

1 Lezione del 27-11-24

1.1 Domini di problemi di programmazione non lineare

Abbiamo studiato finora funzioni non lineari nell'ottica dell'ottimizzazione non vincolata. Vediamo adesso una delle forme dei domini su cui possiamo limitare tali funzioni nel caso dell'ottimizzazione vincolata:

$$D = \{x \in \mathbb{R}^n : g_1(x) \leq 0, g_2(x) \leq 0, \dots, g_m(x) \leq 0, h_1(x) = 0, h_2(x) = 0, \dots, h_p(x) = 0\}$$

dove

$$\begin{cases} g_i : \mathbb{R}^n \rightarrow \mathbb{R}, & i = 1, \dots, m \\ h_j : \mathbb{R}^n \rightarrow \mathbb{R}, & j = 1, \dots, p \end{cases}$$

o in maniera più compatta:

$$D = \{x \in \mathbb{R}^n : g(x) \leq 0, h(x) = 0\}$$

dove

$$\begin{cases} g(x) = (g_1(x), g_2(x), \dots, g_m(x)) \\ h(x) = (h_1(x), h_2(x), \dots, h_p(x)) \end{cases}$$

Notiamo che i poliedri studiati finora non sono che un caso particolare di insiemi in questa forma. Possiamo quindi dire che da qui in poi formuleremo i problemi di NLP come:

$$\begin{cases} \max f(x) \\ g(x) \leq 0 \\ h(x) = 0 \end{cases}$$

dove $f(x)$ è una funzione in n variabili, $g(x)$ è un vettore di m funzioni, e $h(x)$ è un vettore di p funzioni. Vediamo alcune proprietà che ci aspettiamo dal dominio:

- Assumeremo ognuna delle funzioni $g_i(x)$ e $h_j(x)$ come almeno $\in C^2$, in modo da poter calcolare gradiente ed Hessiana;
- Un vantaggio sarà dato dal fatto che l'insieme è **chiuso**: avevamo visto sulla PL come la soluzione stava *sempre* sulla frontiera. Qui, avremo che la soluzione sta *spesso* sulla frontiera, e dovremo quindi includerla in modo da non perdere soluzioni valide;
- L'insieme sarà idealmente **limitato**, così da permetterci di confermare attraverso il teorema di Weierstrass l'esistenza di minimi e massimi. Questo tra l'altro è necessario in quanto computazionalmente porremo come limite ai domini un iperrettangolo $-M \leq x_i \leq M$ con $M \gg 0$. Essendo l'iperrettangolo limitato, e il dominio contenuto interamente al suo interno, risulterà chiaramente che anche il dominio stesso sarà (o dovrà essere) limitato;
- Vorremo poi che l'insieme sia **convesso**, visto che questo facilita la sua esplorazione (spostandoci verso vettori *interni* all'insieme da ogni suo punto saremo sicuri di uscire dall'insieme stesso una e una sola volta). Possiamo dare la seguente caratterizzazione:

Teorema 1.1: caratterizzazione di insieme convesso

Se le g_i sono convesse e le h_j sono lineari, allora D è convesso.

1.1.1 Domini regolari

Una caratteristica che desideriamo dai nostri domini è la **regolarità**. Per giungere a una definizione di regolarità, dobbiamo introdurre due concetti supplementari. Il primo è quello di **cono tangente**, che si ricava dai **vettori tangenti**:

Definizione 1.1: Vettore tangente

Dato un punto $x^* \in \Omega$, un vettore d si dice tangente a Ω in x^* se esistono una successione di vettori $z_k \in \Omega$ e una successione di scalari $t_k \in \mathbb{R}^+$ tali che:

$$\lim_{k \rightarrow \infty} \frac{z_k - x^*}{t_k} = d$$

da cui la definizione:

Definizione 1.2: Cono tangente

L'insieme dei vettori tangenti a Ω in un punto x si dice cono tangente, ed è denotato con $T_\Omega(x)$.

Il cono tangente viene anche detto *cono di Bouligand*, e indicato come:

$$T_\Omega(x) = \left\{ d \in \mathbb{R}^n : \exists d_k \in \mathbb{R}^n, d_k \rightarrow d, \exists t_k \in \mathbb{R}^+, t_k \rightarrow 0 : x^* + d_k t_k \in \Omega \right\}$$

Le due definizioni sono equivalenti in quanto è noto:

$$\lim_{k \rightarrow \infty} \frac{z_k - x^*}{t_k} = d$$

e allora dato d_k tale che:

$$\lim_{k \rightarrow \infty} d_k = d$$

si potrà dire:

$$\lim_{k \rightarrow \infty} \frac{z_k - x^*}{t_k} = \lim_{k \rightarrow \infty} d_k$$

e potremo scegliere d_k in modo tale che soddisfi:

$$\frac{z_k - x^*}{t_k} = d_k \Rightarrow z_k - x^* = d_k t_k \Rightarrow z_k = x^* + d_k t_k$$

Questa è la forma che troviamo nella definizione del cono di Bouligand. Visto che ciò che si imponeva era $x^* + d_k t_k \in \Omega$, e $z_k \in \Omega$ per definizione, le due forme sono equivalenti.

Definiamo quindi il **cono delle direzioni ammissibili** di un punto x sul dominio Ω , sulla base delle funzioni $g(x)$ e $h(x)$ che lo definiscono. Innanzitutto, servirà la definizione dell'insieme dei **vincoli attivi**:

Definizione 1.3: Vincoli attivi

Dato un punto $x^* \in \Omega$, si indicano come $A(x^*)$ gli indici di tutti i vincoli di disuguaglianza che sono stretti in x^* , cioè:

$$A(x^*) = \{i : g(x^*) = 0\}$$

definiamo allora:

Definizione 1.4: Cono delle direzioni ammissibili del primo ordine

Dato un punto $x^* \in \Omega$, indichiamo il cono delle direzioni ammissibili del primo ordine $D(x^*)$:

$$D(x^*) = \left\{ d \in \mathbb{R}^n : \begin{cases} d^T \nabla g_i(x^*) = 0, & i \in A(x^*) \\ d^T \nabla h_j(x^*) = 0, & j = 1, \dots, p \end{cases} \right\}$$

dove notiamo che solo le disequaglianze *attive* figurano, mentre le uguaglianze vengono prese tutte.

Possiamo finalmente dare la definizione di regolarità per un punto:

Definizione 1.5: Regolarità

I vincoli di un dominio Ω si dicono regolari in un punto x^* se vale:

$$T_\Omega(x) = D(x)$$

Abbiamo che il cono tangente deriva solo dalle proprietà geometriche del dominio, mentre il cono delle direzioni ammissibili deriva solo dalle caratteristiche algebriche dei gradienti delle funzioni che definiscono il dominio. Questo significa che imponendo l'uguaglianza fra i due, andiamo a discriminare tutti quei punti di Ω che hanno un comportamento patologico.

1.1.2 Classificazione dei domini regolari

Vediamo quindi una classificazione per domini sicuramente **regolari**:

1. I **poliedri** visti nella PL sono tutti regolari, come conseguenza della linearità delle equazioni che li descrivevano;
2. Sono regolari quei domini per cui vale la **condizione di Slater**: g_i convesse, h_j lineari, $\exists \bar{x} : g_i(\bar{x}) < 0 \ \forall i \in 1, \dots, m$;
3. Infine, sono regolari quei domini dove i gradienti dei vincoli attivi sono fra di loro linearmente indipendenti, cioè:

$$\begin{cases} \nabla g_i(\bar{x}), & i \in A(\bar{x}) \\ \nabla h_j(x) \end{cases} \text{ sono linearmente indipendenti}$$

dove $A(\bar{x})$ dà i **vincoli attivi** $A(\bar{x}) = \{i : g_i(\bar{x}) = 0\}$. Notiamo ancora come si prendono quando attivi solo i vincoli di disequaglianza $g(x)$, mentre i vincoli di uguaglianza $h(x)$ si prendono sempre.