B.1.0

Μεμ-293 Θεωρία Βελτιστοποίησης

Εξ αποστάσεως μάθημα: 1

Πέμπτη 2/4/2020, 1700-1900 (zoom meeting)

Γ. Ζουράρης

Version 2

ΤΟ ΠΡΟΒΛΗΜΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ ΜΕ ΠΕΡΙΟΡΙΣΜΟΥΣ

ΓΕΝΙΚΗ ΜΟΡΦΗ

Έστω $f: \mathbb{R}^n \to \mathbb{R}$, ισοτικοί περιορισμοί $h_i : \mathbb{R}^n \to \mathbb{R}$ για $i=1,\dots,m$

και ανισοτικοί περιορισμοί $g_j : \mathbb{R}^n \to \mathbb{R}$ για $j=1,\dots,p$.

Το πρόβλημα βελτιστοποίησης με περιορισμούς, το οποίο θα μας απασχολήσει διατυπώνεται ως εξής: βρές $x^* \in S$ τ.ω.

$$f(x^*) = \min_{x \in S} f(x)$$

όπου

$$S := \left\{ x \in \mathbb{R}^n : h_i(x) = 0 \text{ για } i=1,\dots,m \text{ και } g_j(x) \leq 0 \text{ για } j=1,\dots,p \right\}$$

<u>ΟΡΙΣΜΟΣ:</u> Κάθε $x \in S$ καλείται <u>εφικτό σημείο</u> του προβλήματος βελτιστοποίησης. Το S καλείται <u>εφικτό σύνολο</u> ή <u>εφικτή περιοχή</u> του προβλήματος βελτιστοποίησης.

<u>ΟΡΙΣΜΟΣ.</u> Για κάθε $x \in S$ ορίζουμε το σύνολο δεικτών

$$J(x) := \left\{ j \in \{1,\dots,p\} : g_j(x) = 0 \right\}$$

Όταν $J(x) = \emptyset$ τότε λέμε ότι το εφικτό σημείο x είναι <u>ανενεργό</u> σημείο των ανισοτικών περιορισμών. Όταν $J(x) \neq \emptyset$ τότε λέμε ότι το εφικτό σημείο x είναι <u>ενεργό</u> σημείο των ανισοτικών περιορισμών. Το σύνολο $J(x)$ καλείται <u>σύνολο των ενεργών ανισοτικών περιορισμών του x.</u>

ΟΡΙΣΜΟΣ 'Εστω ότι: $h_i \in C^1(\mathbb{R}^n)$ για $i=1,\dots,m$ και $g_j \in C^1(\mathbb{R}^n)$ για $j=1,\dots,p$.

'Εστω $x \in S$. Λέμε ότι το x είναι ένα ομαλό σημείο των περιορισμών όταν τα διανύσματα: $\{\nabla h_i(x), i=1,\dots,m\} \cup \{\nabla g_j(x), j \in J(x)\}$ είναι γραμμικώς ανεξάρτητα.

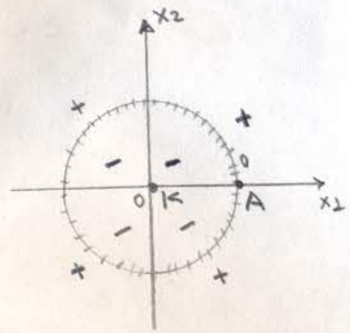
Ας δούμε μερικά παραδείγματα σχετικά με τους παραπάνω ορισμούς.

ΠΑΡΑΔΕΙΓΜΑ 1

$n=2,\ m=0,\ p=1$

$$g_1 : \mathbb{R}^2 \to \mathbb{R} \qquad g_1(x) = x_1^2 + x_2^2 - 1.$$

Τότε: $S = \{x \in \mathbb{R}^2 : g_1(x) \le 0\} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 1\}$



δηλ. η εφικτή περιοχή S είναι ένας κυκλικός δίσκος με ακτίνα 1 και κέντρο το $(0,0)$. Η τιμή $m=0$ σημαίνει ότι δεν υπάρχουν ισοτικοί περιορισμοί.

'Ετσι το σημείο $K=(0,0)$ είναι ανενεργό σημείο των ανισοτικών περιορισμών καθώς: $g_1(K) = -1 < 0$. Το σημείο $A=(1,0)$ είναι ενεργό σημείο των ανισοτικών περιορισμών, καθώς

$g_1(A) = g_1(1,0) = 1^2+0^2-1 = 1-1 = 0$. Εδώ τα ενεργά σημεία των ανισοτικών περιορισμών είναι τα σημεία της περιφέρειας του κύκλου δηλ. τα σημεία του συνόλου:

$$\hat{S} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\} \subsetneq S.$$

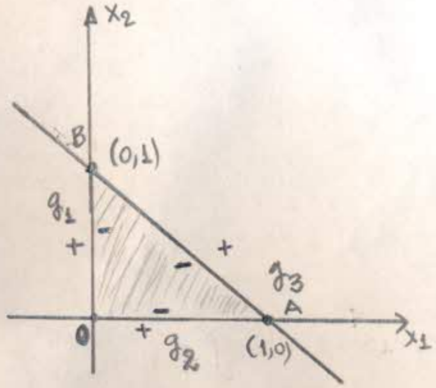
Εύκολα διαπιστώνουμε ότι:

$$\nabla g_1(x) = \left(\partial_{x_1} g_1(x), \partial_{x_2} g(x)\right)$$

$$= (2x_1, 2x_2)$$

$$= 2(x_1, x_2) = 2x$$

Έστω $x \in S$. Όταν το x είναι εσωτερικό σημείο του κυκλικού δίσκου τότε $J(x)=\emptyset$. Σ'αυτή την περίπτωση, αν το x είναι ακρότατο τότε $\nabla f(x)=0$ και επομένως η παρουσία των περιορισμών $(g_i)_{i=1}^{3}$ δεν έχει σημασία. Όταν το x βρίσκεται στην περιφέρεια του κύκλου τότε $x \neq 0$ και επόμενως $\nabla g_1(x)=x\neq 0$, άρα είναι ένα ομαλό σημείο των περιορισμών.



Παράδειγμα 2

$n=2$, $m=0$, $p=3$ $\quad g_1, g_2, g_3 : \mathbb{R}^2 \to \mathbb{R}$ με:

$$g_1(x) = -x_1 \qquad g_2(x) = -x_2 \qquad g_3(x) = x_1 + x_2 - 1$$

Επομένως η εφικτή περιοχή S είναι το τρίγωνο OAB. Τα σημεία O, A, B είναι ενεργά σημεία των ανισοτικών περιορισμών καθώς:

$g_1(0) = 0,\ g_2(0) = 0,\ g_3(0) = -1 < 0$ $\qquad J(0) = \{1,2\}$

$g_2(A) = g_2(1,0) = 0,\ g_3(A) = g_3(1,0) = 1+0-1 = 0,\ g_1(A) = g_1(1,0) = -1 < 0$ $\qquad J(A) = \{2,3\}$

$g_3(B) = g_3(0,1) = 0+1-1 = 0,\ g_1(B) = g_1(0,1) = 0,\ g_2(B) = g_2(0,1) = -1 < 0.$ $\qquad J(B) = \{1,3\}$

Επειδή τα O, A, B είναι κορυφές του τριγώνου το αντίστοιχο σύνολο ενεργών περιορισμών είναι διμελές. Αν Γ είναι κάποιο σημείο σε μία πλευρά, χωρίς να είναι κορυφή, τότε το $J(\Gamma)$ είναι μονομελές. π.χ. $\Gamma = (1/2, 0)$, τότε: $g_1(\Gamma) = g_1(1/2, 0) = -\frac{1}{2} < 0$,
$g_2(\Gamma) = g_2(1/2, 0) = 0$, $g_3(\Gamma) = g_3(1/2, 0) = \frac{1}{2} + 0 - 1 = -\frac{1}{2} < 0$, και επομένως $J(\Gamma) = \{2\}$.

Το σημείο $(1/3, 1/3)$ ανήκει στην εφικτή περιοχή S καθώς

$$g_1(1/3, 1/3) = -\frac{1}{3} < 0 \quad , \quad g_2(1/3, 1/3) = -\frac{1}{3} < 0 , \quad g_3(1/3, 1/3) = \frac{1}{3} + \frac{1}{3} - 1 = -\frac{1}{3} < 0$$

όμως $J(1/3, 1/3) = \phi$, καθώς όλοι οι ανισοτικοί περιορισμοί είναι μη ενεργοί.

Παρατηρώντας ότι $g_1, g_2, g_3 \in C^1(\mathbb{R}^2)$ υπολογίζουμε στη συνέχεια τα $(\nabla g_j)_{j=1}^{3}$.

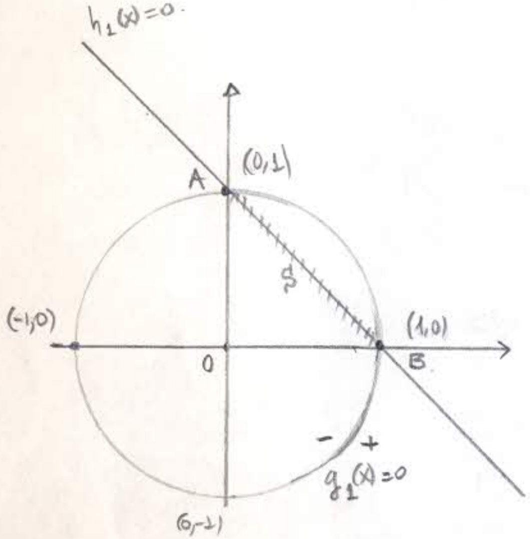
$$\nabla g_1(x) = (-1, 0) \quad , \quad \nabla g_2(x) = (0, -1) \quad , \quad \nabla g_3(x) = (1, 1)$$

Έστω $x \in S$. Αν είναι ενεργό σημείο των ανισοτικών περιορισμών τότε $J(x)$ έχει 1 ή 2 στοιχεία. Επειδή τα $\nabla g_1(x), \nabla g_2(x), \nabla g_3(x)$ είναι μη μηδενικά και ανά δύο γραμμικά ανεξάρτητα τότε τα $\{\nabla g_k(x) : k \in J(x)\}$ είναι γραμμικώς ανεξάρτητα επομένως το x είναι ομαλό σημείο των περιορισμών. Αν το x είναι ανενεργό σημείο των περιορισμών, κάτι που συμβαίνει όταν το x είναι εσωτερικό σημείο του τριγώνου, τότε $J(x) = \phi$. Άρα οι ανισοτικοί περιορισμοί δεν έχουν σημασία. Επιπλέον, όταν το x είναι ακρότατο της f, τότε $\nabla f(x) = 0$. ∎

ΠΑΡΑΔΕΙΓΜΑ 3:

$n=2$, $m=1$, $p=1$, $h_1, g_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ με:

$$h_1(x) = x_1+x_2-1 \qquad g_1(x) = x_1^2+x_2^2-1.$$

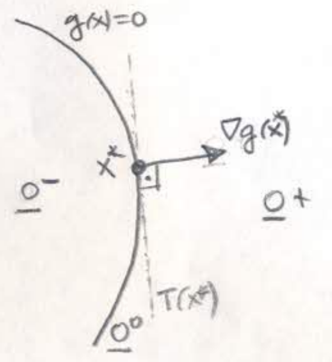


Εδώ η εφικτή περιοχή $\mathcal{S}$ είναι το ευθύγραμμο τμήμα AB, δηλ. το τμήμα της ευθείας $h_1(x)=0$ το οποίο βρίσκεται εντός του κυκλικού δίσκου $g_1(x)\leq 0$. Τα σημεία A, B είναι τα μόνα σημεία του $\mathcal{S}$ τα οποία είναι ενεργά σημεία των ανισοτικών περιορισμών.

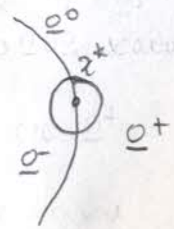
ΠΑΡΑΤΗΡΗΣΗ: Έστω ότι $g \in C^1(\mathbb{R}^n,\mathbb{R})$. Η απαίτηση $g(x)=0$ ορίζει μία υπερεπιφάνεια στο $\mathbb{R}^n$ π.χ. μία καμπύλη όταν $n=2$, μία επιφάνεια όταν $n=3$. Ουσιαστικά χωρίζει τον $\mathbb{R}^n$ σε τρία σύνολα, ξένα μεταξύ τους:

$$\underline{O}^+ := \{x\in\mathbb{R}^n : g(x)>0\}, \quad \underline{O}^0 := \{x\in\mathbb{R}^n : g(x)=0\}, \quad \underline{O}^- := \{x\in\mathbb{R}^n : g(x)<0\}$$

Έστω $x^*\in \underline{O}^0$. Τότε είναι σημείο ολικού ελαχίστου για την g στο $\underline{O}^+\cup\underline{O}^0$, και σημείο ολικού μεγίστου της g στο $\underline{O}^-\cup\underline{O}^0$.



Όταν $d\in\mathbb{R}^n$ ($d\neq 0$) είναι μία εφικτή διεύθυνση του x^* διεύθυνση προς το $\underline{O}^+\cup\underline{O}^0$ έχουμε: $\langle\nabla g(x^*),d\rangle\ge 0$ και όταν είναι εφικτή διεύθυνση του x^* προς το $\underline{O}^-\cup\underline{O}^0$ τότε: $\langle \nabla g(x^*),d\rangle\le 0$.

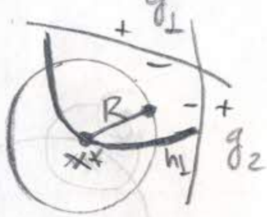


Έστω: $\nabla g(x^*)\neq 0$. Ξέρουμε ότι είναι κάθετο στο εφαπτόμενο επίπεδο της $g(x)=0$ στο x^*, επομένως "δείχνει" προς το $\underline{O}^+$ ή $\underline{O}^-$. Επειδή: $(\nabla g(x^*),\nabla g(x^*)) = |\nabla g(x^*)|^2>0$ συμπεραίνουμε ότι "δείχνει" προς το $\underline{O}^+$ δηλ. είναι εφικτή διεύθυνση του x^* προς το $\underline{O}^+$. Ανάλογα, το $-\nabla g(x^*)$ δείχνει προς το $\underline{O}^-$ δηλ. είναι εφικτή διεύθυνση του x^* προς το $\underline{O}^-$. ■

Έστω ότι $x^* \in S$ είναι σημείο τοπικού ελαχίστου για την f το οποίο, επιπλέον, είναι ομαλό σημείο των περιορισμών. Έχουμε ήδη δεί την περίπτωση $p=0$, δηλ. όταν έχουμε μόνο ισοτικούς περιορισμούς (Πολλαπλασιαστές Lagrange). Έτσι εδώ υποθέτουμε ότι $p \geq 1$.

Περίπτωση 1 $J(x^*) = \phi$

Αυτό σημαίνει ότι: $g_j(x^*) < 0$, $j = 1, \dots, p$, δηλ. το x^* είναι ανενεργό σημείο των ανισοτικών περιορισμών. Επειδή οι συναρτήσεις $(g_j)_{j=1}^{p}$ είναι συνεχείς στο $\mathbb{R}^n$ υπάρχει μια περιοχή του x^* όπου οι $(g_j)_{j=1}^{p}$ διατηρούν το πρόσημο, δηλ. υπάρχει $R>0$ τ.ω $g_j(x) < 0$ για κάθε $x \in \mathbb{R}^n$ με $|x - x^*| \leq R$. Στην περίπτωση $(m=0)$ όπου δεν υπάρχουν ισοτικοί περιορισμοί αυτό σημαίνει ότι το x^* είναι εσωτερικό σημείο του S, άρα $\nabla f(x^*) = 0$.



Όταν υπάρχουν ισοτικοί περιορισμοί $(m \geq 1)$, τότε οι ανισοτικοί περιορισμοί δεν έχουν σημασία δηλ. το x^* μπορεί θεωρηθεί τοπικό ελάχιστο της f υπό τους περιορισμούς $h_i(x) = 0$, για $i = 1, \dots, m$, άρα υπάρχουν $(\lambda_i)_{i=1}^{m} \subset \mathbb{R}$ τ.ω $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) = 0$. Αν $\nabla f(x^*) = 0$, τότε $\lambda_i = 0$, $i = 1, \dots, m$, επειδή τα $(\nabla h_i(x^*))_{i=1}^{m}$ είναι γραμμικώς ανεξάρτητα.

Περίπτωση 2. $J(x^*) \neq \emptyset$

Τότε οι ανισοτικοί περιορισμοί $(g_k)_{k \notin J(x^*)}$ δεν έχουν σημασία, δηλ. μπορούμε να θεωρήσουμε

το x^* ως τοπικό ελάχιστο της f υπό τους ισοτικούς περιορισμούς:

$$h_i(x)=0,\ i=1,\dots,m,\quad g_k(x)=0\ \ \forall k\in J(x^*)$$

Επομένως υπάρχουν $(\lambda_i)_{i=1}^m \subset \mathbb{R}$ και $(\mu_k)_{k\in J(x^*)} \subset \mathbb{R}$ τ.ω

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{k\in J(x^*)} \mu_k \nabla g_k(x^*) = 0$$

Επειδή τα $V := \{ (\nabla h_i(x^*))_{i=1}^m, (\nabla g_k(x^*))_{k\in J(x^*)} \}$ είναι γραμμικώς ανεξάρτητα, όταν

$\nabla f(x^*)=0$ τότε $\lambda_i=0$, για $i=1,\dots,m$, και $\mu_k=0$ για κάθε $k\in J(x^*)$.

Επιπλέον, ισχύει ότι $\mu_k \geq 0$, για κάθε $k\in J(x^*)$. Ας προσπαθήσουμε να το εξηγήσουμε

αυτό με απαγωγή σε άτοπο. Έστω ότι έχουμε ένα μη κενό υποσύνολο $I(x^*) \subset J(x^*)$ όπου $\mu_k<0$ για

κάθε $k\in I(x^*)$. Τότε μπορούμε να βρούμε $v\in \operatorname{span}(V)$ τ.ω

$$(\nabla h_i(x^*), v) = 0 \quad i=1,\dots,m$$
$$(\nabla g_k(x^*), v) = 0 \quad \forall k\in J(x^*)\setminus I(x^*)$$
$$(\nabla g_k(x^*), v) < 0 \quad \forall k\in I(x^*)$$

ο πίνακας είναι θ.ο. άρα σύστημα έχει λύση

Επειδή το v είναι κάθετο στο εφαπτόμενο επίπεδο των $(h_i)_{i=1}^{m}$ στο x^*

έχουμε δείξει ότι θα είναι κάθετο στο $\nabla f(x^*)$. Άρα:

$$\underbrace{(\nabla f(x^*), v)}_{=0} + \sum_{i=1}^{m} \lambda_i \underbrace{(\nabla h_i(x^*), v)}_{=0} + \sum_{k\in J(x^*)\setminus I(x^*)} \mu_k \underbrace{(\nabla g_k(x^*), v)}_{=0} = \underbrace{-\sum_{k\in J(x^*)\setminus I(x^*)} \mu_k (\nabla g_k(x^*), v)}_{<0}$$

Έτσι καταλήξαμε σε άτοπο.

Μ' αυτό τον τρόπο καταλήγουμε στις αναγκαίες συνθήκες Kuhn-Tucker ή Karush-Kuhn-Tucker

$1^{ης}$ τάξης

<u>Θεώρημα</u> Έστω: $f, (h_i)_{i=1}^{m}, (g_j)_{j=1}^{p} \in C^1(\mathbb{R}^n, \mathbb{R})$ και $S = \left\{ x\in\mathbb{R}^n : \begin{matrix} h_i(x)=0 & i=1,\dots,m \\ g_j(x)\le 0, & j=1,\dots,p \end{matrix} \right\}$

<u>Αν</u> $x^* \in S$ είναι ένα τοπικό ελάχιστο της f στο S και ομαλό σημείο των περιορισμών, <u>τότε</u> υπάρχουν $(\lambda_i)_{i=1}^{m} \subset \mathbb{R}$ και $(\mu_i)_{i=1}^{p} \subset [0,+\infty)$ τ.ω.

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(x^*) = 0.$$

Και:

$$\mu_j\, g_j(x^*) = 0 \;, \quad j = 1, \dots, p.$$

<u>Σημ.</u> Όταν ο περιορισμός g_j είναι ανενεργός, δηλ. $g_j(x^*)<0$, τότε: $\mu_j = 0$. Μ' αυτό τον τρόπο ενοποιούμε όλες τις περιπτώσεις που εξετάσαμε.

Θεωρημα (αναγκαίες συνθήκες $2^{ης}$ τάξης)

Έστω: $f, (g_j)_{j=1}^{p}, (h_i)_{i=1}^{m} \in C^2(\mathbb{R}^n; \mathbb{R})$ και $S = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} h_i(x) = 0 & i=1,\dots,m \\ g_j(x) \leq 0 & j=1,\dots,p \end{array} \right\}$

Αν $x^* \in S$ ομαλό σημείο των περιορισμών είναι σημείο τοπικού ελαχίστου της f στο S τότε

υπάρχουν $(\lambda_j)_{j=1}^{m} \subset \mathbb{R}$ και $(\mu_k)_{k=1}^{p} \subset [0, +\infty)$ τ.ω.

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) + \sum_{j=1}^{p} \mu_j \nabla g_j(x^*) = 0.$$

και $$\mu_j\, g_j(x^*) = 0 \quad , j=1,\dots,p.$$

Επιπλέον έχουμε ότι ο πίνακας

$$\Lambda(x^*) := \mathcal{H}f(x^*) + \sum_{i=1}^{m} \lambda_i \mathcal{H}h_i(x^*) + \sum_{j=1}^{p} \mu_j \mathcal{H}g_j(x^*)$$

είναι θετικά ημιορισμένος στο εφαπτόμενο επίπεδο που ορίζεται από τους ισοτικούς περιορισμούς και τους ενεργούς ανισοτικούς περιορισμούς δηλ.

$$M(x^*) = \left\{ y \in \mathbb{R}^n : \begin{array}{ll} (\nabla h_i(x^*), y) = 0 & i=1,\dots,m \\ (\nabla g_k(x^*), y) = 0 & k \in J(x^*) \end{array} \right\}$$

Σημ: Προκύπτει ως συνέπεια των αναγκαίων συνθηκών $2^{ης}$ τάξης στην περίπτωση $p=0$.

ΘΕΩΡΗΜΑ (Ικανές 2ης τάξης)

Έστω $f, (h_i)_{i=1}^m, (g_j)_{j=1}^p \in C^2(\mathbb{R}^n, \mathbb{R})$ και $S = \left\{ x \in \mathbb{R}^n : \begin{array}{ll} h_i(x) = 0 & i = 1, \dots, m \\ g_j(x) \le 0 & j = 1, \dots, p \end{array} \right\}$.

Έστω $x^* \in S$. Το x^* είναι σημείο αυστηρού τοπικού ελαχίστου της f στο S όταν:

υπάρχουν $(\lambda_j)_{j=1}^m \subset \mathbb{R}$ και $(\mu_k)_{k=1}^p \subset [0, +\infty)$ τ.ω.

$$\nabla f(x^*) + \sum_{j=1}^m \lambda_j \nabla h_j(x^*) + \sum_{k=1}^p \mu_k \nabla g_k(x^*) = 0$$

$$\mu_j \, g_j(x^*) = 0 \qquad j = 1, \dots, p$$

και ο πίνακας

$$\Lambda(x^*) = \mathcal{H}f(x^*) + \sum_{j=1}^m \lambda_j \mathcal{H}h_j(x^*) + \sum_{k=1}^p \mu_k \mathcal{H}g_k(x^*)$$

είναι θ.ο. στο

$$M(x^*) = \left\{ y \in \mathbb{R}^n : \begin{array}{l} (\nabla h_i(x^*), y) = 0, \ i = 1, \dots, m \\ (\nabla g_j(x^*), y) = 0 \quad \forall j \in \tilde{J}(x^*) \end{array} \right\}$$

όπου: $\tilde{J}(x^*) = \{ j \in \{1, \dots, p\} : \mu_j > 0 \text{ και } g_j(x^*) = 0 \}$