Metodi Probabilistici per l'Algebra Lineare Numerica

Ivan De Biasi

Indice

1 Fondamenti di probabilità

Definizione [Funzione generatrice dei momenti]: Data X v.a.

$$M_X(\theta) = \mathbb{E}[\exp(\theta X)]$$

Definizione [Funzione generatrice dei cumulanti]: Data X v.a.

$$C_X(\theta) = \log M_X(\theta)$$

Proposizione: Date $X_1, \ldots, X_n$ v.a. indipendenti

$$C_{X_1+\cdots+X_n}(\theta) = C_{X_1}(\theta) + \cdots + C_{X_n}(\theta)$$

Dimostrazione.

... to do ...

□

1.1 Disuguaglianze

Teorema [Jensen]: Data X v.a. e $\varphi : \mathbb{R} \to \mathbb{R}$ convessa

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$$

Dimostrazione. ...

Teorema [Markov]: Data $X \geq 0$ v.a. : $\mathbb{E}[X] < +\infty$

$$\forall \alpha > 0 \quad \mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[X]}{\alpha}$$

Dimostrazione. ...

Teorema [Chebyshev]: Data X v.a. : $\mathbb{E}[X], \mathrm{Var}(X) < +\infty$

$$\forall \alpha > 0 \quad \mathbb{P}(|X - \mathbb{E}[X]| \geq \alpha) \leq \frac{\mathrm{Var}[X]}{\alpha^2}$$

Dimostrazione. ...

Teorema [Chernoff]: Data X v.a.

$$\forall \alpha \quad \forall \theta > 0 \quad \mathbb{P}(X \geq \alpha) \leq \frac{\mathbb{E}[\exp(\theta X)]}{\exp(\theta \alpha)}$$

Dimostrazione. ...

<u>Notazione</u>: $\forall L > 0 \quad g_L(\theta) = \dfrac{\frac{\theta^2}{2}}{1 - \frac{L\theta}{3}}$

Teorema [Bernstein]: Data X v.a. : $\mathbb{E}[X] = 0$

$$X \in [-L, L] \implies \forall \theta \in \left(0, \tfrac{3}{L}\right] \quad C_X(\theta) \leq g_L(\theta)\mathbb{E}[X^2]$$

Dimostrazione. ...

1.2 Proprietà di distribuzioni notevoli

Teorema: Data $X \sim \mathcal{N}(0, I_n)$

$$\forall U \in \mathbb{R}^{n \times n} \text{ ortogonale} \quad UX \sim \mathcal{N}(0, I_n)$$

Dimostrazione. ...

Lemma: Data $\alpha \sim \mathcal{N}(0, 1)$

$$\forall c > 0 \qquad \mathbb{E}\left[\frac{c}{\alpha^2 + c}\right] \leq \sqrt{\frac{\pi c}{2}}$$

Dimostrazione.

... to do ...

□

Proposizione: $X \sim \mathcal{N}(0,1) \implies M_X(\theta) = e^{\frac{\theta^2}{2}}$

Dimostrazione. ...

Proposizione: $X \sim Chi-quadro(1) \implies M_X(\theta) = \dfrac{1}{\sqrt{1-2\theta}}$

Dimostrazione. ...

1.3 Variabili sub-gaussiane e sub-Gamma

Definizione: Data X v.a e $c > 0$

$$X \text{ sub-gaussiana con parametro } c \quad \text{se} \quad \forall \theta \in \mathbb{R} \quad \mathbb{E}[\exp(\theta(X - \mathbb{E}[X])] \leq \exp\left(\frac{c\theta^2}{2}\right)$$

Notazione: $X \sim \mathcal{SG}(c) \iff X$ sub-gaussiana con parametro c

Teorema: Data $X \sim \mathcal{SG}(c)$

$$\forall t > 0 \quad \mathbb{P}(X - \mathbb{E}[X] \geq t) \leq \exp\left(-\frac{t^2}{2c^2}\right)$$

Dimostrazione.
... to do ...

□

Corollario: $\forall t > 0 \quad \mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq 2\exp\left(-\dfrac{t^2}{2c^2}\right)$

Dimostrazione. ...

Teorema [Disuguaglianza di Hoeffding]: Date $X_1, \ldots, X_N$ v.a : $\forall i \quad X_i \sim \mathcal{SG}(c_i)$

$$\forall t > 0 \quad \mathbb{P}\left(\left|\sum_i (X_i - \mathbb{E}[X_i])\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{2(c_1^2 + \cdots + c_N^2)}\right)$$

Dimostrazione.
... to do ...

□

Corollario: $Y = X_1 + \cdots + X_N \sim \mathcal{SG}\left(\sqrt{c_1^2 + \cdots + c_N^2}\right)$

Dimostrazione. ...

Lemma: Data X v.a : $\mathbb{E}[X] = 0$

$$|X| \leq L \implies X \sim \mathcal{SG}(L)$$

Dimostrazione.
... to do ...

□

Definizione: Data X v.a e $\nu, c > 0$

$$X \text{ sub-gamma con parametri } \nu, c \quad \text{se} \quad \forall \theta \in \left(0, \tfrac{1}{c}\right) \quad \mathbb{E}[\exp(\theta(X - \mathbb{E}[X])] \leq \exp\left(\frac{\theta^2 \nu}{2(1 - c\theta)}\right)$$

Notazione: $X \sim \mathcal{S}\Gamma(\nu, c) \iff X$ sub-gamma con parametri ν, c

Lemma: Data X v.a : $\mathbb{E}[X] = 0$

$$X, -X \sim \mathcal{S}\Gamma(\nu, c) \implies \forall t > 0 \quad \mathbb{P}(|X| \geq t) \leq 2\exp\left(-\frac{t^2}{2(\nu + ct)}\right)$$

Dimostrazione.
... to do ...

□

2 Risultati vari

2.1 Metodo delle potenze

Definizione: Data $A \in \mathbb{R}^{n\times n}$ simmetrica semidefinita positiva : $Spec(A) = \{\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n\}$

$$MP_A(y_0) = \left\{ (y_k, \xi_k) \;\middle|\; \begin{array}{l} y_k = Ay_{k-1} \\ \xi_k = \dfrac{y_k^{\mathrm{T}} A y_k}{y_k^{\mathrm{T}} y_k} \end{array} \right\}$$

Proprietà: $\forall y_0 \in \mathbb{R}^n \;:\; y_0 \not\perp u_{\lambda_1} \quad \lim_{k\to+\infty} (y_k, \xi_k) = (u_{\lambda_1}, \lambda_1)$

Teorema: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica semidefinita positiva : $Spec(A) = \{\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n\}$ e $MP_A(y_0) = \{(y_k, \xi_k)\}$ metodo delle potenze

$$y_0 \sim \mathcal{N}(0, I_n) \qquad \mathbb{E}\left[\frac{|\lambda_1 - \xi_k|}{\lambda_1}\right] \leq \sqrt{\frac{(n-1)\pi}{2}} \left(\frac{\lambda_2}{\lambda_1}\right)^k$$

Dimostrazione.
... to do ...

□

3 Stime per la traccia di matrici

3.1 Stimatore di Girard - Hutchinson

Teorema: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X \in \mathbb{R}^n$ v.a.

$$\mathbb{E}[XX^{\mathrm{T}}] = I_n \implies \mathbb{E}[X^{\mathrm{T}} A X] = \mathrm{tr}(A)$$

Dimostrazione.

$$\mathbb{E}[X^{\mathrm{T}} A X] = \mathbb{E}[\mathrm{tr}(X^{\mathrm{T}} A X)] = \mathbb{E}[\mathrm{tr}(AXX^{\mathrm{T}})] = \mathrm{tr}(A\mathbb{E}[XX^{\mathrm{T}}]) = \mathrm{tr}(A)$$

□

Definizione [Stimatore di Girard - Hutchinson]: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X^{(1)}, \ldots, X^{(N)} \in \mathbb{R}^n$ v.a. : $\forall i \quad \mathbb{E}\left[X^{(i)} {X^{(i)}}^{\mathrm{T}}\right] = I_n$

$$\mathrm{tr}_N(A) = \frac{1}{N} \sum_i {X^{(i)}}^{\mathrm{T}} A X^{(i)}$$

Proprietà: $\mathbb{E}[\mathrm{tr}_N(A)] = \mathrm{tr}(A)$

Teorema: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X \sim \mathcal{N}(0, I_n)$

$$\mathrm{Var}\left(X^{\mathrm{T}} A X\right) = 2\|A\|_F^2$$

Dimostrazione.
... to do ...

□

Corollario: $\forall \varepsilon, \delta > 0 \quad N \geq 2\|A\|_F^2 \dfrac{1}{\varepsilon^2} \dfrac{1}{\delta} \implies \mathbb{P}\left(|\,\mathrm{tr}_N(A) - \mathrm{tr}(A)| \geq \varepsilon\right) \leq \delta$

Dimostrazione. ...

Teorema: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X \sim Rademacher(n)$

$$\mathrm{Var}\left(X^{\mathrm{T}} A X\right) = 2\|A - diag(A)\|_F^2$$

Dimostrazione.
... to do ...

□

Corollario: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X^{(1)}, \ldots, X^{(N)} \sim Rademacher(n)$

$$\forall \varepsilon, \delta > 0 \quad N \geq 8n^2 \|A\|_2^2 \frac{1}{\varepsilon^2} \log\left(\frac{2}{\delta}\right) \implies \mathbb{P}\left(|\,\mathrm{tr}_N(A) - \mathrm{tr}(A)| \geq \varepsilon\right) \leq \delta$$

Dimostrazione. ...

Lemma: Date $A \in \mathbb{R}^{n\times n}$ simmetrica, $X \sim \mathcal{N}(0, I_n)$ e $Y = X^{\mathrm{T}} A X - \mathrm{tr}(A)$

$$Y, -Y \sim \mathcal{S}\Gamma\left(2\|A\|_F^2, 2\|A\|_2\right)$$

Dimostrazione.
... to do ...

$\square$

Proposizione: Date $A \in \mathbb{R}^{n\times n}$ simmetrica

$$\forall \varepsilon > 0 \quad \mathbb{P}\left(|\operatorname{tr}_N(A) - \operatorname{tr}(A)| \geq \varepsilon\right) \leq 2\exp\left(-\frac{\varepsilon^2}{4\|A\|_F^2 + 4\varepsilon\|A\|_2}\right)$$

Dimostrazione.
... to do ...

$\square$

Corollario: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X^{(1)}, \dots, X^{(N)} \sim \mathcal{N}(0, I_n)$

$$\forall \varepsilon, \delta > 0 \quad N \geq (4\|A\|_F^2 + 4\varepsilon\|A\|_2)\frac{1}{\varepsilon^2}\log\left(\frac{2}{\delta}\right) \implies \mathbb{P}\left(|\operatorname{tr}_N(A) - \operatorname{tr}(A)| \geq \varepsilon\right) \leq \delta$$

Dimostrazione. ...

Proposizione: Date $A \in \mathbb{R}^{n\times n}$ simmetrica

$$\forall \varepsilon > 0 \quad \mathbb{P}\left(\frac{|\operatorname{tr}_N(A) - \operatorname{tr}(A)|}{\operatorname{tr}(A)} \geq \varepsilon\right) \leq 2\exp\left(-\frac{\varepsilon^2 N}{8}\right)$$

Dimostrazione.
... to do ...

$\square$

Corollario: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X^{(1)}, \dots, X^{(N)} \sim \mathcal{N}(0, I_n)$

$$\forall \varepsilon, \delta > 0 \quad N \geq \frac{8}{\varepsilon^2}\log\left(\frac{2}{\delta}\right) \implies \mathbb{P}\left(\frac{|\operatorname{tr}_N(A) - \operatorname{tr}(A)|}{\operatorname{tr}(A)} \geq \varepsilon\right) \leq \delta$$

Dimostrazione. ...

Teorema [Disuguaglianza di Hanson-Wright]: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X = (X_1, \dots, X_N)$ v.a. : $\forall i \quad X_i \sim \mathcal{SG}(c)$ centrata

$$\forall \varepsilon > 0 \quad \mathbb{P}\left(|X^{\mathrm{T}}AX - \operatorname{tr}(A)| \geq \varepsilon\right) \leq 2\exp\left(-\frac{\varepsilon^2}{80c^4\|A\|_F^2 + 16c^2\varepsilon\|A\|_2}\right)$$

Dimostrazione. ...

Corollario: Dati $A \in \mathbb{R}^{n\times n}$ simmetrica e $X^{(1)}, \dots, X^{(N)} \sim \mathcal{SG}(c)$

$$\forall \varepsilon, \delta > 0 \quad N \geq (80\|A\|_F^2 + 16\varepsilon\|A\|_2)\frac{1}{\varepsilon^2}\log\left(\frac{2}{\delta}\right) \implies \mathbb{P}\left(|\operatorname{tr}_N(A) - \operatorname{tr}(A)| \geq \varepsilon\right) \leq \delta$$

Dimostrazione. ...

3.1.1 Metodi di Krylov per $x^{\mathrm{T}}f(A)x$

- definizione e algoritmo di Lanczos
- definizione di T_m e U_m
- dimostrazione $x^{\mathrm{T}}f(A)x \approx \|x\|^2 f(T_m)_{1,1}$
- rappresentazione integrale di $x^{\mathrm{T}}f(A)x$
- teorema di quadratura gaussiana rispetto alla misura della rappresentazione integrale
- stima errore tramite proprietà della quadratura gaussiana

3.2 Tail bound per matrici

Notazione [Loewner]: Date $A, B \in \mathbb{C}^{n\times n}$ hermitiane

$A \succeq B$ se $A - B$ semidefinita positiva

Proposizione: Date $A, B \in \mathbb{C}^{n\times n}$ hermitiane

$A \preceq B \iff \forall M \in GL(n, \mathbb{C}) \quad MAM^* \preceq MBM^*$

Dimostrazione. ...

Proposizione: Data $A \in \mathbb{C}^{n\times n}$ hermitiana

$\log(I_n + A) \preceq A$

Dimostrazione. ...

Proposizione: Date $A, B \in \mathbb{C}^{n\times n} \ : \ 0 \prec A, B$

- $A \preceq B \iff A^{-1} \succeq B^{-1}$
- $A \preceq B \implies \log(A) \preceq \log(B)$
- $A \preceq B \implies \operatorname{tr}\exp(A) \leq \operatorname{tr}\exp(B)$

Dimostrazione. ...

Lemma [Concavità di Lieb]: Date $H \in \mathbb{C}^{n\times n}$ hermitiana e $X \in \mathbb{C}^{n\times n}$ v.a. hermitiana

$\mathbb{E}[\operatorname{tr}\exp(H + \log X)] \leq \operatorname{tr}\exp(H + \log \mathbb{E}[X])$

Dimostrazione. ...

Corollario: Date $X_1, \dots, X_s \in \mathbb{C}^{n\times n}$ v.a. hermitiane indipendenti

$$\operatorname{tr}\exp C_{X_1+\cdots+X_s}(\theta) \leq \operatorname{tr}\exp\left(\sum_k C_{X_k}(\theta)\right)$$

Dimostrazione.

... to do ...

□

Teorema [Tropp's master inequality]: Date $X_1, \dots, X_s \in \mathbb{C}^{n\times n}$ v.a. hermitiane indipendenti

$\forall t, \theta > 0 \quad \mathbb{P}[\lambda_{\max}(X_1 + \cdots + X_s) \geq t] \leq e^{-\theta t} \operatorname{tr}\exp(C_{X_1}(\theta) + \cdots + C_{X_s}(\theta))$

Dimostrazione.

... to do ...

□

Proposizione: Data $X \in \mathbb{C}^{n\times n}$ v.a. hermitiana centrata

$\|X\| \leq L \implies C_X(\theta) \preceq g_L(\theta)\mathbb{E}[X^2]$

Dimostrazione.

... to do ...

□

Teorema [Bernstein matrix inequality]: Date $X_1, \dots, X_s \in \mathbb{C}^{n\times n}$ v.a. hermitiane indip. centrate e $Y = X_1 + \cdots + X_s$

$\forall k \quad \|X_k\|_2 \in [-L, L] \implies \exists \sigma^2 = \left\|\mathbb{E}[Y^2]\right\|_2 \ : \ \forall t, \theta > 0 \quad \mathbb{P}[\lambda_{\max}(Y) \geq t] \leq e^{-\theta t} n \exp\left(g_L(\theta)\sigma^2\right)$

Dimostrazione.

... to do ...

□

Corollario: $\forall \theta > 0 \quad \mathbb{E}[\lambda_{\max}(Y)] \leq \dfrac{1}{\theta} n \exp\left(g_L(\theta)\sigma^2\right)$

Dimostrazione.

... to do ...

□

Corollario: $\forall t > 0 \quad \mathbb{P}\left[\lambda_{\max}(Y) \geq t\right] \leq n \exp\left(\dfrac{-\frac{t^2}{2}}{\sigma^2 + \frac{Lt}{3}}\right)$

Dimostrazione. ...

Corollario: $\mathbb{E}\left[\lambda_{\max}(Y)\right] \leq \sqrt{2\sigma^2 \log n} + \frac{1}{3} L \log n$

Dimostrazione. ...

Teorema: Date $X_1, \dots, X_s \in \mathbb{C}^{m\times n}$ v.a. indip. centrate $: \forall k \quad \|X_k\|_2 \in [-L, L]$ e $Y = X_1 + \dots + X_s$

$$\exists \sigma^2 = \max\left\{\|\mathbb{E}[Y^*Y]\|_2, \|\mathbb{E}[YY^*]\|_2\right\} \ : \ \begin{cases} \forall t > 0 \quad \mathbb{P}[\|Y\|_2 \geq t] \leq (m+n)\exp\left(\dfrac{-\frac{t^2}{2}}{\sigma^2 + \frac{Lt}{3}}\right) \\ \mathbb{E}[\|Y\|_2] \leq \sqrt{2\sigma^2 \log(m+n)} + \frac{1}{3} L \log(m+n) \end{cases}$$

Dimostrazione.

... to do ...

□

3.2.1 Approsimazione di matrici tramite sampling

<u>Notazione</u>: Data $A = A_1 + \dots + A_d \in \mathbb{C}^{n\times m}$ e $p \in [0,1]^d \ : \ \|p\|_1 = 1$

$$X_A = X_{A,p} \text{ v.a. } : \mathcal{L}(X_A) = \sum_i p^i \delta_{p_i^{-1} A_i}$$

Teorema: Data $A = A_1 + \dots + A_d \in \mathbb{C}^{n\times m}$, $X_1, \dots, X_s \sim X_A$ e $\overline{X} = \dfrac{1}{s}\sum_i X_i$

$$\begin{cases} \|X_A\| \leq L \\ \max\{\mathbb{E}[X^*X], \mathbb{E}[XX^*]\} \leq \sigma^2 \end{cases} \implies \begin{cases} \forall t > 0 \quad \mathbb{P}\left[\left\|\overline{X} - A\right\|_2 \geq t\right] \leq (m+n)\exp\left(\dfrac{-\frac{st^2}{2}}{\sigma^2 + \frac{2Lt}{3}}\right) \\ \mathbb{E}\left[\left\|\overline{X} - A\right\|_2\right] \leq \sqrt{\dfrac{2\sigma^2 \log(m+n)}{s}} + \dfrac{2L\log(m+n)}{3s} \end{cases}$$

Dimostrazione.

... to do ...

□

Corollario: $\forall \varepsilon, \delta > 0 \quad s \geq \log\left(\dfrac{m+n}{s}\right) \dfrac{\sigma^2 + \frac{2L\varepsilon}{3}}{\frac{\varepsilon^2}{2}} \implies \mathbb{P}\left[\left\|\overline{X} - A\right\|_2 \geq \varepsilon\right] \leq \delta$

Dimostrazione. ...

Corollario: $\forall \varepsilon > 0 \quad s \geq \max\left\{\dfrac{2\sigma^2 \log(m+n)}{\varepsilon^2}, \dfrac{2L\log(m+n)}{3\varepsilon}\right\} \implies \mathbb{E}\left[\left\|\overline{X} - A\right\|_2\right] \leq 2\varepsilon$

Dimostrazione. ...

Definizione [Coerence statistic]: Data $A = [\ a_1 \mid \cdots \mid a_m\ ] \in \mathbb{C}^{n\times m}$

$$\mu(A) = m \frac{\max_k \|a_k\|_2}{\|A\|_2}$$

Proposizione: Date $C = [\ c_1 \mid \cdots \mid c_d\ ] \in \mathbb{C}^{n\times d}$, $R = [\ r_1 \mid \cdots \mid r_d\ ] \in \mathbb{C}^{m\times d} \ : \ \|C\|_2 = \|R\|_2 = 1$ e $A = CR^* = \sum_k c_k r_k^*$

$$p = \tfrac{1}{d} 1_d \implies \|X_{A,p}\|_2, \left\|\mathbb{E}\left[X_{A,p}X_{A,p}^*\right]\right\|_2, \left\|\mathbb{E}\left[X_{A,p}^*X_{A,p}\right]\right\|_2 \leq \max\{\mu(C), \mu(R)\}$$

Dimostrazione.

... to do ...

□

Corollario: Date $X_1, \dots, X_s \sim X_{A,p}$ e $\overline{X} = \dfrac{1}{s}\sum_i X_i$

$$\forall \varepsilon > 0 \quad s \geq 2\max\left\{\tfrac{1}{\varepsilon^2}, \tfrac{1}{3\varepsilon}\right\} \max\{\mu(C), \mu(R)\} \log(m+n) \implies \mathbb{E}\left[\left\|\overline{X} - A\right\|_2\right] \leq 2\varepsilon$$

Dimostrazione. ...

Definizione [Stable rank]: $A \in \mathbb{C}^{n\times m}$

$$\operatorname{srank}(A) = \frac{\|A\|_F}{\|A\|_2}$$

Proprietà: $\operatorname{srank}(A) \le \min\{n, m\}$

Proposizione: Date $C = [\, c_1 \mid \cdots \mid c_d \,] \in \mathbb{C}^{n\times d}$, $R = [\, r_1 \mid \cdots \mid r_d \,] \in \mathbb{C}^{m\times d}$: $\|C\|_2 = \|R\|_2 = 1$ e $A = CR^* = \sum_k c_k r_k^*$

$$\forall k \quad p_k = \frac{\|c_k\|_2^2 + \|r_k\|^2}{\|C\|_F^2 + \|R\|_F^2} \implies \|X_{A,p}\|_2, \left\|\mathbb{E}\left[X_{A,p}X_{A,p}^*\right]\right\|_2, \left\|\mathbb{E}\left[X_{A,p}^*X_{A,p}\right]\right\|_2 \le \max\{\operatorname{srank}(C), \operatorname{srank}(R)\}$$

Dimostrazione.
... to do ...

□

Corollario: Date $X_1, \ldots, X_s \sim X_{A,p}$ e $\overline{X} = \frac{1}{s}\sum_i X_i$

$$\forall \varepsilon > 0 \quad s \ge \max\left\{\tfrac{1}{\varepsilon^2}, \tfrac{1}{3\varepsilon}\right\}(m+n)\log(m+n) \implies \mathbb{E}\left[\left\|\overline{X} - A\right\|_2\right] \le 2\varepsilon$$

Dimostrazione. ...

3.3 Multiplicative Chernoff inequality

Lemma: Data X v.a. : $X \in [0,1]$

$$\forall \theta \quad C_X(\theta) \le (e^\theta - 1)\mathbb{E}[X]$$

Dimostrazione.
... to do ...

□

Lemma: Data $X \in \mathbb{R}^{m\times n}$ v.a. simmetrica definita positiva : $\lambda_{\max}(X) \le 1$

$$\forall \theta \quad C_X(\theta) \preceq (e^\theta - 1)\mathbb{E}[X]$$

Dimostrazione.
... to do ...

□

Teorema: Date $X_1, \ldots, X_s \in \mathbb{R}^{n\times n}$ v.a. simmetriche definite positive : $\forall s \quad \lambda_{\max}(X_s) \le 1$ e $Y = X_1 + \cdots + X_s$

- $\forall t > 0 \quad \mathbb{P}\left[\lambda_{\max}(Y) \ge (1+t)\lambda_{\max}(\mathbb{E}[Y])\right] \le n\left(\dfrac{e^t}{(1+t)^{1+t}}\right)^{\lambda_{\max}(\mathbb{E}[Y])}$
- $\forall t \in (0,1) \quad \mathbb{P}\left[\lambda_{\min}(Y) \le (1-t)\lambda_{\max}(\mathbb{E}[Y])\right] \le n\left(\dfrac{e^{-t}}{(1-t)^{1-t}}\right)^{\lambda_{\min}(\mathbb{E}[Y])}$

Dimostrazione. ...

3.3.1 Connessione di un grafo di Erdøs - Rényi

Definizione: Dato $G = (\mathcal{V}_G, \mathcal{E}_G)$ grafo

- $A_G = \left((v_i, v_j) \in \mathcal{E}_G\right)_{i,j}$ matrice di adiacenza di G
- $L_G = diag\left(A_G 1_{\sharp\mathcal{V}_G}\right) - A$ laplaciano di G

Proprietà: $1_{\sharp\mathcal{V}_G} \in \ker(L_G)$

Proposizione: Dati $G = (\mathcal{V}_G, \mathcal{E}_G)$ grafo, $A = A_G$ e $L = L_G$

$$\forall x \in \mathbb{C}^{\sharp\mathcal{V}_G} \quad x^* L x = \sum_{i<j} A_{ij}(x_i - x_j)^2$$

Dimostrazione. ...

Corollario: L semidefinita positiva

Dimostrazione. ...

Corollario: G connesso $\iff \dim\ker(L) = 1$

Dimostrazione. ...

Definizione [Grafo di Erdøs - Rényi]: Dati $n \in \mathbb{N}$ e $p \in (0,1)$

$$G = G(n,p) \ : \ \begin{cases} \sharp \mathcal{V}_G = n \\ \forall i \neq j \quad \mathbb{P}\left[(v_i, v_j) \in \mathcal{E}_G\right] = p \end{cases}$$

Proposizione: Dati $n \in \mathbb{N}$, $p \in (0,1)$ e $\{b_{ij}\} \sim Bernoulli(p)$

$$L_{G(n,p)} = \sum_{i<j} b_{ij} \left(E_{ii} - E_{ij} - E_{ji} + E_{jj}\right)$$

Dimostrazione. ...

Teorema: Dati $n \in \mathbb{N}$, $p \in (0,1)$

$$\mathbb{P}[G(n,p) \text{ sconnesso}] \leq (n-1)e^{-\frac{pn}{2}}$$

Dimostrazione.

... to do ...

□

Corollario: $p > \dfrac{2\log(n-1)}{n} \implies \mathbb{P}[G(n,p) \text{ connesso}] > 0$

Dimostrazione. ...

4 Embedding di sottospazi

<u>Notazione</u>: $Sub_k(\mathbb{R}^n) = \{L \subseteq \mathbb{R}^n \mid L \text{ sottospazio} : \ \dim L = k\}$

Definizione: Dati $\varepsilon > 0$ e $L \in Sub(\mathbb{R}^n)$

$S : \mathbb{R}^n \to \mathbb{R}^s$ ε-embedding per L se $\forall x \in L \quad (1-\varepsilon)\|x\|_2^2 \leq \|Sx\|_2^2 \leq (1+\varepsilon)\|x\|_2^2$

Lemma: Dati $L \in Sub_k(\mathbb{R}^n)$, $U \in R^{n \times k}$ base ortonormale di L, $S \in \mathbb{R}^{s \times n}$ e $\varepsilon > 0$

$$S \ \varepsilon\text{-embedding per } L \iff \begin{cases} \sigma^2_{\min}(SU) \geq 1 - \varepsilon \\ \sigma^2_{\max}(SU) \leq 1 + \varepsilon \end{cases} \iff \left\| I_k - U^{\mathrm{T}} S^{\mathrm{T}} S U \right\|_2 \leq \varepsilon$$

Dimostrazione.

... to do ...

□

Definizione: Dati $k, n \in \mathbb{N}$, $\varepsilon, \delta > 0$ e $S \in \mathbb{R}^{n \times s}$ v.a.

S (k, ε, δ) - oblivius subspace embedding se $\forall L \in Sub_k(\mathbb{R}^n) \quad \mathbb{P}\left[S \ \varepsilon\text{-embedding per } L\right] \geq 1 - \delta$

4.1 Embedding Gaussiani

Lemma: Dati $x \in \mathbb{R}^n$ e $S \in \mathbb{R}^{s \times n}$ v.a : $s_{ij} \sim \mathcal{N}\left(0, \frac{1}{s}\right)$

- $\mathbb{E}\left[\|Sx\|_2^2\right] = \|x\|_2^2$
- $\forall \varepsilon > 0 \quad \mathbb{P}\left(\left|\|Sx\|_2^2 - \|x\|_2^2\right| \geq \varepsilon \|x\|_2^2\right) \leq 2\exp\left(-\dfrac{\varepsilon^2 s}{4 + 4\varepsilon}\right)$

Dimostrazione.

... to do ...

□

Corollario: $\forall \varepsilon, \delta > 0 \quad s \geq \dfrac{4+4\varepsilon}{\varepsilon^2} \log \dfrac{2}{\delta} \implies S \in \mathbb{R}^{s \times n}$ v.a : $s_{ij} \sim \mathcal{N}\left(0, \frac{1}{s}\right)$ $(1, \varepsilon, \delta)$ - oblivius subspace embedding

Dimostrazione. ...

Definizione: Dati (T, d) spazio metrico, $K \subseteq T$ e $\varepsilon > 0$

- $N \subseteq K$ ε-net di K se $\forall x \in K \quad d(x, N) \leq \varepsilon$
- $\mathcal{N}(K, \varepsilon) = \min\left\{ n \in \mathbb{N} \;\middle|\; \exists x_1, \ldots, x_n \in K \ : \ K \subseteq \bigcup_i B_\varepsilon(x_i) \right\}$

Lemma: $\forall n \in \mathbb{N} \quad \forall \varepsilon > 0 \quad \mathcal{N}\left(B_1^n, \varepsilon\right) \leq \left(\frac{2}{\varepsilon} + 1\right)^n$

Dimostrazione.
... to do ...

$\square$

Corollario: $\forall n \in \mathbb{N} \quad \forall \varepsilon > 0 \quad \mathcal{N}\left(\mathbb{S}^{n-1}, \varepsilon\right) \leq \left(\frac{2}{\varepsilon} + 1\right)^n - \left(\frac{2}{\varepsilon} - 1\right)^n$

Dimostrazione. ...

Lemma: Dati $\varepsilon \in \left(0, \frac{1}{2}\right)$, $\mathcal{N} \subseteq \mathbb{S}^{n-1}$ ε-net

$$\forall A \in \mathbb{R}^{n \times n} \text{ simmetrica} \qquad \sup_{x \in \mathcal{N}} |\langle Ax, x\rangle| \leq \|A\|_2 \leq \frac{1}{1-2\varepsilon} \sup_{x \in \mathcal{N}} |\langle Ax, x\rangle|$$

Dimostrazione.
... to do ...

$\square$

Lemma: Data $B \in \mathbb{R}^{s \times k}$ v.a. $:$ $b_{ij} \sim \mathcal{N}\left(0, \frac{1}{s}\right)$

$$\forall \varepsilon > 0 \quad \mathbb{P}\left(\|B^{\mathrm{T}}B - I_k\|_2 \geq \varepsilon\right) \leq 2\exp\left(k\log 9 - \frac{\varepsilon^2 s}{16 + 8\varepsilon}\right)$$

Dimostrazione.
... to do ...

$\square$

Teorema: Dati $L \in Sub_k(\mathbb{R}^n)$, $U \in R^{n \times k}$ base ortonormale di L e $S \in \mathbb{R}^{s \times n}$ v.a $:$ $s_{ij} \sim \mathcal{N}\left(0, \frac{1}{s}\right)$

$$\forall \varepsilon > 0 \quad \mathbb{P}\left(S\ \varepsilon\text{-embedding per } L\right) \geq 1 - 2\exp\left(k\log 9 - \frac{\varepsilon^2 s}{16 + 8\varepsilon}\right)$$

Dimostrazione.
... to do ...

$\square$

Corollario: $\forall k \quad \forall \varepsilon, \delta > 0 \quad s \geq \dfrac{16 + 8\varepsilon}{\varepsilon^2}\left(k\log 9 + \log\dfrac{2}{\delta}\right) \implies S\ (k, \varepsilon, \delta)$ - oblivius subspace embedding

Dimostrazione. ...

Lemma [Johnson - Linderstrauss]: Dati $\varepsilon \in (0, 1)$ e $X \subset \mathbb{R}$ $:$ $\sharp X = N$

$$\forall s \geq \frac{8\log(2N^2)}{\varepsilon^2} \quad \exists S : \mathbb{R}^n \to \mathbb{R}^s \text{ lineare } : \forall u, v \in X \quad (1-\varepsilon)\|u - v\|_2^2 \leq \|Su - Sv\|_2^2 \leq (1+\varepsilon)\|u - v\|_2^2$$

Dimostrazione. ...

4.2 Subsampled Randomized Hadamard Transform (SRHT)

Definizione [Matrice di Hadamard]: $H \in \mathbb{R}^{n \times n}$ $:$ $\begin{cases} \forall i, j \quad h_{ij} = \pm 1 \\ \forall i \neq j \quad H^i, H_i \perp H^j, H_j \end{cases}$

Proposizione [Construzione di Sylvester]: Data $H_1 = [1]$

$$\forall n \quad H_n = \frac{1}{\sqrt{2}}\begin{bmatrix} H_{\frac{n}{2}} & H_{\frac{n}{2}} \\ H_{\frac{n}{2}} & -H_{\frac{n}{2}} \end{bmatrix} \text{ matrice di Hadamard}$$

Dimostrazione.
... to do ...

$\square$

Definizione [SRHT]: Date $C \in \mathbb{R}^{s \times n}$ $:$ $C_i \sim e_{Unif(n)}^{\mathrm{T}}$, $H \in \mathbb{R}^{n \times n}$ matrice di Hadamard riscalata e $D = diag(Rademacher(n))$

$$S = \sqrt{\tfrac{n}{s}}CHD \text{ Subsampled Randomized Hadamard Transform}$$

Lemma: Dati $S = \sqrt{\frac{n}{s}}CHD$ SRHT e $x \in \mathbb{R}^n$

$$\mathbb{E}\left[\|Sx\|_2^2\right] = \|x\|_2^2$$

Dimostrazione.

... to do ...

□

Teorema: Dati $f : \mathbb{R}^n \to \mathbb{R}$ L - Lipschitziana e $r \sim Rademacher(n)$

$$\forall t \quad \mathbb{P}\left(f(r) \geq \mathbb{E}[f(r)] + Lt\right) \leq \exp\left(-\frac{t^2}{8}\right)$$

Dimostrazione. ...

Lemma: Date $U \in \mathbb{R}^{n\times k}$ ortonormale, $H \in \mathbb{R}^{n\times n}$ matrice di Hadamard riscalata e $D \sim diag(Rademacher(n))$

$$W = HDU \text{ ortonormale } : \forall t \quad \mathbb{P}\left(\max_j \left\|e_j^{\mathrm{T}} W\right\|_2 \geq \sqrt{\frac{k}{n}} + \frac{t}{\sqrt{n}}\right) \leq n \exp\left(-\frac{t^2}{8}\right)$$

Dimostrazione.

... to do ...

□

Teorema [Tropp, 2011]: Date $U \in \mathbb{R}^{n\times k}$ ortonormale e $S = \sqrt{\frac{n}{s}}CHD \in \mathbb{R}^{s\times n}$ SRHT

$$\forall \alpha, \xi, \eta \quad s \geq \alpha \log k \left(\sqrt{k} + \sqrt{8\log(nk)}\right)$$

$$\implies \mathbb{P}\begin{pmatrix} \sigma_{\min}(SU) \geq \sqrt{1-\xi} \\ \sigma_{\max}(SU) \leq \sqrt{1+\eta} \end{pmatrix} \geq 1 - \frac{1}{k} - k\left(\frac{e^{-\xi}}{(1-\xi)^{1-\xi}}\right)^{\alpha \log k} - k\left(\frac{e^{\eta}}{(1+\eta)^{1+\eta}}\right)^{\alpha \log k}$$

Dimostrazione. ...

4.3 Sparse random embeddings

Definizione: Dato $p \in (0, 1)$

- $\tilde{s} = \tilde{s}(p)$ v.a. $: \mathcal{L}(\tilde{s}) = (1-p)\delta_0 + \frac{p}{2}\delta_1 + \frac{p}{2}\delta_{-1}$
- $S = S(p) \in \mathbb{R}^{s\times n}$ v.a. $: \forall i, j \quad s_{ij} \sim \frac{1}{\sqrt{ps}}\tilde{s}(p)$

Teorema: $\forall n, k \quad \begin{cases} s \approx (\log n + k)\log k \\ p \approx \frac{\log k}{s} \end{cases} \implies S(p) \in\in \mathbb{R}^{s\times n}$ oblivius subspace embedding

Dimostrazione. ...