Για $g: \mathbb{R}^n \to \mathbb{R}$, C^2, έχουμε την εξίσωση,

$$g(\underline{x}_k) + J_g(\underline{x}_k)(\underline{x} - \underline{x}_k) = 0.$$

όπου $J_g(\underline{x}_k)$ είναι η Jacobian της g στο $\underline{x}_k$, οπότε αν είναι αντιστρέψιμος ($\det J_g \neq 0$) έχουμε:

$$\underline{x}_{k+1} = \underline{x}_k - \left(J_g(\underline{x}_k)\right)^{-1} g(\underline{x}_k)$$

ή

$$\underline{x}_{k+1} = \underline{x}_k - \left(Hf(\underline{x}_k)\right)^{-1} \nabla f(\underline{x}_k)$$

Παρασκευή 23/11/2012

Την προηγούμενη φορά:

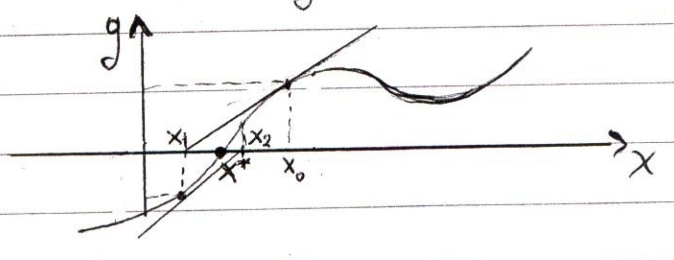
<u>Αλγόριθμος βαθμίδας</u>: x_0, $x_{k+1} = \underline{x}_k + \alpha_k \underline{d}_k$

x_k $\nabla f(x_k)$ d_k

όπου $\underline{d}_k \cdot \nabla f(x_k) < 0$

Επιλέγω α_k τ.ω. $f(\underline{x}_{k+1}) \leq f(\underline{x}_k)$

<u>Newton-Raphson</u>: $g(\underline{x}) = 0$, $g(\underline{x}) \equiv \nabla f(\underline{x})$



$$x_{k+1} = x_k - \left(J_g(x_k)\right)^{-1} g(x_k)$$

για την f: $x_{k+1} = x_k - [Hf(x_k)]^{-1} \nabla f(x_k)$

Αναφέραμε βήμα της μεθόδου N-R : $\eta_f(\underline{x}_k) := -(Hf(\underline{x}_k))^{-1}\nabla f(\underline{x}_k)$.
κ' συγκρίνοντας με τις μεθόδους καθόδου (καθόδου/descent)
$\eta_f(\underline{x}_k)\cdot\nabla f(\underline{x}_k) = -(\nabla f(\underline{x}_k))^T (Hf(\underline{x}_k))^{-1}\nabla f(\underline{x}_k) \leq 0$
αν ο $(Hf(\underline{x}_k))^{-1}$ είναι θετικά ημιορισμένος.

Θεώρημα

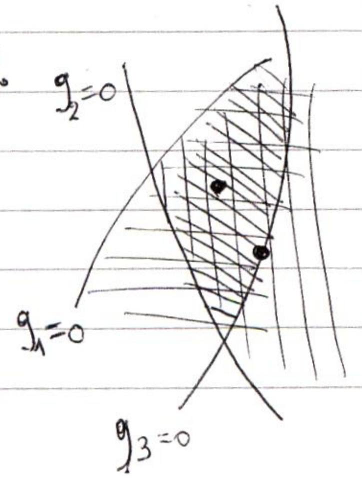
Έστω $f \in C^2(\mathbb{R}^n)$ με Lipschitz-συνεχή Hessian κ' $\underline{x}^* \in \mathbb{R}^n$ είναι ένα **κρίσιμο** σημείο. (η παράγωγος είναι 0).
Αν η $Hf(\underline{x}^*)$ είναι μη ιδιάζουσα τότε υπάρχει περιοχή $B_\rho(\underline{x}^*)$ (μπάλα ακτίνας $\rho>0$, κέντρου $\underline{x}^*$) τ.ω. διαλέγοντας $x_0 \in B_\rho(\underline{x}^*)$ εφαρμόζουμε τον αλγόριθμο N-R κ' δημιουργούμε μια ακολουθία σημείων $(x_k)_{k=1,2,\ldots}$ που είναι καλώς ορισμένη, ανήκει στο $B_\rho(\underline{x}^*)$ κ' τείνει στο $\underline{x}^*$ ως εξής: $\|\underline{x}_{k+1}-\underline{x}^*\| \leq C\|\underline{x}_k-\underline{x}^*\|^2$
σταθερά

Πρόβλημα με περιορισμούς

$$\begin{cases} \min f(\underline{x}), & \underline{x}\in\Omega\subset\mathbb{R}^n \\ \underline{h}(\underline{x})=0, & \underline{g}(\underline{x})\leq 0 \end{cases}$$

Παρατήρηση

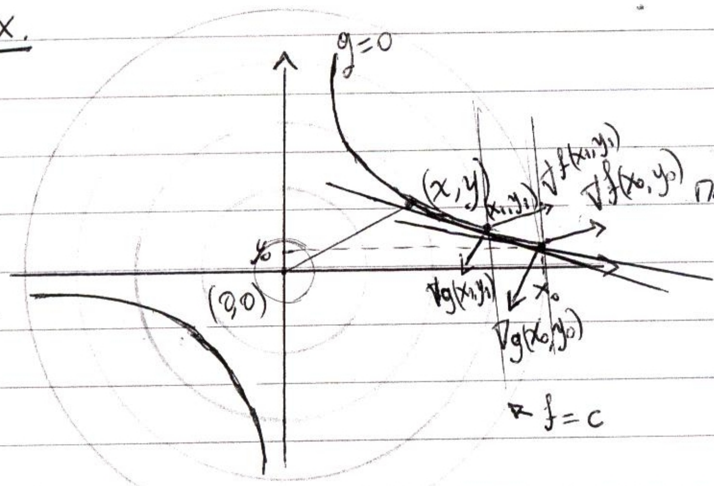
1. Ένα σημείο $\underline{x}$ που ικανοποιεί τους περιορισμούς λέγεται εφικτό.

2. 

Ο περιορισμός $g_i(\underline{x})\leq 0$ λέγεται ενεργός σε ένα εφικτό σημείο $\underline{x}$, αν $g_i(\underline{x})=0$ κ' ανενεργός αν $g_i(\underline{x})<0$

3. Οι τοπικές ιδιότητες στο $\underline{x}$ δεν εξαρτώνται από τους ανενεργούς περιορισμούς.

<u>π.χ.</u>



$x \cdot y = 3$

Ποιο σημείο είναι πιο κοντά στο (0,0);

$$d((0,0),(x,y)) = (x^2+y^2)^{1/2}$$

$\underline{x} = (x,y) \in \mathbb{R}^2$.

$$\min f(x,y) \quad , \quad f(x,y) = x^2+y^2$$
$$g(x,y) = 0 \quad , \quad g(x,y) = x \cdot y - 3$$

Κινούμαστε κατά μήκος του περιορισμού $g(x,y)=0$ κ' θέλουμε να ελαχιστοποιήσουμε τη $f(x,y) = x^2+y^2$.

$c = x^2+y^2 =$ σταθερά, $c \in$ κύκλο.

$\nabla f \perp$ εφαπτομένη του κύκλου

$\nabla g \perp$ εφαπτομένη της g.

Θέλω (x,y) τ.ω. $\nabla g(x,y) \swarrow\!\!\nearrow \nabla f(x,y)$ (παράλληλα κ' αντίρροπα)

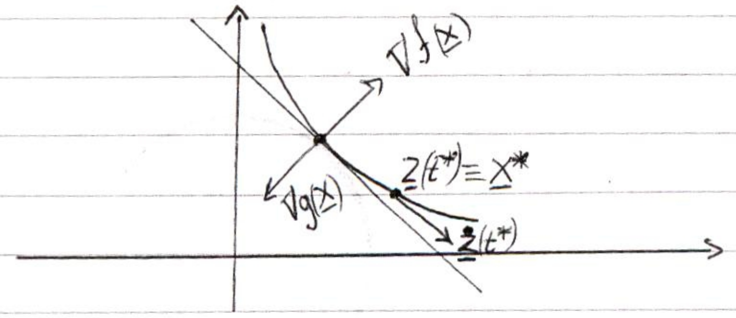
<u>παραμετροποίηση</u> της $g(x,y)=0$

$$t \mapsto z(t) \quad , \quad t \in [\alpha, b]$$

Σε κάθε σημείο $\underline{x}^* = \underline{z}(t^*)$ $\quad \{\underline{x} = (x,y)\}$

το εφαπτόμενο επίπεδο ορίζεται από τα διανύσματα

$$\dot{\underline{z}}(t^*) \equiv \frac{d}{dt} z(t)\Big|_{t=t^*}$$



Έχουμε: $\frac{d}{dt} f(\underline{z}(t)) = \nabla f(\underline{z}(t)) \cdot \dot{\underline{z}}(t)$

Αν $\underline{x}^*$ είναι κρίσιμο σημείο της f, τότε $0 = \frac{d}{dt} f(\underline{z}(t))\Big|_{t=t^*} = \nabla f(\underline{z}(t)) \cdot \dot{\underline{z}}(t)\Big|_{t=t^*}$

Άρα $\nabla f(\underline{x}^*) \perp \dot{\underline{z}}(t^*)$

διάνυσμα που ορίζει την εφαπτομένη της g.

Επίσης, έχουμε ότι $\nabla g \perp \dot{\underline{z}}(t)$

Εν γένει πάντα κάθετο στην εφαπτομένη της g

Επομένως, ψάχνουμε το σημείο $\underline{x}^*$, όπου $\nabla f(\underline{x}^*) \,//\, \nabla g(\underline{x}^*)$

Στο συγκεκριμένο παράδειγμα έχουμε: $\nabla f(\underline{x}) = \lambda \nabla g(\underline{x})$

$$\Leftrightarrow \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = \lambda \left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) \Leftrightarrow \begin{cases} 2x = \lambda y \\ 2y = \lambda x \end{cases} \quad \text{κ'} \quad x \cdot y = 3$$

(← σύστημα 3 εξισώσεων)

$$\begin{cases} 2x = \lambda y \\ 2y = \lambda x \\ x \cdot y = 3 \end{cases} \Leftrightarrow \begin{bmatrix} 2 & -\lambda \\ -\lambda & 2 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x \cdot y = 3$$

$$\det \begin{pmatrix} 2 & -\lambda \\ -\lambda & 2 \end{pmatrix} = 0 \Rightarrow \lambda^2 = 4 \Rightarrow \lambda = \pm 2$$

α) $\lambda = 2 \Rightarrow x = y$, $x \cdot y = 3 \Rightarrow x = \pm\sqrt{3}$, $y = \pm\sqrt{3}$

β) $\lambda = -2 \Rightarrow x = -y$, $x \cdot y = 3 \Rightarrow -x^2 = 3$ αδύνατο.