

Elementi di Teoria dell'Ottimizzazione: Metodi Analitici per la PM non Lineare

Docente Prof. Francesco Amato



Condizioni necessarie di minimo locale

- **Teorema [Condizione necessaria del I ordine].**

Data una funzione

$$f : x \in X \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$$

derivabile in X , condizione necessaria affinché x^* sia un minimo locale di f è che per ogni direzione d ammissibile

$$(\text{grad}_x f(x^*))^T d \geq 0$$



- La condizione necessaria del primo ordine enunciata prima vale sia nel caso di problemi vincolati che di problemi non vincolati.
- In quest'ultimo caso è possibile dare una versione della condizione più semplice; ciò è dovuto al fatto che ogni direzione è ammissibile.
- Questo è ovviamente anche il caso in cui il punto x è *interno* all'insieme X .



- **Teorema [Condizione necessaria del I ordine – problema NV].** Data una funzione

$$f : x \in R^n \rightarrow R$$

derivabile in R^n , condizione necessaria affinché x^* sia un minimo locale di f è che risulti

$$(\text{grad}_x f(x^*)) = 0$$



- **Teorema [Condizione necessaria del II ordine].**
Data una funzione

$$f : x \in X \subseteq R^n \rightarrow R$$

due volte derivabile in X , condizione necessaria affinché x^* sia un minimo locale di f è che

i) per ogni direzione d ammissibile

$$(\text{grad}_x f(x^*))^T d \geq 0$$

ii) per ogni vettore d per cui la a) è soddisfatta con il segno di uguaglianza risulti

$$d^T H_f(x^*) d \geq 0$$



- **Teorema [Condizione necessaria del II ordine – problema NV].** Data una funzione

$$f : x \in R^n \rightarrow R$$

due volte derivabile in R^n , condizione necessaria affinché x^* sia un minimo locale di f è che risulti

- i) $(grad_x f(x^*)) = 0$
- ii) $H_f(x^*)$ semidefinita positiva



- **Teorema [Condizione sufficiente di minimo locale – problema NV].** Data una funzione

$$f : x \in R^n \rightarrow R$$

due volte derivabile in R^n , condizione sufficiente affinché x^* sia un minimo locale di f è che risulti

- i) $(grad_x f)(x^*) = 0$
- ii) $H_f(x^*)$ definita positiva



- In definitiva per risolvere un problema non vincolato si trovano prima tutti i punti che soddisfano la condizione necessaria del I ordine.
- Quindi si studia la matrice Hessiana della funzione f valutata in tali punti.
- Se, in un dato punto, l'Hessiana risulta definita positiva siamo in presenza di un minimo locale; se risulta semidefinita positiva non si può concludere nulla; negli altri casi il punto *non* è di minimo locale.
- Ovviamente il minimo globale risulta essere il più piccolo tra tutti i minimi locali.



- Ricordiamo che le condizioni esaminate nel caso non vincolato sono valide anche nel caso vincolato laddove si considerino punti interni all'insieme X .
- Dunque, per i problemi vincolati, si possono applicare gli stessi ragionamenti dei problemi NV per quanto riguarda i *punti interni* all'insieme X .
- Occorre trattare a parte i punti appartenenti alla frontiera di X (dove non tutte le direzioni sono ammissibili).
- Nel seguito vedremo come trattare alcuni problemi vincolati di particolare importanza nell'ambito dei controlli multivariabile.



Esempio.

- Calcolare

$$\min_x f(x)$$

Con

$$f(x) = x_1^3 - x_1^2 x_2 + 2x_2^2$$



Problemi vincolati: Vincoli di uguaglianza (VU)

- Consideriamo il seguente problema con vincoli di uguaglianza.

$$\min_x f(x)$$

$$s. a$$

$$g(x) = 0$$

$$x \in R^n, \quad g(x) = [g_1(x) \quad g_2(x) \quad \cdots \quad g_m(x)]^T$$

$$m < n$$



- Ovviamente in un problema con vincoli di uguaglianza tutti i punti che soddisfano la condizione di vincolo sono di frontiera, e quindi i risultati precedenti *non* si possono applicare.
- La strada da seguire consiste nel partizionare il vettore x come segue

$$x = \begin{bmatrix} y \\ w \end{bmatrix} \quad y \in R^{n-m} \quad w \in R^m$$

E quindi di esprimere il vettore w in funzione del vettore y .

Sia S la superficie individuata dall'equazione vincolare $g(x)=0$.



- **Teorema [Condizione necessaria di minimo locale del I ordine – Problema VU]** Condizione necessaria affinché un punto x^* sia un minimo locale di f soggetto ai vincoli $g(x)=0$ è che per ogni vettore infinitesimo dx appartenente al piano tangente ad S in x^* , cioè tale che

$$\frac{\partial g}{\partial x}(x^*)dx = 0$$

sia

$$(grad_x f(x^*))^T dx = 0$$



- Queste condizioni si possono riscrivere:

$$a) g(x^*) = 0$$

$$b) \left(\text{grad}_y f(x^*) \right)^T dy + \left(\text{grad}_w f(x^*) \right)^T dw = 0$$

$$\forall dx = \begin{bmatrix} dy \\ dw \end{bmatrix} : \quad \frac{\partial g}{\partial y}(x^*) dy + \frac{\partial g}{\partial w}(x^*) dw = 0$$



- Per trovare una condizione di minimo locale più agevole ricaviamo dalla condizione b) l'espressione di w in funzione di y ; si ottiene

$$dw = - \left[\frac{\partial g}{\partial w}(x^*) \right]^{-1} \frac{\partial g}{\partial y}(x^*) dy$$

- A valle di opportuni passaggi, la condizione necessaria di minimo locale per un problema VU si può riscrivere

$$\left(\text{grad}_y f(x^*) \right)^T - \left(\text{grad}_w f(x^*) \right)^T \left[\frac{\partial g}{\partial w}(x^*) \right]^{-1} \frac{\partial g}{\partial y}(x^*) = 0$$

$$g(x^*) = 0$$



- Alla condizione precedente si può arrivare passando per un problema fittizio non vincolato.
- Introduciamo a questo proposito la funzione Lagrangiana

$$l(x, \lambda) := f(x) + \lambda^T g(x)$$

dove

$$\lambda = [\lambda_1 \quad \lambda_2 \quad \cdots \quad \lambda_m]^T$$

si chiama vettore dei moltiplicatori di Lagrange.

Faremo ora vedere che il calcolo di un minimo locale per il problema non vincolato, avente per funzione obiettivo la Lagrangiana, restituisce le stesse condizioni del teorema precedente.



- Ricordiamo che la condizione necessaria del primo ordine per un problema non vincolato richiede che il gradiente della funzione obiettivo sia nullo nel punto di minimo.
- Nel nostro caso, partizionando in modo opportuno il vettore x si ha:

$$\text{grad}_y l(x^*, \lambda^*) = \text{grad}_y f(x^*) + \left[\frac{\partial g}{\partial y}(x^*) \right]^T \lambda^* = 0$$

$$\text{grad}_w l(x^*, \lambda^*) = \text{grad}_w f(x^*) + \left[\frac{\partial g}{\partial w}(x^*) \right]^T \lambda^* = 0$$

$$\text{grad}_\lambda l(x^*, \lambda^*) = g(x^*) = 0$$



- Ricavando dalla seconda equazione λ^* e sostituendo nella prima si ha

$$\text{grad}_y f(x^*) - \left[\frac{\partial g}{\partial y}(x^*) \right]^T \left[\frac{\partial g}{\partial w}(x^*) \right]^{-T} \text{grad}_w f(x^*) = 0$$

che restituisce la condizione precedente.

In definitiva il nostro problema con vincoli di uguaglianza si può ricondurre ad un problema non vincolato avente come funzione obiettivo la Lagrangiana. Ciò è riassunto nel prossimo teorema.



- **Teorema [Condizioni necessarie del primo ordine utilizzando la Lagrangiana – Problema VU]** Condizione necessaria affinché un punto x^* sia un minimo locale di f soggetto ai vincoli $g(x)=0$ è che esista un λ^* in R^m tale che

$$\text{grad}_{\lambda} l(x^*, \lambda^*) = g(x^*) = 0$$

$$\text{grad}_x l(x^*, \lambda^*) = \text{grad}_x f(x^*) + \left[\frac{\partial g}{\partial x}(x^*) \right]^T \lambda^* = 0$$



- La determinazione delle soluzioni della equazione del teorema precedente richiede la valutazione dei moltiplicatori di Lagrange.
- Si può quindi concludere che il relativo aggravio di complessità dal punto di vista computazionale è il prezzo da pagare per aver sostituito un problema di ottimizzazione vincolata con uno di ottimizzazione non vincolata.
- Al solito, si può dimostrare che, introducendo gli Hessiani delle funzioni coinvolte, si ottengono condizioni necessarie del secondo ordine e condizioni sufficienti.



- **Teorema [Condizione necessaria del secondo ordine per la soluzione del Problema VU]**
- Condizione necessaria affinché un punto x^* sia un minimo locale di f soggetto ai vincoli $g(x)=0$ è che

$$g(x^*) = 0$$

$$\exists \lambda^* \in R^m \quad : \quad \text{grad}_x f(x^*) + \left[\frac{\partial g}{\partial x}(x^*) \right]^T \lambda^* = 0$$

$$H_f(x^*) + \sum_{i=1}^m \lambda_i^* H_{g_i}(x^*) \geq 0 \quad \text{sul piano tangente in } x^* \text{ alla}$$

iper - superficie definita dai vincoli.



- **Teorema [Condizione sufficiente del secondo ordine per la soluzione del Problema VU]**

Basta richiedere la definitezza positiva in luogo della semidefinitezza nell'ultimo teorema.



- **Esempio.** Calcolare

$$\min_x x_1 x_2 + x_2 x_3 + x_1 x_3$$

s. a

$$x_1 + x_2 + x_3 = 3$$



- In generale, le $n+m$ equazioni che si ottengono scrivendo le condizioni necessarie del primo ordine sono non lineari. Ciò limita fortemente l'utilizzo di tali condizioni analitiche.
- Tuttavia, come vedremo, ci sono delle applicazioni di interesse nell'ambito dei controlli automatici, nelle quali queste condizioni risultano molto utili.



Un primo problema VU: Funzione obiettivo di tipo quadratico e funzioni di vincolo di tipo lineare

- Di particolare importanza nell'ambito dei problemi di controllo sono i problemi di PM in cui la funzione obiettivo è di tipo quadratico e le funzioni di vincolo sono lineari.

$$\min_x x^T Q x$$

$$s.a$$

$$F x = b$$

con $x \in R^n$, $F \in R^{m \times n}$, $m < n$, $Q > 0$; F è di pieno rango riga ($=m$).



- Si noti che se nel problema precedente si pone $Q=I$ il problema diventa quello di minimizzare la norma di x soggetta al vincolo $Fx=b$.
- Poiché il problema in esame presenta solo vincoli di uguaglianza si può utilizzare per la sua soluzione la teoria sviluppata in precedenza.
- In questo caso la Lagrangiana risulta essere definita come segue

$$l(x, \lambda) = x^T Qx + \lambda^T (Fx - b)$$



- Si ha:

$$\text{grad}_x l(x, \lambda) = 2Qx + F^T \lambda$$

$$\text{grad}_\lambda l(x, \lambda) = Fx - b$$

Dunque i punti di minimo si possono trovare imponendo le condizioni

$$2Qx + F^T \lambda = 0$$

$$Fx = b$$



- Alla fine si ottiene, come possibile punto di minimo,

$$x^* = Q^{-1} F^T (F Q^{-1} F^T)^{-1} b =: F_Q^{RM} b$$

Ora ponendo

$$F = \begin{bmatrix} f_1^T \\ f_2^T \\ \vdots \\ f_m^T \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

si ha che effettivamente x^* è un minimo globale, infatti

$$H_{(x^T Q x)} + \sum_{i=1}^m \lambda_i H_{(f_i^T x - b_i)} = H_{(x^T Q x)} = 2Q > 0$$



- La matrice F_Q^{RM} è detta *pseudoinversa minima a destra (PMD)*.
- L'aggettivo “minima” ricorda che tale matrice è legata al calcolo del minimo della forma quadratica $x^T Q x$.
- L'aggettivo “pseudoinversa a destra” discende dal fatto che

$$FF_Q^{RM} = FQ^{-1}F^T \left(FQ^{-1}F^T\right)^{-1} = I$$



Una applicazione della PMD ad un problema di raggiungibilità

- Si consideri il sistema lineare a tempo discreto, supposto essere raggiungibile, descritto dalle equazioni

$$x(k+1) = Ax(k) + Bu(k), \quad x(0) = x_0, \quad x(k) \in R^n, \quad u(k) \in R^m$$

Un classico problema è quello di trovare la legge di controllo che in $N > n$ passi porti lo stato del sistema in uno stato desiderato minimizzando l'energia impiegata dal controllore, portata in conto attraverso un indice di costo del tipo

$$\sum_{i=0}^{N-1} u^T(i) Q_i u(i), \quad Q_i > 0, \quad i = 0, \dots, N-1$$



- Dalla teoria della raggiungibilità per i sistemi a tempo discreto si ha che

$$x_N = A^N x_0 + \begin{bmatrix} A^{N-1}B & A^{N-2}B & \dots & B \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{bmatrix}$$



- L'ultima equazione può essere riscritta:

$$Fu = b$$

dove

$$x_N - A^N x_0 = b$$

$$\begin{bmatrix} A^{N-1}B & A^{N-2}B & \dots & B \end{bmatrix} = F$$

$$\begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{bmatrix} = u \in R^{Nm}$$



- Dunque il problema di ottimizzazione può essere formulato come:

$$\min u^T Q u$$

$$s.a$$

$$F u = b$$

$$Q = \begin{bmatrix} Q_0 & 0 & \cdots & 0 \\ 0 & Q_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & Q_{N-1} \end{bmatrix} > 0$$

Per quanto si è detto prima la soluzione è $u^* = F_Q^{RM} b$



- Esempio. Dato il sistema:

$$x(k+1) = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} x(k) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(k) \quad x(0) = 0$$

Portare il sistema in $N=4$ passi nello stato

$$x_N = \begin{pmatrix} 10 \\ 1 \end{pmatrix}$$

in modo che la somma dei quadrati dei valori in ingresso sia minima.



Un secondo problema VU: Pseudoinversa minima a sinistra

- Supponiamo di dover risolvere il sistema di equazioni

$$Fx = b$$

con $x \in \mathbb{R}^n$, $F \in \mathbb{R}^{m \times n}$, $m > n$, dove F è di pieno rango colonna ($=n$).

Questo problema non è risolubile, a meno che b non appartenga al range di F .



- In molte situazioni pratiche, piuttosto che trovare la soluzione, siamo interessati a ottenere il vettore x che più si ‘avvicina’ al soddisfacimento dell’equazione.
- Questo obiettivo si può impostare come problema di ottimizzazione come segue:

$$\min_{x, \varepsilon} \varepsilon^T Q \varepsilon$$

$$s.a$$

$$Fx = b + \varepsilon$$

con $Q > 0$.

Si noti che il vettore ε permette di ‘rilassare’ il problema originale, introducendo n variabili di ottimizzazione fittizie.



- Si noti il problema può essere ricondotto nella forma standard ponendo

$$v := \begin{bmatrix} x^T & \varepsilon^T \end{bmatrix}^T \in R^{n+m}, \tilde{Q} = \begin{bmatrix} 0 & 0 \\ 0 & Q \end{bmatrix} \in R^{(n+m) \times (n+m)}, \tilde{F} = \begin{bmatrix} F & -I \end{bmatrix} \in R^{m \times (m+n)}$$

e risolvendo

$$\min_v v^T \tilde{Q} v$$

s.a

$$\tilde{F}v = b$$

Questo problema non può essere risolto con le tecniche di pseudoinversione studiate precedentemente, in quanto la matrice dei pesi è solo semidefinita positiva.



- Applicando la condizione necessaria del primo ordine si ottiene

$$l(v, \lambda) = l(x, \varepsilon, \lambda) = \varepsilon^T Q \varepsilon + \lambda^T (Fx - \varepsilon - b)$$

$$\begin{aligned} \text{grad}_x l(x, \varepsilon, \lambda) &= \text{grad}_x (\lambda^T Fx) \\ &= F^T \lambda = 0 \end{aligned}$$

$$\begin{aligned} \text{grad}_\varepsilon l(x, \varepsilon, \lambda) &= \text{grad}_\varepsilon (\varepsilon^T Q \varepsilon - \lambda^T \varepsilon) \\ &= 2Q\varepsilon - \lambda = 0 \end{aligned}$$

$$\text{grad}_\lambda l(x, \varepsilon, \lambda) = Fx - \varepsilon - b = 0$$



- Un punto candidato ad essere minimo risulta

$$x^* = \left(F^T Q F \right)^{-1} F^T Q b$$

La condizione sufficiente di minimo richiede che

$$H_{v^T \tilde{Q} v} + \sum_{i=1}^m \lambda_i^* H_{\tilde{f}_i^T v - b_i} = H_{v^T \tilde{Q} v} = 2\tilde{Q}$$

sia definita positiva per ogni vettore

$$y \in \left\{ y \in R^{n+m} : \frac{\partial(\tilde{F}v - b)}{\partial v} \Big|_{v=v^*} y = 0 \right\} = \{ y \in R^{n+m} : \tilde{F}y = 0 \}$$

Dunque deve essere

$$y^T \tilde{Q} y > 0 \quad \forall y \in R^{n+m} : \tilde{F}y = 0$$



- Partizionando il vettore y come

$$y = \begin{bmatrix} y_1^T & y_2^T \end{bmatrix}^T$$

si ottiene

$$y_1^T F^T Q F y_1 > 0 \quad \forall y_1 \in R^n$$

L'ultima condizione è verificata poiché Q è definita positiva e F è di rango pieno colonna; quindi il punto x^* è effettivamente un punto di minimo.



- La matrice

$$F_Q^{LM} = \left(F^T Q F\right)^{-1} F^T Q$$

è detta *pseudoinversa minima a sinistra (PMS)*.
Infatti

$$F_Q^{LM} F = \left(F^T Q F\right)^{-1} F^T Q F = I$$



Una applicazione della PMS: stima parametrica a minimi quadrati

- Consideriamo una relazione causa – effetto tra due grandezze x (vettoriale) e y (scalare). Supponiamo che sia nota la struttura della dipendenza tra x e y :

$$y = k_1 f_1(x) + k_2 f_2(x) + \cdots + k_n f_n(x)$$

$$k = [k_1 \quad k_2 \quad \cdots \quad k_n]^T \in R^n$$



- Siano (x_i, y_i) , $i=1, \dots, N$, i risultati di N esperimenti; poiché, ovviamente, il modello fornito dalla precedente equazione non descrive perfettamente il fenomeno in esame, si ha

$$\hat{y}_i = k_1 f_1(x_i) + k_2 f_2(x_i) + \dots + k_n f_n(x_i) \neq y_i, \quad i = 1, \dots, N$$



- Il problema della stima ai minimi quadrati è quello di trovare il vettore dei parametri k che minimizzi la norma pesata dello scarto tra gli effetti misurati e gli effetti calcolati dal modello teorico , cioè

$$\min_{k, \varepsilon} \varepsilon^T Q \varepsilon$$

$$s.a$$

$$Fk = b - \varepsilon$$

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}, \quad \varepsilon_i := y_i - \hat{y}_i$$

$$F = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_n(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_n(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_N) & f_2(x_N) & \cdots & f_n(x_N) \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$



- Sulla base dei risultati precedenti, la soluzione è data da

$$k^* = F_Q^{LM} b$$



Una applicazione della PMS: identificazione della risposta armonica

- Supponiamo di aver determinato sperimentalmente la risposta armonica di un sistema lineare che esibisca un comportamento riconducibile a quello di un sistema del primo ordine:

$$W(\omega) = \frac{G}{1 + j\omega\tau}$$

Il problema è determinare G e τ in modo tale che la funzione ricavata teoricamente sia il più possibile simile a quella ricavata sperimentalmente.



- La funzione di risposta armonica può essere scritta

$$\Re(W(\omega)) + j\Im(W(\omega)) = \frac{G}{1 + j\omega\tau}$$

Dunque possiamo scrivere:

$$\begin{bmatrix} \Re(W(\omega)) \\ \Im(W(\omega)) \end{bmatrix} = \begin{bmatrix} 1 & \omega\Im(W(\omega)) \\ 0 & -\omega\Re(W(\omega)) \end{bmatrix} \begin{bmatrix} G \\ \tau \end{bmatrix}$$



- Quindi possiamo esprimere G e τ in funzione dei dati sperimentali:

$$\begin{bmatrix} G \\ \tau \end{bmatrix} = \begin{bmatrix} 1 & \omega \Im(W(\omega)) \\ 0 & -\omega \Re(W(\omega)) \end{bmatrix}^{-1} \begin{bmatrix} \Re(W(\omega)) \\ \Im(W(\omega)) \end{bmatrix}$$

Poiché l'inversa esiste, con una unica misura sperimentale della risposta armonica si riuscirebbero ad ottenere i valori di G e τ e quindi $W(\omega)$.



- Tuttavia, poiché nella pratica la risposta armonica non è misurata con precisione assoluta, si effettuano più misure che portano al sistema di equazioni :

$$\begin{bmatrix} \Re(W(\omega_1)) \\ \Im(W(\omega_1)) \\ \Re(W(\omega_2)) \\ \Im(W(\omega_2)) \\ \vdots \\ \Re(W(\omega_N)) \\ \Im(W(\omega_N)) \end{bmatrix} = \begin{bmatrix} 1 & \omega_1 \Im(W(\omega_1)) \\ 0 & -\omega_1 \Re(W(\omega_1)) \\ 1 & \omega_2 \Im(W(\omega_2)) \\ 0 & -\omega_2 \Re(W(\omega_2)) \\ \vdots & \vdots \\ 1 & \omega_N \Im(W(\omega_N)) \\ 0 & -\omega_N \Re(W(\omega_N)) \end{bmatrix} \begin{bmatrix} G \\ \tau \end{bmatrix}$$

Questo è un sistema sovraspecificato, la cui soluzione può essere trovata con le tecniche di pseudoinversione descritte precedentemente.



Una applicazione della PMS: assegnamento dell'autostruttura

- Ricordiamo che il problema dell'assegnamento dell'autostruttura a ciclo chiuso con retroazione di stato può schematizzarsi come segue.
- Trovare una matrice K tale che, assegnata una n -pla di numeri complessi simmetrici $\lambda_1, \lambda_2 \dots \lambda_n$ e una corrispondente n -pla di vettori complessi simmetrici $u_1, u_2 \dots u_n$, si abbia

$$(A + BK)u_i = \lambda_i u_i \quad i = 1, \dots, n$$

e quindi

$$[\lambda_i I - (A + BK)]u_i = 0 \quad i = 1, \dots, n$$



- Dal punto di vista matematico , questo problema è equivalente a trovare una *n-pla* di vettori v_i tali che

$$u_i = (\lambda_i I - A)^{-1} B v_i \quad i = 1, \dots, n$$

L'ultima equazione non può essere risolta in maniera *esatta*, a meno che u_i non appartenga al sottospazio m -dimensionale generato da $(\lambda_i I - A)^{-1} B$.

Dal punto di vista ingegneristico, più che essere interessati ad assegnare esattamente u_i , ci interessa ottenere un autovettore che sia il più possibile vicino a quello desiderato.



- Dunque il problema, in riferimento all'autovettore i -esimo, si può riformulare nel modo seguente.

$$\min_{v_i, \varepsilon} \varepsilon^T Q \varepsilon$$

s.a

$$(\lambda_i I - A)^{-1} B v_i = u_i + \varepsilon$$

Dove $Q > 0$ è una matrice di pesi da scegliere in modo opportuno se si desidera pesare in modo diverso le componenti dell'autovettore desiderato.



Problemi con vincoli di uguaglianza e disuguaglianza (VUD)

- Si consideri il seguente problema di ottimizzazione:

$$\min_x f(x)$$

s. a

$$g_i(x) = 0 \quad i = 1, \dots, m$$

$$h_i(x) \leq 0 \quad i = 1, \dots, p$$



- **Definizione [Vincolo attivo]**
- Un vincolo di disuguaglianza $h_i(x) \leq 0$ è detto essere attivo in un punto x^* se risulta $h_i(x^*) = 0$; inattivo se risulta $h_i(x^*) < 0$.
- Si noti che i vincoli attivi in un punto sono gli unici che limitano le direzioni di spostamento ammissibili in quel punto.
- Ovviamente i vincoli di uguaglianza sono per definizione sempre attivi qualunque sia $x \in R^n$.



- **Teorema [Condizione necessaria del primo ordine per la soluzione del Problema VUD]**
- Condizione necessaria affinché un punto x^* sia un minimo locale di f è che esista un vettore λ in R^m e un vettore μ in R^p tali che

$$i) \operatorname{grad}_x f(x^*) + \left(\frac{\partial g}{\partial x}(x^*) \right)^T \lambda^* + \left(\frac{\partial h}{\partial x}(x^*) \right)^T \mu^* = 0$$

$$ii) \mu_j h_j(x^*) = 0 \quad j = 1, \dots, p$$

$$iii) \mu_j \geq 0 \quad j = 1, \dots, p$$

$$iv) g(x^*) = 0 \quad h(x^*) \leq 0$$



- **Teorema [Condizione necessaria del secondo ordine per la soluzione del Problema VUD]**

Condizione necessaria affinché un punto x^* sia un minimo locale di f è che esista un vettore λ in R^m e un vettore μ in R^p tali che

$$i) \operatorname{grad}_x f(x^*) + \left(\frac{\partial g}{\partial x}(x^*) \right)^T \lambda^* + \left(\frac{\partial h}{\partial x}(x^*) \right)^T \mu^* = 0$$

$$ii) \mu_j h_j(x^*) = 0 \quad j = 1, \dots, p$$

$$iii) \mu_j \geq 0 \quad j = 1, \dots, p$$

$$iv) g(x^*) = 0 \quad h(x^*) \leq 0$$

$$v) H_f(x^*) + \sum_{i=1}^m \lambda_i^* H_{g_i}(x^*) + \sum_{i=1}^p \mu_i^* H_{h_i}(x^*) \text{ è semidefinita positiva}$$

sul piano tangente ai vincoli attivi in x^*



- **Teorema [Condizione sufficiente per la soluzione del Problema VUD]**

Condizione sufficiente affinché un punto x^* sia un minimo locale di f è che esista un vettore λ in R^m e un vettore μ in R^p tali che

$$i) \operatorname{grad}_x f(x^*) + \left(\frac{\partial g}{\partial x}(x^*) \right)^T \lambda^* + \left(\frac{\partial h}{\partial x} \right)^T \mu^* = 0$$

$$ii) \mu_j h_j(x^*) = 0 \quad j = 1, \dots, p$$

$$iii) \mu_j \geq 0 \quad j = 1, \dots, p$$

$$iv) g(x^*) = 0 \quad h(x^*) \leq 0$$

$$v) H_f(x^*) + \sum_{i=1}^m \lambda_i^* H_{g_i}(x^*) + \sum_{i=1}^p \mu_i^* H_{h_i}(x^*) \text{ è definita positiva}$$

sul piano tangente ai vincoli attivi in x^*



- Esempio. Risolvere il problema delle n centrali elettriche interconnesse descritto in precedenza con i seguenti valori dei parametri:

$$n = 2$$

$$c_1(p_1) = p_1^2 + 1$$

$$c_2(p_2) = \frac{1}{2} p_2^2 + 2$$

$$P_{tot} = 15$$

$$p_{1M} = 10$$

$$p_{2M} = 10$$

