MPALN: Metodi Probabilistici per l'Algebra Lineare Numerica

A.A. 2025–2026 Primo semestre

Insegnanti:	Federico Poloni (federico.poloni@unipi.it)
	Alice Cortinovis (alice.cortinovis@unipi.it)
Orario:	Martedì & Mercoledì 9:00 – 11:00 (aula O1)

Obiettivi del corso

Acquisire conoscenze sui principali algoritmi probabilistici per risolvere problemi di algebra lineare numerica e sugli strumenti utili per l'analisi di tali metodi.

Prerequisiti: Calcolo Scientifico, Elementi di Probabilità e Statistica

Organizzazione del corso: Lezioni teoriche e sperimentazione numerica in Matlab.

Modalità di esame: L'esame può essere svolto in due modalità, a scelta dello studente: un seminario su un articolo legato ai contenuti del corso, oppure un esame orale tradizionale.

Contenuti del corso

I principali argomenti trattati saranno:

- metodi probabilistici per il calcolo di autovalori di matrici;
- algoritmi per l'approssimazione della traccia di (funzioni di) matrici, con tecniche di riduzione della varianza;
- disuguaglianze di Chernoff e Bernstein per matrici;
- tecniche di embedding per sottospazi e loro applicazione a problemi ai minimi quadrati sovradeterminati;
- algoritmi probabilistici per trovare approssimazioni di rango basso di matrici;
- interpretazione probabilistica di algoritmi su alberi binari.

Il corso evidenzierà alcuni vantaggi degli algoritmi probabilistici rispetto alle loro controparti deterministiche: per esempio, possono essere usati per approssimare delle quantità che sono troppo costose da calcolare esattamente quando le matrici sono di dimensioni molto grandi, oppure possono essere usati per ridurre la dimensione di alcuni problemi senza (quasi) perdere informazioni. Riportiamo qui, brevemente, due algoritmi che studieremo nel corso: la SVD randomizzata e lo stimatore di Hutchinson per la traccia di matrici.

Esempio 1: approssimazioni di rango basso di una matrice

Data una matrice $A \in \mathbb{R}^{n\times n}$, la migliore approssimazione di rango r rispetto alla norma di Frobenius si ottiene, per il teorema di Eckart-Young, troncando la decomposizione ai valori singolari (SVD) di A: per calcolarla, però, servono $\mathcal{O}(n^3)$ operazioni, un costo proibitivo per matrici di grandi dimensioni. Consideriamo invece il procedimento seguente, chiamato *SVD randomizzata*:

1. sia $\Omega \in \mathbb{R}^{n\times r}$ una matrice con entrate i.i.d prese dalla distribuzione Gaussiana standard;

2. calcoliamo $Y := A\Omega \in \mathbb{R}^{n\times r}$;
3. calcoliamo (con fattorizzazione QR) una base ortonormale Q dello span delle colonne di Y;
4. consideriamo l'approssimazione $A \approx Q(Q^T A)$ (mantenendola in forma fattorizzata).

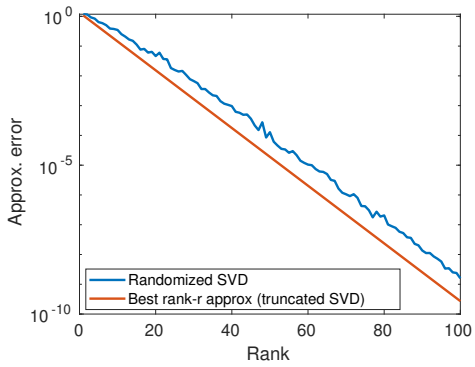
Il costo di questo algoritmo è al più $\mathcal{O}(n^2 r)$, ben più economico della SVD, e l'errore dell'approssimazione ottenuta è (tipicamente) di poco più grande di quello ottenuto dalla SVD classica (Figura (a)).

Esempio 2: stima della traccia di una matrice

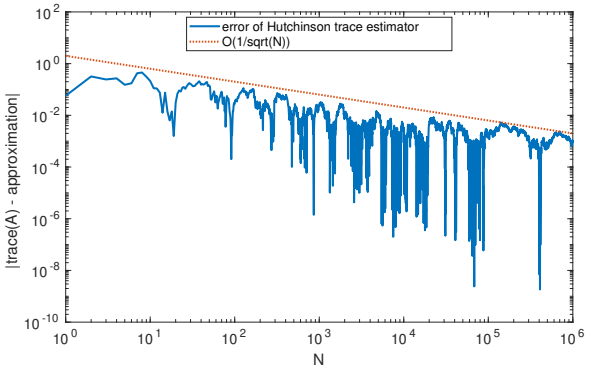
Assumiamo di voler calcolare la traccia di una matrice simmetrica $A \in \mathbb{R}^{n\times n}$, tale per cui l'unica operazione che siamo autorizzati a fare è moltiplicare A per uno o più vettori a nostra scelta. (Nel corso vedremo perché questa assunzione apparentemente stramba ha senso!)
Se X è un vettore Gaussiano standard di lunghezza n, si ha che $\mathbb{E}[X^T AX] = \text{traccia}(A)$ (provate a dimostrarlo!). Questo suggerisce che si possa approssimare la traccia di A prendendo N vettori $X^{(1)}, \dots, X^{(N)}$ Gaussiani standard, e calcolando

$$\text{traccia}(A) \approx \frac{1}{N}\sum_{i=1}^{N} (X^{(i)})^T A X^{(i)}. \tag{1}$$

Questo oggetto si chiama *stimatore di Hutchinson* e nella Figura (b) potete vedere il comportamento dell'errore di questo stimatore per una realizzazione di (1) con $N = 1, 2, \dots, 1000000$.



(a) Grafico raffigurante l'errore delle approssimazioni di rango r (compreso tra 1 e 100) di una matrice A ottenute con la SVD "standard" e la SVD randomizzata (più veloce). Nel corso vedremo perché l'algoritmo probabilistico funziona quasi bene quanto quello deterministico, in quali casi è conveniente usarlo, e delle varianti per matrici simmetriche definite positive.



(b) Grafico raffigurante l'errore dello stimatore di Hutchinson quando il numero di vettori (N) aumenta. La scala è semilogaritmica su entrambi gli assi. Nel corso studieremo il comportamento dello stimatore, daremo stime sul numero di vettori che dobbiamo prendere per avere un errore ε con almeno il 99% di probabilità, e discuteremo alcune tecniche per ridurre la varianza dello stimatore.

Letture consigliate

- Randomized algorithms for matrix computations, J. Tropp https://tropp.caltech.edu/notes/Tro20-Randomized-Algorithms-LN.pdf
- An introduction to matrix concentration inequalities, J. Tropp https://tropp.caltech.edu/books/Tro14-Introduction-Matrix-FnTML-rev.pdf
- High-dimensional probability: an introduction with applications in data science, R. Vershynin https://www.math.uci.edu/~rvershyn/papers/HDP-book/HDP-book.html#
- Randomized numerical linear algebra: a perspective on the field with an eye to software, R. Murray et al. https://arxiv.org/pdf/2302.11474