Elementi di Teoria dell'Ottimizzazione: Introduzione

Docente Prof. Francesco Amato

- Studiando le tecniche di controllo con retroazione di stato, si è visto come queste richiedano, per la sintesi del controllore, la soluzione di un opportuno problema di ottimizzazione.
- Ad esempio, il problema dell'allocazione dei poli in assegnate regioni del piano complesso può essere ridotto ad un problema di esistenza di una soluzione (matriciale) ammissibile soggetto a LMI.
- I problemi di esistenza di una soluzione ammissibile soggetti a disuguaglianze lineari, rientrano nella più ampia categoria dei problemi di ottimizzazione convessa.
- D'altro canto, anche la tecnica di controllo basata sull'assegnamento dell'autostruttura conduce ad un problema di ottimizzazione.
- Infatti, si è visto che l'assegnamento di tutte le componenti dell'autovettore conduce ad un problema sovra-specificato e che quindi non ammette soluzione.
- Dunque, in genere, si segue la strategia di trovare un autovettore ammissibile che abbia distanza minima da quello desiderato (nel senso dello scarto quadratico).

- Inoltre, ricordiamo che le tecniche di controllo basate sulla retroazione di stato finora investigate permettono di allocare in posizione opportuna i poli del sistema a ciclo chiuso e quindi di agire sulla velocità di risposta e sulla sovraelongazione.
- Più in generale, le tecniche di assegnamento dell'autostruttura permettono anche di assegnare (parzialmente) l'insieme degli autovettori del sistema a ciclo chiuso, e quindi di agire sulla posizione degli zeri e/o introdurre un certo grado di disaccoppiamento tra i diversi canali del ciclo.
- Tuttavia, non essendo agevole la traduzione di alcune specifiche, quali la moderazione delle grandezze in gioco (variabili di controllo e di stato), nel dominio del tempo, tali tecniche non consentono di portare in conto all'atto della sintesi tali specifiche.

- Per ovviare a tali inconvenienti è necessario portare in conto, all'atto della sintesi del controllore, tutte le specifiche richieste attraverso la definizione di una funzione obiettivo e/o di opportuni vincoli.
- Il progetto della legge di controllo va dunque effettuato in modo che essa ottimizzi tale indice e contemporaneamente soddisfi, se presenti, preassegnati vincoli sulle variabili di interesse.
- Come vedremo, questa visione del problema di controllo nel dominio del tempo, porterà allo sviluppo della teoria del controllo ottimo.
- Ancora una volta, per poter sviluppare tale teoria, è opportuno avere a disposizione alcuni risultati di teoria dell'ottimizzazione.
- Da queste considerazioni, discende immediatamente la necessità di dedicare una parte del corso alla illustrazione dei fondamenti della teoria dell'ottimizzazione.

- Data una funzione obiettivo *f (x)* e due funzioni vettoriali *g(x)* e *h(x)*, un problema di ottimizzazione può essere formalizzato, in generale, come segue:

$$\min_x f(x)$$
$$\text{s.a}$$
$$x \in X \subseteq R^n$$

$$X = \left\{ x \in R^n, g(x) = 0, h(x) \leq 0 \right\}$$

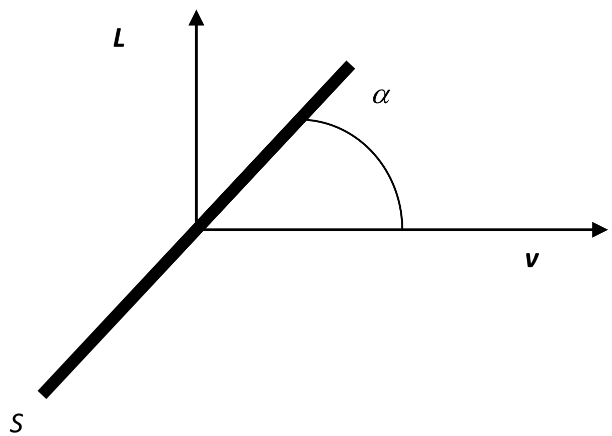
L’insieme delle metodologie orientate alla soluzione dei problemi di questo tipo va sotto il nome di *Programmazione Matematica (PM)*.

- Un problema di programmazione matematica si dice lineare se le funzioni g e h sono lineari. Altrimenti si dice non lineare.
- Un problema di programmazione matematica si dice *non vincolato* se $X=R^n$, vincolato nel caso contrario.
- Ad esempio, i problemi di ammissibilità soggetti a LMI rientrano nella categoria dei problemi di programmazione lineare.
- I problemi sovraspecificati (come ad esempio l'assegnamento dell'autostruttura) rientrano, invece, nei problemi di programmazione non lineare.

- Vedremo che la maggior parte dei problemi di controllo ricade nella categoria dei problemi non lineari.
- Per questo motivo, nel seguito focalizzeremo l’attenzione su tali problemi.
- Le tecniche di programmazione non lineare si dividono in *analitiche* e *numeriche*.
  - Le prime si utilizzano quando si è in presenza di poche variabili di ottimizzazione e avendo a disposizione l’espressione analitica delle funzioni coinvolte.
  - Le seconde entrano in gioco per problemi di grosse dimensioni e/o quando la funzione obiettivo o i vincoli non sono esprimibili in forma chiusa.

Formulazione di alcuni problemi decisionali in termini di PM: Esempio 1- ottimizzazione della portanza di una superficie mobile

- Si consideri una superficie di sezione *S* che si muova in un fluido di tipo Newtoniano con velocità costante ***v*** ; indichiamo con α l'angolo formato dalla superficie *S* con il vettore ***v*** .
- La superficie *S* sarà sottoposta ad una forza che si oppone al movimento e diretta in direzione opposta al vettore ***v*** (*resistenza*) e in una forza diretta verso l'alto ortogonalmente a ***v*** (*portanza* ***L***).



- Si intuisce che L sarà nulla tanto per $\alpha=0$ che per $\alpha=\pi/2$. In effetti, sotto opportune ipotesi semplificative si può dimostrare che risulta

$$|L| = \mathrm{k}\,\mathrm{sen}^2\alpha\,\cos\alpha$$

dove k è una costante che dipende da S, $|v|$ e dal tipo di fluido in cui è immersa la superficie.

Il problema è allora quello di trovare il valore di α che massimizzi il modulo di L. La soluzione di questo problema verrà fornita successivamente.

Esempio 2 - Gestione ottimale di centrali elettriche interconnesse

- Il fabbisogno di potenza elettrica di un utente può porsi nella forma

$$P = P_r + P_p(t)$$

dove P_r è detta potenza di base e può ritenersi pressoché costante durante il giorno, mentre $P_p(t)$, detta potenza di picco, dipende dalla richiesta dell'utente all'istante t e quindi è di natura stocastica e tempo-variante.

- Supponiamo che per generare la potenza di base P_r si ricorra al contributo di *n* centrali elettriche interconnesse.
- La centrale *i-esima* eroga una potenza p_i e ha un costo di gestione $c_i(p_i)$. In genere le funzioni $c_i(p_i)$ si possono approssimare come segue:

$$c_i(p_i) = a_i p_i^2 + b_i$$

L'obiettivo è quello di minimizzare la spesa complessiva delle *n* centrali, soddisfacendo la richiesta dell'utenza.

- Denotando con $p=(p_1 \ p_2 \ \dots \ p_n)^T$, il problema può formularsi in questo modo:

$$\min_{\boldsymbol{p}} \sum_{i=1}^{n} c_i(p_i)$$

$$s.a$$

$$\sum_{i=1}^{n} p_i = P_r$$

$$0 \le p_i \le p_{iM}$$

Esempio 3 – Soluzione di un sistema di equazioni lineari

- Si supponga di dover risolvere il seguente sistema di equazioni lineari

$$f_1(x_1, x_2) = 0$$
$$f_2(x_1, x_2) = 0$$

Innanzitutto, si osservi che tutte le soluzioni della prima equazione sono anche soluzioni dell’equazione:

$$f_1^2(x_1, x_2) = 0$$

- Ovviamente una soluzione $x^*=(x_1\ x_2)^T$ dell’ultima equazione è anche minimo globale di f_1^2.
- Pertanto il problema di trovare le soluzioni del sistema di equazioni è equivalente al seguente problema di ottimizzazione:

$$\begin{aligned}&\min_{\boldsymbol{x}}\ f_1^2(x_1, x_2)\\ &s.a\\ &f_2(x_1, x_2) = 0\end{aligned}$$

Alcune definizioni preliminari

- **Minimo locale e globale.** Data la funzione

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

si dice che un punto x^* è di *minimo locale* per f se esiste un intorno I_{x^*} del punto x^* tale che

$$f(x^*) \leq f(x) \qquad \forall x \in I_{x^*}$$

Il minimo si dice globale se la disuguaglianza precedente vale per ogni x in X.

Gradiente. Si chiama *gradiente di* f il vettore colonna

$$grad_{\mathbf{x}} f = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \dfrac{\partial f}{\partial x_2} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix}^T$$

- **Matrice Hessiana. S**i chiama *matrice Hessiana di* f la seguente matrice quadrata

$$\boldsymbol{H}_f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \dfrac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Sotto opportune ipotesi la matrice Hessiana è simmetrica.

- **Jacobiano.** Data la funzione vettoriale

$$g : x \in X \subseteq R^n \rightarrow R^m$$

si chiama Jacobiano di *g* la seguente matrice

$$\frac{\partial g}{\partial x} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \dots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \dots & \frac{\partial g_2}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \dots & \frac{\partial g_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \left(grad_x g_1\right)^T \\ \left(grad_x g_2\right)^T \\ \dots \\ \left(grad_x g_m\right)^T \end{bmatrix}$$

- Data la funzione vettoriale *g*, l'insieme dei punti

$$S = \{x \in R^n : g(x) = 0\}$$

definisce una iper-superficie in R^n.

Si assuma che la funzione vettoriale *g* sia derivabile in *x*; si definisce *piano tangente* in *x* alla iper - superficie *S* l'insieme dei punti

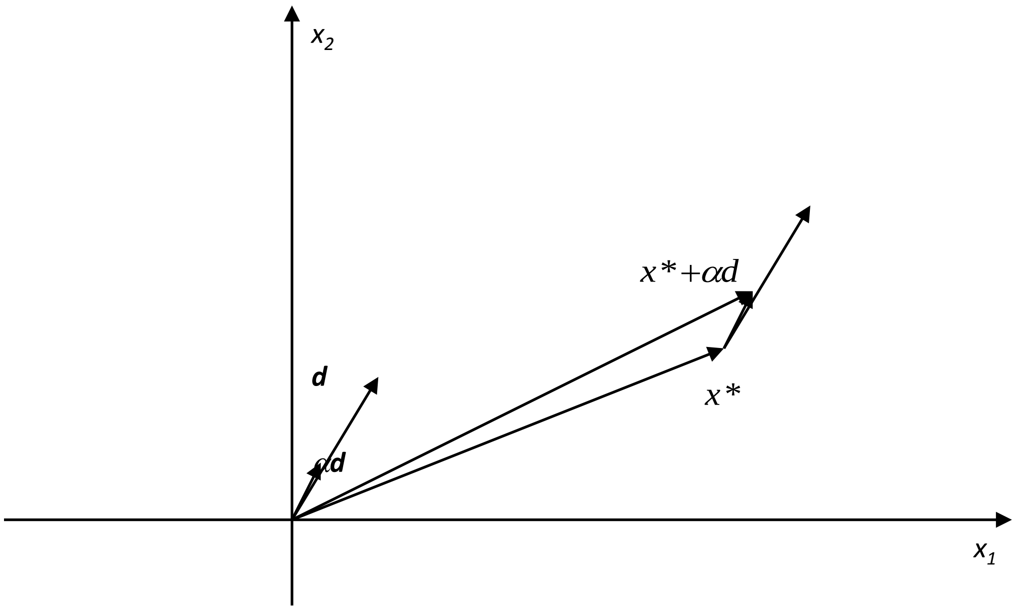
$$M = \left\{ y \in R^n : \frac{\partial g}{\partial x}(x) y = 0 \right\}$$

- In altri termini il piano tangente è composto da tutti i vettori $y \in R^n$ che sono ortogonali ai gradienti delle funzioni $g_i(x)$, $i=1,\ldots,m$, che definiscono la iper – superficie S .
- Un ruolo importante nei problemi di PM lo giocano le matrici definite, che sono state introdotte precedentemente.
- Chiudiamo questa sezione di definizioni, introducendo il concetto di direzione ammissibile.

- **Direzione ammissibile.** Dato un insieme X in R^n e un punto x^* in X, di dice che il vettore d in R^n *individua una direzione ammissibile in* x^* se esiste uno scalare $\alpha^*>0$ tale che

$$x^*+\alpha d \in X \qquad \forall \alpha \in [0,\alpha^*]$$

Come si può notare dalla figura, *d* individua un segmento orientato che parte dal punto x^* e si svolge parallelamente a *d*.



- Si noti che:
  - Se $X=R^n$ ogni vettore d in R^n individua una direzione ammissibile in x, qualunque sia x.
  - Se X è un sottoinsieme di R^n e x è un punto interno a X allora ogni vettore d in R^n individua una direzione ammissibile in x.

- Il concetto di direzione ammissibile è molto utile nel calcolo del minimo di una funzione.
- Infatti per dimostrare che un punto x è di minimo locale basta far vedere che muovendosi lungo una qualsiasi direzione ammissibile in x la funzione obiettivo $f(x)$ è non decrescente.
- Ciò ci consente di studiare il comportamento della funzione $f(x)$ come se essa dipendesse di volta in volta da una sola variabile.