

Metodi di ottimizzazione su varietà differenziabili per la media di Fréchet su disco di Poincaré

Candidato: Luca Moroni
Relatore: Bruno Iannazzo

Tesi Informatica

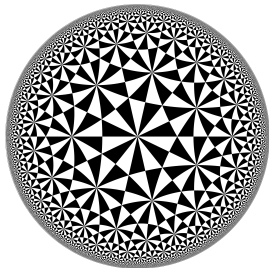
Anno Accademico 2020/2021

Introduzione

Tratteremo metodi di ottimizzazione su varietà, ovvero tecniche di ottimizzazione di funzioni non lineari definite su spazi topologici chiamati varietà differenziabili, i quali rappresentano una generalizzazione dello spazio euclideo.

Introduzione

Parleremo nello specifico delle tecniche di ottimizzazione utilizzate per trovare il centroide di p punti in una varietà differenziabile chiamata **disco di Poincaré**.



Ottimizzazione euclidea

Data $f : \mathbb{R}^n \rightarrow \mathbb{R}$, con f due volte differenziabile, cercheremo di trovare un punto di minimo di f . Condizioni necessarie per dire se x è un punto di minimo:

- ▶ $\nabla f(x) = 0$
- ▶ $\nabla^2 f(x)$ definita positiva

Ottimizzazione euclidea

I metodi di ottimizzazione trattati sono del tipo iterative descent.

- ▶ Partire da un vettore iniziale x^0
- ▶ Creare una sequenza (x^1, x^2, \dots) , tale per cui $f(x^{k+1}) < f(x^k)$, la quale sperabilmente converga ad un punto di minimo

Ottimizzazione euclidea

I metodi iterative descent hanno bisogno di:

- ▶ **direzione di discesa:** dato $g = \nabla f(x)$ allora ogni vettore direzione d tale che $g'd < 0$ è una direzione di discesa
- ▶ **selezione del passo** α

I metodi di ottimizzazione iterativi sono così generalizzati:

$$x^{k+1} = x^k - \alpha^k D^k \nabla f(x^k).$$

con D^k matrice definita positiva.

Metodo a passo fisso

Metodo a passo fisso: fissato $\alpha \in (0, 1]$ la formula di iterazione è la seguente:

$$x^{k+1} = x^k - \alpha \nabla f(x^k)$$

Non è garantita per ogni α la convergenza del metodo.

Metodo di Armijo

Tale metodo seleziona α^k garantendo una decrescita necessaria per la convergenza globale del metodo sotto opportune ipotesi, siano fissati s , γ e σ , con $0 < \gamma < 1$, $0 < \sigma < 1$ e $s > 0$, allora $\alpha^k = \sigma^m s$, dove m è il naturale più piccolo per cui vale:

$$f(x^k) - f(x^k + \sigma^m \lambda d^k) \geq -\gamma \sigma^m \lambda \nabla f(x^k)' d^k.$$

Metodi quasi Newton

I metodi quasi newton sono una famiglia di metodi i quali agiscono nella scelta della matrice D^k tale che approssima la matrice hessiana.

Definiti:

- ▶ $p^k = x^{k+1} - x^k$
- ▶ $q^k = \nabla f(x^{k+1}) - \nabla f(x^k)$

Allora D^{k+1} deve soddisfare $D^{k+1}p^k = q^k$ la quale prende il nome di **equazione delle secanti**.

Ottimizzazione su varietà

La funzione da ottimizzare è $f : M \rightarrow \mathbb{R}$. Dove M è una varietà differenziabile.

Una **retrazione** R_x su x in M è una mappa che va da $T_x M$ in M , con una condizione di rigidità locale che preserva il gradiente in x .

$T_x M$ rappresentante lo spazio tangente ad x di M .

Ottimizzazione su varietà

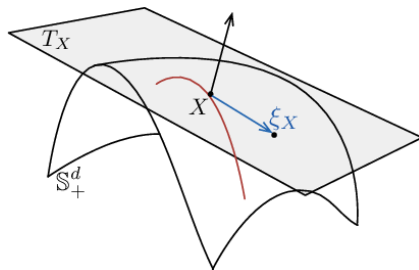
Un caso particolare di retrazione è la **mappa esponenziale**.
Definita su x in M e v in $T_x M$ ha proprietà di retrazione e rappresenta la curva **geodetica** calcolata nel punto 1 originaria in x con derivata in 0 pari a v .

Ottimizzazione su varietà

I metodi iterativi trattati nel caso euclideo hanno una controparte per il caso generale su varietà, la cui formulazione generale è,

$$x^{k+1} = R_{x^k}(t^k \eta^k).$$

Dove appartiene η^k appartenente a $T_{x^k}M$ e t^k è uno scalare.
I metodi sono analoghi al caso euclideo.



Disco di Poincaré

Il disco di Poincaré è uno dei modelli più famosi per la descrizione dello spazio iperbolico.

$$\mathbb{D}^n = \{x \in \mathbb{R}^n \mid \|x\| < 1\}.$$

Il disco di Poincaré è una **varietà riemanniana**.

Lo spazio tangente è rappresentato da \mathbb{R}^n in ogni punto.

La **metrica** nello spazio tangente rispetto a p nel disco:

$$g^{\mathbb{D}} = \lambda_p^2 g^E; \text{ con } \lambda_p = \frac{2}{1 - \|p\|^2} \text{ e con } g^E \text{ la metrica euclidea.}$$

Sono ben note le formule necessarie per effettuare ottimizzazione ovvero, la **distanza** e la **mappa esponenziale**.

Iperboloide

É un altro modello per la rappresentazione dello spazio iperbolico.

$$\mathbb{H}^n = \{x \in \mathbb{R}^{n+1} \mid \langle x, x \rangle_{n:1} = -1; \quad x_{n+1} > 0\}.$$

L'iperboloide è una **varietà riemanniana**.

Lo spazio tangente è rappresentato da,

$$T_p \mathbb{H}^n \cong \{x \in \mathbb{R}^{n:1} \mid \langle p, x \rangle_{n:1} = 0\}.$$

La **metrica** è definita dal **prodotto di Minkowski**.

$$\langle u, v \rangle_{n:1} = \sum_{i=1}^n u_i v_i - u_{n+1} v_{n+1}; \quad u, v \in \mathbb{R}^{n+1}.$$

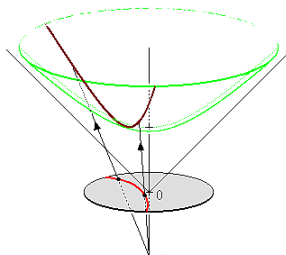
Sono ben note la **distanza** e la **mappa esponenziale**.

Modelli conformi

Le due varietà appena definite sono legate da un **diffeomorfismo conforme** (mantiene gli angoli ma non le distanze), la mappa stereografica ρ , definita come segue.

$$\rho : \mathbb{H}^n \rightarrow \mathbb{D}^n \mid x \rightarrow \frac{1}{x_{n+1} + 1}(x_1, \dots, x_n),$$

$$\rho^{-1} : \mathbb{D}^n \rightarrow \mathbb{H}^n \mid y \rightarrow \frac{2}{1-r}(y_1, \dots, y_n, \frac{1+r}{2}) ; \text{ con } r = \|y\|^2.$$



Modelli conformi

In funzione della mappa conforme che lega le due varietà presentate vale $d_{\mathbb{D}^n}(a, b) = d_{\mathbb{H}^n}(\rho^{-1}(a), \rho^{-1}(b))$, per a e b appartenenti al disco.

Con:

$$d_{\mathbb{D}^n}(a, b) = \operatorname{arccosh}\left(1 + 2 \frac{\|a - b\|^2}{(1 - \|a\|^2)(1 - \|b\|^2)}\right),$$

$$d_{\mathbb{H}^n}(\rho^{-1}(a), \rho^{-1}(b)) = \operatorname{arccosh}(-\langle \rho^{-1}(a), \rho^{-1}(b) \rangle_{n:1}).$$

Media di Fréchet e problema del centroide

Dati p punti in una varietà la media di Fréchet di p punti è l'argomento che minimizza la funzione **somma delle distanze al quadrato**.

$$f(\Theta) = \frac{1}{p} \sum_{i=1}^p d^2(\Theta, x^{(i)}).$$

Definendo la funzione f su disco se esiste un valore di minimo allora esiste un valore di minimo anche per la medesima funzione \hat{f} definita su iperboloide.

Media di Fréchet e problema del centroide

Essendo inoltre $f(x) = \hat{f}(\rho^{-1}(x))$ allora cercare il minimo di f equivale a cercare il minimo di $\hat{f} \circ \rho^{-1}$.

Possiamo perciò trasportare il problema del calcolo della media di Fréchet definito su disco di Poincaré nella varietà iperboloide e viceversa.

Algoritmi

Gli esperimenti che descriveremo si basano sull'utilizzo di tre differenti algoritmi, implementati nella versione riemanniana sia su disco che su iperboloide.

Gli algoritmi implementati sono:

- ▶ Algoritmo a **passo fisso**
- ▶ Algoritmo che si basa sul metodo di **Armijo**
- ▶ Algoritmo di **Barzilai-Borwein**: metodo **quasi Newton**

Esperimenti

Capiremo se le due varietà sono equivalenti, non solo da un punto di vista geometrico ma anche da un punto di vista computazionale.

Abbiamo creato un dataset contenente varie (duecento) istanze del problema della media di Fréchet su disco di Poincaré, per ogni istanza ne abbiamo calcolato il limite.

Esperimenti

Abbiamo utilizzato un sottoinsieme del dataset per definire i parametri “liberi” degli algoritmi.

- ▶ α : dell'algoritmo a passo fisso
- ▶ s : per l'algoritmo di Armijo

Esperimenti

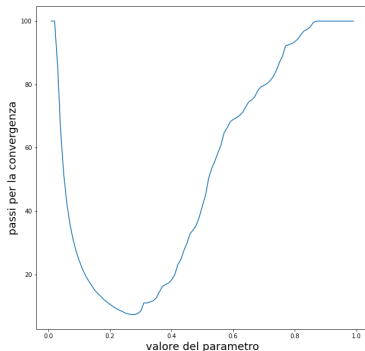
Abbiamo applicato gli algoritmi con un parametro incrementale su ogni istanza del sotto-dataset.

Per ogni esecuzione abbiamo calcolato il numero di passi per la convergenza, generando per ogni istanza una sequenza ed alla fine ne abbiamo calcolato una media.

Abbiamo selezionato i parametri che minimizzano le sequenze medie sia per l'implementazione su disco che su iperboloide.

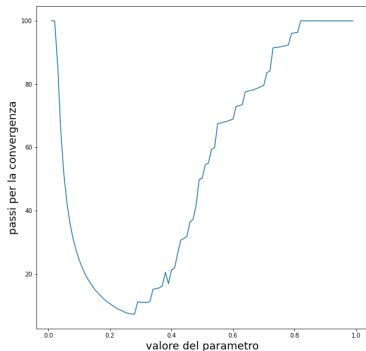
Scelta del parametro passo fisso

Disco di Poincaré



Punto di minimo: 0.27, valore di convergenza minimo: 7.35.

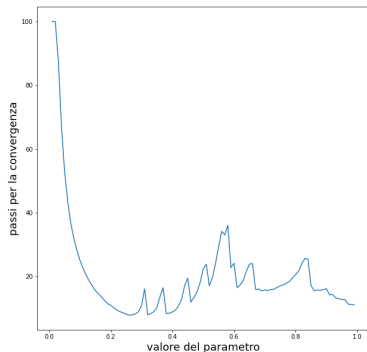
Iperboloide



Punto di minimo: 0.27, valore di convergenza minimo: 7.35.

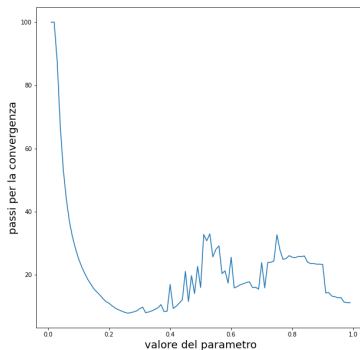
Scelta del parametro metodo di Armijo

Disco di Poincaré



Punto di minimo: 0.26, valore di convergenza minimo: 7.9.

Iperboloide



Punto di minimo: 0.26, valore di convergenza minimo: 7.9.

Esperimenti

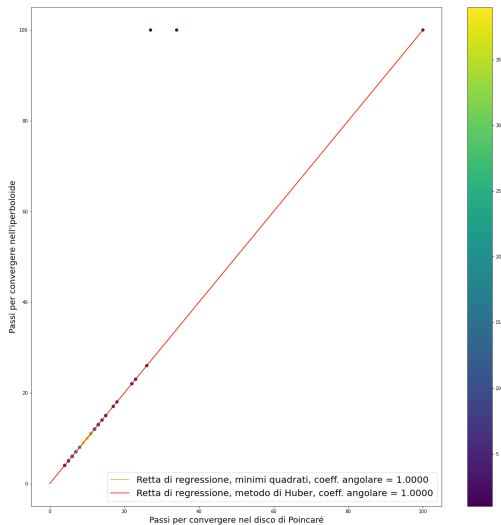
Per ogni algoritmo lo abbiamo eseguito su ogni istanza del dataset totale creando una serie di punti $(a, b) \in \mathbb{R}^2$.

- ▶ a : numero di passi per la convergenza su disco
- ▶ b : il numero di passi per la convergenza su iperboloide

Dalla nuvola di punti si sono eliminati quei punti con una molteplicità bassa.

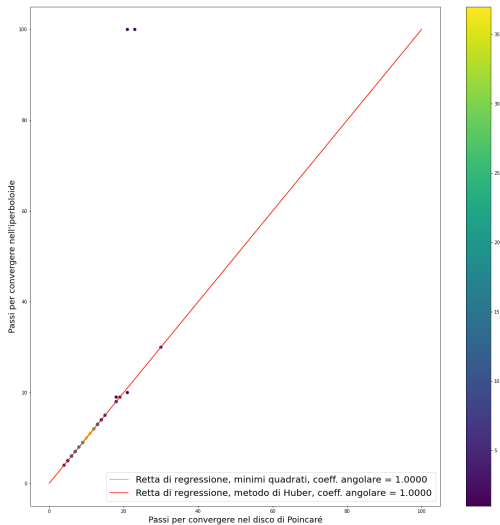
Si sono calcolate due rette di regressione, la prima tramite il metodo dei **minimi quadrati**, la seconda tramite l'algoritmo di **Huber**.

Esperimenti



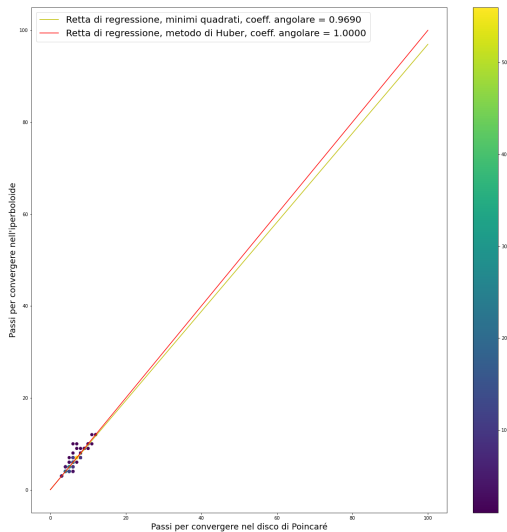
Nuvola di punti relativa all'algoritmo a passo fisso.

Esperimenti



Nuvola di punti relativa all'algoritmo di Armijo.

Esperimenti



Nuvola di punti relativa all'algoritmo di Barzilai-Borwein.

Esperimenti

Confronto passi di convergenza medi per gli algoritmi implementati.

	Passo fisso	Armijo	B.B.
Disco di Poincaré	10.5	10.4	6.13
Iperboloide	11.2	11.2	6.8

Conclusioni

Dai risultati è possibile notare come l'algoritmo di Barzilai-Borwein si comporta meglio rispetto agli algoritmi a passo fisso e di ricerca lineare (Armijo).

Possiamo congetturare che, è possibile trasferire un problema tra due varietà conformi, deve comunque essere preservata l'esistenza di una soluzione ottima per la funzione costo.

FINE.
Grazie per l'attenzione.