1 Lezione del 04-12-24

1.1 Insiemi di livello

Per la ricerca di massimi e minimi vincolati può essere utile la definizione di insiemi di livello:

Definizione 1.1: Insieme di livello

Definiamo l'insieme di livello di una funzione f per una certa costante k come:

$$\text{lev}_{\leq k} f(x) = \{x \in \mathbb{R}^n : f(x) \leq k\}$$

o alternativamente:

$$\text{lev}_{\geq k} f(x) = \{x \in \mathbb{R}^n : f(x) \geq k\}$$

Si ha che, se l'insieme $\text{lev}_{\leq k} f(x)$ di f esiste ed è compatto per ogni k, allora f ammette minimo, mentre se l'insieme $\text{lev}_{\geq k} f(x)$ esiste ed è compatto, sempre per ogni k, allora f ammette massimo.

1.2 Definizione generale di un algoritmo di PNL

Abbiamo quindi che vorremmo che un algoritmo di PNL, generalmente, fosse costituito da:

- Una **regola** per la scelta delle direzioni di crescita d_k;
- Una **regola** per la scelta del passo t;
- Un **teorema** che dimsotra che le regole di scelta portano sempre a un massimo.

Purtroppo, se non in alcuni casi particolari, questo teorema non esiste. Di base, arriviamo sicuramente ad un *punto stazionario*, cioè un punto $\bar{x}$ dove $\nabla \bar{x} = 0$, cioè vale il teorema:

Teorema 1.1: sui punti di accumulazione di successioni crescenti

Supponiamo $\text{lev}_{\geq x}$ compatti. Allora la successione $x_k \in \mathbb{R}^n$ ha punti di accumulazione, ognuno dei quali è stazionario.

Notiamo che, nel caso di funzioni convesse, significa che il massimo effettivamente viene sempre raggiunto (e l'unico punto a gradiente nullo), come di contro nel caso di funzioni concave il minimo viene sempre raggiunto.

1.2.1 Criterio di stop

Visto che non è assicurato che la serie raggiunga il massimo in un numero finito di passi, introduciamo un **criterio di stop**, cioè che decide quando il risultato trovato finora dall'algoritmo e valido e so può estrarre la soluzione approssimata che ha trovato:

1. Innanzitutto possiamo limitare le iterazioni: prendi x_k con $k < M$ elevato a piacere;

2. Possiamo scegliere il punto di scomparsa del gradiente sotto un certo ε, cioè $|\nabla f(x_k)| \leq \varepsilon$;

3. Posssiamo pensare in termini di *guadagni* veri e propri su ogni passo, cioè $\Delta f = f(x_{k+1}) - f(x_k)$, e porre $\Delta f < \varepsilon$ per qualche ε.

Vediamo quindi alcuni algoritmi di PNL di questo tipo:

- **Metodo del gradiente a passo costante**
  Si pone come regola:
  $$\begin{cases} a_k = \nabla f(x_k) \\ t_k = \eta \end{cases}$$
  con $\eta \in 0 < \eta < \frac{2}{L}$, dove L, detto *numero di Lipschitz*, è il masssimo delle derivate seconde.

  Si ha quindi la relazione di ricorrenza:
  $$x_{k+1} = x_k + \eta \nabla f(x_k)$$
  Il teorema di convergenza sarà allora, semplicemente, che x_k ha punto di accumulazione su f e ogni punto di accumulazione $\bar{x}$ è tale che $\nabla f(\bar{x}) = 0$, $\forall x_0$ posizione di partenza:

- **Metodo del gradiente a passo ideale**
  Si pone come regola:
  $$\begin{cases} d_k = \nabla f(x_k) \\ t_k \in \text{argmax}_{t \geq 0} f(x_k + t\, d_k) \end{cases}$$
  dove si deve avere H invertibile. La relazione di ricorrenza è:
  $$x_{k+1} = x_k + t_k \nabla f(x_k)$$
  Il teorema di convergenza è come sopra;

- **Metodo di Newton**
  Si pone come regola:
  $$\begin{cases} d_k = Hf(x_k)^{-1} \nabla f(x_k) \\ t_k = 1 \end{cases}$$
  da cui la relazione di ricorrenza:
  $$x_{k+1} = x_k - Hf(x_k)^{-1} \nabla f(x_k)$$
  Il teorema di convergenza è come sopra.