0.1 Modelli quadratici

Algoritmi di ottimizzazione che approssimano localmente f con modelli quadratici:

$$\min f(x) = \frac{1}{2} x^T Q x - b^T x, \text{t.c.} c \in \mathbb{R}^n$$

dove Q è una matrice quadrata di ordine n.

0.1.1 Casi Possibili

- Q non è semi-definita positiva: f non ha un minimo. - Q è definita positiva: $x^* = Q^{-1} b$ è l'unico minimo locale. Il punto x^* è il **punto di ottimo globale**. - Q è definita semi-positiva: - Q non è singolare: $x^* = Q^{-1} b$ è l'unico minimo globale. - Q è singolare: - non ho soluzioni. - ho infinite soluzioni.

0.1.2 Esempio

$$f(x, y) = \frac{1}{2}(ax^2 + by^2) - x$$

Riscrivo nei termini della formula per l'algoritmo:

$$f(x) = \frac{1}{2} \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

0.2 Introduzione agli algoritmi

Metodi di ottimizzazione continua:

- Dato un punto di inizio x_0, generiam una sequenza $x_k{}_{k=0}^{\infty}$. - Terminato l'algoritmo, quando le condizioni necessarie sono soddisfatte con una certa precisione, per esempio $\|\nabla f(x_k)\| \leq \epsilon$ - Monotone algorithms requires that $f(x_k) < f(x_{k-1}) \forall k$

0.2.1 Quanto è buono un algoritmo di ottimizzazione?

Un algoritmo è **decente** se **converge**.

Definizione 0.2.1 (Convergente globalmente) *Un algoritmo è chiamato convergente globalmente se converge a un punto* x^*

// PERSE COSE DA SLIDE QUI

Un algoritmo è **buono** se **converge rapidamente**

Chiamando x_k una sequenza in $\mathbb{R}^n$ che converge a x^*. La **convergenza** è chiamata:

- **Q-lineare** se $\exists r \in (0, 1) \text{s.t.} \frac{\| x_{k+1-x^*} \|}{\| x_k - x^* \|} \leq r, \text{for} k \geq \overline{k}$

- **Q-superlineaee** se $\lim_{k \to \infty} \frac{\| x_{k+1-x^*} \|}{\| x_k - x^* \|} = 0$

- **Q-quadratica** se $\exists C > 0 \text{s.t.} \frac{\| x_{k+1-x^*} \|}{\| x_k - x^* \|^2} \leq C, \text{for} k \geq \overline{k}$

Q-quadratica $\Rightarrow$ superlineare $\Rightarrow$ lineare.

0.2.2 Come determiniamo un punto di minimo

Line search

Dato il punto corrente determino la direzione e dopo di che determino di quanto muovermi.

Trust Region

Costruisco un modello quadratico in base alle informazioni locali, quindi scelgo un parametro Δk, un raggio, e scelgo la direzione risolvendo un problema di ottimizzazione vincolato al parametro Δ_k.

0.2.3 Condizioni di Wolfe

Per essere efficiente, la **linesearch inesatta** richiede alcune condizioni:

Decremento sufficiente

Sono accettabili solo i valori di $\phi(\alpha)$ che siano inferiori a quelli della funzione.

$$f(\boldsymbol{x} + \alpha \boldsymbol{d}) \leq f(\boldsymbol{x}) + c_1 \alpha \nabla f(\boldsymbol{x})^T \boldsymbol{d}, c \in (0,1)$$

$$\phi(\alpha) \leq \phi(0) + \alpha c_1 \phi'(0)$$

0.2.4 Condizione di curvatura

$$\nabla f(\boldsymbol{x} + \alpha \boldsymbol{d})^T \boldsymbol{d} \geq c_2 \nabla f(x)^T$$

STUDIARE BENE TEOREMA DI ZOUTENDIJK CON DIMOSTRAZIONE