PROB. DI OTTIMIZZAZIONE CON INSIEME AMMISSIBILE CONVESSO

$\min f(x)$
$x \in F$
$F \subseteq \mathbb{R}^n$
convesso

- Dato $\bar{x} \in F$, $d = x - \bar{x}$ con $x \in F$ è una direzione ammissibile

- $x^* \in F$ minimo locale $\Rightarrow \nabla f(x^*)^T (x - x^*) \geq 0 \quad \forall x \in F$

NON in programma

METODO DI FRANK-WOLFE

Supponiamo che $F \subseteq \mathbb{R}^n$ compatto e convesso, $f \in C^1(\mathbb{R})$

Il metodo di F-W ad ogni iterazione k genera a partire dal punto $x^k \in F$ una direzione ammissibile di discesa in x^k, risolvendo

$$d^k = \begin{cases} d^k = \hat{x} - x^k & \text{(dir. ammissibile)} \\ \nabla f(x^k)^T d^k < 0 & \text{(dir. di discesa)} \end{cases}$$

$$(P^k) \quad \min_{x \in F} \nabla f(x^k)^T (x - x^k)$$

$$\nabla f(x^k)^T (x - x^k) = \underbrace{\nabla f(x^k)^T x}_{c^T x} - \underbrace{\nabla f(x^k)^T x^k}_{d} \quad \text{è lineare!}$$

(P^k) ha f.o lineare
F convesso e compatto
$\Rightarrow (P^k)$ ammette soluzione per Weierstrass!

Sia $\hat{x}^k$ la soluzione del problema (P^k) $\min_{x \in F} \nabla f(x^k)^T (x - x^k)$ $P(x)$

Se $\hat{x}^k$ è t.c. $p(\hat{x}^k) = 0$ si ha che

$$0 = \nabla f(x^k)^T (\hat{x}^k - x^k) \leq \nabla f(x^k)^T (x - x^k) \quad \forall x \in F$$

poiché $\hat{x}^k$ è il punto di minimo globale di (P^k)

$\rightarrow P(\hat{x}^k) \leq P(x)$

posso concludere che x^k è un punto che $x^k \in F$ soddisfa le condizioni necessarie di ottimalità per problemi con insieme ammissibile convesso.

Altrimenti se $\nabla f(x^k)^T (\hat{x}^k - x^k) \neq 0 \quad (< 0)$

si definisce $d^k = \hat{x}^k - x^k$ come direzione ammissibile di discesa

e il nuovo punto x^{k+1} sarà definito come $x^{k+1} = x^k + \alpha^k d^k$

con $\alpha^k \in (0,s)$ che realizza $f(x^{k+1}) < f(x^k)$

possiamo determinarla attraverso una procedura di linesearch.

SCHEMA del METODO DI FRANK-WOLFE

1) $x^0 \in F$

2) $\hat{x}^k := \underset{x \in F}{\text{argmin}}\ \nabla f(x^k)^T(x - x^k)$

punto che risolve il problema

$$\min_{x \in F} \nabla f(x^k)^T(x - x^k)$$

valore f.o

3) se $\nabla f(x^k)^T(\hat{x}^k - x^k) = 0$ STOP (x^k e' un punto stazionario)

4) Altrimenti $d^k = \hat{x}^k - x^k$

5) Calcolare $\alpha^k > 0$ t.c $f(x^k + \alpha^k d^k) < f(x^k)$

6) Porre $x^{k+1} = x^k + \alpha^k d^k$, $k = k+1$

Tornare al passo 2

TEOREMA:

Sia $f: \mathbb{R}^n \to \mathbb{R}$, $f \in C^1(\mathbb{R}^n)$, $F \subseteq \mathbb{R}^n$ compatto e convesso

Sia $\{x^k\} \subseteq F$ la successione di punti prodotta dal metodo di Frank-Wolfe

Allora, o $\exists$ un indice $\bar{k} \geq 0$ t.c $x^{\bar{k}}$ e' un punto stazionario, oppure viene prodotta una successione infinita t.c ogni punto di accumulazione di $\{x^k\}$ e' un punto stazionario.

OSS:

Se f e' convessa ogni punto di accumulazione di $\{x^k\}$ e' un punto di minimo globale di $\min_{x \in F} f(x)$

CASO PARTICOLARE POLIEDRO

$F = \{ x \in \mathbb{R}^n \mid Ax \le b \}$

$$A = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{pmatrix} \qquad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

F poliedro $\Rightarrow$ (P^k) $\quad \min \nabla f(x^k)^T (x - x^k)$

$$x \in F$$

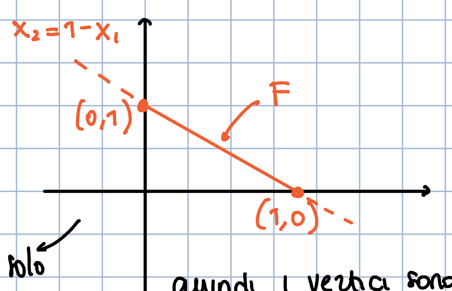
è un problema di programmazione lineare

Consideriamo un poliedro particolare : **IL SIMPLESSO STANDARD**

$F = \{ x \in \mathbb{R}^n \mid \underbrace{e^T x = 1}_{\parallel} ; \ x \ge 0 \}$ con $e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$

$$\sum_{i=1}^{n} x_i = 1 \ = \ x_1 + x_2 \ldots + x_n = 1$$

• in $\mathbb{R}^2$ $\quad F = \{ x \in \mathbb{R}^2 \mid x_1 + x_2 = 1, \ x_1 \ge 0, \ x_2 \ge 0 \}$



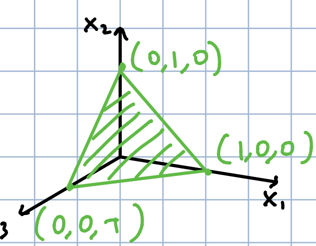
Il poliedro è solo il segmento

quindi i vertici sono $(0,1), (1,0)$

• in $\mathbb{R}^3$

$F = \{ x \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 1, \ x_i \ge 0, \ i = 1,2,3 \}$

I vertici sono $(1,0,0), (0,1,0), (0,0,1)$



Si può verificare che i vertici del simplesso standard

$$\{ x \in \mathbb{R}^n \mid e^T x = 1, \ x \ge 0 \}$$

sono i vettori della base canonica $e_i \in \mathbb{R}^n \quad i = 1, \ldots, n$

$$\underset{\uparrow e_1}{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}, \ \underset{\uparrow e_2}{\begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}} \ \ldots \ \underset{\uparrow e_n}{\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}}$$

Supponiamo di applicare il metodo di F-W ad un problema $\min f(x)$

$x \in F$ con F simplesso standard

Il sottoproblema (P^k) e' dato da

$$\left.\begin{array}{l}\min \nabla f(x^k)^T(x-x^k) \\ x \in \underbrace{F}_{\text{simplesso standard}}\end{array}\right] \text{prob di P.L}$$

Essendo (P^k) un problema di programmazione concava, i punti di minimo globale si trovano sulla frontiera di F (e in particolare sui vertici)

lo dimostreremo

Per risolvere (P^k) e' sufficiente valutare la funzione obiettivo

$$\nabla f(x^k)^T(x-x^k) \quad \text{sui vettori } e_i \in \mathbb{R}^n \quad i=1,\dots,n$$

$$\nabla f(x^k)^T(e_i - x^k) = \nabla f(x^k)^T e_i - \nabla f(x^k)^T x^k = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x^k) \\ \frac{\partial f}{\partial x_2}(x^k) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x^k) \end{pmatrix}^T \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} - \nabla f(x^k)^T x^k =$$

$\to$ i-esima componente

$$= \frac{\partial f}{\partial x_i}(x^k) - \nabla f(x^k)^T x^k$$

$\uparrow$ dipende dal vertice e_i

La soluzione di (P^k) sara' il vertice di F (simplesso standard) per $\frac{\partial f}{\partial x_i}(x^k)$ e' minimo

ovvero

$$\hat{x}^k = \underbrace{e_{\hat{i}}}_{\text{vettore della base canonica}} \quad \text{t.c.} \quad \frac{\partial f}{\partial x_{\hat{i}}}(x^k) = \underbrace{\min_{j=1,\dots,n} \frac{\partial f}{\partial x_j}(x^k)}_{\text{minima componente del } \nabla f(x^k)}$$

Quindi

$$d^k = e_{\hat{i}} - x^k$$

direzione che da x^k punta verso $e_{\hat{i}}$

Esempio in $\mathbb{R}^3$

$$\min f(x)$$
$$e^T x = 1$$
$$x \geq 0$$

Applichiamo il metodo F-W al problema

con $f(x) = x^T Q x$

con $Q = \begin{pmatrix} 3 & 0 & 3 \\ 0 & 1.5 & 1.5 \\ 3 & 1.5 & 5 \end{pmatrix}$

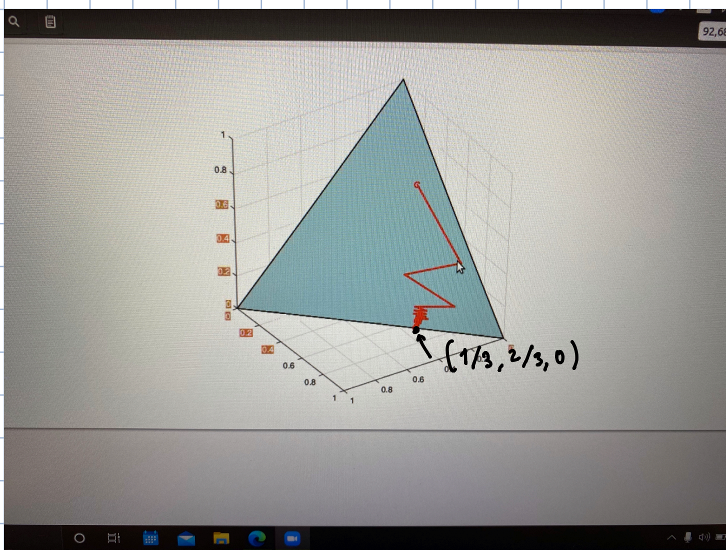
si può verificare che Q è semidefinita positiva

$\Downarrow$

f è convessa !!

$x^0 \in F$

$x^0 \in = \begin{pmatrix} 0.1 \\ 0.3 \\ 0.6 \end{pmatrix}$ manca un pezzo



esercizio: (KKT)

$$\min f(x) \qquad \nabla f(x) = Qx$$

$$\left.x_1 + x_2 + x_3 = 1\right\} \quad h(x) = x_1 + x_2 + x_3 - 1$$

$$\left.\begin{array}{l} x_1 \geq 0 \\ x_2 \geq 0 \\ x_3 \geq 0 \end{array}\right\} \begin{array}{l} g_1(x) = -x_1 \\ g_2(x) = -x_2 \\ g_3(x) = -x_3 \end{array} \qquad \nabla g_1(x) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \nabla g_2(x) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \nabla g_3(x) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\nabla h(x) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{cases} \nabla f(x) + \sum_{i=1}^{3} \lambda_i \nabla g_i(x) + \mu h(x) = 0 \\ x_1 + x_2 + x_3 = 1 \\ x_i \geq 0 \qquad i = 1,2,3 \\ \lambda_i \geq 0 \\ \lambda_i g_i(x) = 0 \end{cases}$$

condizioni KKT

Quindi il sistema diventa:

$$\begin{cases} 3x_1 + 3x_3 - \lambda_1 + \mu = 0 \\ \frac{3}{2} x_2 + \frac{3}{2} x_3 - \lambda_2 + \mu = 0 \\ 3x_1 + \frac{3}{2} x_2 + 5x_3 - \lambda_3 + \mu = 0 \\ \lambda_1, \lambda_2, \lambda_3 \geq 0 \\ e^T x = 1, \quad x \geq 0 \\ -\lambda_1 x_1 = 0 \\ -\lambda_2 x_2 = 0 \\ -\lambda_3 x_3 = 0 \end{cases}$$

partendo da qui e studiamo i casi $\lambda_3 = 0$, $x_3 = 0$

Supponiamo $x_3 = 0$, e $\lambda_3 \geq 0$, otteniamo:

$$\left.\begin{array}{l} 3x_1 - \lambda_1 + \mu = 0 \\ \frac{3}{2} x_2 - \lambda_2 + \mu = 0 \\ 3x_1 + \frac{3}{2} x_2 - \lambda_3 + \mu = 0 \end{array}\right\} \longrightarrow \begin{cases} \mu = \lambda_1 - 3x_1 \\ \frac{3}{2} x_2 - \lambda_2 + \lambda_1 - 3x_1 = 0 \\ \cancel{3x_1} + \frac{3}{2} - \lambda_3 + \lambda_1 - \cancel{3x_1} = 0 \end{cases}$$

$$\longrightarrow x_2 = \frac{2}{3}(\lambda_3 - \lambda_1)$$

$\lambda_1, \lambda_2, \lambda_3 \geq 0$

$-\lambda_1 x_1 = 0$

$-\lambda_2 x_2 = 0$

$x_1 + x_2 = 1 \; ; \quad x_1 \geq 0, \; x_2 \geq 0$

sostituendo nella seconda equazione otteniamo

$$x_1 = \frac{1}{3}(\lambda_3 - \lambda_2)$$

quindi $x_3 = 0$ otteniamo:

$$\begin{cases} \mu = \lambda_1 - \lambda_3 + \lambda_2 \\ x_2 = \frac{2}{3}(\lambda_3 - \lambda_1) \\ x_1 = \frac{1}{3}(\lambda_3 - \lambda_2) \\ \lambda_i \geq 0 \\ -\lambda_i x_i = 0 \\ e^T x = 1 \; ; \; x_i \geq 0 \quad i = 1, \dots, 3 \end{cases}$$

Supponendo $x_1 > 0$, $x_2 > 0 \implies \lambda_1 = 0; \; \lambda_2 = 0$

il sistema diventa

$$\begin{cases} \mu = -\lambda_3 \\ \boxed{\begin{array}{l} x_2 = \frac{2}{3}\lambda_3 \\ x_1 = \frac{1}{3}\lambda_3 \end{array}} \end{cases} \quad \text{con } \lambda_3 \geq 0$$

$\underbrace{x_1 + x_2 = 1}_{\lambda_3 = 1}, \; x_1 \geq 0, \; x_2 \geq 0$

$$\Downarrow$$

$$\left(\begin{pmatrix} 1/3 \\ 2/3 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ; -1 \right)$$

e' punto KKT per il nostro problema

e quindi ottimo globale poiché il problema è convesso

Se avessi voluto trovare <u>tutte</u> le soluzioni del problema dovevo vedere anche i casi $(x_3 \geq 0 \text{ e } \lambda_3 = 0)$ $(\lambda_2 \geq 0, \; x_2 = 0)$, $(x_2 \geq 0, \; \lambda_2 = 0)$ ecc.