

# **Introduzione all'ottimizzazione nonlineare**

# Definizione del problema

$$\begin{aligned} \min \quad & f(x) \\ & x \in \mathbb{R}^n \\ & c_i(x) = 0 \quad i \in K_1 \\ & c_i(x) \geq 0 \quad i \in K_2 \end{aligned}$$

# Definizioni

$f$  viene detta **funzione obiettivo**.

L'insieme

$$S = \left\{ \begin{array}{l} x \in \mathbb{R}^n \\ c_i(x) = 0 \quad i \in K_1 \\ c_i(x) \geq 0 \quad i \in K_2 \end{array} \right\}$$

viene detto **regione ammissibile** del problema.

# Ipotesi

Supporremo che:

- $f, c_i \in \mathcal{C}^2$  (ipotesi di differenziabilità);
- almeno una tra le funzioni  $f, c_i, i \in K_1 \cup K_2$  è **non** lineare (ipotesi di nonlinearità).

Il caso in cui  $f$  e  $c_i, i \in K_1 \cup K_2$  sono tutte funzioni lineari corrisponde alla classe dei problemi di **Programmazione Lineare**.

# Altre definizioni

**Minimo globale:** un punto  $x^* \in S$  tale che:

$$f(x^*) \leq f(x) \quad \forall x \in S.$$

**Minimo locale forte:** un punto  $x^*$  tale che per qualche  $\delta > 0$ :

$$f(x^*) < f(x) \quad \forall x \in S \cap \{x : \|x - x^*\|_2 \leq \delta\}, \quad x \neq x^*.$$

**Minimo locale debole:** un punto  $x^*$  tale che per qualche  $\delta > 0$ :

$$f(x^*) \leq f(x) \quad \forall x \in S \cap \{x : \|x - x^*\|_2 \leq \delta\}$$

e  $x^*$  non è un minimo locale forte.

# Casi possibili

- non esistono minimi locali e/o globali perché la funzione è illimitata inferiormente. Esempio:  $f(x) = x^3$ ,  $S = \mathbb{R}$ ;
- esistono minimi locali ma non minimi globali perché la funzione è illimitata inferiormente. Esempio:  
 $f(x) = x^3 - x$ ,  $S = \mathbb{R}$  dove  $x' = 1/\sqrt{3}$  è minimo locale ma non globale.
- non esistono minimi locali e/o globali anche se la funzione è limitata inferiormente. Esempio:  $f(x) = e^{-x}$ ,  $S = \mathbb{R}$
- un minimo globale (e quindi anche locale) esiste certamente se:
  - $S$  è compatto (chiuso e limitato);
  - $f$  continua.(teorema di Weierstrass).

# Complessità

Possiamo aspettarci che l'identificazione di ottimi globali sia un problema non semplice.

Ma la teoria ci dice che le cose sono anche più complicate.

# Complessità

Introduciamo il problema  $\mathcal{NP}$ - completo SUBSET SUM:

dati gli interi positivi  $d_0, d_1, \dots, d_n$ , ci chiediamo se esiste una soluzione del sistema

$$\begin{aligned} \sum_{j=1}^n d_j y_j &= d_0 \\ y_j &\in \{0, 1\} \quad j = 1, \dots, n \end{aligned}$$

In altre parole, si cerca di stabilire se esiste un sottinsieme degli interi  $d_1, \dots, d_n$  la cui somma è pari a  $d_0$ .



# Individuazione ottimo globale

Dai dati del subset problem deriviamo il seguente problema con funzione obiettivo quadratica e vincoli box:

$$\begin{aligned} \min \quad & \left( \sum_{j=1}^n d_j y_j - d_0 \right)^2 + \sum_{j=1}^n y_j (1 - y_j) \\ & 0 \leq y_j \leq 1 \qquad j = 1, \dots, n \end{aligned}$$

Si ha che l'ottimo globale di questo problema ha valore dell'obiettivo pari a 0 se e solo se il corrispondente problema subset sum ammette soluzione.

Quindi, calcolare il minimo globale di un problema non lineare è un problema difficile anche nel caso di funzioni quadratiche e vincoli box.

# Ottimalità locale e illimitatezza

Se la difficoltà di individuare minimi globali era prevedibile, un po' meno lo è quella dei seguenti problemi:

- nei problemi con e senza vincoli, stabilire se un dato punto **non** è un punto di minimo locale è  $\mathcal{NP}$ -completo;
- nei problemi con e senza vincoli, stabilire se l'obiettivo del problema è illimitato sulla regione ammissibile è  $\mathcal{NP}$ -completo.

Le cose vanno un po' meglio se si impongono restrizioni sulle funzioni  $f$  e  $c_i$ .

# Definizione

Data una matrice simmetrica  $A$  di ordine  $n$ , diciamo che questa è **semidefinita positiva** se:

$$x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n.$$

Diciamo che questa è **definita positiva** se:

$$x^T A x > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Una matrice  $A$  è semidefinita (definita) positiva se e solo se tutti i suoi autovalori sono non negativi (positivi) [*data una matrice  $A$ , i suoi autovalori sono le radici della seguente equazione  $\det(A - \lambda I) = 0$  dove  $I$  è la matrice identica*].

Infine, una matrice  $A$  è semidefinita positiva (definita positiva) se tutti i suoi minori principali sono non negativi (positivi), dove i minori principali sono i determinanti di tutte le sottomatrici quadrate ottenute da  $A$  rimuovendo un sottinsieme delle sue righe con le corrispondenti colonne.

# Funzioni convesse e concave

Una funzione  $f$  si dice **convessa** se (condizioni equivalenti):



$$\forall x_1, x_2 \in \mathbb{R}^n, \forall \lambda \in (0, 1) : f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2);$$



$$\forall x_1, x_2 \in \mathbb{R}^n : f(x_2) \geq f(x_1) + \nabla f(x_1)^T (x_2 - x_1);$$



$$\forall x_1 \in \mathbb{R}^n : \nabla^2 f(x_1) \text{ è semidefinita positiva.}$$

Nel caso in cui le diseguaglianze siano strette si parla di funzione **strettamente convessa**.

Condizione sufficiente (ma non necessaria, vedi  $f(x) = x^4$ ) perché ci sia stretta convessità è che l'Hessiano sia definito positivo.

Definiamo una funzione  $f$  **concava** (**strettamente concava**) se  $-f$  è convessa (strettamente convessa).

# Programmazione Convessa (PC)

I problemi di *Programmazione Convessa (PC)* hanno la seguente forma

$$\begin{aligned} \min \quad & f(x) \\ & x \in \mathbb{R}^n \\ & c_i(x) \geq 0 \quad i \in K_2 \end{aligned}$$

con  $f$  convessa e le  $c_i$ ,  $i \in K_2$ , concave. Si noti che i problemi di PL sono una sottoclasse dei problemi di PC.

# Osservazione

I problemi PC appartengono alla classe  $\mathcal{P}$ . Gli algoritmi di complessità polinomiale che hanno consentito di stabilire questo sono gli stessi (elissoide, punto interno) che hanno permesso di catalogare nella classe  $\mathcal{P}$  i problemi di PL.

# Problemi trattati

Nell'ambito del corso non potremo effettuare una trattazione completa di tutti gli aspetti dei problemi di ottimizzazione non lineare. Per questa ragione, nel seguito concentreremo la nostra attenzione sul problema dell'individuazione di minimi locali e, dal punto di vista algoritmico, ci restringeremo ulteriormente ai problemi non vincolati.

Questo chiaramente rappresenta una restrizione rispetto ai problemi generali introdotti sopra, sia perché non si discute di possibili approcci (esatti o euristici) per l'individuazione di minimi globali, sia perché si tralascia dal punto di vista algoritmico la trattazione dei vincoli (per eventuali approfondimenti su tali argomenti si rimanda a testi specializzati).

Tuttavia già la trattazione degli argomenti citati rappresenta un'utile introduzione alle problematiche dell'ottimizzazione non lineare, che possono poi venire utili anche per la trattazione dei casi non trattati in questi appunti.



# Il caso non vincolato

Nel caso non vincolato gli insiemi  $K_1$  e  $K_2$  devono essere entrambi vuoti.

Vedremo per prima cosa condizioni necessarie e sufficienti di ottimalità locale per questo caso.

In seguito, passeremo alla descrizione di alcuni approcci risolutivi.

# Alcuni risultati sui gradienti

Data  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , si ha che

$$\nabla f(x) = \begin{cases} c & \text{se } f(x) = c^T x + c_0 \\ Hx & \text{se } f(x) = \frac{1}{2}x^T Hx, \quad H \text{ simmetrica} \end{cases}$$

Dati  $x_i : \mathbb{R} \rightarrow \mathbb{R}^{n_i}$ ,  $i = 1, \dots, k$ , e  $f : \mathbb{R}^{\sum_{i=1}^k n_i} \rightarrow \mathbb{R}$ , definiamo

$$g(\lambda) = f(x_1(\lambda), \dots, x_k(\lambda)).$$

Per le regole delle derivate di funzioni composte si ha che:

$$\frac{d}{d\lambda} g = \sum_{i=1}^k \left[ \frac{d}{d\lambda} x_i \right]^T \nabla_{x_i} f$$

# In particolare ...

... per  $k = 1$  e  $x(\lambda) = x_0 + \lambda s$ , ovvero

$$g(\lambda) = f(x_0 + \lambda s),$$

si ha

$$\frac{d}{d\lambda}g = \left[ \frac{d}{d\lambda}x \right]^T \nabla_x f = s^T \nabla_x f$$

# Condizioni di ottimalità

Non è possibile la verifica di ottimalità **locale** sulla base della definizione.

⇒ vengono ricavate delle **condizioni di ottimalità** locale.

# Condizione necessaria del primo ordine

Se  $x^*$  è un minimo locale, allora

$$\nabla f(x^*) = 0$$

(condizione di stazionarietà).

**Non** è una condizione sufficiente. Esempio:

$$f(x) = -x^2,$$

dove  $\bar{x} = 0$  è un punto stazionario ma non è un minimo locale (è un punto di massimo).

# Dimostrazione

Espansione in forma di Taylor attorno a  $x^*$ :

$$f(x^* + \varepsilon p) = f(x^*) + \varepsilon \nabla f(x^*)^T p + o(\varepsilon).$$

Per assurdo, sia  $\nabla f(x^*) \neq 0$ .

Allora esiste  $p$  tale che  $\nabla f(x^*)^T p < 0$  (esempio:  $p = -\nabla f(x^*)$ ).

Quindi, per  $\varepsilon$  sufficientemente piccolo:

$$f(x^* + \varepsilon p) < f(x^*),$$

il che contraddice l'ottimalità locale di  $x^*$ .

# Condizione necessaria del secondo ordine

Se  $x^*$  è un minimo locale, allora:

- $\nabla f(x^*) = 0$ ;
- $\nabla^2 f(x^*)$  è semidefinita positiva.

# Dimostrazione

$\nabla f(x^*) = 0$  già dimostrato.

Quindi:

$$f(x^* + \varepsilon p) = f(x^*) + \frac{1}{2}\varepsilon^2 p^T \nabla^2 f(x^*) p + o(\varepsilon^2).$$

Per assurdo, sia  $\nabla^2 f(x^*)$  non semidefinita positiva. Allora,  
 $\exists p$  tale che

$$p^T \nabla^2 f(x^*) p < 0.$$

Quindi, per  $\varepsilon$  sufficientemente piccolo:

$$f(x^* + \varepsilon p) < f(x^*),$$

il che contraddice l'ottimalità locale di  $x^*$ .



# La condizione non è sufficiente

Esempio:  $f(x) = x^3$ .

In  $\bar{x} = 0$ , la condizione è soddisfatta ma il punto non è di minimo.

# Condizione sufficiente

Un punto  $x^*$  che soddisfi le seguenti condizioni, è un minimo locale *forte*:

- $\nabla f(x^*) = 0$ ;
- $\nabla^2 f(x^*)$  è definita positiva.

# Dimostrazione

Per ogni matrice definita positiva  $A$ , si ha che:

$$x^T A x \geq \lambda_{\min}(A) \|x\|_2^2,$$

dove  $\lambda_{\min}(A) > 0$  è il più piccolo autovalore di  $A$ .

Da:

$$f(x^* + \varepsilon p) = f(x^*) + \frac{1}{2} \varepsilon^2 p^T \nabla^2 f(x^*) p + o(\varepsilon^2).$$

segue che per ogni  $p$  tale che  $\|p\|_2 = 1$  e per ogni  $\varepsilon > 0$  sufficientemente piccolo:

$$f(x^* + \varepsilon p) \geq f(x^*) + \frac{1}{2} \varepsilon^2 \lambda_{\min}(\nabla^2 f(x^*)) + o(\varepsilon^2) > f(x^*).$$

# Condizione non necessaria

Esempio:  $f(x) = x^4$ .

Il punto  $x^* = 0$  è di minimo locale forte, ma  $\nabla^2 f(0)$  non è definita positiva.

# Condizione necessaria e sufficiente

Nel caso convesso la condizione di stazionarietà:

$$\nabla f(x^*) = 0,$$

è necessaria **e** sufficiente perché  $x^*$  sia un minimo locale.

# Minimi locali $\equiv$ minimi globali

Nel caso convesso, ogni minimo locale è anche globale.

Dimostrazione per assurdo: sia  $x^*$  un minimo locale non globale, ovvero  $\exists \bar{x}$  tale che  $f(\bar{x}) < f(x^*)$ .

Allora  $\forall \lambda \in (0, 1)$ :

$$f(\lambda x^* + (1 - \lambda)\bar{x}) \leq \lambda f(x^*) + (1 - \lambda)f(\bar{x}) < f(x^*),$$

il che contraddice l'ottimalità locale di  $x^*$ .

# Unicità del minimo

Nel caso *strettamente convesso*, se esiste un minimo globale, esso è anche l'unico.

Dimostrazione per assurdo: siano  $x^*, \bar{x}$ ,  $x^* \neq \bar{x}$ , due minimi globali distinti. Si avrà che  $f(\bar{x}) = f(x^*)$ .

Inoltre,  $\forall \lambda \in (0, 1)$ :

$$f(\lambda x^* + (1 - \lambda)\bar{x}) < \lambda f(x^*) + (1 - \lambda)f(\bar{x}) = f(x^*),$$

il che contraddice l'ottimalità globale di  $x^*$ .

# Il caso quadratico

Sia data una funzione quadratica:

$$f(x) = \frac{1}{2}x^T Qx + c^T x,$$

dove  $Q$  può essere sempre presa simmetrica.

Si ha:

$$\nabla f(x) = Qx + c \quad \nabla^2 f(x) = Q.$$

Quindi:

- $f$  convessa  $\Leftrightarrow Q$  semidefinita positiva;
- $f$  strettamente convessa  $\Leftrightarrow Q$  definita positiva.



# Ottimo unico del caso str. convesso

Nel caso quadratico strettamente convesso per individuare l'unica soluzione ottima, basta imporre la condizione di stazionarietà:

$$\nabla f(x^*) = 0 \quad \text{ovvero} \quad Qx^* + c = 0,$$

da cui:

$$x^* = -Q^{-1}c,$$

è l'unico ottimo locale e globale per questo problema.

# Algoritmi caso non vincolato

Due grandi categorie di metodi:

- metodi linesearch;
- metodi trust region.

Entrambi i metodi sono iterativi, ovvero generano una sequenza  $\{x_k\}$  di punti.

# Metodi linesearch

Schema generale:

- Si individui una direzione di discesa  $d_k$  (ovvero,  $\nabla f(x_k)^T d_k < 0$ );
- Si individui, tramite una ricerca lungo la semiretta con origine  $x_k$  e direzione  $d_k$ , uno scalare  $\alpha_k > 0$  tale che

$$f(x_k + \alpha_k d_k) < f(x_k);$$

- si ponga  $x_{k+1} = x_k + \alpha_k d_k$  e  $k = k + 1$ .

Diverse scelte di  $d_k$  conducono ad algoritmi con proprietà di convergenza diverse.

# Antigradiente

Nel metodo dell'antigradiente si sceglie come direzione di discesa quella che garantisce la rapidità di discesa massima, ovvero:

$$d_k = -\nabla f(x_k).$$

# Newton

Nel metodo di Newton, si prende come direzione di discesa il minimo (se esiste) della funzione quadratica ottenuta dall'espansione in forma di Taylor di  $f$  attorno a  $x_k$  troncata al secondo ordine, ovvero

$$d_k \in \arg \min_{d \in \mathbb{R}^n} [f(x_k) + \nabla f(x_k)^T d + \frac{1}{2} d^T \nabla^2 f(x_k) d]$$

o, equivalentemente:

$$d_k = -[\nabla^2 f(x_k)]^{-1} \nabla f(x_k).$$

**NOTA BENE:** non è detto che  $d_k$  sia definita e, qualora lo sia, che sia una direzione di discesa. Ma è certamente una direzione di discesa nelle vicinanze di un minimo locale che soddisfa le condizioni sufficienti del secondo ordine.

# Quasi-Newton

Si utilizza un'approssimazione  $B_k$  dell'Hessiano  $\nabla^2 f(x_k)$ :

$$d_k = -[B_k]^{-1} \nabla f(x_k).$$

A ogni iterazione, l'approssimazione viene aggiornata in modo che soddisfi alcune proprietà.

# Condizione quasi-newtoniana

Sia  $s_k = x_{k+1} - x_k$  e

$$\nabla f(x_{k+1}) \approx \nabla f(x_k) + \nabla^2 f(x_k) s_k,$$

l'espansione in forma di Taylor, troncata al primo ordine, del *gradiente* della  $f$  attorno a  $x_k$ .

Indicando con

$$\Delta_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

si richiederà quindi che:

$$B_{k+1} s_k = \Delta_k$$

# Inoltre...

... viene richiesto che:

- $B_{k+1}$  sia simmetrica (come lo è l'Hessiano);
- $B_{k+1}$  sia definita positiva (si garantisce che la direzione di ricerca sia di discesa).

Tipicamente si cercano aggiornamenti "il più piccolo possibile":  $B_{k+1} - B_k$  è una matrice a rango molto basso (uno o due).

Una scelta usuale è porre  $B_0 = I$  (matrice identica), ovvero il primo passo è lungo la direzione dell'antigradiente.



# Un esempio: BFGS

$$B_{k+1} = B_k + \frac{\nabla f(x_k) \nabla f(x_k)^T}{\nabla f(x_k)^T d_k} - \frac{\Delta_k \Delta_k^T}{\alpha_k \Delta_k^T d_k}$$

Esiste anche una versione limited-memory BFGS, utile per problemi su larga scala, dove la matrice  $B_k$  **non** può essere memorizzata per problemi di occupazione di memoria, ma vengono memorizzati solo vettori che identificano un certo numero limitato di aggiornamenti quasi-newtoniani rispetto alla matrice identica.

# Convergenza globale

Dato un punto iniziale  $x_0$ , tutti i punti di accumulazione della sequenza  $\{x_k\}$  sono **punti stazionari**, ovvero

$$\|\nabla f(x_k)\| \rightarrow 0.$$

Si dimostra che vale se:

- l'insieme di livello:

$$L(f(x_0)) = \{x \in \mathbb{R}^n : f(x) \leq f(x_0)\}$$

è chiuso e limitato;

- a ogni iterazione si ha una "sufficiente" decrescita della funzione, cioè  $f(x_k) - f(x_{k+1})$  è sufficientemente grande;
- il passo compiuto  $x_{k+1} - x_k$  è "sufficientemente" distante dall'ortogonalità rispetto al gradiente  $\nabla f(x_k)$ .

# Condizioni di Armijo-Goldstein

Sono condizioni di "sufficiente decrescita".

Indicata con  $d$  la direzione di ricerca scelta, sia  $\bar{\alpha}$  il più piccolo valore per cui  $f(x_k + \bar{\alpha}d) = f(x_k)$ .

Per garantire una sufficiente decrescita, si vuole evitare un passo troppo vicino a 0 o troppo vicino a  $\bar{\alpha}$ .

Si fissi un  $\rho \in (0, 1/2)$ . Non troppo vicino a 0:

$$f(x_k + \alpha d) \geq f(x_k) + \alpha(1 - \rho)\nabla f(x_k)^T d$$

Non troppo vicino a  $\bar{\alpha}$ :

$$f(x_k + \alpha d) \leq f(x_k) + \alpha\rho\nabla f(x_k)^T d$$

Quindi il passo  $\alpha_k$  soddisfa:

$$-\rho\alpha_k\nabla f(x_k)^T d \leq f(x_k) - f(x_k + \alpha_k d) \leq -(1 - \rho)\alpha_k\nabla f(x_k)^T d$$

# Non ortogonalità rispetto al gradiente

Sia  $\theta_k$  l'angolo tra il gradiente in  $x_k$  e il passo  $s_k = x_{k+1} - x_k$ .

Si ha:

$$\cos(\theta_k) = \frac{\nabla f(x_k)^T s_k}{\|\nabla f(x_k)\| \|s_k\|}.$$

Ipotesi:

$$\sum_{k=1}^{\infty} \cos^2(\theta_k) = +\infty,$$

cioè  $\{\theta_k\}$  può convergere a  $\pi/2$  ma deve farlo in modo "sufficientemente" lento.

# Velocità di convergenza locale

Dato un minimo locale  $\bar{x}$  e un punto  $x_0$  "sufficientemente" vicino a  $\bar{x}$ , la velocità di convergenza locale può essere:

- lineare con coefficiente  $r < 1$ :

$$\|\mathbf{x}_{k+1} - \bar{\mathbf{x}}\| / \|\mathbf{x}_k - \bar{\mathbf{x}}\| \rightarrow r$$

- superlineare:

$$\|\mathbf{x}_{k+1} - \bar{\mathbf{x}}\| / \|\mathbf{x}_k - \bar{\mathbf{x}}\| \rightarrow 0$$

- quadratica:

$$\|\mathbf{x}_{k+1} - \bar{\mathbf{x}}\| / \|\mathbf{x}_k - \bar{\mathbf{x}}\|^2 \rightarrow A > 0.$$

# Antigradiente

Il metodo dell'antigradiente ha velocità di convergenza locale lineare.

Se applicato a una funzione quadratica strettamente convessa  $f(x) = c^T x + 1/2 x^T H x$ , con  $H$  simmetrica definita positiva e minimo in  $x^*$ , si dimostra che

$$| f(x_{k+1}) - f(x^*) | \leq \left( \frac{\lambda_{\max} - \lambda_{\min}}{\lambda_{\max} + \lambda_{\min}} \right)^2 | f(x_k) - f(x^*) |$$

dove  $\lambda_{\max}$  e  $\lambda_{\min}$  sono, rispettivamente, il massimo e il minimo autovalore di  $H$ .

Si noti che se  $\lambda_{\max}/\lambda_{\min}$  è molto grande, la convergenza è lineare ma molto lenta.

# Esempio

Si consideri la funzione

$$f(x, y) = Mx^2 + y^2,$$

con  $M > 0$ . Si può facilmente dimostrare che il punto di ottimo locale e globale di questa funzione è l'origine  $(0, 0)$ .

Prendiamo ora il punto

$$\mathbf{x}_0 = \alpha(1/M, 1).$$

Il gradiente è  $\nabla f(x, y) = (2Mx, 2y)$  e quindi nel punto  $\mathbf{x}_0$  la direzione di ricerca è l'antigradiente  $-\alpha(2, 2)$  e il nuovo punto viene cercato lungo la semiretta:

$$\alpha(1/M, 1) - \lambda \alpha(2, 2) \quad \lambda \geq 0.$$

# Continua

Possiamo facilmente trovare il minimo della funzione:

$$q(\lambda) = f(\alpha(1/M - 2\lambda), \alpha(1 - 2\lambda)) = M\alpha^2(1/M - 2\lambda)^2 + \alpha^2(1 - 2\lambda)^2 \quad \lambda \geq 0.$$

Il minimo è il valore  $\lambda_0 = \frac{1}{M+1}$ , indipendente da  $\alpha$  e il nuovo punto  $\mathbf{x}_1$  è il seguente:

$$\mathbf{x}_1 = \alpha \frac{M-1}{M+1} (-1/M, 1).$$



# Continua

Con il nuovo punto si può ripetere quanto visto sopra e si ottiene:

$$\mathbf{x}_2 = \alpha \left( \frac{M-1}{M+1} \right)^2 (1/M, 1),$$

Notando che

$$\mathbf{x}_2 = \left( \frac{M-1}{M+1} \right)^2 \mathbf{x}_0,$$

alla  $2k$ -esima iterazione avremo:

$$\mathbf{x}_{2k} = \left( \frac{M-1}{M+1} \right)^{2k} \mathbf{x}_0, \quad \forall k \in \mathbb{N}$$

# Continua

Questo ci dice che ad ogni iterazione la distanza dall'ottimo (l'origine  $(0, 0)$ ) viene ridotta del fattore

$$\frac{M - 1}{M + 1}$$

che al crescere di  $M$  tende a 1, confermando quindi che il metodo dell'antigradiente ha convergenza locale lineare con un coefficiente  $r$  che può essere reso arbitrariamente vicino a 1.

# Newton e quasi-Newton

Sotto opportune ipotesi (in particolare si richiede che nel minimo locale  $\bar{x}$  l'Hessiano sia definito positivo), il metodo di Newton ha velocità di convergenza locale quadratica.

D'altro canto non si ha alcuna garanzia di convergenza globale (in certe iterazioni la direzione di discesa può non essere definita).

Per i metodi quasi-newtoniani si riesce di solito a dimostrare la convergenza superlineare.

Per il metodo BFGS si riesce a dimostrare:

- la convergenza in al più  $n$  iterazioni su funzioni quadratiche, **se** si fanno ricerche lineari esatte;
- la convergenza globale per funzioni convesse, ma **non** per funzioni più generali .

# Caratterizzazione di Dennis-Morè

Data una sequenza  $\{x_k^{\mathcal{A}}\}$  generata da un metodo  $\mathcal{A}$ ,  
indichiamo con

$$s_k^{\mathcal{A}} = x_{k+1}^{\mathcal{A}} - x_k^{\mathcal{A}},$$

il passo all'iterazione  $k$  e con  $s_k^{\mathcal{N}}$  il passo che verrebbe  
compiuto dal metodo di Newton.

Si dimostra che  $\mathcal{A}$  ha convergenza superlineare se e solo  
se

$$s_k^{\mathcal{A}} = s_k^{\mathcal{N}} + o(\|s_k^{\mathcal{N}}\|),$$

ovvero il passo è asintoticamente pari a quello di Newton.

# Metodi trust-region

Metodi iterativi in cui l'iterato successivo  $x_{k+1}$  viene determinato risolvendo un opportuno sottoproblema basato sull'iterato corrente  $x_k$ .

Nel sottoproblema si minimizza un modello quadratico  $M_k(x)$  che approssima  $f$  in una sfera centrata in  $x_k$  e con un certo raggio  $\rho_k$ :

$$\tilde{x}_k \in \arg \min_{x : \|x - x_k\|_2 \leq \rho_k} M_k(x).$$

Quindi si pone

$$x_{k+1} = \begin{cases} x_k & \text{se } f(\tilde{x}_k) \geq f(x_k) \\ \tilde{x}_k & \text{altrimenti} \end{cases}$$

# Modello quadratico

Una tipica scelta per il modello quadratico è:

$$M_k(x) = f(x_k) + \nabla f(x_k)^T (x - x_k) + \frac{1}{2} (x - x_k)^T S_k (x - x_k),$$

dove  $S_k$  coincide con  $\nabla^2 f(x_k)$  o è una sua approssimazione (come nei metodi quasi-newtoniani).

# Aggiornamento raggio

A ogni iterazione il raggio  $\rho_k$  viene aggiornato in un nuovo valore  $\rho_{k+1}$  secondo le regole seguenti.

Sia

$$r_k = \frac{f(\tilde{x}_k) - f(x_k)}{M_k(\tilde{x}_k) - M_k(x_k)},$$

il rapporto tra la riduzione effettiva e quella sul modello.

Dati  $0 < \delta_1 < \delta_2 < 1$ ,  $0 < \gamma_1 < 1 < \gamma_2$  si pone:

$$\rho_{k+1} = \begin{cases} \gamma_1 \|\tilde{x}_k - x_k\| & \text{se } r_k < \delta_1 \\ \gamma_2 \rho_k & \text{se } r_k > \delta_2 \text{ e } \|\tilde{x}_k - x_k\| = \rho_k \\ \rho_k & \text{altrimenti} \end{cases}$$

In pratica, se il modello è affidabile, tendiamo a estendere la regione in cui applicarlo, se non lo è la riduciamo.

# Risultati di convergenza

Sia  $\{x_k\}$  la sequenza generata da un metodo trust region con  $S_k = \nabla^2 f(x_k)$ .

Se  $\{x_k\} \subseteq B$ , dove  $B$  è un insieme limitato (ad esempio, se l'insieme di livello  $L(f(x_0))$  è limitato), allora esiste un punto di accumulazione della sequenza che soddisfa le condizioni **necessarie** del secondo ordine.

Se il punto di accumulazione soddisfa le condizioni **sufficienti** del secondo ordine, allora la velocità di convergenza locale è quadratica,  $r_k \rightarrow 1$  e il vincolo  $\|x - x_k\| \leq \rho_k$  diventa inattivo, ovvero

$$\|\tilde{x}_k - x_k\| < \rho_k$$

per  $k$  sufficientemente grande.



# Il caso vincolato

# Direzioni ammissibili

Sia  $x'$  un punto ammissibile e  $\{x_k\}$  una sequenza di punti ammissibili convergente a  $x'$  con

$$x_k - x' = \delta_k s_k \quad \|s_k\| = 1, \quad \delta_k \rightarrow 0.$$

Sia  $s_k \rightarrow s$ . La direzione  $s$  viene detta **direzione ammissibile** rispetto a  $x'$ .

L'insieme delle direzioni ammissibili rispetto a  $x'$  viene indicato con

$$\mathcal{F}(x').$$

# Vincoli linearizzati

Sia

$$\mathcal{A}(x') = \{i \in K_2 : c_i(x') = 0\},$$

l'insieme dei **vincoli attivi** in  $x'$  (vincoli di disequaglianza soddisfatti come uguaglianza in  $x'$ ).

Se linearizziamo i vincoli con un'espansione di Taylor troncata al primo ordine attorno a  $x'$ , questi avranno la seguente forma:

$$\nabla c_i(x')^T (x - x') = 0 \quad i \in K_1$$

$$c_i(x') + \nabla c_i(x')^T (x - x') \geq 0 \quad i \in K_2$$

L'insieme  $F(x')$  delle direzioni ammissibili dei vincoli linearizzati è dato da tutte le direzioni  $s$  tali che

$$\nabla c_i(x')^T s = 0 \quad i \in K_1$$

$$\nabla c_i(x')^T s \geq 0 \quad i \in K_2 \cap \mathcal{A}(x')$$

# Relazione tra $F(x')$ e $\mathcal{F}(x')$

In generale si ha

$$F(x') \supseteq \mathcal{F}(x').$$

con la possibilità che non valga l'uguaglianza.

Esempio con soli vincoli di uguaglianza:

$$\begin{aligned} c_1(x_1, x_2) &= (x_1 - 1)^2 + x_2^2 - 1 \\ c_2(x_1, x_2) &= (x_1 + 1)^2 + x_2^2 - 1 \end{aligned}$$

Per  $x' = (0, 0)$  ogni vettore  $(0, \delta)$  è in  $F(x')$  ma essendo  $x'$  l'unico punto ammissibile, si ha che  $\mathcal{F}(x')$  è vuoto.

# Constraint qualification

Sotto opportune condizioni, dette **constraint qualification**, vale l'uguaglianza.

Possibili condizioni:

- tutti i vincoli  $c_i$  sono lineari;
- i gradienti  $\nabla c_i(x')$  dei vincoli in  $K_1 \cup \mathcal{A}(x')$  sono tra loro linearmente indipendenti.

Nell'esempio:

$$\nabla c_1(0, 0) = (-2, 0), \quad \nabla c_2(0, 0) = (2, 0)$$

non linearmente indipendenti.

# Condizione necessaria

Insieme direzioni di discesa in  $x'$ :

$$\mathcal{D}(x') = \{s : \nabla f(x')^T s < 0\}.$$

Se  $x^*$  è un minimo locale, allora

$$\mathcal{F}(x^*) \cap \mathcal{D}(x^*) = \emptyset.$$

Se in  $x^*$  valgono le constraint qualification, allora:

$$F(x^*) \cap \mathcal{D}(x^*) = \emptyset.$$

# Lemma di Farkas

Dati i vettori  $a_1, \dots, a_m$  e  $g$  si ha che

$$U = \{s : s^T g < 0, s^T a_i \geq 0, i = 1, \dots, m\}$$

è vuoto **se e solo se**  $\exists$  moltiplicatori  $\lambda_i \geq 0$  tali che

$$g = \sum_{i=1}^m \lambda_i a_i.$$

# Dimostrazione

**Se:** facile  $\rightarrow$

$$s^T g = \sum_{i=1}^m \underbrace{\lambda_i}_{\geq 0} \underbrace{s^T a_i}_{\geq 0} \geq 0$$

**Solo se:** sia

$$C = \{v : v = \sum_{i=1}^m \lambda_i a_i, \lambda_i \geq 0, i = 1, \dots, m\}$$

un cono poliedrale e si ipotizzi che  $g \notin C$ .

Allora esiste un iperpiano  $s^T x = 0$  (con normale  $s$ ) che separa  $g$  da  $C$ , ovvero:

$$s^T a_i \geq 0, \quad s^T g < 0,$$

come richiesto.



# Esistenza iperpiano separatore

Si consideri  $\min_{x \in C} \|g - x\|_2^2$ . Dato  $x_1 \in C$  si può imporre, senza perdere soluzioni ottime, il vincolo  $\|g - x\|_2 \leq \|g - x_1\|_2$  che rende la regione ammissibile limitata.

Essendo l'obiettivo una funzione continua, il teorema di Weierstrass garantisce l'esistenza di un punto di minimo  $\hat{g}$ .

Dal momento che  $\lambda \hat{g} \in C \forall \lambda \geq 0$ , si ha che  $\|\lambda \hat{g} - g\|_2^2$  ha un minimo in  $\lambda = 1$ .

Quindi,

$$\frac{d}{d\lambda} \|\lambda \hat{g} - g\|_2^2 \big|_{\lambda=1} = \hat{g}^T (\hat{g} - g) = 0.$$

# Continua

Sia  $x \in C$ ; allora, per convessità,  $\forall \theta \in (0, 1)$ :

$$\hat{g} + \theta(x - \hat{g}) \in C$$

e quindi:

$$\|\theta(x - \hat{g}) + \hat{g} - g\|_2^2 \geq \|\hat{g} - g\|_2^2.$$

e:

$$\theta^2 \|x - \hat{g}\|_2^2 + \theta(x - \hat{g})^T (\hat{g} - g) \geq 0$$

Prendendo il limite per  $\theta \rightarrow 0$ , si ha

$$(x - \hat{g})^T (\hat{g} - g) = x^T (\hat{g} - g) \geq 0.$$

Si prenda ora  $s = \hat{g} - g \neq 0$  poiché  $g \notin C$ . Si ha

$$s^T x \geq 0 \quad \forall x \in C, \quad s^T g = -s^T s < 0,$$

come si voleva dimostrare.

# Conseguenza

$$\mathcal{D}(x^*) \cap F(x^*) = \left\{ s : \begin{array}{ll} s^T \nabla f(x^*) < 0 \\ s^T \nabla c_i(x^*) = 0 & i \in K_1 \\ s^T \nabla c_i(x^*) \geq 0 & i \in \mathcal{A}(x^*) \end{array} \right\}$$

è vuoto **se e solo se**:

$$\left\{ s : \begin{array}{ll} s^T \nabla f(x^*) < 0 \\ s^T \nabla c_i(x^*) \geq 0, \quad -s^T \nabla c_i(x^*) \geq 0 & i \in K_1 \\ s^T \nabla c_i(x^*) \geq 0 & i \in \mathcal{A}(x^*) \end{array} \right\}$$

# Continua

se e solo se (lemma di Farkas)  $\exists \lambda_i^+, \lambda_i^- \geq 0, i \in K_1, \mu_i^* \geq 0, i \in K_2$  (con  $\mu_i^* = 0$  se  $i \notin \mathcal{A}(x^*)$ ), tali che:

$$\nabla f(x^*) = \sum_{i \in K_1} (\lambda_i^+ - \lambda_i^-) \nabla c_i(x^*) + \sum_{i \in K_2} \mu_i^* \nabla c_i(x^*),$$

o, equivalentemente, esistono  $\lambda_i^*, i \in K_1$  e  $\mu_i^* \geq 0, i \in K_2$  tali che

$$\nabla f(x^*) = \sum_{i \in K_1} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in K_2} \mu_i^* \nabla c_i(x^*).$$

# Condizioni KKT

Condizioni di Karush-Kuhn-Tucker (KKT) (necessarie del primo ordine): indicata con

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_{i \in K_1} \lambda_i c_i(x) - \sum_{i \in K_2} \mu_i c_i(x),$$

la **funzione Lagrangiana**, se  $x^*$  è un minimo locale *in cui è soddisfatta una constraint qualification*, allora esistono **moltiplicatori di Lagrange**  $\lambda^*, \mu^*$ , tali che

- $\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*) = 0$ ;
- $c_i(x^*) = 0, i \in K_1$ ;
- $c_i(x^*) \geq 0, i \in K_2$ ;
- $\mu^* \geq 0$ ;
- $\mu_i^* c_i(x^*) = 0, \forall i \in K_2$  (condizioni di complementarità).

# Interpretazione moltiplicatori Lagrange

Solo per il caso di vincoli di uguaglianza (simile per gli altri).

Sia  $x^*$  un minimo del problema e  $\lambda^*$  il relativo moltiplicatore di Lagrange. Si consideri la perturbazione

$$c_i(x) = \varepsilon_i$$

dell' $i$ -esimo vincolo di uguaglianza.

Lagrangiana

$$\mathcal{L}(x, \lambda, \varepsilon) = f(x) - \sum_{i \in K_1} \lambda_i (c_i(x) - \varepsilon_i).$$

Siano  $x(\varepsilon)$ ,  $\lambda(\varepsilon)$  la soluzione e il moltiplicatore di Lagrange del problema "perturbato" (ovviamente  $x(0) = x^*$ ,  $\lambda(0) = \lambda^*$ ).

# Continua

Si ha  $f(x(\varepsilon)) = \mathcal{L}(x(\varepsilon), \lambda(\varepsilon), \varepsilon)$  (valore ottimo del problema perturbato) .

Applicando la regola della derivata delle funzioni composte, si ha:

$$\frac{d f}{d \varepsilon_i} \Big|_{\varepsilon_i=0} = \frac{d \mathcal{L}}{d \varepsilon_i} = \frac{\partial x^T}{\partial \varepsilon_i} \underbrace{\nabla_x \mathcal{L}}_{=0} + \frac{\partial \lambda^T}{\partial \varepsilon_i} \underbrace{\nabla_\lambda \mathcal{L}}_{=0} + \frac{\partial \mathcal{L}}{\partial \varepsilon_i} = \lambda_i^*$$

Quindi, i moltiplicatori di Lagrange misurano la rapidità di variazione del valore ottimo in corrispondenza di perturbazioni dei vincoli.

# Problemi convessi

Se:

- $f$  convessa;
- $K_1 = \emptyset$ ;
- $c_i, i \in K_2$  sono funzioni concave (equivalentemente:  $-c_i$  sono funzioni convesse);

si parla di problema di **programmazione convessa**.



# Programmazione convessa

I problemi di programmazione convessa sono caratterizzati dal fatto che:

- ogni minimo locale è anche globale;
- se  $f$  è strettamente convessa, un minimo globale (se esiste) è unico;
- le condizioni KKT sono *necessarie e sufficienti* per identificare i minimi locali (globali).

# Condizioni secondo ordine

Sotto l'ipotesi del primo ordine

$$\nabla f(x^*) = \sum_{i \in K_1} \lambda_i^* \nabla c_i(x^*) + \sum_{i \in K_2} \mu_i^* \nabla c_i(x^*),$$

se  $\mu_j^* > 0$ , allora le direzioni ammissibili  $s$  per cui  $\nabla c_j(x^*)^T s > 0$ , sono di crescita:

$$\nabla f(x^*)^T s = \sum_{i \in K_1} \lambda_i^* \nabla c_i(x^*)^T s + \sum_{i \in K_2} \mu_i^* \nabla c_i(x^*)^T s \geq \mu_j^* \nabla c_j(x^*)^T s > 0.$$

Quindi, sotto l'ipotesi che valgano delle constraint qualification in  $x^*$ , possiamo restringerci alle direzioni

$$G(x^*) = \{s : \nabla c_i(x^*)^T s = 0, i \in K_1 \text{ o } \mu_i^* > 0, \nabla c_i(x^*)^T s \geq 0 \text{ } i \in \mathcal{A}(x^*) \text{ e } \mu_i^* = 0\}.$$

**Se**  $x_k = x^* + \delta_k s_k$  **con**  $\delta_k \rightarrow 0$ ,  $s_k \rightarrow s \in G(x^*)$  **e**

$$c_i(x_k) = 0 \quad \forall i : i \in K_1 \text{ o } \mu_i^* > 0,$$

si ha

$$f(x^* + \delta_k s_k) = \mathcal{L}(x^* + \delta_k s_k, \lambda^*, \mu^*)$$

Si ha anche:

$$\begin{aligned} \mathcal{L}(x^* + \delta_k s_k, \lambda^*, \mu^*) = \\ \underbrace{\mathcal{L}(x^*, \lambda^*, \mu^*)}_{=f(x^*)} + \delta_k \underbrace{\nabla_x \mathcal{L}(x^*, \lambda^*, \mu^*)^T s_k}_{=0} + \frac{1}{2} \delta_k^2 s_k^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) s_k + o(\delta_k^2) \end{aligned}$$

# Continua

Se  $x^*$  è un minimo locale, allora  $f(x_k) \geq f(x^*)$ , per  $k$  sufficientemente grande. Quindi, per  $k \rightarrow \infty$ :

$$s^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) s \geq 0 \quad \forall s \in G(x^*),$$

ovvero, se indichiamo con  $Z(x^*)$  una matrice le cui colonne formano una base dell'insieme di direzioni  $G(x^*)$ , si ha la condizione necessaria:

$$Z(x^*)^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) Z(x^*) \quad \text{semidefinita positiva.}$$

# Condizione sufficiente del secondo ordine

Se  $x^*, \lambda^*, \mu^*$  soddisfano le condizioni KKT e

$$Z(x^*)^T \nabla_{xx}^2 \mathcal{L}(x^*, \lambda^*, \mu^*) Z(x^*) \quad \text{definita positiva,}$$

allora  $x^*$  è un minimo locale del problema.

# Duale di Wolfe

Per problemi di programmazione convessa si definisce il duale di Wolfe:

$$\begin{aligned} \max_{x, \mu} \quad & \mathcal{L}(x, \mu) \\ & \nabla_x \mathcal{L}(x, \mu) = 0 \\ & \mu \geq 0 \end{aligned}$$

Se:

- $x^*$  è una soluzione del problema di programmazione convessa con i corrispondenti moltiplicatori di Lagrange  $\mu^*$ ;
- in  $x^*$  vale una constraint qualification;

allora  $(x^*, \mu^*)$  risolve il duale di Wolfe e i valori ottimi dei due problemi coincidono.

# Esempio: problema lineare

$$\begin{aligned} \min \quad & c^T x \\ & A^T x \geq b \\ & x \geq 0 \end{aligned}$$

Duale di Wolfe:

$$\begin{aligned} \max_{x, \mu, \gamma} \quad & c^T x - \mu^T (A^T x - b) - \gamma^T x \\ & c - A\mu - \gamma = 0 \\ & \mu, \gamma \geq 0 \end{aligned}$$

Equivalentemente, osservando che  $c = A\mu + \gamma$ :

$$\begin{aligned} \max_{\mu} \quad & \mu^T b \\ & A\mu \leq c \\ & \mu \geq 0 \end{aligned}$$

Si noti che questo coincide con il duale già visto per questi problemi di PL e che i moltiplicatori di Lagrange coincidono con le variabili del duale.

Dalla soluzione ottima  $\mu^*$  del duale, si ricava immediatamente la soluzione ottima del primale a partire dalle condizioni di complementarità:

$$\gamma^{*T} x^* = (A\mu^* - c)^T x^* = 0 \quad \mu^{*T} (A^T x^* - b) = 0$$



# Esempio: problema quadratico

$$\begin{aligned} \min_x \quad & \frac{1}{2}x^T Qx + c^T x \\ & A^T x \geq b \end{aligned}$$

$Q$  simmetrica definita positiva.

Duale di Wolfe:

$$\begin{aligned} \max_{x,\mu} \quad & \frac{1}{2}x^T Qx + c^T x - \mu^T (A^T x - b) \\ & Qx + c - A\mu = 0 \\ & \mu \geq 0 \end{aligned}$$

Osservando che  $x = Q^{-1}(A\mu - c)$ , si ha

$$\max_{\mu} \quad -\frac{1}{2}\mu^T (A^T Q^{-1} A)\mu + \mu^T (b + A^T Q^{-1} c) - \frac{1}{2}c^T Q^{-1} c$$
$$\mu \geq 0$$

Dalla soluzione ottima  $\mu^*$  del duale, si ricava immediatamente la soluzione ottima del primale

$$x^* = Q^{-1}(A\mu^* - c).$$