SCHEMA SULLE CONDIZIONI DI OTTIMALITA`

C. NECESSARIE

Se $\bar{x} \in F$ e' minimo locale ALLORA

NON VINCOLATO $F = \mathbb{R}^n$	VINCOLATO $F = \{x \in \mathbb{R}^n \mid g(x) \le 0,\ h(x) = 0\}$
PRIMO ORDINE $\nabla f(x) = 0$	• $\bar{x}$ e' un punto di F-J • Se valgono ipotesi di regolarita` sui vincoli (g,h lineari, MFCQ, LICQ) $\Rightarrow$ $\bar{x}$ e' punto di KKT
SECONDO ORDINE $\nabla f(\bar{x}) = 0$ $\nabla^2 f(\bar{x})$ e' semidefinita positiva $(d^T \nabla^2 f(\bar{x}) d \ge 0 \quad \forall x \in \mathbb{R}^n)$	Sotto l'ipotesi di LICQ $\bar{x}$ e' punto KKT e $d^T \nabla^2 \mathcal{L}(\bar{x}, \lambda, \mu) d \ge 0 \quad \forall d \in T(\bar{x}, \lambda)$ insieme delle direzioni ammissibili che fa int. i vincoli strettamente attivi

C. SUFFICIENTI

NON VINCOLATO $F = \mathbb{R}^n$	VINCOLATO $F = \{x \in \mathbb{R}^n \mid g(x) \le 0,\ h(x) = 0\}$
$\nabla f(\bar{x}) = 0$ $d^T \nabla^2 f(\bar{x}) d > 0 \quad \forall d \ne 0$ $\Updownarrow$ $\nabla^2 f(\bar{x})$ e' definita positiva	$\bar{x}$ e' punto di F-J e $d^T \nabla^2 \tilde{\mathcal{L}}(\bar{x}, \lambda_0, \lambda, \mu) d > 0$ $\forall d \in T(\bar{x}, \lambda)$; $d \ne 0$

Allora $\bar{x} \in F$ e' punto di minimo locale (stretto)

N.B per avere condizioni sufficienti e' necessario utilizzare info del 2° ordine

Esempio:

$$\min \; x_1^2 + x_2^2 - 4x_1 - 4x_2$$
$$x_1^2 \leq x_2 \qquad (P)$$
$$x_1 + x_2 \leq 2$$

Le condizioni KKT per (P) portano al sistema

$$\begin{cases} 2x_1 + 2\lambda x_1 + \lambda_2 - 4 = 0 \\ 2x_2 - \lambda_1 + \lambda_2 - 4 = 0 \\ \lambda_1 \geq 0, \; \lambda_2 \geq 0 \\ \lambda_1 (x_1^2 - x_2) = 0 \; ; \; \lambda_2 (x_1 + x_2 - 2) = 0 \\ (x_1^2 - x_2) \leq 0 \; ; \; (x_1 + x_2 - 2) \leq 0 \end{cases}$$

$\longrightarrow$ Si può verificare che $(\bar{x}, \bar{\lambda}) = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right)$ è KKT

$$\begin{cases} \lambda_1 = 0 \\ \lambda_2 = 2 > 0 \end{cases}$$

Vogliamo verificare che $(\bar{x}, \bar{\lambda}) = \left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right)$ sia minimo locale per (P) sfruttando le condizioni sufficienti.

$$F = \left\{ x \in \mathbb{R}^2 \;\middle|\; \begin{matrix} x_1^2 \leq x_2 \\ x_1 + x_2 \leq 2 \end{matrix} \right\} \qquad \begin{matrix} g_1(x) = x_1^2 - x_2 \\ g_2(x) = x_1 + x_2 - 2 \end{matrix}$$

$\bar{I}(\bar{x}) = \{2\}$

poiché $g_2(\bar{x}) = 0$ e $\lambda_2 > 0$

$|\bar{I}(\bar{x})| = 1 < 2 = n$ — dimensione del problema

ho un vincolo attivo

cardinalità dell'insieme dei vincoli strettamente attivi

$\Rightarrow$ Non posso applicare il corollario perché non ho soddisfatte le ipotesi (n vettori l.i. tra $\nabla g_i(\bar{x})$ con $i \in \bar{I}(x)$)

Cerchiamo di verificare le ipotesi del teorema:

$(\bar{x}, \bar{\lambda})$ KKT $\Rightarrow$ $(\bar{x}, \lambda_0 = 1, \bar{\lambda})$ è un punto di F-J

Vogliamo verificare

$$d^T \nabla^2 \tilde{\mathcal{L}}(\bar{x}, \lambda_0, \bar{\lambda}) d > 0 \qquad \forall d \in T(\bar{x}, \bar{\lambda}) \text{ con } d \neq 0$$

con $\tilde{\mathcal{L}}(x, \lambda_0, \lambda) = \lambda_0 f(x) + \sum_{i=1}^{m} \lambda_i g_i(x)$

$$\nabla_x \tilde{\mathcal{L}}(x, \lambda_0, \lambda) =$$

Ricordiamo che $T(\bar{x},\bar{\lambda}) = \left\{ d \in \mathbb{R}^2 \;\middle|\; \begin{array}{ll} \nabla g_i(x)^T d = 0 & \forall i \in \bar{I}(\bar{x}) \\ \nabla g_i(x)^T d \le 0 & \forall i \in I(\bar{x}) \setminus \bar{I}(\bar{x}) \end{array} \right\}$

$\bar{I}(\bar{x}) = \{2\} \quad \bar{I}(\bar{x}) = \{1\}$

$$= \left\{ d \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} \nabla g_2(\bar{x})\, d = 0 \\ \nabla g_1(\bar{x})^T d \le 0 \end{array} \right\} = \left\{ d \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} 2d_1 - d_2 = 0 \\ d_1 + d_2 \le 0 \end{array} \right\} =$$

$$= \left\{ d \in \mathbb{R}^2 \;\middle|\; \begin{array}{l} d_2 = 2d_1 \\ d_1 + 2d_2 \le 0 \end{array} \right\} = \left\{ \begin{pmatrix} d_1 \\ 2d_1 \end{pmatrix} ;\ d_1 \le 0 \right\}$$

Calcoliamo $\nabla^2 \tilde{\mathcal{L}}(\bar{x}, \lambda_0 = 1, \bar{\lambda})$

$$\nabla_x \tilde{\mathcal{L}}(x, \lambda_0, \lambda) = \begin{pmatrix} 2\lambda_0(x_1 - 2) + 2\lambda_1 x_1 + \lambda_2 \\ 2\lambda_0(x_2 - 2) - \lambda_1 + \lambda_2 \end{pmatrix}$$

$$\nabla^2 \tilde{\mathcal{L}}(x, \lambda_0, \lambda) = \begin{pmatrix} 2\lambda_0 + 2\lambda_1 & 0 \\ 0 & 2\lambda_0 \end{pmatrix}$$

$$\nabla^2 \tilde{\mathcal{L}}(\bar{x}, \lambda_0 = 1, \bar{\lambda}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \implies \nabla^2 \tilde{\mathcal{L}}(\bar{x}, \lambda_0 = 1, \bar{\lambda}) \text{ e' definita positiva su } \mathbb{R}^2$$

$$\implies d^T \nabla^2 \mathcal{L}(\bar{x}, \lambda_0 = 1, \bar{\lambda})\, d > 0 \quad \forall d \in T(\bar{x}, \bar{\lambda}),\ d \neq 0$$

ALGORITMI P.O VINCOLATA

Cercare di determinare DIREZIONI AMMISSIBILI DI DISCESA (cioè direzioni che riescono a far diminuire la funzione obiettivo rispettando i vincoli)

Metodo delle penalità sequenziali

Risolvere una successione di problemi di ottimizzazione NON VINCOLATA in cui la fu. obiettivo tenga conto in qualche modo dei vincoli del prob. di partenza

TEOREMA:

Sia $f, g, h \in C^1(\mathbb{R}^n)$

$f: \mathbb{R}^n \to \mathbb{R}$; $g: \mathbb{R}^n \to \mathbb{R}^m$; $h: \mathbb{R}^n \to \mathbb{R}^p$

$$\min f(x) \quad g(x) \le 0 \quad h(x) = 0$$

Sia $\bar{x} \in F$. Se $\bar{x}$ NON è punto di F-J allora i seguenti insiemi:

$$D_s(\bar{x}) = \{ d \in \mathbb{R}^n \mid \nabla f(\bar{x})^T d < 0 \}$$

$$F_s(\bar{x}) = \{ d \in \mathbb{R}^n \mid \nabla g_i(\bar{x})^T d < 0 ; i \in I(\bar{x}) \}$$

$$H_s(\bar{x}) = \{ d \in \mathbb{R}^n \mid \nabla h_j(\bar{x})^T d = 0 ; j = 1, \dots, p \}$$

hanno intersezione non vuota ($\neq \emptyset$)

TEOREMA

$f, g_i, h_j \in C^1(\mathbb{R}^n)$ $\quad i = 1, \dots, m ; j = 1, \dots, p$

Sia $\bar{x} \in F$ t.c $D_s(\bar{x}) \cap F_s(\bar{x}) \cap H_s(\bar{x}) \neq \emptyset$

Allora $\forall d \in D_s(\bar{x}) \cap F_s(\bar{x}) \cap H_s(\bar{x})$

$\exists \; \delta > 0$ t.c $\forall \alpha \in (0, \delta)$ si ha

$$f(\bar{x} + \alpha d) < f(\bar{x})$$

$$g_i(\bar{x} + \alpha d) \le 0 \qquad i = 1, \dots, m$$

$$h_j(\bar{x} + \alpha d) = r_{h_j}(\bar{x}, \alpha d) \qquad j = 1, \dots, p$$

con $\lim_{\alpha \to 0} \frac{r_{h_j}(\bar{x}, \alpha d)}{\alpha} = 0$ e $r_{h_j}(\bar{x}, \alpha d) = 0$ se h_j è lineare

DIMOSTRAZIONE:

Sfruttiamo il t. della media

- $f(\bar{x}+\alpha d) = f(\bar{x}) + \alpha \left(\underbrace{\nabla f(\bar{x})^T d}_{<0} + \frac{\overbrace{r_f(\bar{x},\alpha d)}^{0 \;\alpha\to 0}}{\alpha} \right)$

  con $\lim_{\alpha \to 0} \frac{r_f(\bar{x},\alpha d)}{\alpha} = 0$ poiché $d \in D_s(\bar{x})$

$\Rightarrow \exists\, \delta_1 > 0$ t.c $\forall \alpha \in (0,\delta_1)$ $f(\bar{x}+\alpha d) < f(\bar{x})$

- $h_j(\bar{x}+\alpha d) = \underbrace{\cancel{h_j(\bar{x})}}_{=0} + \alpha \underbrace{\cancel{\nabla h_j(\bar{x})^T d}}_{=0} + r_{h_j}(\bar{x},\alpha d)$ $j = 1,\dots,p$

  poiché $\bar{x} \in F$ poiché $d \in H_s(\bar{x})$

Per $g_i(\bar{x}+\alpha d)$ distinguiamo due casi:

- Se $i \in I(\bar{x})$

  $g_i(\bar{x}+\alpha d) = \overbrace{g_i(\bar{x})}^{0} + \alpha \left(\underbrace{\nabla g_i(\bar{x})^T d}_{<0} + \frac{r_{g_i}(\bar{x},\alpha d)}{\alpha} \right)$ poiché $i \in I(x)$

  con $\lim_{\alpha\to 0} \frac{r_{g_i}(\bar{x},\alpha d)}{\alpha} = 0$ poiché $d \in F_s(\bar{x})$

  Quindi $\exists\, \delta_2 > 0$ t.c $\forall \alpha \in (0,\delta_2)$

  $g_i(\bar{x}+\alpha d) < 0$

- Sia $i \notin I(\bar{x}) \Rightarrow g_i(\bar{x}) < 0$

  $g_i(\bar{x}+\alpha d) = g_i(\bar{x}) + \alpha \left(\nabla g_i(\bar{x})^T d + \frac{r_{g_i}(\bar{x},\alpha d)}{\alpha} \right)$

  $\exists\, \delta_3 > 0$ t.c $\forall \alpha \in (0,\delta_3)$

  $|g_i(\bar{x})| > \left| \alpha \left(\nabla g_i(\bar{x})^T d + \frac{r_{g_i}(\bar{x},\alpha d)}{\alpha} \right) \right|$

  Quindi

  $g_i(\bar{x}+\alpha d) \le 0$

Scegliendo $\delta = \min\{\delta_1, \delta_2, \delta_3\}$ ho la mia tesi del teorema ☒