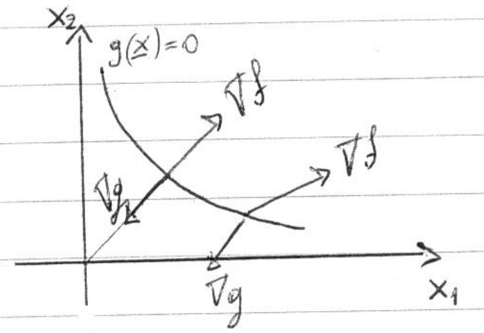
Δευτέρα 26/11/2012

$$\min f(x)$$
$$g(x)=0 \qquad (1)$$
$$x\in\Omega\subset\mathbb{R}^n$$

π.χ.



$\underline{x}=(x_1,x_2)$

Έχουμε $\nabla g \perp$ εφαπτόμενο επίπεδο
Το κρίσιμο σημείο της f θα είναι εκεί που $\nabla f \perp$ εφαπτόμενο επίπεδο.
Άρα $\nabla g \parallel \nabla f$.

Οπότε το πρόβλημα (1) μετασχηματίστηκε στο πρόβλημα:

$$\begin{cases} \dfrac{\partial f}{\partial x_1}(x) = \lambda \dfrac{\partial g}{\partial x_1}(x) \\[2ex] \dfrac{\partial f}{\partial x_2}(x) = \lambda \dfrac{\partial g}{\partial x_2}(x) \\[2ex] g(x_1,x_2)=0 \end{cases} \qquad (2)$$

όπου $x=(x_1,x_2)$

Η (2) μπορεί να γραφεί κ' ως εξής:

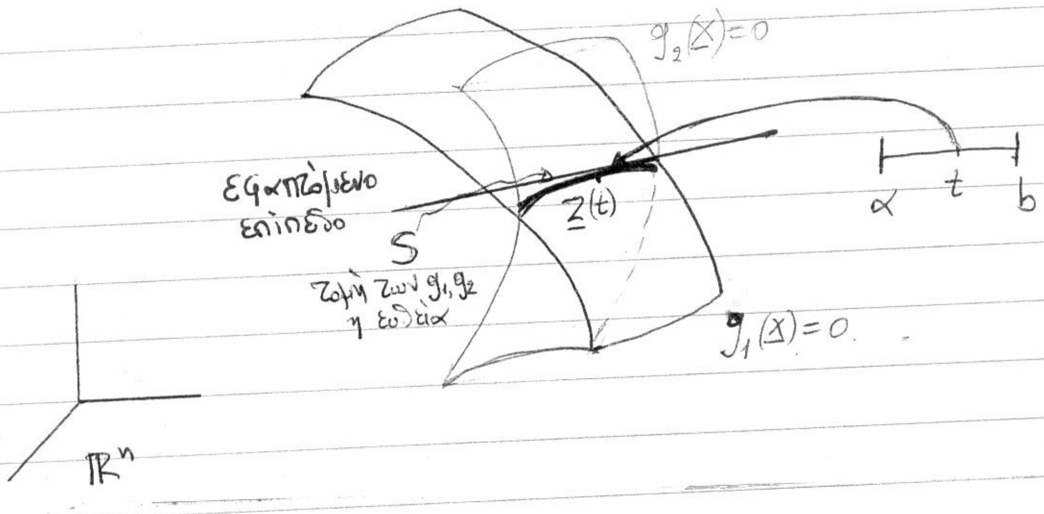
Ορίζουμε τη Lagrangian: $\mathcal{L}(x_1,x_2,\lambda) := f(x_1,x_2)+\lambda g(x_1,x_2)$

η (2) είναι ισοδύναμη με

$$\nabla L = \begin{pmatrix} \frac{\partial f}{\partial x_1} + \frac{\partial g}{\partial x_1}\lambda \\ \frac{\partial f}{\partial x_2} + \frac{\partial g}{\partial x_2}\lambda \\ g(x_1, x_2) \end{pmatrix} = 0.$$

$$\begin{cases} \min f(\underline{x}) \\ \underline{g}(\underline{x}) = 0 \\ \underline{x} \in \mathbb{R}^n \end{cases} \Leftrightarrow \{g_1(\underline{x})=0, \dots, g_m(\underline{x})=0\}$$

Κάθε περιορισμός $g_i(\underline{x})=0$, $i=1,\dots,m$ ορίζει μια υπερεπιφάνεια $S \subset \mathbb{R}^n$, η οποία είναι <u>λεία</u> αν $g_i(\underline{x}) \in C^1$.



Αν οι m περιορισμοί είναι ανεξάρτητοι, η υπερεπιφάνεια που ορίζουν έχει διάσταση $n-m$.
Μια καμπύλη πάνω στην S είναι το σύνολο των σημείων $\underline{z}(t) \in S$ με $t \in [\alpha, b]$.
Η καμπύλη είναι παραγωγίσιμη, αν $\frac{dz}{dt}$ υπάρχει,
δύο φορές παραγωγίσιμη, αν $\frac{d^2z}{dt^2}$ υπάρχει,
κ.ο.κ.

Το εφαπτόμενο επίπεδο στο σημείο $\underline{x}^* \equiv \underline{z}(t^*)$, για κάποιο $t^* \in [\alpha, b]$, ορίζεται ως το σύνολο $M^S_{\underline{x}^*} = \{\dot{\underline{z}}(t) \mid \underline{z}(t) \in S, \underline{z}(t^*) = \underline{x}^*\}$

Ένα σημείο $\underline{x}$ που ικανοποιεί τους περιορισμούς $\underline{g}(\underline{x}) = 0$, λέγεται κανονικό σημείο, αν τα διανύσματα $\nabla g_1, \dots, \nabla g_m$ είναι γραμμικά ανεξάρτητα.

Περιμέναμε να δούμε ότι $\nabla f(\underline{x}^*) \cdot v = 0 \quad , \forall v \in M^S_{\underline{x}^*}$

Επιπλέον, θα θέλαμε να δούμε ότι το ∇f είναι γραμμικός συνδυασμός των $\nabla g_1, \dots, \nabla g_m$ (αφού επίσης $\nabla g_i \cdot v = 0 \; , \forall i$)

Αυτό όμως ισχύει μόνο αν τα ∇g_i span τον υπόχωρο $(M^S_{\underline{x}^*})^{\perp}$, ("ορίζουν")

που με τη σειρά του ισχύει, αν το $\underline{x}$ είναι κανονικό σημείο.

Οπότε έχουμε:

<u>Θεώρημα 1</u>

Σε ένα κανονικό σημείο $\underline{x}$ της επιφάνειας S που ορίζεται από τις σχέσεις $\underline{g}(\underline{x}) = \underline{0}$, το εφαπτόμενο επίπεδο δίνεται από $\{\underline{v} : \nabla g(\underline{x}) \cdot \underline{v} = 0\}$

<u>Θεώρημα 2</u>

Έστω $\underline{x}^*$ τοπικό ακρότατο της f, υπό τους περιορισμούς $\underline{g}(\underline{x}) = 0$. Τότε $\forall \underline{y}$ στο εφαπτόμενο επίπεδο στο $\underline{x}^*$ ισχύει: $\nabla f(\underline{x}^*) \cdot \underline{y} = 0$

<u>Απόδειξη</u>

Έστω κάποιο διάνυσμα $\underline{y} \in M^S_{\underline{x}^*}$, οπότε υπάρχει $\underline{z}(t)$:

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

Αφού $\underline{x}^*$ είναι τοπικό ακρότατο (κινούμενοι κατά μήκος της καμπύλης $\underline{z}(t)$) στο σημείο $\underline{z}(t^*) = \underline{x}^*$, για κάποιο t^*, θα ισχύει $\frac{d}{dt} f(\underline{z}(t))\big|_{t=t^*} = 0$

$$\Leftrightarrow \nabla f(\underline{x}^*) \cdot \frac{d\underline{z}(t)}{dt}\Big|_{t=t^*} = 0 \quad \Leftrightarrow \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}\right) = 0$$

$$\Leftrightarrow \nabla f(\underline{x}^*) \cdot \dot{\underline{z}}(t^*) = 0.$$

Θεώρημα 3
(Αναγκαίες Συνθήκες 1ης τάξης)
Έστω $\underline{x}^*$ ένα τοπικό ακρότατο της f υπό τους περιορισμούς $g(\underline{x}) = \underline{0}$.
Υποθέτοντας ότι το $\underline{x}^*$ είναι κανονικό σημείο (regular point) των περιορισμών, τότε $\exists\ \lambda \in \mathbb{R}^m$ τ.ω. $\nabla f(\underline{x}^*) + \lambda^T \nabla g(\underline{x}^*) = 0$.

$$\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}, \quad \nabla g = \begin{pmatrix} \nabla g_1 \\ \vdots \\ \nabla g_m \end{pmatrix}$$

Σημείωση
Οι παραπάνω αναγκαίες συνθήκες, μαζί με τους περιορισμούς $g(\underline{x}) = \underline{0}$ μας δίνουν $n+m$ εξισώσεις με n (για το $\underline{x}^*$) κ' m (για τα $\lambda_1, \ldots, \lambda_m$) αγνώστους.

Απόδειξη
Από Θεώρημα 1 έχουμε: $\nabla g(\underline{x}^*) \cdot \underline{y} = 0 \quad (*)$
κ' από Θεώρημα 2, αν ισχύει το $(*)$, τότε $\nabla f(\underline{x}^*) \cdot \underline{y} = 0$.
Επομένως, έστω το Π.Γ.Π:

$$\begin{cases} \max\limits_{\underline{y}} \nabla f(\underline{x}^*) \cdot \underline{y} \\ \nabla g(\underline{x}^*) \cdot \underline{y} = 0 \end{cases}$$

Ξέρουμε ότι αν $\nabla g(\underline{x}^*) \cdot \underline{y} = 0$ τότε κ' $\nabla f(\underline{x}^*) \cdot \underline{y} = 0, \ \forall \underline{y}$.
Άρα $\max\limits_{\underline{y}} \nabla f(\underline{x}^*) \cdot \underline{y} = 0$.

Επομένως, αφού έχει λύση, από το δυϊκό θεώρημα του Π.Γ.Π. έχουμε ότι κ' το δυϊκό πρόβλημα έχει λύση (μη κενή)

Δυϊκό : $\begin{cases} \min \lambda^T \cdot 0 \\ \underline{\lambda}^T \nabla g(\underline{x}^*) = \nabla f(\underline{x}^*) \end{cases}$

Οπότε $\nabla f(\underline{x}^*) = \lambda^T \nabla \underline{g}(\underline{x}^*)$.

Τετάρτη 28/11/2012

<u>Θεώρημα 1</u> $\boxed{f \in C^2(\mathbb{R}^n)}$

Έστω $\underline{x}^*$ ένα τοπικό ακρότατο της f υπό τους περιορισμούς $g(\underline{x})=0$. Υποθέτοντας ότι το $\underline{x}^*$ είναι κανονικό σημείο των περιορισμών, τότε $\exists$ $\underline{\lambda} \in \mathbb{R}^n$ τω. $\nabla f(\underline{x}^*) + \underline{\lambda}^T \nabla g(\underline{x}^*) = 0$.

<u>Θεώρημα 2</u>

Αν το $\underline{x}^*$ είναι τοπικό ελάχιστο, τότε $L(\underline{x}^*) = Hf(\underline{x}^*) - \underline{\lambda}^T H \underline{g}(\underline{x}^*)$ είναι θετικά ημιορισμένη στο $\{\underline{y} : \nabla g(\underline{x}^*) \cdot \underline{y} = 0\} = M^{g}_{\underline{x}^*}$

<u>Σημείωση</u>

Για $\underline{g}=(g_1, \ldots, g_m)$ έχουμε $\nabla \underline{g} = \begin{pmatrix} \nabla g_1 \\ \vdots \\ \nabla g_m \end{pmatrix} \in M_{m \times n}$

κ' για κάθε ∇g_i, $i=1,\ldots,m$ μπορούμε να βρούμε την Hessian.

Η πραγματική συνάρτηση $\underline{\lambda}^T \underline{g}$ έχει gradient $\underline{\lambda}^T \nabla \underline{g}$ κ' Hessian $\lambda^T Hg = \sum_{i=1}^{m} \lambda_i Hg_i$.

<u>Απόδειξη 2</u>

Το $\underline{x}^*$ είναι σημείο τοπικού ελαχίστου, οπότε για την παραμετρικοποίηση $\{t \to z(t)\}$, με $\underline{z}(t^*) = \underline{x}^*$, έχουμε $\frac{d^2}{dt^2} f(z(t))\big|_{t=t^*} \geq 0$.

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

$\Rightarrow 0 \leq \frac{d^2}{dt^2} f(\underline{z}(t)) = \dot{\underline{z}}(t)^T Hf(\underline{z}(t)) \cdot \dot{\underline{z}}(t) + \nabla f(\underline{z}(t)) \cdot \ddot{\underline{z}}(t)$ (1)