

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Corso di Laurea in Fisica

Tesi di Laurea

Un approccio al trasporto ottimale

basato sulla meccanica statistica

Relatore

Prof. Sandro Azaele

Correlatore

Dr. Giacomo Barzon

Laureanda

Miriam Zara

Anno Accademico 2023/2024

Indice

	Introduzione	1
1	Trasporto ottimale tra insiemi discreti	3
1.1	Il problema di Kantorovich discreto	3
1.2	La regolarizzazione entropica	4
1.3	Il formalismo dell'energia libera	6
1.4	Test degli algoritmi sviluppati: trasporto ottimale tra gaussiane 1D	12
2	Un'applicazione alle neuroscienze: quantificare il costo cognitivo	15
2.1	Descrizione del dataset EEG	18
2.2	Analisi del dataset EEG	18
2.3	Descrizione del dataset fMRI	21
2.4	Analisi del dataset fMRI	21
3	Conclusioni	25

Introduzione

La teoria del trasporto ottimale (spesso abbreviato come OT, dall'inglese *Optimal Transport*) nacque con Gaspard Monge nel XVIII secolo, il quale si confrontò con il problema pratico di trasferire beni o risorse da un luogo a un altro, minimizzando il costo complessivo del trasporto. Questo concetto assunse una formulazione più generale ed astratta grazie a Leonid Kantorovich nel XX secolo, che diede vita a un fruttuoso filone di ricerca in matematica e statistica. L'importanza che la teoria dell'OT ha tuttora in queste discipline si evince dai prestigiosi riconoscimenti ricevuti dai suoi principali contributori, come le recenti Medaglie Fields assegnate a Cedric Villani nel 2013 e a Alessio Figalli nel 2018.

Sebbene le sue origini siano radicate in economia e logistica, la teoria del trasporto ottimale ora abbraccia una vasta gamma di campi, tra cui matematica, informatica (in particolare nell'apprendimento automatico e nella visione artificiale), fisica e altri settori. Uno dei suoi punti di forza, per quanto riguarda le applicazioni, è la capacità di confrontare distribuzioni di probabilità in modo matematicamente rigoroso, definendo una distanza sullo spazio delle distribuzioni che soddisfa tutte le proprietà di una metrica. Questo ha portato allo sviluppo di numerose misure statistiche e metriche, rendendo l'OT uno strumento di grande utilità per le Data Sciences, con applicazioni in costante crescita.

Recentemente, l'interesse per il trasporto ottimale si è esteso anche alla meccanica statistica, che ha contribuito allo sviluppo di nuovi algoritmi per la risoluzione dell'OT. Questi nuovi approcci riducono significativamente il costo computazionale rispetto ai tradizionali algoritmi della programmazione lineare, e forniscono un nuovo sguardo teorico sul problema.

Nella prima parte di questa tesi, esploreremo gli aspetti della teoria del trasporto ottimale più legati alla meccanica statistica, concentrandoci sull'utilizzo di una barriera entropica per facilitare l'ottimizzazione e sulla sostituzione della funzione di costo con una formalmente analoga a un'energia libera. Questi approcci saranno descritti nel loro formalismo matematico ed implementati. Nella seconda parte della tesi, presenteremo un'applicazione nel campo delle neuroscienze, utilizzando le proprietà metriche del trasporto ottimale per definire una funzione di costo basata sui dati sperimentali dell'attività cerebrale umana. Cercheremo di verificare se questa funzione di costo sia utile a quantificare la difficoltà associata all'esecuzione di determinate attività cognitive.

1

Trasporto ottimale tra insiemi discreti

1.1 Il problema di Kantorovich discreto

Partiamo da un esempio. Supponiamo che in Veneto ci sia una società di N produttori di latte, che rifornisce M soggetti acquirenti (supermercati, negozi di alimentari, bar e caffetterie..). Chiamiamo a_1, \dots, a_N i litri di latte prodotti da ciascun socio e b_1, \dots, b_M i litri richiesti da ciascun acquirente. Supponiamo che non ci sia avanzo di produzione: complessivamente, i litri di latte prodotti sono esattamente uguali a quelli richiesti

$$\sum_{i=1}^N a_i = \sum_{j=1}^M b_j$$

Un socio produttore rifornisce in generale più di un acquirente, così come un acquirente, allo scopo di soddisfare il suo bisogno, può rivolgersi a più di un socio. Adesso, supponiamo che l'associazione dei produttori abbia affidato il trasporto del latte ad un'unica ditta. Questa ditta ha stabilito per ogni coppia, produttore i -esimo/ acquirente j -esimo, un prezzo per litro di latte trasportato: $C_{i,j}$ (ad esempio basandosi sulla distanza tra i due) e si fa pagare $H = G_{i,j} \cdot C_{i,j}$ euro per il trasporto di $G_{i,j}$ litri. L'associazione dei produttori di latte, a cui spetta sostenere il costo dei trasporti, complessivamente paga $U = \sum_{i=1}^N \sum_{j=1}^M C_{i,j} \cdot G_{i,j}$. Come organizzare i trasporti in modo da minimizzare il costo totale, facendo arrivare ad ogni acquirente la quantità di litri che ha richiesto? Questo è il problema di Kantorovich.

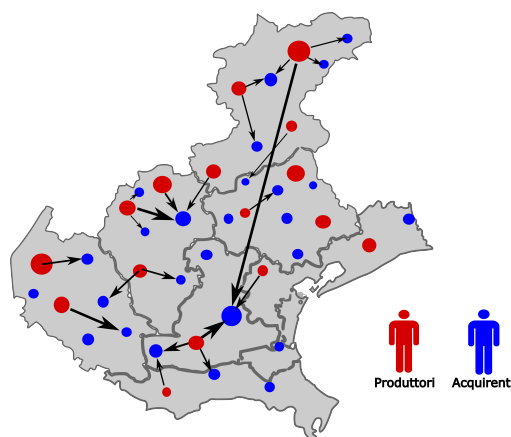


Figura 1.1: Rappresentazione qualitativa di un piano di trasporto di Kantorovich: i cerchi hanno dimensione proporzionale ai a_i , b_j e lo spessore delle frecce è proporzionale ai litri $G_{i,j}$ trasportati. Solo alcune frecce sono disegnate.

Formalizzando, siano:

- A e B due insiemi qualsiasi di N elementi, ovvero il dominio di partenza e il dominio di arrivo (solo per semplificare la notazione, assumiamo che A e B abbiano lo stesso numero di punti)

- VETTORI DI MASSA: $\mathbf{a} = (a_1, \dots, a_N)$ con $a_i \geq 0 \quad \forall i$ e $\mathbf{b} = (b_1, \dots, b_N)$, $b_j \geq 0 \quad \forall j$ le "masse" associate a ciascuno degli elementi di A e di B rispettivamente. Per semplicità, assumiamo d'ora in avanti che la massa totale di ciascun insieme sia unitaria;

$$\sum_{i=1}^N a_i = \sum_{j=1}^N b_j = 1$$

- MATRICE DI COSTO, C: una matrice di numeri reali non negativi, in cui l'entrata $C_{i,j} \geq 0$ rappresenta il costo del trasporto di una massa unitaria dall'elemento i-esimo di A all'elemento j-esimo di B.
- PIANO DI TRASPORTO, G: la matrice $N \times N$ di elementi reali non negativi, in cui l'entrata $G_{i,j}$ rappresenta la quantità di massa trasportata dall'elemento i-esimo di A all'elemento j-esimo di B. Per avere senso fisico, G deve soddisfare le $N^2 + 2N$ condizioni:

$$0 \leq G_{i,j} \leq 1 \quad \forall i, j \quad \sum_{i=1}^N G_{i,j} = b_j \quad \forall j \quad \sum_{j=1}^N G_{i,j} = a_i \quad \forall i$$

- INSIEME DEI PIANI DI TRASPORTO, $\mathcal{G}(\mathbf{a}, \mathbf{b})$: l'insieme delle matrici che definiscono le condizioni appena scritte;

Il problema di Kantorovich è trovare i punti di minimo G_{OPT} della funzione COSTO DI TRASPORTO

$$U(G) := \sum_{i,j=1}^N G_{i,j} \cdot C_{i,j} \equiv \langle G, C \rangle$$

L'esistenza della soluzione è garantita, mentre non lo è la sua unicità [Peyré e Cuturi 2019].

Un caso interessante è quello in cui il dominio di partenza A e il dominio di arrivo B sono contenuti in uno stesso spazio metrico (\mathcal{M}, D) e i vettori di massa \mathbf{a} e \mathbf{b} rappresentano due distribuzioni discrete di probabilità associate agli elementi. Se la matrice di costo è indotta dalla metrica D , il costo del trasporto ottimale a sua volta definisce una funzione che ha tutte le proprietà di una metrica: fornisce quindi un modo per quantificare in modo matematicamente rigoroso la somiglianza tra distribuzioni discrete di probabilità.

Proposizione 1.1 [Peyré e Cuturi 2019, capitolo 2.4] :

Se $C_{i,j} = [D_{i,j}]^p$ per qualche intero $p \geq 1$, allora, detto $U_{OPT} = U(G_{OPT})$ l'ottimo di Kantorovich, $W_p(\mathbf{a}, \mathbf{b}) := \sqrt[p]{U_{OPT}}$ ha le proprietà di una metrica, ovvero:

$$\begin{cases} W_p(\mathbf{a}, \mathbf{b}) \geq 0 \quad \text{e} \quad W_p(\mathbf{a}, \mathbf{b}) = 0 \iff \mathbf{a} = \mathbf{b} \\ W_p(\mathbf{a}, \mathbf{b}) = W_p(\mathbf{b}, \mathbf{a}) \quad \forall \mathbf{a}, \mathbf{b} \in \mathcal{M} \\ W_p(\mathbf{a}, \mathbf{b}) \leq W_p(\mathbf{a}, \mathbf{c}) + W_p(\mathbf{c}, \mathbf{b}) \quad \forall \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathcal{M} \end{cases}$$

W_p è detta p - distanza di Wasserstein.

1.2 La regolarizzazione entropica

Per risolvere numericamente il problema di Kantorovich sono applicabili diversi algoritmi che ormai sono metodi standard nell'ambito della programmazione lineare. In tempi abbastanza recenti, tuttavia, è stato proposto di risolvere il problema da un altro punto di vista, che d'ora in avanti chiameremo REGOLARIZZAZIONE ENTROPICA. Si tratta di considerare una funzione di costo alternativa $U_\epsilon(G)$, dipendente dal parametro reale $\epsilon \geq 0$ definita come:

$$U_\epsilon(G) := U(G) + \epsilon \sum_{i,j=1}^N G_{i,j} \cdot [\ln G_{i,j} - 1] = U(G) - \epsilon S(G)$$

$S(G)$ rappresenta l'entropia di Shannon associata alla matrice G . Un vantaggio della regolarizzazione entropica è il fatto che definisce un problema strettamente convesso, per cui la soluzione è unica, a differenza di quanto succede per il problema di Kanthorovich originario.

Proposizione 1.2 [Peyré e Cuturi 2019, capitolo 4] :

$\forall \epsilon > 0$, esiste un'unico punto di minimo $G^\epsilon \in \mathcal{G}(\mathbf{a}, \mathbf{b})$ per U_ϵ .

Per $\epsilon \rightarrow 0^+$, $G^\epsilon \rightarrow \left\{ \arg \min_G S(G) : G \in \mathcal{G}(\mathbf{a}, \mathbf{b}) \text{ e } \langle G; C \rangle = U_{\text{OPT}} \right\}$. La soluzione ad ϵ fissato tende nel limite di $\epsilon = 0$ ad una delle soluzioni del problema di Kanthorovich. In particolare, tra tutte le soluzioni, tende ad una soluzione che massimizza l'entropia $S(G)$.

Per $\epsilon \rightarrow +\infty$, $G^\epsilon \rightarrow \mathbf{a} \mathbf{b}^T$ (Nota: se x e y sono variabili indipendenti distribuite secondo \mathbf{a} e \mathbf{b} rispettivamente, la coppia di variabili (x, y) è distribuita secondo $\mathbf{a} \mathbf{b}^T$).

Un altro vantaggio sta nel fatto che diventa possibile trovare la soluzione ad ϵ fissato attraverso uno schema molto semplice, che è detto ALGORITMO DI SINKHORN e consente di abbassare il costo computazionale del problema rispetto agli algoritmi standard della programmazione lineare ($\sim \mathcal{O}(N^2)$, rispetto a $\mathcal{O}(N^3)$). Lo svantaggio principale, invece è che per valori di ϵ prossimi allo zero è di fatto impossibile non incorrere in errori causati da instabilità numeriche. Alcune raffinazioni dell'algoritmo di Sinkhorn sono state sviluppate per arginare almeno in parte questo problema [vedi Peyré e Cuturi 2019, capitolo 4.4 "Stability and Log Domain Computations"].

Qui descriviamo la versione base dell'algoritmo di Sinkhorn. La condizione di non negatività delle entrate $G_{i,j}$ è già implicitamente presa in considerazione per la presenza della funzione logaritmo in $S(G)$. Il problema che ci poniamo è minimizzare $U_\epsilon(G)$ sotto i $2N$ vincoli:

$$\sum_{i=1}^N G_{i,j} = b_j \quad \forall j \quad \sum_{j=1}^N G_{i,j} = a_i \quad \forall i$$

Per fare questo usiamo il metodo dei moltiplicatori di Lagrange.

METODO DEI MOLTIPLICATORI DI LAGRANGE

Sia $f(x) : A \in \mathbb{R}^n \rightarrow \mathbb{R}$ con A aperto una funzione differenziabile e sia $\mathcal{M} \in \mathbb{R}^n$ una varietà differenziale di dimensione $m \leq n$ descritta in forma vincolare da:

$$\mathcal{M} = \left(x \in \mathbb{R}^n : \begin{cases} g_1(x) = 0, \\ g_2(x) = 0, \\ \vdots \\ g_{n-m}(x) = 0 \end{cases} \right) \quad (1.1)$$

per il teorema di Lagrange, $x_0 \in \mathcal{M}$ è un punto stazionario per f in \mathcal{M} se e solo se

$$\exists \Lambda = (\lambda_1, \dots, \lambda_{n-m}) \neq 0 \text{ t.c. } \nabla f(x_0) = \lambda_1 \cdot \nabla g_1(x_0) + \dots + \lambda_{n-m} \cdot \nabla g_{n-m}(x_0) \quad (1.2)$$

Per trovare i punti stazionari di f in \mathcal{M} si deve quindi risolvere per $x \in \mathbb{R}^n$ il sistema di $(n-m) + n$ equazioni in $(n-m) + n$ variabili definite da 1.1 e 1.2. Questo sistema coincide evidentemente con

$$\begin{cases} \nabla_x \mathcal{L} = 0, \\ \nabla_\Lambda \mathcal{L} = 0 \end{cases} \quad \text{definita } \mathcal{L}(x, \Lambda) := f(x) - \left\langle \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_{n-m} \end{pmatrix}, \begin{pmatrix} g_1(x) \\ \vdots \\ g_{n-m}(x) \end{pmatrix} \right\rangle$$

Definiamo le $2N$ variabili ausiliarie $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_N)$ e $\boldsymbol{\mu} = (\mu_1, \dots, \mu_N)$ e la funzione Lagrangiana \mathcal{L}_ϵ :

$$\mathcal{L}_\epsilon(G, \boldsymbol{\lambda}, \boldsymbol{\mu}) := U_\epsilon(G) - \left\langle \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_N \end{pmatrix}, \begin{pmatrix} \sum_{j=1}^N G_{1,j} - a_1 \\ \vdots \\ \sum_{j=1}^N G_{N,j} - a_N \end{pmatrix} \right\rangle - \left\langle \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix}, \begin{pmatrix} \sum_{i=1}^N G_{i,1} - b_1 \\ \vdots \\ \sum_{i=1}^N G_{i,N} - b_N \end{pmatrix} \right\rangle$$

Un punto stazionario di \mathcal{L}_ϵ deve soddisfare $\vec{\nabla} \mathcal{L}_\epsilon = 0$, ovvero:

$$C_{k,l} + \epsilon \cdot [1 + \ln G_{k,l}] - \lambda_k - \mu_l = 0 \quad \forall k, l; \quad \sum_{i=1}^N G_{i,j} = b_j \quad \forall j; \quad \sum_{j=1}^N G_{i,j} = a_i \quad \forall i$$

Le prime N^2 equazioni si riscrivono come:

$$\ln G_{k,l} = \left(\frac{\lambda_k}{\epsilon} + \frac{1}{2} \right) + \left(\frac{\mu_l}{\epsilon} - \frac{1}{2} \right) - \left(\frac{C_{k,l}}{\epsilon} \right)$$

.Definendo:

$$A_\epsilon(k) := \exp\left(\frac{\lambda_k}{\epsilon} - \frac{1}{2}\right), \quad B_\epsilon(l) := \exp\left(\frac{\mu_l}{\epsilon} - \frac{1}{2}\right), \quad K_\epsilon(k, l) := \exp\left(-\frac{C(k, l)}{\epsilon}\right)$$

Il sistema di $N^2 + 2N$ equazioni si riscrive in forma matriciale compatta come:

$$\begin{cases} G = \text{diag } A_\epsilon K \text{ diag } B_\epsilon & \text{eq. 1} \\ \mathbf{A}_\epsilon \otimes K \mathbf{B}_\epsilon = \mathbf{a} & \text{eq. 2} \\ \mathbf{B}_\epsilon \otimes K^T \mathbf{A}_\epsilon = \mathbf{b} & \text{eq. 3} \end{cases}$$

dove con $\text{diag } A_\epsilon$ si intende la matrice diagonale ottenuta dal vettore A_ϵ e " \otimes " è il prodotto di Handamard $[A \otimes B]_{i,j} := A_{i,j} \cdot B_{i,j}$. Il sistema composto dalle equazioni 2 + 3 si risolve rispetto ad \mathbf{A}_ϵ e \mathbf{B}_ϵ in modo iterativo. Una volta trovati \mathbf{A}_ϵ e \mathbf{B}_ϵ con la precisione desiderata, si calcola la matrice piano di trasporto G usando (eq.1): il numero di variabili indipendenti del problema si è così ridotto da N^2 (le entrate di G) a $2N$ (le entrate di \mathbf{A}_ϵ e \mathbf{B}_ϵ).

ALGORITMO 1: SINKHORN AD ϵ FISSATO

Input: vettori \mathbf{a} e \mathbf{b} , matrice di costo C , parametro di regolarizzazione ϵ , parametro di convergenza δ

Inizializzazione: calcolo di K , inizializzazione di $\mathbf{B}_\epsilon = \mathbf{1}$

Loop:

1. aggiorna A: $\mathbf{A}^{(n+1)} = \mathbf{a} \oslash (K \mathbf{B}^{(n)})$
2. aggiorna B: $\mathbf{B}^{(n+1)} = \mathbf{b} \oslash (K^T \mathbf{A}^{(n+1)})$
3. controlla convergenza: **if** $\|\mathbf{B}_\epsilon^{(n+1)} - \mathbf{B}_\epsilon^{(n)}\| \leq \delta$ **stop**

Output: piano di trasporto $G_\epsilon = \text{diag } A_\epsilon^{(n+1)} K \text{ diag } B_\epsilon^{(n+1)}$ e costo di trasporto $U_\epsilon(G_\epsilon)$

\oslash è la divisione di Handamard, definita in modo analogo al prodotto.

Guardando le definizioni date per \mathbf{A}_ϵ , \mathbf{B}_ϵ e K_ϵ risulta chiaro quanto detto prima a proposito delle instabilità numeriche del SINKHORN: l'argomento di $\exp\left(-\frac{C}{\epsilon}\right)$ diverge per $\epsilon \rightarrow 0^+$. Quindi, nella pratica, è impossibile avvicinarsi indefinitamente alla soluzione esatta senza incorrere nell'errore di overflow.

1.3 Il formalismo dell'energia libera

Descriviamo ora un altro formalismo recentemente proposto [Koehl, Delarue e Orland 2019] che, come quello della regolarizzazione entropica, permette di trovare una soluzione approssimata per il piano di trasporto, dipendente da un parametro.

Il punto di partenza è l'osservazione che la funzione di costo regolarizzata $U_\epsilon(G) = U(G) - \epsilon \cdot S(G)$ assomiglia formalmente ad una energia libera di Helmholtz $F = U - T S$, a patto di vedere $S(G)$ come un'entropia e il parametro di regolarizzazione ϵ come una temperatura T . Dalla meccanica statistica ricordiamo che sistema in equilibrio alla temperatura T esibisce numerosi microstati. La densità di

probabilità del microstato è funzione dell'energia del sistema e il più probabile coincide con quello ad energia minima. Consideriamo l'insieme dei piani di trasporto $G \in \mathcal{G}(\vec{a}, \vec{b})$ come i possibili microstati di un ensemble canonico a temperatura fissata $T \equiv \epsilon =: \frac{1}{\beta}$.

sistema	gas ideale	trasporti
macrostato	β, V, N	β, N
microstato	$\mathbf{x} = \{(\mathbf{q}_i, \mathbf{p}_i)\}_{i=1, \dots, N} \in \mathbb{R}^{6N}$	$G \in \mathcal{G}(\mathbf{a}, \mathbf{b})$
funzione energia	$H(\mathbf{x}) = \sum_{i=1}^N \frac{p_i^2}{2m}$	$U(G) = \sum_{i,j} C_{i,j} G_{i,j}$
densità di probabilità del microstato	$p(\mathbf{x} \beta) = \frac{1}{Z(\beta, V, N)} \cdot e^{-\beta H(\mathbf{x})}$	$p(G \beta) = \frac{1}{Z(\beta, N)} \cdot e^{-\beta \cdot U(G)}$
energia interna termodinamica	$\mathbb{E}[U] = -\frac{\partial}{\partial \beta} \log Z(\beta, V, N)$	$\mathbb{E}[U] = -\frac{\partial}{\partial \beta} \log Z(\beta, N)$
energia libera di Helmholtz	$F = -\frac{1}{\beta} \log Z(\beta, V, N)$	$F = -\frac{1}{\beta} \log Z(\beta, N)$

Tabella 1.1: Corrispondenze formali tra l'ensemble dei piani di trasporto e l'ensemble canonico "standard" del gas ideale.

Avendo associato al piano di trasporto G la densità di probabilità $p(G) = e^{-\beta \cdot U(G)}$, il problema di calcolare il punto di minimo di $U(G)$ equivale a quello di calcolare il valore di aspettazione del piano di trasporto $\mathbb{E}[G]$ nel limite di $\beta \rightarrow +\infty$.

A valore di β fissato, finito, il valore di aspettazione $\mathbb{E}[G]$ è dato da:

$$\begin{aligned} \mathbb{E}[G_{i,j}] &:= \int_{\mathcal{G}} d\mu \ G_{i,j} \cdot p(G) = \int_{\mathcal{G}} d\mu \ G_{i,j} \cdot \frac{e^{-\beta \sum_{k,l} C_{k,l} G_{k,l}}}{Z_{\beta}} = \frac{1}{Z_{\beta}} \int_{\mathcal{G}} d\mu -\frac{1}{\beta} \frac{\partial}{\partial C_{i,j}} \left[e^{-\beta \sum_{k,l} C_{k,l} G_{k,l}} \right] \\ &= -\frac{1}{\beta} \cdot \frac{1}{Z_{\beta}} \cdot \frac{\partial}{\partial C_{i,j}} \left[\int_{\mathcal{G}} d\mu \ e^{-\beta \cdot U(G)} \right] = -\frac{1}{\beta} \cdot \frac{\partial}{\partial C_{i,j}} \ln(Z_{\beta}) \end{aligned}$$

La funzione di partizione $Z(\beta)$ non è analiticamente calcolabile. È però possibile approssimarla attraverso il metodo di punto sella nel limite $\beta \rightarrow +\infty$ che ci interessa. Esplicitando i vincoli imposti alle matrici $G \in \mathcal{G}$, la funzione di partizione è:

$$Z_{\beta}(\mathbf{a}, \mathbf{b}) = \int_0^1 dG_{1,1} \cdots \int_0^1 dG_{N,N} \left[e^{-\beta \cdot U(G)} \cdot \prod_{l=1}^N \delta \left(\sum_{i=1}^N G_{i,l} - b_l \right) \cdot \prod_{k=1}^N \delta \left(\sum_{j=1}^N G_{k,j} - a_k \right) \right]$$

dove le δ sono delta di Dirac. Cominciamo esprimendo le delta di Dirac attraverso il loro sviluppo in serie di Fourier, $\delta(x) = \int_{-\infty}^{+\infty} dt e^{-itx}$:

$$Z_{\beta}(\mathbf{a}, \mathbf{b}) = \int_0^1 dG_{1,1} \cdots \int_0^1 dG_{N,N} \left[e^{-\beta U(G)} \cdot \prod_{l=1}^N \int_{-\infty}^{+\infty} d\mu_l e^{-i\mu_l [\sum_i G_{i,l} - b_l]} \cdot \prod_k \int_{-\infty}^{+\infty} d\lambda_k e^{-i\lambda_k [\sum_j G_{k,j} - a_k]} \right]$$

Abbiamo introdotto $2N$ variabili mute di integrazione, $\{\lambda_1, \dots, \lambda_N$ e $\mu_1, \dots, \mu_N\}$. Adesso facciamo innanzitutto un cambio di variabile $\beta \tilde{\lambda}_k = i \lambda_k$, $\beta \tilde{\mu}_l = i \mu_l \rightarrow \tilde{\lambda}_k = \frac{i}{\beta} \cdot \lambda_k$, $\tilde{\mu}_l = \frac{i}{\beta} \cdot \mu_l$.

$$\begin{aligned} Z_{\beta}(\mathbf{a}, \mathbf{b}) &= \left(\frac{\beta}{i} \right)^{2N} \int_0^1 \prod_{k,l} dG_{k,l} e^{-\beta \cdot U(G)} \cdot \int_{-i\infty}^{+i\infty} d\mu_l e^{-\beta \cdot \mu_l \cdot [\sum_{n=1}^N G(n,l) - b_l]} \\ &\quad \cdot \int_{-i\infty}^{+i\infty} d\lambda_k e^{-\beta \cdot \lambda_k \cdot [\sum_{n=1}^N G(k,n) - a_k]} \end{aligned}$$

Adesso riorganizziamo i termini:

$$Z_{\beta}(\mathbf{a}, \mathbf{b}) = \beta^{2N} \cdot (-1)^N \int \prod_l d\mu_l e^{\beta \mu_l \cdot b_l} \int \prod_k d\lambda_k e^{\beta \lambda_k \cdot a_k} \int_0^1 dG_{k,l} e^{-\beta \cdot [G_{k,l} \cdot C_{k,l} + \lambda_k + \mu_l]}$$

facendo gli integrali in G_{kl} , otteniamo:

$$Z_\beta(\mathbf{a}, \mathbf{b}) = \int \prod_k d\lambda_k \int \prod_l d\mu_l \left[\frac{1}{\beta[G_{k,l} \cdot C_{k,l} + \lambda_k + \mu_l]} \cdot (1 - \exp -\beta[G_{k,l} \cdot C_{k,l} + \lambda_k + \mu_l]) \right]$$

che possiamo riscrivere in forma compatta come:

$$Z_\beta(\mathbf{a}, \mathbf{b}) = \int \prod_k d\lambda_k \int \prod_l d\mu_l e^{-\beta F_\beta(\lambda, \mu)}$$

definita $F_\beta(\lambda, \mu)$:

$$F_\beta(\lambda, \mu) := - \sum_k \lambda_k \cdot a_k - \sum_l \mu_l \cdot b_l - \frac{1}{\beta} \sum_{k,l} \ln \left[\frac{1 - e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}}{\beta \cdot [C_{k,l} + \lambda_k + \mu_l]} \right]$$

L'integrale non dipende dal cammino di integrazione. Per $\beta \rightarrow +\infty$ (limite di soluzione esatta al problema di Kanthorovich) si può approssimare con il metodo di punto-sella.

Proposizione 1.4: La funzione $F_\beta(\lambda, \mu)$ è debolmente concava in \mathbb{R}^{2N} ; in particolare, l'autovalore nullo ha molteplicità geometrica 1. La funzione $\tilde{F}_\beta(\lambda, \mu_1, \dots, \mu_{N-1}) := F_\beta(\lambda, \mu_1, \dots, \mu_{N-1}; \mu_N = 0)$ è strettamente concava in \mathbb{R}^{2N-1} . Di conseguenza esiste un unico punto di massimo (globale).

Il punto di massimo per \tilde{F}_β , $(\lambda^*, \mu^*) \in \mathbb{R}^{2N-1}$. (λ^*, μ^*) deve soddisfare le $2N - 1$ equazioni:

$$\frac{\partial \tilde{F}_\beta}{\partial \lambda_k} = 0 \quad \forall k = 1, \dots, N \quad \frac{\partial \tilde{F}_\beta}{\partial \mu_l} = 0 \quad \forall l = 1, \dots, N - 1$$

Si vede immediatamente che

$$\frac{\partial \tilde{F}_\beta}{\partial \lambda_k}(\lambda_1, \dots, \mu_{N-1}) \equiv \frac{\partial F_\beta}{\partial \lambda_k}(\lambda_1, \dots, \mu_{N-1}; \mu_N = 0)$$

e similmente per $\partial \mu_l$ con $l = \{1, \dots, N - 1\}$ e per tutte le derivate di ordine superiore.

Sviluppiamo i conti:

$$\begin{aligned} \frac{\partial F_\beta}{\partial \lambda_k} &= -a_k - \frac{1}{\beta} \cdot \sum_l \frac{\partial}{\partial \lambda_k} \ln \left[\frac{1 - e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}}{\beta \cdot [C_{k,l} + \lambda_k + \mu_l]} \right] = -a_k - \frac{1}{\beta} \sum_l \left(\frac{\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}{1 - e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}} \right) \\ &\quad \cdot \left(\frac{\beta^2 \cdot [C_{k,l} + \lambda_k + \mu_l] \cdot e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]} - \beta \cdot (1 - e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]})}{\beta^2 \cdot [C_{k,l} + \lambda_k + \mu_l]^2} \right) \\ &= -a_k - \sum_l \left(-\frac{1}{\beta [C_{k,l} + \lambda_k + \mu_l]} + \frac{e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}}{[1 - e^{-\beta \cdot [C_{k,l} + \lambda_k + \mu_l]}} \right) \\ &= -a_k + \sum_{l=1}^N \left[\frac{1}{\beta [C(k,l) + \lambda_k + \mu_l]} + \frac{e^{-\beta [C(k,l) + \lambda_k + \mu_l]}}{[e^{-\beta [C(k,l) + \lambda_k + \mu_l]} - 1]} \right] \\ &= -a_k + \sum_{l=1}^N \Phi(\beta [C_{k,l} + \lambda_k + \mu_l]) \quad \text{con} \quad \Phi(x) := \frac{1}{x} + \frac{e^{-x}}{e^{-x} - 1} \end{aligned}$$

I conti per le derivate parziali $\partial/\partial \mu$ sono identici. Il punto di massimo $(\lambda^*, \mu^*) \in \mathbb{R}^{2N-1}$ per \tilde{F}_β risolve il set di $2N - 1$ equazioni:

$$\begin{cases} a_k &= \sum_{l=1}^{N-1} \Phi(\beta [C_{k,l} + \lambda_k^* + \mu_l^*]) + \Phi(\beta [C_{k,N} + \lambda_k^*]) \quad \forall k = \{1, \dots, N\} \\ b_l &= \sum_{k=1}^N \Phi(\beta [C_{k,l} + \lambda_k^* + \mu_l^*]) \quad \forall l = \{1, \dots, N - 1\} \end{cases} \quad (1.3)$$

Adesso calcoliamo la matrice hessiana $H_{F_\beta}(\boldsymbol{\lambda}, \boldsymbol{\mu})$:

$$\begin{aligned} \frac{\partial^2 F_\beta}{\partial \lambda_i \partial \lambda_j} &= \delta_{i,j} \cdot \beta \sum_{l=1}^N \Phi'(\beta[C_{i,l} + \lambda_i + \mu_l]) & \frac{\partial^2 F_\beta}{\partial \lambda_i \partial \mu_j} &= \beta \cdot \Phi'(\beta[C_{i,j} + \lambda_i + \mu_j]) \\ \frac{\partial^2 F_\beta}{\partial \mu_i \partial \mu_j} &= \delta_{i,j} \cdot \beta \sum_{k=1}^N \Phi'(\beta[C_{k,j} + \lambda_k + \mu_i]) & \text{con } \Phi'(x) &= \frac{d\Phi}{dx}(x) = \frac{e^{-x}}{(e^{-x} - 1)^2} - \frac{1}{x^2} \end{aligned}$$

La matrice hessiana ha una struttura a blocchi:

$$H_{F,\beta}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{bmatrix} \partial \lambda_i \partial \lambda_j F_\beta & \partial \lambda_i \partial \mu_j F_\beta \\ \partial \mu_i \partial \lambda_j F_\beta & \partial \mu_i \partial \mu_j F_\beta \end{bmatrix}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{bmatrix} D_\lambda & G' \\ (G')^T & D_\mu \end{bmatrix}(\boldsymbol{\lambda}, \boldsymbol{\mu}) \quad (1.4)$$

con D_λ e D_μ blocchi $N \times N$ diagonali e $G'_{k,l} := \beta \cdot \Phi'(\beta[C_{k,l} + \lambda_k + \mu_l])$. La matrice hessiana di \tilde{F}_β è semplicemente la sottomatrice di $H_{F,\beta}$ ottenuta togliendo l'ultima riga e l'ultima colonna, calcolata in $\mu_N = 0$. In simboli,

$$H_{\tilde{F},\beta}(\boldsymbol{\lambda}, \mu_1 \cdots \mu_{N-1}) = [H_{F,\beta}]^{N,N}(\boldsymbol{\lambda}, \mu_1 \cdots \mu_{N-1}; \mu_N = 0)$$

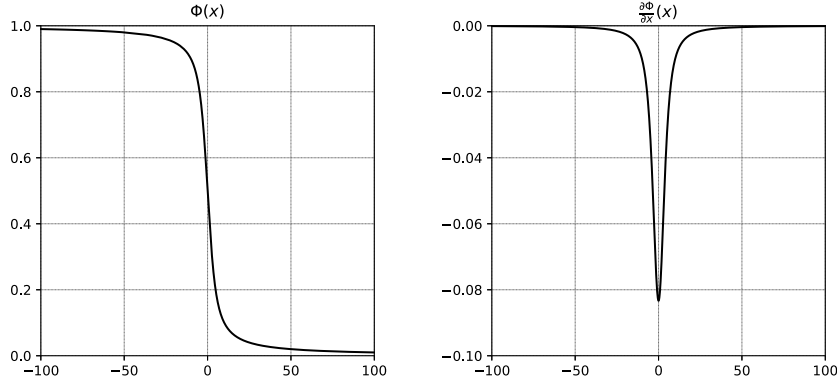


Figura 1.2: Le funzioni Φ e Φ' . Osserviamo che il codominio di Φ è contenuto nell'intervallo $[0, 1]$, coerentemente con la condizione $G_{i,j} \in [0, 1]$. La derivata Φ' è negativa e limitata.

Osservando che Φ' è negativa, è semplice verificare che $H_{F,\beta}$ è una matrice semidefinita negativa, ovvero, da definizione, che

$$Q(\mathbf{x}_1, \mathbf{x}_2) := \left\langle \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}, \begin{pmatrix} D_\lambda & G' \\ (G')^T & D_\mu \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} \right\rangle \leq 0 \quad \forall \mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^{2N}$$

Facciamo i conti:

$$\begin{aligned} Q(\mathbf{x}_1, \mathbf{x}_2) &= \langle \mathbf{x}_1, D_\lambda \mathbf{x}_1 \rangle + \langle \mathbf{x}_1, G' \mathbf{x}_2 \rangle + \langle \mathbf{x}_2, (G')^T \mathbf{x}_1 \rangle + \langle \mathbf{x}_2, D_\mu \mathbf{x}_2 \rangle \\ &= \sum_{i=1}^N x_{1,i}^2 D_{\lambda,i,i} + \sum_{i=1}^N \sum_{j=1}^N x_{1,i} (G')'_{i,j} x_{2,j} + \sum_{i=1}^N \sum_{j=1}^N x_{2,j} (G')^T_{j,i} x_{1,i} + \sum_{j=1}^N x_{2,j}^2 D_{\mu,j,j} \\ &= \sum_{i,j} [x_{1,i}^2 + 2 \cdot x_{1,i} x_{2,j} + x_{2,j}^2] G'_{i,j} = \sum_{i,j} [x_{1,i} + x_{2,j}]^2 G'_{i,j} \\ &= \beta \cdot \sum_{i,j} [x_{1,i} + x_{2,j}]^2 \cdot \Phi'[\beta(C_{i,j} + \lambda_i + \mu_j)] \\ &\leq 0 \quad \forall (\mathbf{x}_1, \mathbf{x}_2) \in \mathbb{R}^{2N} \end{aligned}$$

Inoltre, $Q(\mathbf{x}) = 0 \iff \mathbf{x}_1 = -\mathbf{x}_2$, quindi l'autovalore nullo è assunto e ha molteplicità geometrica uguale a 1. Questo dimostra la prima parte della [proposizione 1.4]. Fissare una delle variabili a zero permette di eliminare l'autovalore nullo, pertanto la matrice hessiana di $\tilde{F}_\beta(\boldsymbol{\lambda}, \mu_1, \dots, \mu_{N-1})$ è definita negativa.

Detto $(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) \in \mathbb{R}^{2N-1}$ il punto di massimo globale di \tilde{F}_β , l'approssimazione di punto-sella di $Z_\beta(\mathbf{a}, \mathbf{b})$ è:

$$Z_\beta(\mathbf{a}, \mathbf{b}) \simeq \sqrt{\left(\frac{2\pi}{\beta}\right)^{2N-1}} \cdot \frac{e^{-\beta F_\beta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)}}{\sqrt{|\det(H_{F_\beta})(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)|}} \quad \text{per } \beta \rightarrow +\infty$$

Il determinante dell'hessiana [eq. 1.4] ha l'andamento asintotico

$$|\det H_\beta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)| \sim \beta^{-2N-1} \cdot \text{funzione di } (C, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$$

Quindi, per il valore di aspettazione $\mathbb{E}(G)$ [Eq. 1.5] si ottiene:

$$\mathbb{E}(G_{i,j}) = -\frac{1}{\beta} \cdot \frac{\partial}{\partial C_{i,j}} \ln(Z_\beta) \simeq -\frac{1}{\beta} \cdot \frac{\partial}{\partial C_{i,j}} [-\beta F_\beta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) - \ln[\text{cost} \cdot \text{funzione di } (C, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)]]$$

trascurando il termine sub-dominante,

$$\mathbb{E}[G]_\beta(k, l) \simeq \frac{\partial}{\partial C_{i,j}} F_\beta(\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \frac{1}{\beta[C_{k,l} + \lambda_k^* + \mu_l^*]} + \frac{e^{-\beta[C_{k,l} + \lambda_k^* + \mu_l^*]}}{e^{-\beta[C_{k,l} + \lambda_k^* + \mu_l^*]} - 1} = \Phi(\beta[C_{k,l} + \lambda_k^* + \mu_l^*])$$

Da queste equazioni si può computare, a β fissato, un piano ottimo G_{OPT}^β e un costo associato $U_{OPT}^\beta(\mathbf{a}, \mathbf{b})$.

$$G_{OPT}^\beta \equiv \mathbb{E}[G](\beta) \simeq \Phi(\beta[C_{k,l} + \lambda_k^* + \mu_l^*]), \quad U_{OPT}^\beta(\mathbf{a}, \mathbf{b}) := \langle G_{OPT}^\beta, C \rangle \quad (1.5)$$

Il piano ottimo a β fissato, G_{OPT}^β , approssima una delle soluzioni del problema di Kantorovich originale. Inoltre, se il costo ottimo di Kantorovich ha proprietà metriche, queste sono conservate anche dal costo ottimo a β fissato $U_{OPT}^\beta(\mathbf{a}, \mathbf{b})$:

Proposizione 1.3: [Koehl, Delarue e Orland 2019]

Siano $\mathbf{a}, \mathbf{b} \in (\mathcal{M}, D)$ spazio metrico, con D metrica.

1. $\forall \beta > 0$, se $C = [D_{i,j}]^p$, allora $U_{OPT}^\beta(\mathbf{a}, \mathbf{b})$ è simmetrica e soddisfa la disuguaglianza triangolare;
2. U_{OPT}^β è una funzione monotona decrescente di β e per $\beta \rightarrow +\infty$ converge all'ottimo di Kantorovich.

Il formalismo dell'energia libera offre un metodo alternativo alla regolarizzazione entropica per approssimare una soluzione esatta del problema di Kantorovich quando la temperatura T tende a zero. Questo è l'utilizzo che ne faremo nel presente lavoro di tesi. Tuttavia, osserviamo brevemente che la sua potenzialità non si limita a questo scenario specifico ($T \rightarrow 0^+$): potremmo anche utilizzare questo formalismo per trovare il piano ottimale a una temperatura finita, superiore a zero. In questa situazione, consideriamo il sistema dei piani di trasporto immerso in un'ambiente esterno che interferisce con esso (ad esempio, fornendo rumore stocastico) impedendo così al sistema di raggiungere l'ottimo di Kantorovich. La soluzione ottimale a T fissata diventa dunque un sub-ottimo. Questa metodologia potrebbe risultare utile in applicazioni dell'OT in cui la scelta della funzione di costo da minimizzare è in una certa misura arbitraria e l'introduzione del parametro di regolarizzazione $T > 0$ potrebbe modellare la presenza di un eventuale rumore stocastico.

Calcolo numerico del punto di sella

Per trovare la soluzione delle equazioni [Eq. 1.3] implementiamo il metodo iterativo di Newton-Raphson. Introdotto il campo vettoriale $A : \mathbb{R}^{2N-1} \rightarrow \mathbb{R}^{2N-1}$:

$$A_k := -\frac{\partial \tilde{F}_\beta}{\partial \lambda_k} \quad \forall k = 1, \dots, N \quad A_l := -\frac{\partial \tilde{F}_\beta}{\partial \mu_l} \quad \forall l = 1, \dots, N-1$$

dobbiamo trovare $\mathbf{X} = (\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ tale che $A(\mathbf{X}) = 0$. Partiamo da un guess iniziale \mathbf{X}_0 : se la soluzione \mathbf{X} si trova all'intorno di \mathbf{X}_0 , cioè $\mathbf{X} = \mathbf{X}_0 + \delta\mathbf{X}$, allora

$$\mathbf{A}(\mathbf{X}) = 0_{2N-1} \simeq \mathbf{A}(\mathbf{X}_0) + J_{\mathbf{A}}(\mathbf{X}_0) \cdot (\mathbf{X} - \mathbf{X}_0) \quad \rightarrow \quad J_{\mathbf{A}}(\mathbf{X}_0) \cdot \delta\mathbf{X} = -\mathbf{A}(\mathbf{X}_0)$$

Lo jacobiano $J_{\mathbf{A}}$ per costruzione è l'opposto della matrice hessiana di \tilde{F}_β (1.4)

$$J_{\mathbf{A}}(\mathbf{X}) = -H_{\tilde{F}_\beta}(\mathbf{X}) = -\begin{bmatrix} D_\lambda & G' \\ G'^T & D_\mu \end{bmatrix}(\mathbf{X})$$

qualche passaggio algebrico per riscrivere il sistema $J_{\mathbf{A}}(\mathbf{X}_0) \cdot \delta\mathbf{X} = -\mathbf{A}(\mathbf{X}_0)$ in una forma conveniente:

$$\begin{cases} \text{(I)} & D_\lambda \cdot \delta\lambda + G' \cdot \delta\mu = +\mathbf{A}_\lambda \\ \text{(II)} & (G')^T \cdot \delta\lambda + D_\mu \cdot \delta\mu = +\mathbf{A}_\mu \end{cases} \quad \rightarrow \quad \begin{cases} \text{(I)} & (I - G' D_\mu^{-1} \cdot \text{II}) \quad [D_\lambda - G' D_\mu^{-1} (G')^T] \delta\lambda = \mathbf{A}_\lambda - G' D_\mu^{-1} \mathbf{A}_\mu \\ \text{(II)} & (G')^T \cdot \delta\lambda + D_\mu \cdot \delta\mu = +\mathbf{A}_\mu \end{cases}$$

L'equazione in $\delta\lambda$,

$$- [D_\lambda - G' D_\mu^{-1} (G')^T] \delta\lambda = G' D_\mu^{-1} \mathbf{A}_\mu - \mathbf{A}_\lambda$$

è del tipo $Mx = b$ con M matrice definita positiva e simmetrica. Infatti, $S := [D_\lambda - G' D_\mu^{-1} (G')^T]$ è il complemento di Schur del blocco D_μ in $H_{F,\beta}$, ed è definito positivo in conseguenza del fatto che $J_{\mathbf{A}}(\mathbf{X})$ lo sia (vedere, ad esempio, queste note [Gallier 2010]). Si verifica facilmente che S è anche simmetrico:

$$S^T = [D_\lambda - G' D_\mu^{-1} (G')^T]^T = D_\lambda - ((G')^T)^T (D_\mu^{-1})^T (G')^T = D_\lambda - G' D_\mu^{-1} (G')^T = S$$

Il metodo standard di risoluzione numerica di un'equazione di questo tipo ($Mx = b$ con M matrice definita positiva e simmetrica) è il seguente:

1. si trova la sua decomposizione di Cholesky: $M = U^T U$ con U matrice triangolare superiore
2. si risolve di $U^T y = b$
3. si risolve di $Ux = y$

Se si desidera sfruttare direttamente i moduli della libreria LAPACK-BLAS, è essenziale tenere presente questi passaggi. Tuttavia, un'opzione più rapida consiste nell'utilizzare librerie come SciPy (per Python) [Virtanen et al. 2020] o Armadillo (C++) [Sanderson, Curtin and Curtin 2016]. Queste librerie sono costruite sopra i moduli LAPACK e includono risolutori universali, che selezionano automaticamente la decomposizione da eseguire in base al tipo di matrice.

Noto $\delta\lambda$, si calcola $\delta\mu$ dalla (II), che invece è del tipo $Dx = y$ con D matrice diagonale, quindi non presenta particolari difficoltà di calcolo numerico. A questo punto avremo una nuova stima $\mathbf{X}_1 = \mathbf{X}_0 + \delta\mathbf{X}$ per la soluzione, e potremo iterare fino ad ottenere la precisione desiderata su \mathbf{X} .

ALGORITMO 2: FREEENERGY A β FISSATO

Input: vettori \mathbf{a} e \mathbf{b} , matrice di costo C , parametro β , parametro di convergenza δ

Inizializzazione: inizializzazione dei vettori $\boldsymbol{\lambda}_0$, $\boldsymbol{\mu}_0$ e calcolo del campo vettoriale $\mathbf{A}(\boldsymbol{\lambda}_0, \boldsymbol{\mu}_0)$

Loop:

1. calcolo delle variazioni $\delta\lambda^n$, $\delta\mu^n$
2. aggiornamenti $\boldsymbol{\lambda}^{n+} = \delta\lambda^n$, $\boldsymbol{\mu}^{n+} = \delta\mu^n$
3. ricalcolo del campo vettoriale \mathbf{A}^n
4. controllo della convergenza: **if** $\|\mathbf{A}^n\| \leq \delta$ **stop**

Output: piano di trasporto $G_{i,j}^{OPT} = \Phi(\beta[C_{k,l} + \lambda_k + \mu_l])$ e costo di trasporto $U_{OPT} = \langle G_{OPT}^\beta, C \rangle$

1.4 Test degli algoritmi sviluppati: trasporto ottimale tra gaussiane 1D

Per controllare il corretto funzionamento degli algoritmi descritti e delle relative implementazioni, li abbiamo utilizzati per computare il piano di trasporto ottimale tra distribuzioni di probabilità gaussiane (discretizzate). Il motivo di questa scelta è che il trasporto ottimale tra gaussiane è analiticamente risolvibile, quindi abbiamo modo di confrontare la soluzione numerica con quella esatta.

Proposizione 1.5 [Santambrogio 2015, capitolo 2], [Peyré e Cuturi 2019, capitolo 2.6] :

Siano $f, g : \mathbb{R} \rightarrow \mathbb{R}$ due distribuzioni di probabilità gaussiane 1D, di media e deviazione standard m_1, σ_1 e m_2, σ_2 rispettivamente. Sia $c(x, y) = |x - y|^2$ il costo. Allora, la soluzione al problema di Kantorovich esiste ed è unica. Il piano di trasporto ottimale è $(x, f(x)) \rightarrow (y = T(x), g(x))$ con $T(x) := m_2 + \left(\frac{\sigma_2}{\sigma_1}\right) \cdot (x - m_1)$. Il costo ottimale corrispondente a tale soluzione è dato da:

$$U_{opt} = [m_1 - m_2]^2 + [\sigma_1 - \sigma_2]^2$$

Questa proposizione ci dice che esiste un piano ottimale piuttosto semplice da realizzare: è sufficiente spostare la massa di probabilità $f(x)$ alla locazione x "spalmandola" all'intorno della locazione $y = T(x)$, che è una retta di intercetta e pendenza determinate dai parametri delle gaussiane. Il costo di tale trasporto, invece, è semplicemente la norma euclidea quadrata tra i punti (m_1, σ_1) e (m_2, σ_2) in un piano 2D.

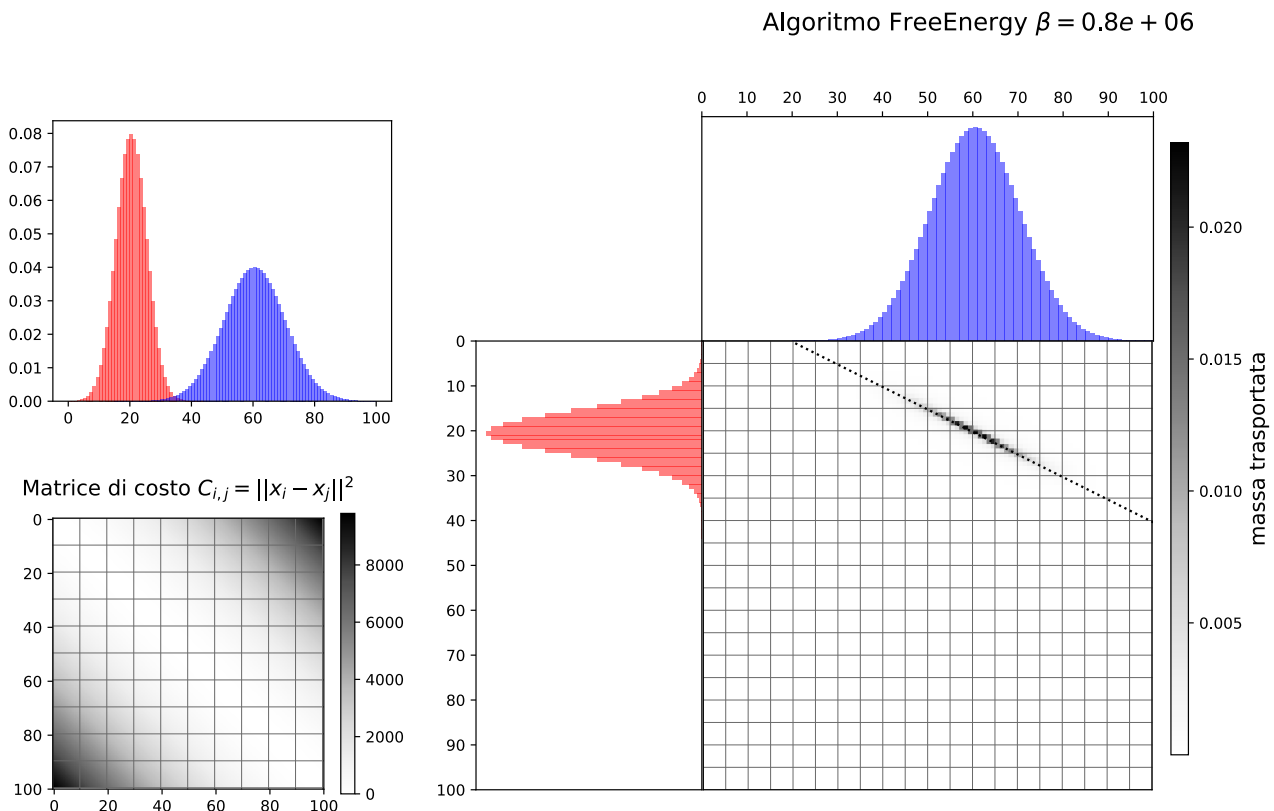


Figura 1.3: A sinistra: distribuzioni gaussiane tra cui calcoliamo il piano di trasporto. La gaussiana rossa, di parametri $(m_1 = 20, \sigma_1 = 5)$, è la distribuzione di partenza, mentre la gaussiana blu, di parametri $(m_2 = 60, \sigma_2 = 10)$, è quella di arrivo. A destra: piano di trasporto ottenuto per il valore massimo di β raggiunto dall'algoritmo FREEENERGY. Come previsto, il piano di trasporto viene indotto dalla mappa lineare $T(x) = m_2 + \left(\frac{\sigma_2}{\sigma_1}\right) \cdot (x - m_1) = 60 + 2 \cdot (x - 20)$. Con x si intende l'asse della gaussiana rossa.

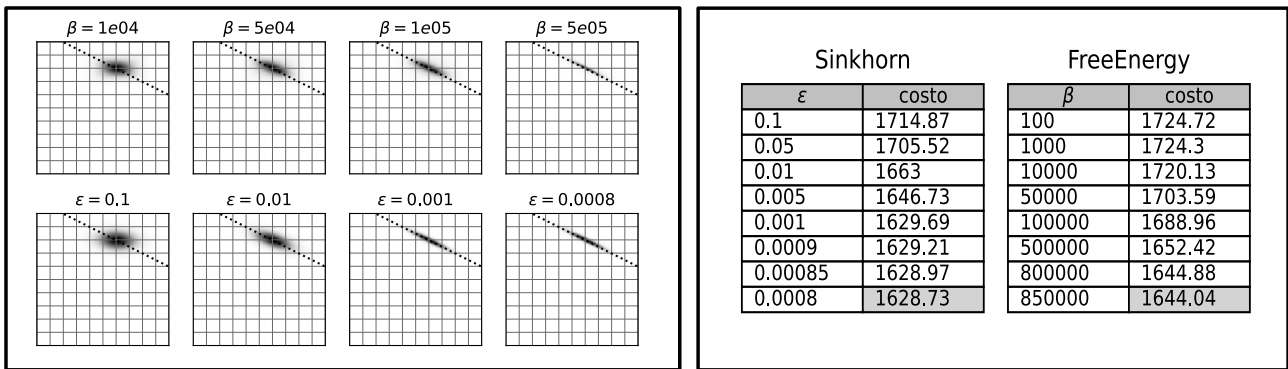


Figura 1.4: A sinistra: il piano di trasporto al variare della regolarizzazione: aumentando β (\equiv diminuendo ϵ) diventa meno sparsificato. L'andamento è quello previsto secondo la [Proposizione 1.2]: per $\epsilon \gg 0$, la soluzione approssima la gaussiana $2D$ data da \mathbf{ab}^T . A destra: costi ottimali ottenuti con i due algoritmi implementati. Il costo ottimale esatto è $U_{opt} = 1625$. L'approssimazione ottenuta con SINKHORN ha un errore relativo dello 0.2%, quella ottenuta con FREEENERGY dell' 1.1%. Nella nostra implementazione, il SINKHORN notevolmente più veloce del FREEENERGY. Per FREE ENERGY, il numero massimo di iterazioni è fissato a 1000. L'algoritmo si ferma quando raggiunge il criterio di convergenza, oppure quando raggiunge il numero massimo di iterazioni, oppure quando la matrice chiamata M al paragrafo precedente non risulta più simmetrica a causa dell'insorgenza di instabilità numeriche.

2

Un'applicazione alle neuroscienze: quantificare il costo cognitivo

Come abbiamo detto, una delle caratteristiche più utili del trasporto ottimale ai fini delle sue applicazioni è il fatto che permette di definire una *metrica* sullo spazio delle distribuzioni di probabilità, e quindi fornisce uno strumento per quantificare la "somiglianza" tra due distribuzioni in modo matematicamente rigoroso. In questo capitolo proponiamo un'applicazione di questo strumento nel campo delle neuroscienze.

È ormai noto in letteratura che il cervello umano, sia a riposo, sia durante lo svolgimento di attività, è organizzato come una rete, in cui zone macroscopiche spazialmente dislocate ma funzionalmente correlate si attivano in sincronia. Le tecniche di EEG (ElectroEncephaloGram) e fMRI (functional Magnetic Resonance Imaging) evidenziano l'esistenza di *pattern* globali ricorrenti di attività neurale, chiamati *microstati* [Tognoli e Scott Kelso 2014]. L'attività celebrale su larga scala si può descrivere come una successione temporale di microstati. Si è osservato che la dinamica di tali microstati, descritta attraverso la *vita media* e la *frequenza di occorrenza* di ciascuno, è modulata dall'attività cognitiva svolta [JM e RA. 2018]. Diversi studi che si sono concentrati sullo studio della dinamica dei microstati durante l'attività spontanea del cervello a riposo (*resting state*) hanno anche mostrato che tale dinamica risulta alterata in soggetti con malattie neuropsichiatriche, come la schizofrenia [Christoph M. Michel 2018].

Detta \mathbf{a} la distribuzione di probabilità dei microstati nel soggetto a riposo e \mathbf{b} la distribuzione durante lo svolgimento della task T, associamo a T il costo $U_{OPT}(\mathbf{a}, \mathbf{b})$ del piano ottimale di trasporto da \mathbf{a} a \mathbf{b} , previa opportuna definizione di una matrice di costo metrica C.

Intendiamo mettere in relazione il costo ottimale $U_{OPT}(\mathbf{a}, \mathbf{b})$ con lo sforzo cognitivo percepito dal soggetto durante l'esecuzione dell'attività. La nostra ipotesi è che lo sforzo cognitivo e il costo ottimale siano positivamente correlati. Tale ipotesi è giustificata dal fatto che il costo ottimale è una misura metrica del grado di somiglianza tra le due distribuzioni \mathbf{a} e \mathbf{b} , quindi un costo maggiore indica una più marcata riconfigurazione globale dell'attività celebrale rispetto allo stato di riposo (che supponiamo essere quello a dispendio cognitivo minimo).

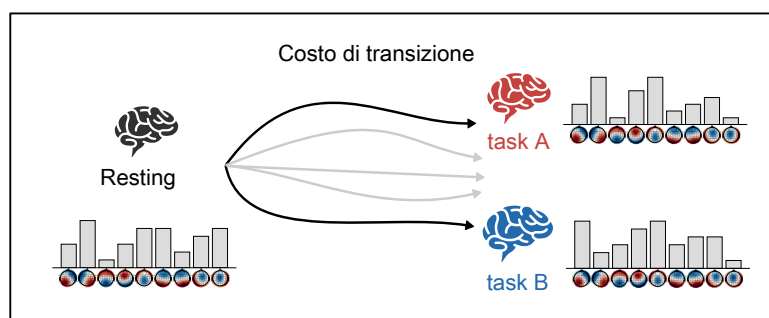


Figura 2.1: Schema qualitativo del metodo proposto.

In sintesi, se traduciamo il problema nel linguaggio del trasporto ottimale che abbiamo introdotto: detti $\{s_1, \dots, s_N\}$ i soggetti sperimentali e $\{T_1, T_2, \dots\}$ le task,

- i microstati $\{x_1, \dots, x_K\}$ sono punti che vivono in uno spazio d-dimensionale \mathbb{R}^d . Per le misure di elettroencefalogramma, d è il numero di canali di presa dati, ovvero il numero di elettrodi posizionati sullo scalpo. Per le misure fMRI, invece, d è il numero di regioni cerebrali macroscopiche (*parcels*) per cui si registra il segnale *BOLD*. In entrambi i casi, ciò che ora e nel seguito verrà chiamato *microstato* è il centroide di un cluster. Maggiori dettagli riguardo la definizione dei microstati vengono dati nel prossimo paragrafo.
- le distribuzioni di probabilità discrete $\mathbf{a}(x_i | s, T_\alpha)$, $\mathbf{b}(x_i | s, T_\beta)$ sono le *frequenze di occorrenza* dei microstati, specifiche del compito svolto e del soggetto s ;
- Il costo della transizione dalla task T_α alla task T_β , per il soggetto s , è identificato con il costo del piano di trasporto ottimale U_{opt} tra le distribuzioni \mathbf{a} e \mathbf{b}
- la scelta della matrice di costo C è a priori arbitraria, in mancanza di un modello teorico. Procedendo per via euristica, due opzioni ci sembrano ragionevoli. La prima è definire il costo tra due microstati come la p-norma euclidea del segmento d-dimensionale che li congiunge

$$C(x_i, x_j) := \|x_i - x_j\|^p$$

L'ipotesi sottostante è che punti vicini nello spazio d-dimensionale corrispondano a stati di attività celebrale simili dal punto di vista funzionale. Una matrice di costo così definita induce una metrica sullo spazio delle distribuzioni. In particolare, il costo ottimale che ne deriva è simmetrico: $U_{OPT}(\mathbf{a}, \mathbf{b}) \equiv U_{OPT}(\mathbf{b}, \mathbf{a})$.

La seconda opzione è definire la matrice di costo a partire dalle *probabilità di transizione* tra microstati nel resting state, dove la dinamica celebrale è spontanea in quanto non finalizzata all'esecuzione di una task:

$$C(x_i, x_j) := \text{funzione decrescente di } (Q_{i,j}) \quad \text{con} \quad Q_{i,j} = \text{probabilità}(x_{t+1} = x_j | x_t = x_i)$$

L'idea sottostante questa scelta è che la dinamica spontanea del cervello nella condizione di riposo sia quella che minimizza il dispendio energetico. Questa matrice di costo, a differenza della precedente, non è necessariamente simmetrica. Questo fatto ci obbliga a rinunciare alle proprietà metriche del costo ottimale, ma potrebbe essere interessante proprio per questo (vedi [Sezione 2.4]).

EEG e fMRI: tecnica di misura, pre-processing e definizione dei microstati

EEG (ELECTROENCEPHALOGRAM): consiste nella registrazione del potenziale elettrico associato all'attività neuronale spontanea attraverso un set di N elettrodi, posizionati a coprire interamente lo scalpo. Il punto di forza dell'EEG sta nel fatto di essere una misura *diretta* dell'attività elettrica dei neuroni, ad alta risoluzione temporale (range di frequenze $0.1 - 100 \text{ Hz}$). Il suo svantaggio principale, però, è la scarsa risoluzione spaziale delle misure, dovuta al fatto che gli elettrodi sono posizionati sulla superficie dello scalpo e la mappa del segnale è una proiezione bidimensionale dell'attività tridimensionale del cervello.

I dati grezzi di un EEG sono gli andamenti temporali del voltaggio V misurato per ciascuno degli N elettrodi (anche detti canali). All'istante temporale t fissato, quindi, il segnale è rappresentato da un vettore N dimensionale $V_1(t), \dots, V_N(t)$. Alternativamente, il segnale si può visualizzare attraverso una mappa 2D in cui i punti vengono colorati in base al voltaggio (*scalp topography map*). Ciascuna mappa rimane stabile mediamente per $100 - 200 \text{ ms}$ e poi cambia quasi istantaneamente nella successiva.

Il primo passo del pre-processing consiste nel selezionare gli istanti temporali in cui è più alto il rapporto segnale-rumore *GFP* (Global Field Power): si ottiene una successione temporale discreta di mappe topografiche. Successivamente, si usa un algoritmo di clustering per raggruppare mappe topografiche simili in k cluster: i *microstati*. I cluster necessari a descrivere la gran parte delle mappe sono solitamente pochi, tipicamente nel range $2 - 15$. Per l'attività spontanea del cervello a riposo,

molti esperimenti concordano nell'identificazione di 5 cluster, che vengono denominati con le lettere A, B, C, D, E [Michael et al, 2018].

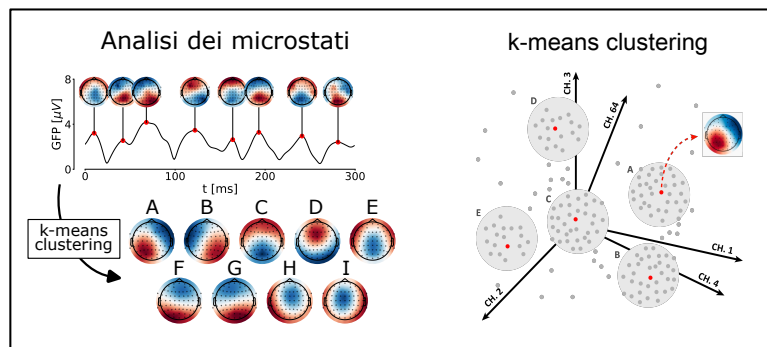


Figura 2.2: Schema del preprocessing di dati EEG. A sinistra: un'ipotetica time-series, da cui vengono estratte le mappe in corrispondenza dei picchi locali del rapporto segnale-rumore, quantificato dal GFP. A destra: ogni mappa è rappresentabile da un punto in uno spazio N -dimensionale. Sulla sequenza discreta estratta dalla time series viene effettuata la clusterizzazione.

fMRI (FUNCTIONAL MAGNETIC RESONANCE IMAGING): Il segnale misurato (detto *BOLD signal*, "Blood Oxygenation Level Dependent") è principalmente sensibile alla concentrazione locale di deossiemoglobina (HbR). Una regione cerebrale manda un segnale BOLD positivo quando la sua concentrazione di HbR diminuisce, cioè quando in essa aumenta l'ossigenazione del sangue. Le regioni attive vengono dunque identificate con quelle che presentano un'aumentata ossigenazione. La spiegazione classica vuole che l'incremento di ossigeno avvenga per fornire energia ai neuroni che si sono attivati. Tuttavia, nonostante la fMRI sia ormai in uso da più di vent'anni, i meccanismi neurofisiologici alla base della risposta emodinamica sono ancora poco compresi [Hillman 2014]. La risoluzione temporale di un segnale BOLD è scarsa se paragonata a quella di un segnale EEG: la risposta emodinamica si attiva solitamente in $\sim 500\text{ ms}$ dall'insorgenza dello stimolo e raggiunge il massimo dopo circa $3 - 5\text{ s}$, anche per stimoli brevi di durata inferiore ad 1 s [Hillman 2014]. La risoluzione spaziale è invece molto elevata: la corteccia viene suddivisa in cubetti (*voxels*) di volume $\simeq 1\text{ mm}^3$.

I dati grezzi di *MRI* hanno quindi alta dimensionalità, ed è indispensabile ridurla per procedere all'analisi: ciò si ottiene raggruppando i voxels in regioni più grandi (*parcels*). Il segnale dei voxels afferenti alla stessa parcella viene mediato: in questo modo si riesce ad aumentare notevolmente il rapporto segnale-rumore. La suddivisione della corteccia cerebrale in parcella (*parcellation*) è un'operazione delicata e la ricerca del metodo migliore non si può dire conclusa (vedere ad esempio [Arslan et al. 2018] per una panoramica). I metodi più recenti usano in combinazione informazioni di tipo anatomico (citoarchitettoniche) sia di tipo funzionale (vengono raggruppati quei voxels che presentano profili temporali di attivazione simili in misure MRI). La parcellazione dei dati usati nel presente lavoro è quella proposta in [Schaefer et al. 2018] in 100 zone corticali +16 zone subcorticali. All'istante t fissato, il segnale BOLD è rappresentato da un vettore $\mathbf{x}(t) = (x_1, \dots, x_N)$ dove N è il numero di parcella e $x_i(t)$ l'ampiezza del segnale nella parcella i -esima. I passi successivi del pre-processing sono gli stessi descritti per i dati EEG: dalla serie temporale $\mathbf{x}(t)$ si estrae una sequenza discreta di mappe $\{\mathbf{x}(t_1), \mathbf{x}(t_2), \dots\}$ cercando i massimi del rapporto segnale-rumore, infine e si raggruppano mappe simili in k cluster.

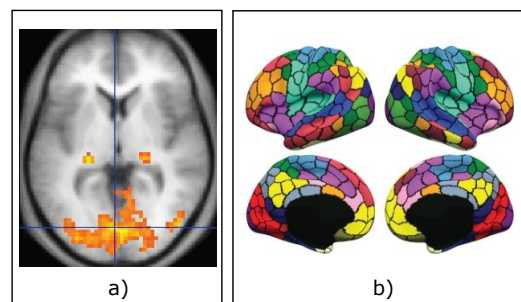


Figura 2.3: a) scan di fMRI, in cui le zone gialle-arancioni indicano le zone attive (figura tratta da https://en.wikipedia.org/wiki/Functional_magnetic_resonance_imaging; b) parcellazione, figura riadattata da [Schaefer et al. 2018].

2.1 Descrizione del dataset EEG

Abbiamo analizzato un set di dati recentemente collezionati in uno studio svolto al dipartimento di Neuroscienze dell'Università degli studi di Padova [Tafuro e Vallesi 2020]. Si tratta di registrazioni EEG multicanale ($N = 64$) eseguite su una coorte di 44 soggetti partecipanti. L'attività celebrale è stata registrata sia a riposo (*resting state*) sia durante lo svolgimento di *Stroop tasks* di tipo spaziale.

La Stroop task è un paradigma sperimentale frequentemente utilizzato per indagare l'abilità di controllo cognitivo e di risoluzione dell'interferenza in presenza di informazioni discordanti. In una Stroop task, il partecipante viene sottoposto ad uno stimolo con due caratteristiche, di cui una è rilevante e l'altra è irrilevante ai fini del compito. Lo stimolo è detto congruente quando le due caratteristiche sono concordi, incongruente quando sono discordi. Al partecipante viene richiesto di compiere un'azione in base alla sola caratteristica rilevante dello stimolo, ignorando quella irrilevante. Solitamente, al partecipante viene richiesto di completare un blocco contenente qualche decina di *tasks*, che vengono presentate in successione. Si trova che i partecipanti sono più accurati e in media più veloci nell'esecuzione della task quando lo stimolo è di tipo congruente. Ciò è in relazione con il maggior grado di controllo cognitivo necessario a sopprimere l'informazione distraente nel caso di stimolo incongruente.

Nella versione usata nel presente esperimento, rappresentata in figura 2.4 lo stimolo è costituito da una freccia, che appare in uno dei quattro angoli di uno schermo (caratteristica irrilevante) e punta in una delle stesse quattro direzioni (caratteristica rilevante). Uno stimolo di tipo congruente, ad esempio, è una freccia che appare in alto a destra e punta nella direzione in alto a destra. Uno incongruente, una freccia che appare in alto a destra ma punta in alto a sinistra. Ai partecipanti viene chiesto di spostare il mouse nella zona dello schermo indicata dalla direzione della freccia, ignorando la posizione in cui la freccia appare.

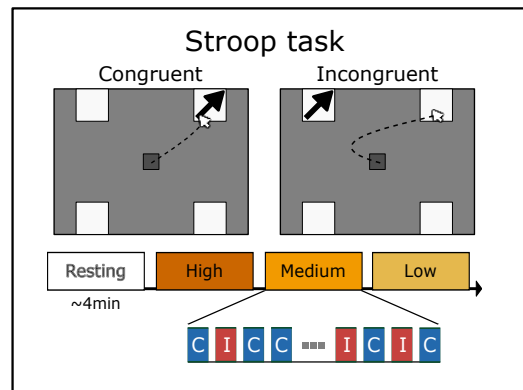


Figura 2.4: La linea tratteggiata rappresenta una tipica traiettoria del mouse: nel caso di stimolo incongruente (destra), l'azione motoria iniziale viene corretta.

Questo dataset si presta bene per testare l'approccio che abbiamo proposto perchè la difficoltà della singola task è ben quantificabile. Essa è modulata da due fattori, indipendenti:

- il tipo di stimolo (congruente C o incongruente I)
- la percentuale di coerenza (PC, *Proporzione di Congruenza*) del blocco in cui la task appare, definita come la percentuale di stimoli coerenti sul totale.

I soggetti dell'esperimento si sono confrontati con tre blocchi di *tasks*, somministrati in successione a difficoltà decrescente: blocco di difficoltà alta (H, *high*, $PC = 25\%$), media (M, *medium*, $PC = 50\%$), bassa (L, *low*, $PC = 75\%$). Dobbiamo quindi distinguere 6 *tasks* diverse. Le denominiamo nel seguito con la sigla (PC/tipo) : HC, MC, LC, HI, MI, LI. Il risultato che auspichiamo di trovare è che il costo ottimale di trasporto sia significativamente influenzato sia dal tipo di stimolo (costo minore per stimolo congruente) che dalla PC (costo minore ad alta percentuale di coerenza).

2.2 Analisi del dataset EEG

Il pre-processing dei dati è stato svolto dal dr. Giacomo Barzon del LPh lab dell'Università degli studi di Padova. La descrizione del metodo seguito può essere trovata in [Barzon et al, 2023]. Sono stati identificati 9 microstati che comprendono quasi l'80% delle mappe topografiche: i 5 microstati ovunque trovati in letteratura nell'analisi dell'attività spontanea a riposo, denominati (A,B,C,D,E), più 4 altri microstati che potrebbero essere specifici delle task svolte.

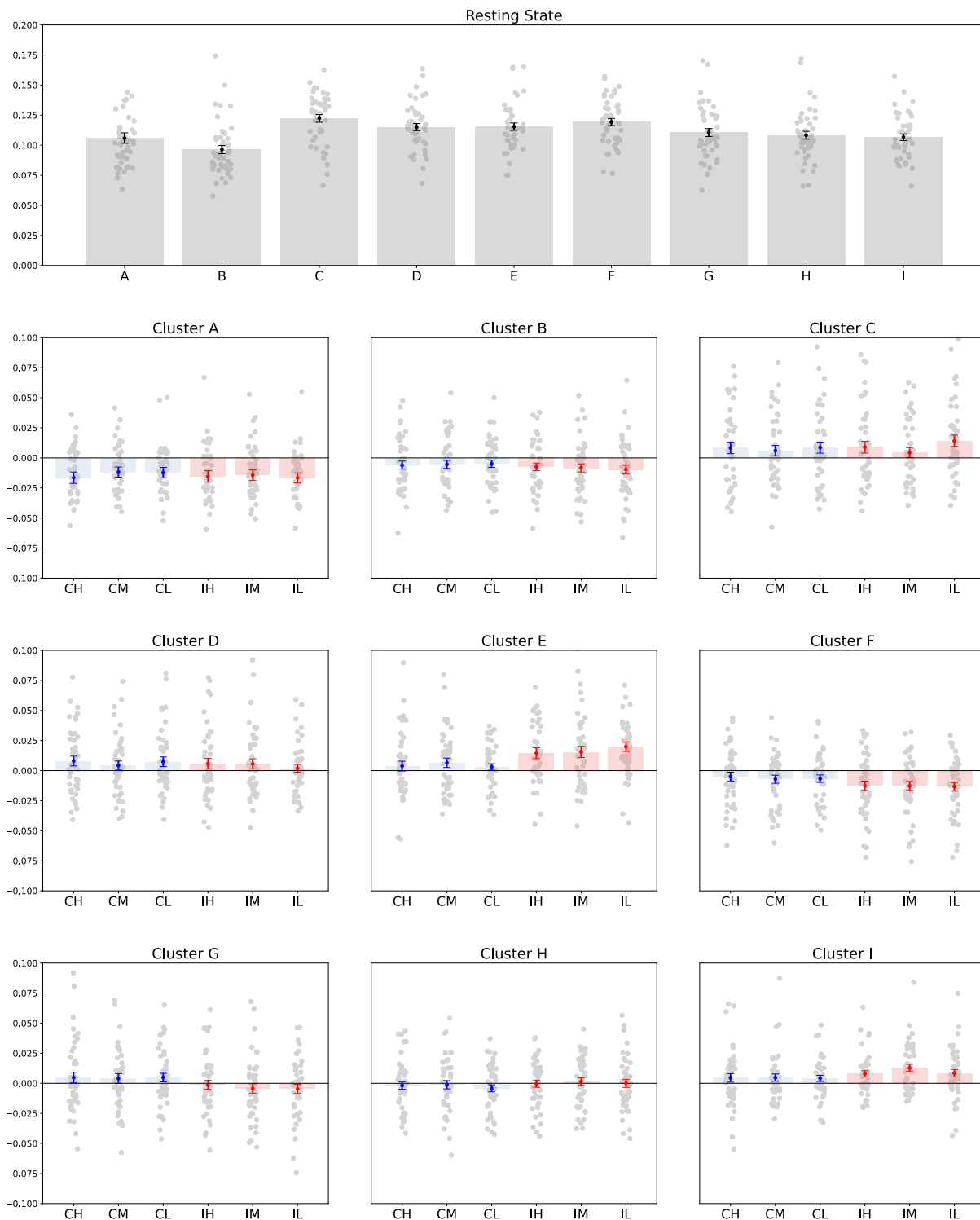


Figura 2.5: Nella figura in alto ('Resting State'): frequenze di occorrenza di ciascun cluster nella condizione di resting $f_{REST}(X)$. Ogni punto grigio corrisponde al dato di un soggetto sperimentale. In nero è riportata la media delle frequenze, con l'errore associato. Nelle altre figure ('Cluster X') è riportata la differenza tra la frequenza di occorrenza del cluster durante le task e la frequenza di occorrenza al resting $\Delta f_i(X) = f_{TASK,i}(X) - f_{REST}(X)$. Ogni punto grigio rappresenta il dato di uno dei 44 soggetti. In blu (per le task di tipo congruente) e in rosso (per le task di tipo incongruente) è riportata la media di $\Delta f_i(X)$ con l'errore associato. I dati mostrano che il microstato A è significativamente soppresso durante l'esecuzione delle tasks, mentre la frequenza dei microstati E ed F è significativamente modulata dal tipo di stimolo. I rimanenti microstati non mostrano modulazioni significative. Si rimanda a Barzon et al. [2023](#) per la discussione dei test statistici eseguiti.

Come già detto, la scelta della matrice di costo è in linea di principio arbitraria. Una scelta che appare naturale è usare un costo basato sulla metrica euclidea, perchè è ragionevole assumere che cambiare un microstato x_i con un'altro x_j sia tanto più costoso quanto più dissimili (\equiv lontani nello spazio \mathbb{R}^N) sono x_i e x_j . Nello specifico, abbiamo scelto come costo la norma euclidea *quadrata* del segmento che congiunge due centroidi: $C_{i,j} := \|x_i - x_j\|^2$. Essendo strettamente convessa, tale funzione di costo garantisce l'esistenza e l'unicità della soluzione al problema di Kanthorovich [Santambrogio 2015, capitolo 2]. In più, essendo una matrice metrica, garantisce le proprietà metriche del costo ottimale che ne deriva [vedi Proposizione 1.1]. Per facilitare la convergenza dell'algoritmo SINKHORN, le entrate di C sono state normalizzate ad 1 (questo è irrilevante, perchè il valore assoluto del costo non ci interessa). Di seguito i passi seguiti:

1. per ciascuno dei $N = 44$ soggetti, separatamente, abbiamo calcolato il costo ottimale di trasporto $U_{OPT}(\mathbf{a}, \mathbf{b})$ tra la distribuzione dei microstati al resting, \mathbf{a} e quella durante ciascuna task. \mathbf{b} (vedi [Fig. 2.6]). Nota: essendo $C_{i,j}$ una matrice metrica, tale costo è simmetrico: $U(\mathbf{a}, \mathbf{b}) = U(\mathbf{b}, \mathbf{a})$;
2. abbiamo calcolato le medie aritmetiche $\langle U_{OPT}(\mathbf{a}, \mathbf{b}) \rangle$ con relativo errore $\sigma_{\langle U_{OPT} \rangle} = \frac{\sigma_{U_{OPT}}}{\sqrt{N}}$, dove $\sigma_{U_{OPT}}$ è la deviazione standard sperimentale e N è il numero di soggetti;
3. abbiamo eseguito il test t di Student per confrontare le medie di ciascuna coppia di tasks, con confidence level assunto del 95%

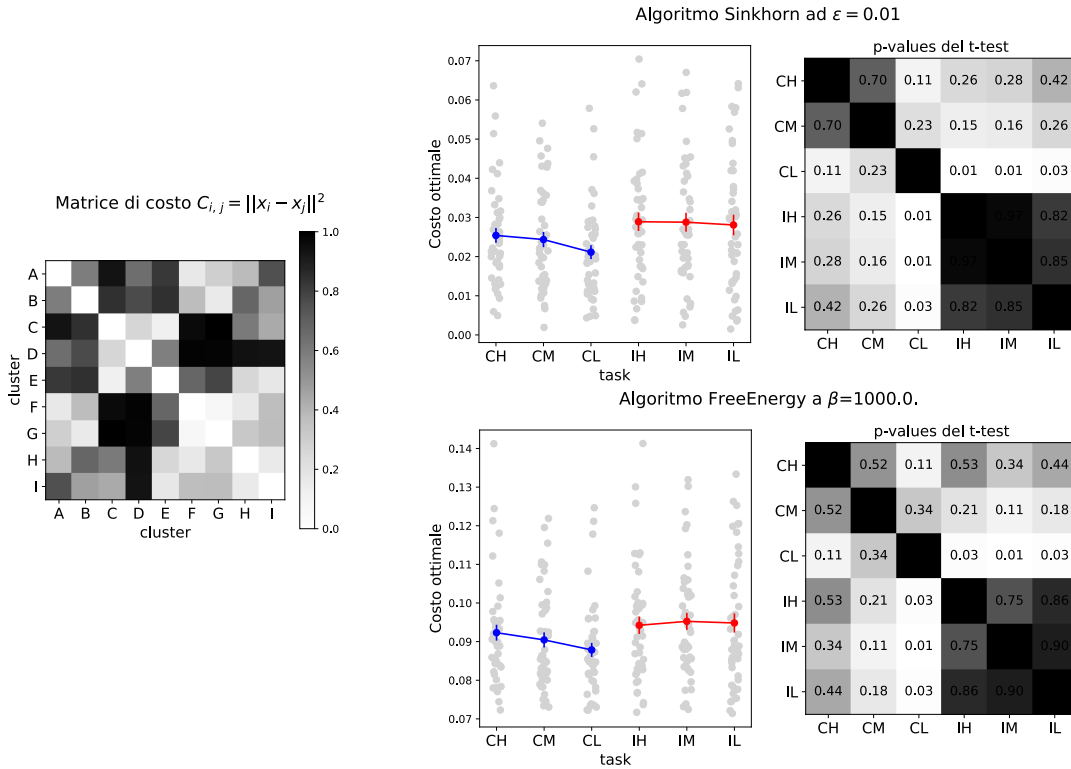


Figura 2.6: A sinistra: matrice di costo C . A destra: risultati ottenuti con l'algoritmo SINKHORN e FREEENERGY. Ogni punto grigio corrisponde al costo ottimale U_{OPT} di un soggetto. I punti rossi rappresentano la media aritmetica sugli N soggetti. Le barre di errore associate sono l'errore sulla media, calcolato nell'ipotesi che i dati seguano una distribuzione gaussiana. A destra: risultati del test di student per il confronto tra le medie, per ciascuna coppia di tasks.

I risultati, riportati in 2.6 mostrano che il costo della task CL (stimolo congruente, blocco di difficoltà bassa) è significativamente più basso del costo di tutte e tre le task con stimolo incongruente (pvalue del t-test ≤ 0.03 , two-sided). Invece, non risulta significativa la differenza di costo tra le task con stimolo dello stesso tipo ma diversa PC. Assumendo un confidence level del 95%, l'ipotesi nulla è rigettata per tutte le coppie (CL, IX). Il test ANOVA-two-way restituisce un pvalue=0.002 per la modulazione da parte del tipo di stimolo (C,I), un pvalue=0.24 per la modulazione da parte dell'intensità del blocco (H, M, L) e un pvalue=0.01 per la modulazione combinata da parte dei due fattori.

2.3 Descrizione del dataset fMRI

I dati utilizzati provengono dal database HCP ("Human Connectome Project") [Van Essen e Smith 2013], pubblicamente disponibile online. Abbiamo selezionato un sottoinsieme del database, contenente le registrazioni fMRI di 44 soggetti a riposo e durante lo svolgimento di 7 *tasks* diverse ((emotion, gambling, language, motor, relational, social, and working memory). A differenza delle Stroop tasks del dataset EEG, queste tasks sono qualitativamente diverse e non è possibile a priori fare ipotesi sulla precisa scala di difficoltà attesa. In aggiunta, purtroppo, il dataset non include un report di autovalutazione dei soggetti sperimentali sulla difficoltà percepita. Possiamo perlomeno attenderci, per senso comune, che una task in cui sono messe alla prova le abilità motorie sia più facile di una task di memoria o di ragionamento (working memory, gambling), etc. Ma l'analisi che facciamo è da intendersi come esplorativa ed è principalmente mirata a verificare se la funzione di costo che abbiamo proposto è effettivamente in grado di distinguere tra le tasks, piuttosto che a stabilire una precisa gerarchia di difficoltà. Dati ulteriori in cui le tasks hanno difficoltà oggettivamente diverse sarebbero richiesti per stabilire la generalità degli eventuali risultati.

2.4 Analisi del dataset fMRI

Il preprocessing dei dati è stato eseguito dal dr. Giacomo Barzon seguendo i metodi descritti in [Satterthwaite et al. 2013]. Il numero di cluster ottimale è stato individuato pari a $k = 8$, in accordo con quanto trovato da Kawakita et al. 2022.

Per il calcolo dei costi, abbiamo utilizzato due matrici di costo:

- la matrice di entrate $C_{i,j} = \|x_i - x_j\|^2$, cioè la distanza euclidea al quadrato tra i centroidi di ciascuna coppia di cluster;
- la matrice di entrate $C_{i,j} = -\log P_{i,j}$, dove $P_{i,j}$ è la probabilità di transizione al resting tra il microstato i e il microstato j (vedi introduzione al capitolo 2). Le probabilità di auto-transizione ($x(t) = x_i \rightarrow x(t+1) \equiv x_i$), $P_{i,i}$ sono notevolmente alte rispetto alle altre. Ciò riflette il fatto che i microstati sono metastabili: nella sequenza discreta $\{x(t_1), x(t_2), \dots\}$ estratta dalla time series ciascun microstato viene mantenuto per un numero significativo di passi prima di cambiare nel successivo. La funzione logaritmo assegna un costo prossimo allo zero alle auto-transizioni e un costo rapidamente crescente alle transizioni poco probabili, con asintoto a $+\infty$ per $P_{i,j} \rightarrow 0$. Ci sembra quindi adeguata a catturare in modo semplice la dinamica, ma altre scelte sono sicuramente possibili.

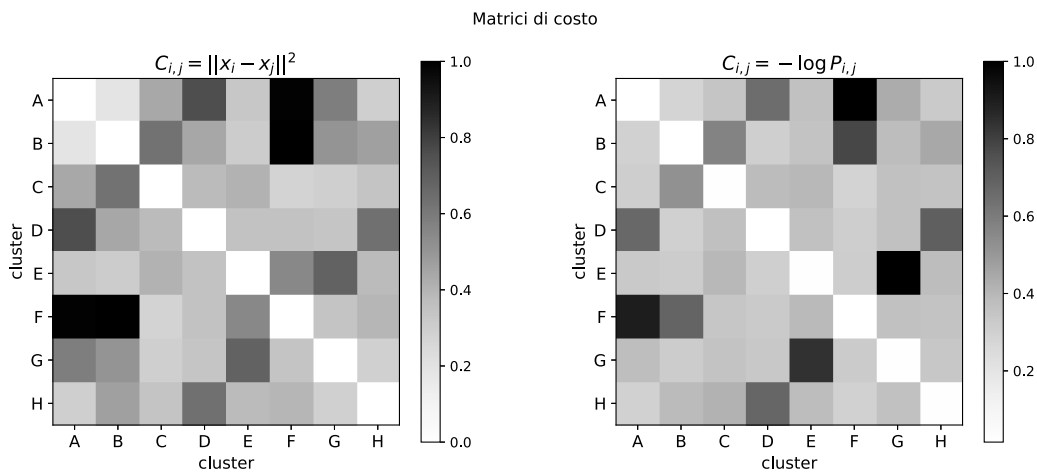


Figura 2.7: Le matrici di costo sono simili. Vuol dire che la traiettoria $\mathbf{x}(t)$ nello spazio 116-dimensionale è continua e non fa "salti" da un cluster all'altro in zone spazialmente non contigue. Anche la matrice basata sulle probabilità di transizione appare abbastanza simmetrica.

I passi seguiti sono gli stessi del dataset EEG : abbiamo calcolato il costo ottimale della transizione resting \rightarrow task per ciascun soggetto e ciascuna task, la media e l'errore sulla media. Poi abbiamo eseguito il test di student su ciascuna coppia di task, con l'ipotesi nulla che i campioni provengano da popolazioni con la stessa media.

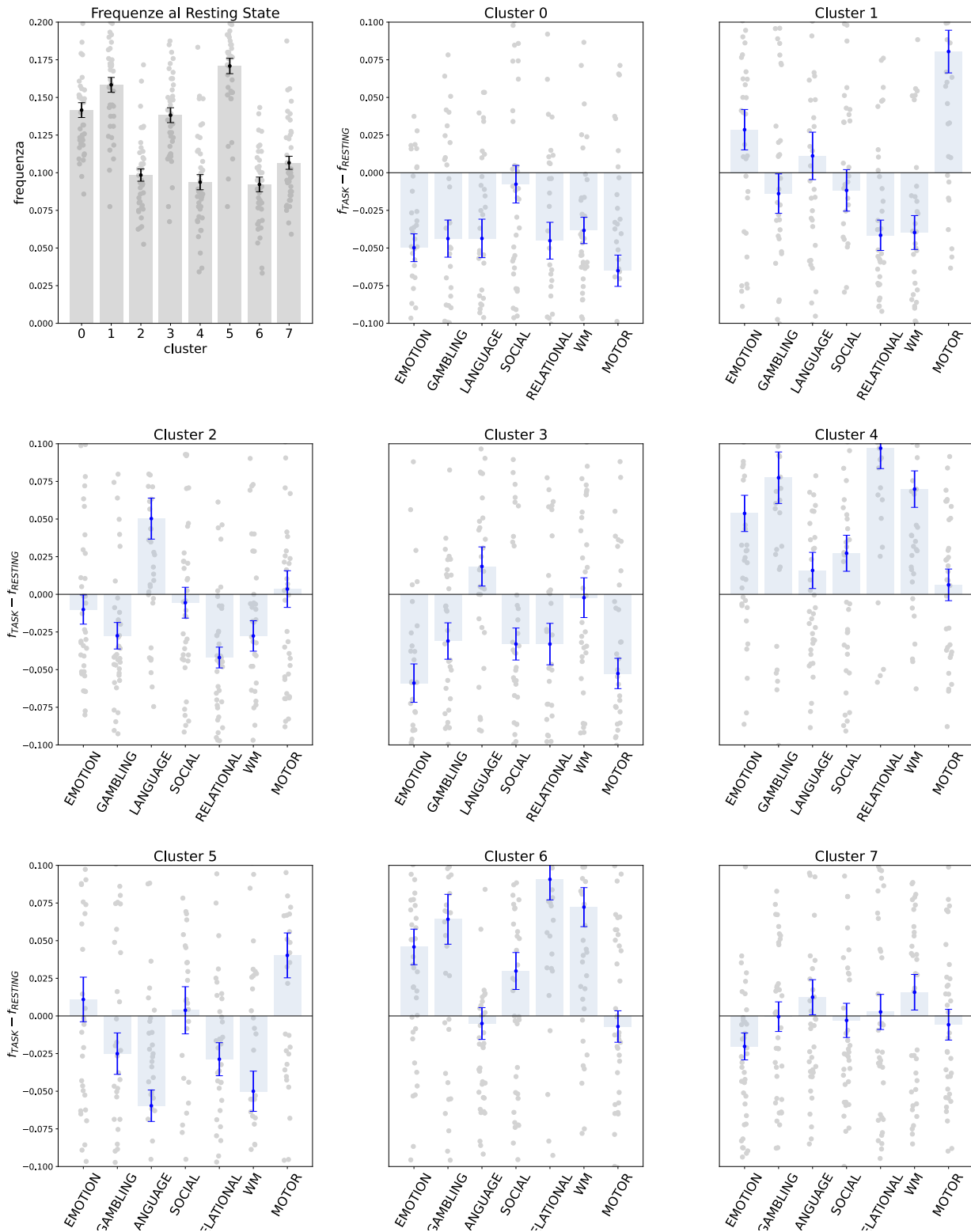


Figura 2.8: Come per i dati EEG (vedi figura [2.5](#)) riportiamo qui la distribuzione delle frequenze dei cluster al resting (figura in alto a sinistra) e la differenza tra le frequenze di occorrenza $f_{TASK} - f_{RESTING}$. Rispetto ai dati EEG, qui si vede più chiaramente la modulazione della frequenza da parte della tasks: ad esempio, il cluster 1 sembra essere specifico della task motoria, il cluster 0 specifico del resting.

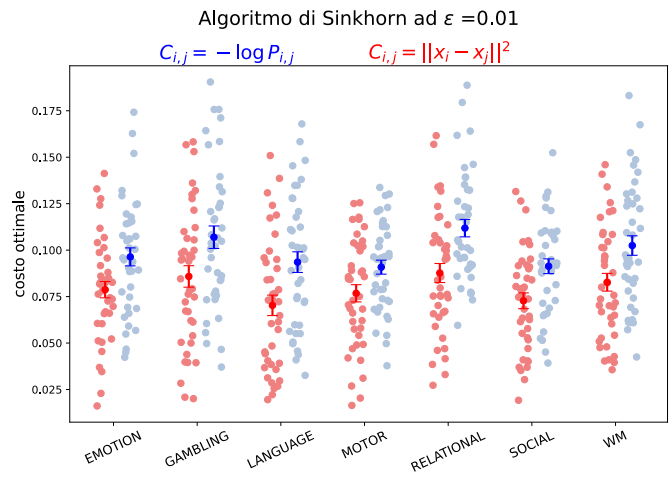


Figura 2.9: Scatterplot dei costi ottimali. In alto: matrice di costo euclidea, in basso: matrice di costo derivata dalle probabilità di transizione. Ciascun punto grigio corrisponde ad un soggetto. I punti in rosso (in alto) e in blu (in basso) rappresentano le medie aritmetiche con errore associato. I risultati riportati sono ottenuti con l'algoritmo di Sinkhorn. L'algoritmo FREEENERGY porta a risultati del tutto analoghi, per cui la figura è omessa.

Task in ordine di costo crescente e risultati del test di Student. Algoritmo FreeEnergy $\beta=10000.0$

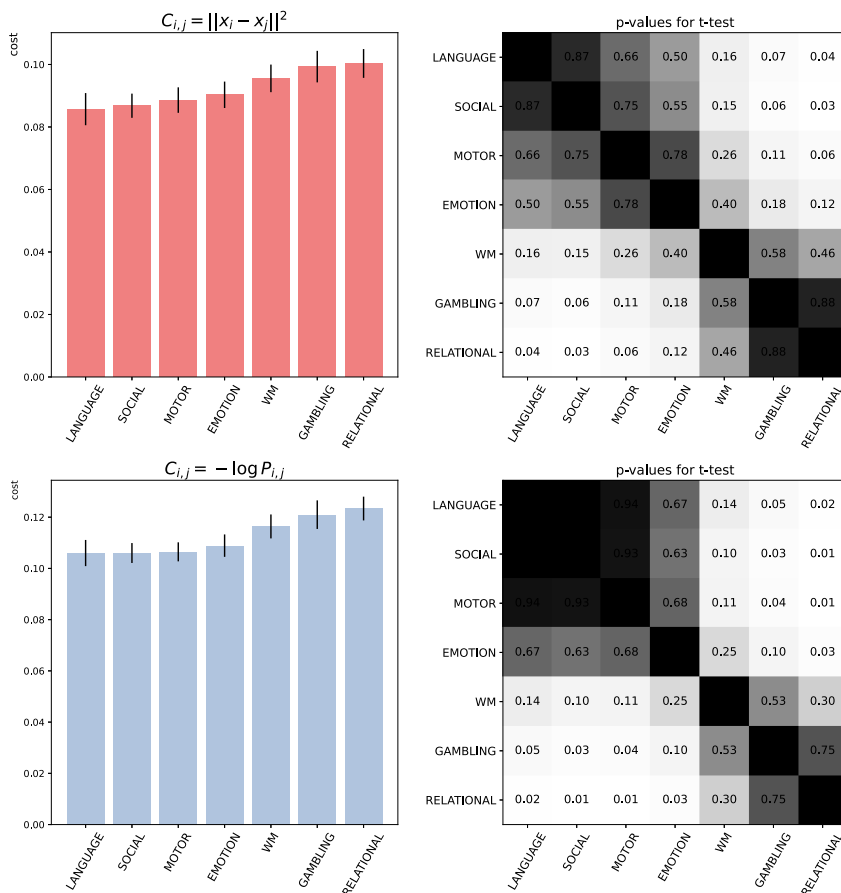


Figura 2.10: Task ordinate per costo crescente: istogramma delle medie con errore della media associato. I costi si distribuiscono approssimativamente su due livelli distinti: quello più basso è occupato dalle task [LANGUAGE, SOCIAL, MOTOR, EMOTION], il più alto da [WORKING MEMORY, GAMBLING, RELATIONAL]. I risultati del test di Student indicano un potere discriminante leggermente superiore per la matrice di costo ottenuta dalle probabilità di transizione.

Asimmetria del costo

I risultati che abbiamo trovato (riportati in figura 2.10) non indicano sostanziali differenze tra la matrice di costo euclidea e quella ottenuta dalle probabilità di transizione. Tuttavia, ci attendiamo di trovare differenze tra le due matrici se ci occupiamo di analizzare i costi delle generiche transizioni task $T_i \rightarrow$ task T_j .

Come già detto nell'introduzione al presente capitolo, dette \mathbf{a} la distribuzione associata alla task i e \mathbf{b} quella associata alla task j , il costo $U_{OPT}(\mathbf{a}, \mathbf{b})$ è simmetrico se C è una matrice metrica, come lo è la matrice delle distanze euclidee al quadrato che abbiamo utilizzato

$$C_{i,j} = \|x_i - x_j\|^2 \text{ (matrice metrica)} \rightarrow U_{OPT}(\mathbf{a}, \mathbf{b}) \equiv U_{OPT}(\mathbf{b}, \mathbf{a})$$

$$\text{costo (task } i \rightarrow \text{task } j) \equiv \text{costo (task } j \rightarrow \text{task } i)$$

Questa relazione non è invece necessariamente verificata se usiamo la matrice di costo derivata dalle probabilità di transizione, che non è metrica. L'asimmetria del costo, in realtà, potrebbe essere una proprietà desiderabile se vogliamo che la funzione di costo proposta misuri effettivamente la *difficoltà cognitiva* di una task. Se la task i -esima è più difficile della j -esima (ovvero, per definizione, se la transizione (resting \rightarrow task i -esima) ha un costo maggiore della transizione (resting \rightarrow task j -esima)), vorremmo anche trovare che il costo della transizione $i \rightarrow j$ è maggiore del costo della transizione inversa $j \rightarrow i$ (fare prima una cosa semplice e poi una cosa difficile richiede più sforzo rispetto al passaggio contrario). Verifichiamo se la matrice di costo $C_{i,j} = -\log P_{i,j}$ è in grado di catturare questa caratteristica. Calcoliamo, per ciascuna coppia di tasks, la quantità

$$\Delta_{i,j} := \frac{U_{OPT}(i,j) - U_{OPT}(j,i)}{U_{OPT}(i,j)}$$

Vogliamo testare la relazione:

$$\Delta_{i,j} > 0 \iff U_{OPT}(\text{resting}, j) > U_{OPT}(\text{resting}, i)$$

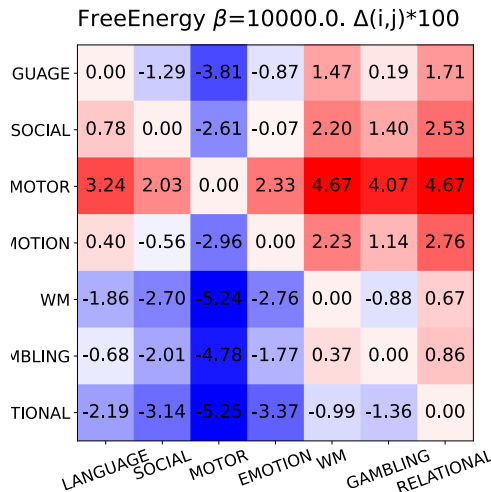


Figura 2.11: Asimmetria dei costi, $\Delta_{i,j}$, in percentuale. I risultati trovati riflettono quelli riportati in [Fig. 2.10]: la relazione risulta verificata per tutte le transizioni che avvengono tra tasks in "blocchi" di difficoltà diversi (blocco 1=[LANGUAGE, SOCIAL, MOTOR, EMOTION], blocco 2=[WORKING MEMORY, GAMBLING, RELATIONAL]), mentre $\Delta_{i,j}$ non ha segno definito per le transizioni tra tasks appartenenti allo stesso blocco.

3

Conclusioni

In questo lavoro di tesi abbiamo innanzitutto presentato alcuni aspetti matematici del trasporto ottimale tra insiemi discreti. Ci siamo soffermati sulla descrizione e l'implementazione di due algoritmi per la risoluzione del problema che sono stati ideati prendendo in prestito idee e tecniche dalla meccanica statistica. Abbiamo verificato che, nella nostra implementazione, i due algoritmi hanno prestazioni comparabili in termini di accuratezza, mentre in termini di velocità il primo si è mostrato più performante.

In futuro si potrebbero fare test più estesi sugli algoritmi, come quelli riportati in [Koehl, Delarue e Orland [2019](#)]. Potrebbe essere valutata l'ottimizzazione dell'algoritmo FreeEnergy attraverso l'utilizzo diretto di moduli LAPACK e l'implementazione di metodi più efficienti per la ricerca delle soluzioni delle equazioni SPA, come alternative al metodo Newton-Raphson. Inoltre, si potrebbe migliorare ulteriormente l'algoritmo Sinkhorn implementando tecniche di stabilizzazione [Peyré e Cuturi [2019](#)] per aumentare la precisione della soluzione approssimata.

Nella seconda parte della tesi, abbiamo applicato il framework teorico del trasporto ottimale al problema della quantificazione del costo cognitivo. Abbiamo usato un approccio interamente basato su dati sperimentali dell'attività celebrale umana per proporre la definizione di misura quantitativa della difficoltà percepita durante l'esecuzione di un'attività cognitiva. Abbiamo analizzato due dataset con caratteristiche sostanzialmente diverse. Il primo è un dataset di misure di EEG per l'attività a riposo e durante l'esecuzione di tasks tra loro molto simili, e contraddistinte da livelli di difficoltà oggettivamente diversi. La nostra analisi è stata mirata a verificare se la funzione di costo proposta fosse in grado di riprodurre l'ordine di difficoltà delle tasks. I risultati confermano almeno parzialmente la nostra ipotesi e sono concordi con quelli già riportati da [Barzon et al. [2023](#)], in cui si usava una diversa funzione di costo. Il secondo dataset comprende misure di resting-state MRI e functional MRI. A differenza del primo caso, le tasks sono qualitativamente diverse e non è possibile stimare a priori quali siano le più difficili. I risultati concordano in buona sostanza con quelli trovati da [Kawakita et al. [2022](#)], che avevano usato un approccio e una funzione di costo diversa dalla nostra. Le differenze riscontrate possono essere imputabili alla ridotta dimensione del campione che abbiamo utilizzato (44 soggetti, rispetto ai 937 soggetti usati dagli autori) e a differenze nel preprocessing dei dati.

Studi futuri sono assolutamente necessari per verificare la robustezza e l'applicabilità dei risultati qui ottenuti. In particolare, sarebbe utile analizzare un altro dataset di misure fMRI che includa tasks con livelli di difficoltà oggettivamente diversi oppure, almeno, la presenza di un report di autovalutazione dei soggetti sulla difficoltà percepita. In caso di risultati positivi, il passo successivo sarebbe verificare se la misura del costo cognitivo definita sia in grado di distinguere soggetti sani da soggetti affetti da malattie neuropsichiatriche, prevedendo un costo maggiore per i soggetti malati.

Trasporto ottimale

- Gallier, Jean (2010). *The Schur complement and symmetric positive semidefinite (and definite) matrices*. Department of Computer and Information Science, University of Pennsylvania.
- Koehl, Patrice, Marc Delarue e Henri Orland (lug. 2019). «Optimal transport at finite temperature». In: *Phys. Rev. E* 100 (1), p. 013310. DOI: [10.1103/PhysRevE.100.013310](https://doi.org/10.1103/PhysRevE.100.013310). URL: <https://link.aps.org/doi/10.1103/PhysRevE.100.013310>.
- Peyré, Gabriel e Marco Cuturi (2019). «Computational Optimal Transport». In: *Foundations and Trends in Machine Learning* 11.5-6, pp. 355–607.
- Sanderson, Conrad, Ryan Curtin e Ryan Curtin (2016). «Armadillo: a template-based C++ library for linear algebra.» In: *Journal of Open Source Software* Vol.1.
- Santambrogio, Filippo (2015). *Optimal Transport for Applied Mathematicians*. Birkhäuser Cham.
- Virtanen, Pauli et al. (2020). «SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python». In: *Nature Methods* 17, pp. 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).

Neuroscienze

- Arslan, S. et al. (2018). «Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex». In: *NeuroImage* 170, pp. 5–30. DOI: [10.1016/j.neuroimage.2017.04.014](https://doi.org/10.1016/j.neuroimage.2017.04.014).
- Barzon, Giacomo et al. (2023). «EEG microstate transition cost correlates with task demands». In: *bioRxiv*. DOI: [10.1101/2023.12.07.570625](https://doi.org/10.1101/2023.12.07.570625).
- Christoph M. Michel, Thomas Koenig (2018). «Eeg microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review». In: *Neuroimage* 180.
- Hillman, EM. (2014). «Coupling mechanism and significance of the BOLD signal: a status report.» In: *Annu Rev Neurosci*. DOI: [10.1146/annurev-neuro-071013-014111](https://doi.org/10.1146/annurev-neuro-071013-014111).
- JM, Shine e Poldrack RA. (2018). «Principles of dynamic network reconfiguration across diverse brain states.» In: *NeuroImage* 180(Pt B), pp. 396–405. DOI: [10.1016/j.neuroimage.2017.08.010](https://doi.org/10.1016/j.neuroimage.2017.08.010).
- Kawakita, Genji et al. (feb. 2022). «Quantifying brain state transition cost via Schrödinger Bridge». In: *Network Neuroscience* 6.1, pp. 118–134. ISSN: 2472-1751. DOI: [10.1162/netn_a_00213](https://doi.org/10.1162/netn_a_00213).
- Satterthwaite, Theodore D. et al. (2013). «An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data». In: *NeuroImage* 64, pp. 240–256. DOI: [10.1016/j.neuroimage.2012.08.052](https://doi.org/10.1016/j.neuroimage.2012.08.052). URL: www.sciencedirect.com/science/article/pii/S1053811912008609.
- Schaefer, A. et al. (2018). «Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI». In: *Cerebral Cortex* 28 (9), pp. 3095–3114. DOI: [10.1093/cercor/bhx179](https://doi.org/10.1093/cercor/bhx179).
- Tafuro, A. e A. Vallesi (2020). «Cognitive brakes in interference resolution: A mouse-tracking and EEG co-registration study.» In: *Cortex*. DOI: [10.1016/j.cortex.2020.09.024](https://doi.org/10.1016/j.cortex.2020.09.024).
- Tognoli, Emmanuelle e J.A. Scott Kelso (2014). «The Metastable Brain». In: *Neuron* 81 (1), pp. 35–48. DOI: [10.1016/j.neuron.2013.12.022](https://doi.org/10.1016/j.neuron.2013.12.022).
- Van Essen, D. C. e S. M. Smith (2013). «The WU-Minn Human Connectome Project: An Overview.» In: *NeuroImage*. DOI: [10.1016/j.neuroimage.2013.05.041](https://doi.org/10.1016/j.neuroimage.2013.05.041).

Dati e codici

I codici e i dati preprocessati utilizzati sono disponibili alla repository GitHub https://github.com/miriamzara/Optimal_Transport