Δευτέρα 10/12/2012

Οι συνθήκες Karush-Kuhn-Tucker για ΠΜΓΠ με περιορισμούς.

Έστω το πρόβλημα:

$$(1)\begin{cases} \min f(x) \\ h_1(x)=0, \dots, h_m(x)=0. \\ g_1(x)\le 0, \dots, g_p(x)\le 0 \\ x\in\mathbb{R}^n \end{cases}$$

Για να βρούμε μία λύση μπορούμε να εξετάσουμε όλους τους συνδυασμούς ενεργών περιορισμών.

Για κάποιο $\underline{x}^*$, ορίζουμε το σύνολο ενεργών περιορισμών

$J=\{j : g_j(\underline{x}^*)=0\}$

Το $\underline{x}^*$ λέγεται κανονικό σημείο των περιορισμών αν τα ∇h_i, $i=1,\dots,m$, ∇g_j, $j\in J$ είναι γραμμικά ανεξάρτητα.

Θεώρημα

Έστω $\underline{x}^*$ ένα τοπικό ελάχιστο του (1) κ' έστω ότι είναι κανονικό σημείο των περιορισμών.

Τότε υπάρχει $\underline{\lambda}\in\mathbb{R}^m$ κ' $\underline{\mu}\in\mathbb{R}^p$ με $\underline{\mu}\ge 0$ τ.ω.

$\nabla f(\underline{x}^*)+\underline{\lambda}^T\nabla \underline{h}(\underline{x}^*)+\underline{\mu}^T\nabla \underline{g}(\underline{x}^*)=0$ (2) κ'

$\underline{\mu}^T\cdot \underline{g}(\underline{x}^*)=0$ (3)

((2), (3), $\underline{\mu}\ge 0$: συνθήκες πρώτης τάξης)

Σημείωση

Αφού $\underline{\mu}\ge 0$ κ' $\underline{g}(\underline{x}^*)\le 0$, η (3) είναι ισοδύναμη με την εξής

Πρόταση

Μια συνιστώσα του $\underline{\mu}$ μπορεί να είναι θετική, μόνο αν ο αντίστοιχος περιορισμός είναι 0.

Δηλαδή, $g_i(\underline{x}^*) < 0 \Rightarrow \mu_i = 0$.

$\mu_i > 0 \Rightarrow g_i(\underline{x}^*) = 0$.

<u>Σημείωση</u> (συνθήκες 2ης τάξης)

Επίσης, αν οι $f, h, g \in C^2$, ο πίνακας $L(\underline{x}^*) = H f(\underline{x}^*) + \underline{\lambda}^T H \underline{h}(\underline{x}^*) + \underline{\mu}^T H \underline{g}(\underline{x}^*)$ είναι θετικά ημιορισμένος στο εφαπτόμενο επίπεδο των ενεργών περιορισμών στο $\underline{x}^*$: $\{\underline{y} : \nabla \underline{h}(\underline{x}^*) \cdot \underline{y} = 0,\ \nabla g_j(\underline{x}^*) \cdot \underline{y} = 0,\ j \in J\}$

<u>π.χ.</u>

$$\begin{cases} \min\ 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ x_1^2 + x_2^2 \le 5 \\ 3x_1 + x_2 \le 6 \end{cases}$$

$$\nabla f(\underline{x}) + \underline{\mu}^T \nabla g(\underline{x}) = 0 \Rightarrow \begin{cases} 4x_1 + 2x_2 - 10 + \mu_1 2x_1 + \mu_2 \cdot 3 = 0 \\ 2x_1 + 2x_2 - 10 + \mu_1 2x_2 + \mu_2 \cdot 1 = 0 \end{cases}$$

$$\mu_1 \ge 0,\ \mu_2 \ge 0$$

$$\mu_1 \cdot (x_1^2 + x^2 - 5) = 0$$

$$\mu_2 (3x_1 + x_2 - 6) = 0$$

Για την επίλυση εξετάζουμε όλους τους δυνατούς συνδυασμούς κ' βρίσκουμε το πρόσημο των αντίστοιχων πολλ/στών Lagrange.

Υποθέτοντας ότι ο 1ος περιορισμός είναι ενεργός, ενώ ο 2ος όχι, έχουμε $\mu_2 = 0$, οπότε έχουμε: $4x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0$

$$2x_1 + 2x_2 - 10 + 2\mu_1 x_1 = 0$$

$$x_1^2 + x_2^2 - 5 = 0$$

Με λύση: $x_1 = 1,\ x_2 = 2,\ \mu_1 = 1$

που ικανοποιεί κ' το 2ο περιορισμό $(3x_1 + x_2 = 5 < 6)$

Οπότε αφού $\mu_1 = 1 > 0$ ικανοποιούνται οι αναγκαίες συνθήκες.

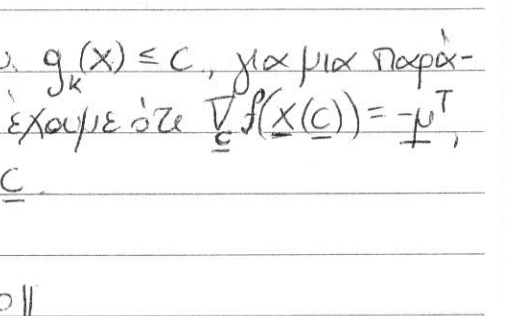
Συνεχίζουμε με τις υπόλοιπες 3 υποθέσεις.

<u>Παρατηρήσεις</u>

1. Η (2) ισχύει με $\mu_j = 0$ αν $g_j(\underline{x}^*) \neq 0$, αφού το $\underline{x}^*$ είναι ακρότατο για όποιους από τους περιορισμούς ισχύει ότι $g(\underline{x}^*) = 0$

2. Γιατί $\underline{\mu} \geq 0$? (Σχέση κανονικού-δυϊκού προβλήματος)
Ιδέα: Έστω ότι $\mu_k < 0$ για κάποιο k.
Αν "χαλαρώσουμε" τον k-οστό περιορισμό τ.ω. $g_k(x) \leq c$, για μια παράμετρο $c \geq 0$, από την ανάλυση ευαισθησίας έχουμε ότι $\nabla_{\underline{c}} f(\underline{x}(\underline{c})) = -\underline{\mu}^T$, όπου $\underline{x}(\underline{c})$ είναι η λύση του $\min f$, $\underline{g}(\underline{x}) \leq \underline{c}$.
$\boxed{\nabla_{\underline{c}} f(\underline{x}(\underline{c})) = \left(\frac{\partial f}{\partial c_1}, \dots, \frac{\partial f}{\partial c_m}\right)}$
Άλλα, $\nabla_{\underline{c}} f(\underline{x}(\underline{c})) \sim \left[f(\underline{x}(\underline{c})) - f(\underline{x}(\underline{0}))\right] / \|\underline{c} - \underline{0}\|$
κ' για την κατεύθυνση k έχουμε ότι $f(\underline{x}(\underline{c})) - f(\underline{x}(\underline{0})) \sim c_k(-\mu_k) > 0$

$\min f(x)$	$\min f(x)$
$g(x) \leq 0$	$g(x) \leq c,\ c > 0$
$\min f(x) = f(x^*) = f(x^*(0))$	$\min f(x) = f(x^*(c))$



Καταλήξαμε σε άτοπο.
Διότι, για $c_k > 0$ έχουμε μεγαλύτερη εφικτή περιοχή σε πρόβλημα ελαχίστου.