναι σαγματικό σημείο για την f.

Η μελέτη μας μπορεί να συνοψισθεί στον ακόλουθο πίνακα:

x	y	$a = f_{xx}$	$\beta = f_{xy}$	$\gamma = f_{yy}$	$\Delta = a\gamma - \beta^2$	Κατάταξη
0	-2	-4	0	-2	8	Τοπικό μέγιστο
1	$-\frac{3}{2}$	3	2	-2	-10	Σαγματικό σημείο
-4	6	-12	-8	-2	-40	Σαγματικό σημείο

2) Μελετήστε τα κρίσιμα σημεία της συνάρτησης $f(x, y) = x^2 y^2 - 5x^2 - 8xy - 5y^2$, $(x, y) \in R^2$. (Πρβλ το παράδειγμα (5) μετά το θεώρημα 13.2)

Λύση $\frac{\partial f}{\partial x} = 2xy^2 - 10x - 8y$ και $\frac{\partial f}{\partial y} = 2x^2 y - 8x - 10y$

Έτσι έχουμε να επιλύσουμε το σύστημα $\begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases}$ δηλαδή το $\begin{cases} 2xy^2 - 10x - 8y = 0 \\ 2x^2 y - 8x - 10y \end{cases}$ (1)

Παρατηρούμε ότι αν $x = 0$ τότε (και μόνο τότε) $y = 0$.

Άρα μία λύση είναι η $(x,y)=(0,0)$. Υποθέτοντας ότι $x\neq 0$ (και άρα $y\neq 0$) πολλαπλασιάζουμε την δεύτερη εξίσωση με $-y$ και την πρώτη με x και καταλήγουμε στο $\left\{\begin{array}{l} 2x^2y^2-10x^2-8xy=0 \\ -2x^2y^2+8xy+10y^2=0 \\ x\neq 0 \end{array}\right\} \Leftrightarrow x^2=y^2$

Συνεπώς με την υπόθεση ότι $x\neq 0$ (και $y\neq 0$) το αρχικό σύστημα ισοδυναμεί με το

$$\left\{\begin{array}{l} x^2=y^2 \\ 2x^2y-8x-10y=0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} x=1 \text{ και } y=-1 \\ x=-1 \text{ και } y=1 \\ x=3 \text{ και } y=3 \\ x=-3 \text{ και } y=-3 \end{array}\right\}$$

Έτσι οι λύσεις του συστήματος (1) και άρα τα κρίσιμα σημεία της f είναι τα ακόλουθα $(0,0),(1,-1),(-1,1),(3,3),(-3,-3)$.

Οι δεύτερες μερικές παράγωγοι της f είναι:

$\dfrac{\partial^2 f}{\partial x^2}=2y^2-10$, $\dfrac{\partial^2 f}{\partial x\partial y}=4xy-8$ και $\dfrac{\partial^2 f}{\partial y^2}=2x^2-10$

Η Εσσιανή της f στο κρίσιμο σημείο (x_0,y_0) είναι: $Q(x_0,y_0)(h_1,h_2)=\dfrac{\partial^2 f}{\partial x^2}(x_0,y_0)h_1^2+2\dfrac{\partial^2 f}{\partial x\partial y}(x_0,y_0)h_1h_2+\dfrac{\partial^2 f}{\partial y^2}(x_0,y_0)h_2^2$.

(I) Στο κρίσιμο σημείο $(0,0)$ έχουμε:

$a=\dfrac{\partial^2 f}{\partial x^2}(0,0)=-10$, $\beta=\dfrac{\partial^2 f}{\partial x\partial y}(0,0)=-8$ και $\gamma=\dfrac{\partial^2 f}{\partial y^2}(0,0)=-10$

Συνεπώς $a=-10<0$ και $a\gamma-\beta^2=(-10)(-10)-(-8)^2=100-64=36>0$.

Άρα το $(0,0)$ είναι τοπικό μέγιστο για την f.

(II) Στο κρίσιμο σημείο $(1,-1)$ έχουμε: $a=\dfrac{\partial^2 f}{\partial x^2}(1,-1)=-8$, $\beta=\dfrac{\partial^2 f}{\partial x\partial y}(1,-1)=-12$, $\gamma=\dfrac{\partial^2 f}{\partial y^2}(1,-1)=-8$.

Συνεπώς ($a=-8$) και $a\gamma-\beta^2=(-8)(-8)-(-12)^2=64-144<0$. Άρα το $(1,-1)$ είναι σαγματικό σημείο για την f.

(III) Στο κρίσιμο σημείο $(-1,1)$ έχουμε: $a=\dfrac{\partial^2 f}{\partial x^2}(-1,1)=-8$, $\beta=\dfrac{\partial^2 f}{\partial x\partial y}(-1,1)=-12$, $\gamma=\dfrac{\partial^2 f}{\partial y^2}(-1,1)=-8$.

Συνεπώς ($a=-8$) και $a\gamma-\beta^2=(-8)(-8)-(-12)^2=64-144<0$. Άρα το $(-1,1)$ είναι σαγματικό σημείο για την f.

(IV) Στο κρίσιμο σημείο $(3,3)$ έχουμε: $a = \frac{\partial^2 f}{\partial x^2}(3,3) = 2\cdot 3^2 - 10 = 8$,
$\beta = \frac{\partial^2 f}{\partial x \partial y}(3,3) = 4\cdot 3\cdot 3 - 8 = 28$ και $\gamma = \frac{\partial^2 f}{\partial y^2}(3,3) = 2\cdot 3^2 - 10 = 8$
Συνεπώς ($a = 8$ και) $a\gamma - \beta^2 = 8\cdot 8 - 28^2 = 64 - 28^2 < 0$. Έτσι το $(3,3)$ είναι επίσης σαγματικό σημείο για την f.

(V) Στο κρίσιμο σημείο $(-3,-3)$ έχουμε: $a = \frac{\partial^2 f}{\partial x^2}(-3,-3) = 2\cdot(-3)^2 - 10 = 8$,
$\beta = \frac{\partial^2 f}{\partial x \partial y}(-3,-3) = 4\cdot(-3)\cdot(-3) - 8 = 28$ και $\gamma = \frac{\partial^2 f}{\partial y^2}(-3,-3) = 2\cdot(-3)^2 - 10 = 8$.
Έπεται όπως προηγουμένως ότι $a\gamma - \beta^2 = 8\cdot 8 - 28^2 = 64 - 28^2 < 0$.
Έτσι και το σημείο $(-3,-3)$ είναι επίσης σαγματικό σημείο για την f.

3)Να μελετηθούν τα κρίσιμα σημεία της συνάρτησης

$$f(x,y,z) = x^3 - 3x - y^3 + 9y + z^2.$$

Λύση Οι μερικές παράγωγοι πρώτης τάξης της f είναι οι, $\frac{\partial f}{\partial x} = 3x^2 - 3$, $\frac{\partial f}{\partial y} = -3y^2 + 9$, $\frac{\partial f}{\partial z} = 2z$.

Οι μερικές παράγωγοι της f δεύτερης τάξης είναι οι $\frac{\partial^2 f}{\partial x^2} = 6x$, $\frac{\partial^2 f}{\partial y^2} = -6y$, $\frac{\partial^2 f}{\partial z^2} = 2$ όλες οι υπόλοιπες μικτές παράγωγοι της f δεύτερης τάξης είναι ταυτοτικά μηδέν ($\frac{\partial^2 f}{\partial x \partial y} = 0$, $\frac{\partial^2 f}{\partial x \partial z} = 0$, κτλ).

Έπεται ότι ο Εσσιανός πίνακας της f στο (x_0, y_0, z_0) είναι ο

$$\left[\mathrm{H}f(x_0,y_0,z_0)\right] = \begin{pmatrix} 6x_0 & 0 & 0 \\ 0 & -6y_0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Η Εσσιανή $Q = \mathrm{H}f(x_0,y_0,z_0)$ είναι η τετραγωνική μορφή:

$$Q(h_1,h_2,h_3) = \frac{\partial^2 f}{\partial x^2}(x_0,y_0,z_0)h_1^2 + \frac{\partial^2 f}{\partial y^2}(x_0,y_0,z_0)h_2^2 + \frac{\partial^2 f}{\partial z^2}(x_0,y_0,z_0)h_3^2 +$$
$$2\left(\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0,z_0)h_1h_2 + \frac{\partial^2 f}{\partial x \partial z}(x_0,y_0,z_0)h_1h_3 + \frac{\partial^2 f}{\partial y \partial z}(x_0,y_0,z_0)h_2h_3\right)$$
$$= 6x_0h_1^2 - 6y_0h_2^2 + 2h_3^2,\ h = (h_1,h_2,h_3) \in R^3.$$

Τα κρίσιμα σημεία (x_0, y_0, z_0) της f είναι οι λύσεις του συστήματος

$$\begin{cases} \dfrac{\partial f}{\partial x} = 0 \\ \dfrac{\partial f}{\partial y} = 0 \\ \dfrac{\partial f}{\partial z} = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 - 1 = 0 \\ y^2 - 3 = 0 \\ z = 0 \end{cases} \Leftrightarrow \begin{cases} x = \pm 1 \\ y = \pm\sqrt{3} \\ z = 0 \end{cases}$$

Συνεπώς τα κρίσιμα σημεία της f είναι τα ακόλουθα: (α) $\left(1, \sqrt{3}, 0\right)$, (β) $\left(1, -\sqrt{3}, 0\right)$, (γ) $\left(-1, -\sqrt{3}, 0\right)$ και (δ) $\left(-1, \sqrt{3}, 0\right)$.

Έτσι έχουμε:

(α) Το χαρακτηριστικό πολυώνυμο του πίνακα $Hf\left(1, \sqrt{3}, 0\right)$ είναι το $\varphi(\lambda) = (6-\lambda)\left(-6\sqrt{3}-\lambda\right)(2-\lambda)$, που έχει ρίζες τις $\lambda_1 = 6$, $\lambda_2 = -6\sqrt{3}$, $\lambda_3 = 2$. Επειδή οι ρίζες αυτές είναι ετερόσημες το κρίσιμο σημείο $\left(1, \sqrt{3}, 0\right)$ είναι σαγματικό σημείο της f.

(β) Το χαρακτηριστικό πολυώνυμο του πίνακα $Hf\left(1, -\sqrt{3}, 0\right)$ είναι το $\varphi(\lambda) = (6-\lambda)\left(6\sqrt{3}-\lambda\right)(2-\lambda)$, που έχει ρίζες τις $\lambda_1 = 6$, $\lambda_2 = 6\sqrt{3}$, $\lambda_3 = 2$ οι οποίες είναι θετικές, άρα το κρίσιμο σημείο $\left(1, -\sqrt{3}, 0\right)$ είναι τοπικό ελάχιστο της f.

(γ) Το χαρακτηριστικό πολυώνυμο του πίνακα $Hf\left(-1, -\sqrt{3}, 0\right)$ είναι το $\varphi(\lambda) = (-6-\lambda)\left(6\sqrt{3}-\lambda\right)(2-\lambda)$, που έχει ρίζες τις $\lambda_1 = -6$, $\lambda_2 = 6\sqrt{3}$, $\lambda_3 = 2$. Επειδή οι ρίζες αυτές είναι ετερόσημες το κρίσιμο σημείο $\left(-1, -\sqrt{3}, 0\right)$ είναι σαγματικό σημείο της f.

(δ) Το χαρακτηριστικό πολυώνυμο του πίνακα $Hf\left(-1, \sqrt{3}, 0\right)$ είναι το $\varphi(\lambda) = (-6-\lambda)\left(-6\sqrt{3}-\lambda\right)(2-\lambda)$, που έχει ρίζες τις $\lambda_1 = -6$, $\lambda_2 = -6\sqrt{3}$, $\lambda_3 = 2$. Επειδή οι ρίζες αυτές είναι ετερόσημες το κρίσιμο σημείο $\left(-1, \sqrt{3}, 0\right)$ είναι σαγματικό σημείο της f.

4) Έστω $f(x, y) = \dfrac{ax^2 + by^2}{e^{x^2+y^2}}, (x, y) \in R^2$, όπου $a > b > 0$. Αποδείξτε ότι η συνάρτηση f έχει ολικό μέγιστο στο σημείο $(x_0, y_0) = (1, 0)$.

Λύση $$\frac{\partial f}{\partial x} = \frac{\dfrac{\partial}{\partial x}\left(ax^2 + by^2\right) \cdot e^{x^2+y^2} - \left(ax^2 + by^2\right)\dfrac{\partial}{\partial x}\left(e^{x^2+y^2}\right)}{e^{2\left(x^2+y^2\right)}} = \frac{2x\left(a - ax^2 - by^2\right)}{e^{x^2+y^2}},$$

$(x, y) \in R^2$.

Έπεται ότι (λόγω συμμετρίας): $\dfrac{\partial f}{\partial y} = \dfrac{2y\left(b - ax^2 - by^2\right)}{e^{x^2+y^2}}$, $(x, y) \in R^2$.

Η συνάρτηση $f(x,y)$ είναι βέβαια $C^{\infty}-$ διαφορίσιμη στο R^2 ως πηλίκο των $C^{\infty}-$ διαφορίσιμων συναρτήσεων ax^2+by^2 (πολυωνυμική) και της $e^{x^2+y^2}$.

Λύνουμε τώρα το σύστημα $\begin{cases} \dfrac{\partial f}{\partial x}=0 \\ \dfrac{\partial f}{\partial y}=0 \end{cases}$ για να βρούμε τα κρίσιμα σημεία της f.

Έτσι έχουμε το αλγεβρικό σύστημα:

$$\begin{cases} \dfrac{2x\left(a-ax^2-by^2\right)}{e^{x^2+y^2}}=0 \\ \dfrac{2y\left(b-ax^2-by^2\right)}{e^{x^2+y^2}}=0 \end{cases} \Leftrightarrow \left.\begin{cases} x\left(ax^2+by^2-a\right)=0 \\ y\left(ax^2+by^2-b\right)=0 \end{cases}\right\}(1).$$

Μια προφανής λύση του συστήματος (1) είναι η $x=0=y$.

Οι υπόλοιπες λύσεις του (1) είναι οι λύσεις των συστημάτων $\begin{cases} x=0 \\ by^2-b=0 \end{cases}$ και $\begin{cases} y=0 \\ ax^2-a=0 \end{cases}$, οι οποίες είναι οι $x=0, y=1$ και $x=0, y=-1$ για το πρώτο και $x=1, y=0$ και $x=-1, y=0$ για το δεύτερο.

Έτσι οι λύσεις του (1) και άρα τα κρίσιμα σημεία της f είναι οι ακόλουθες: $(0,0)$, $(0,1)$, $(0,-1)$, $(1,0)$, $(-1,0)$.

Παρατηρούμε ότι: $f(0,0)=0$, $f(0,1)=f(0,-1)=\dfrac{b}{e}$ και $f(1,0)=f(-1,0)=\dfrac{a}{e}$, άρα η μεγαλύτερη τιμή είναι η $f(1,0)=f(-1,0)=\dfrac{a}{e}$ (2)

Ισχυρισμός: $\lim\limits_{\|(x,y)\|\to+\infty} f(x,y)=0$

Απόδειξη ισχυρισμού. Προφανώς ισχύει $0\le f(x,y)\le \dfrac{a\left(x^2+y^2\right)}{e^{x^2+y^2}}$, $(x,y)\in R^2$. Θέτουμε $z=x^2+y^2$ και παρατηρούμε ότι $z\to+\infty \Leftrightarrow \sqrt{z}=\|(x,y)\|\to+\infty$.

Όπως είναι γνωστό από τον Απειροστικό Λογισμό μιας μεταβλητής $\dfrac{z}{e^z}\xrightarrow[z\to+\infty]{}0 \Rightarrow \dfrac{az}{e^z}\xrightarrow[z\to+\infty]{}0$ για κάθε $a\in R$.

Έπεται ότι, $\dfrac{a\left(x^2+y^2\right)}{e^{x^2+y^2}}\xrightarrow[\|(x,y)\|\to+\infty]{}0 \Rightarrow f(x,y)=\dfrac{ax^2+by^2}{e^{x^2+y^2}}\xrightarrow[\|(x,y)\|\to+\infty]{}0$.

Έστω τώρα $\delta>0$ αρκετά μεγάλο ώστε: $\|(x,y)\|>\delta \Rightarrow 0\le f(x,y)<\dfrac{a}{e}$.

(Η ύπαρξη τέτοιου δ έπεται από τον ισχυρισμό). Είναι τότε σαφές ότι, $(1,0)\in \widehat{\mathrm{B}}(0,\delta)$.

Έπεται τότε ότι: $\dfrac{a}{e}=f(1,0)\le \sup\limits_{\|(x,y)\|\le\delta} f(x,y)=\sup\limits_{(x,y)\in R^2} f(x,y)$ (3)

Από την (3) συμπεραίνουμε ότι για να βρούμε το $\sup\limits_{(x,y)\in R^2} f(x,y)$ αρκεί να περιοριστούμε στον κλειστό δίσκο $\widehat{\mathrm{B}}(0,\delta)$.

Επειδή η f συνεχής και $\widehat{\mathrm{B}}(0,\delta)$ συμπαγές σύνολο η f επιτυγχάνει μέγιστη τιμή (x_0,y_0) στον δίσκο $\widehat{\mathrm{B}}(0,\delta)$ άρα και στον R^2. Το σημείο (x_0,y_0) θα είναι ένα από τα κρίσιμα σημεία της f, δηλαδή ένα από τα $(0,0)$, $(0,1)$, $(0,-1)$, $(1,0)$, $(-1,0)$.

Επειδή από την (2) την μεγαλύτερη τιμή την παίρνει στο $(1,0)$,(και το $(-1,0)$) και η τιμή αυτή είναι $f(1,0)=f(-1,0)=\dfrac{a}{e}$, έπεται ότι στο $(1,0)$ (και στο $(-1,0)$) η f παίρνει την μέγιστη τιμή της.

Συμπέρασμα: Αν $a>b>0$ τότε,

$$0\le f(x,y)=\frac{ax^2+by^2}{e^{x^2+y^2}}\le\frac{a}{e}=f(1,0)=f(-1,0) \text{ για κάθε } (x,y)\in R^2.$$

<u>Παρατήρηση</u> Το παράδειγμα 4 μπορεί να αντιμετωπιστεί και ως εξής: Η συνάρτηση $g(z)=\dfrac{z}{e^z}, z\ge 0$ παίρνει την μέγιστη τιμή της στο $z=1$ (όπως μπορεί να ελεγχθεί με παραγώγους ή με ένα επιχείρημα συμπάγειας όπως στο παράδειγμα 4). Θέτομε $z=x^2+y^2$ και παρατηρούμε ότι: $\dfrac{ax^2+by^2}{e^{x^2+y^2}}\le a\dfrac{x^2+y^2}{e^{x^2+y^2}}\le a\dfrac{z}{e^z}\le\dfrac{a}{e}, z\ge 0$.

Επειδή, $f(1,0)=\dfrac{a}{e}$, έπεται ότι το $(1,0)$ είναι θέση ολικού μεγίστου για την f