Día 2 - Matrices aleatorias

Un montón de aplicaciones

Hoy estudiaremos matrices aleatorias con entradas sub-gaussianas.

Recordemos que la norma espectral de una matriz $A \in \mathbb{R}^{n\times n}$ simétrica es igual a

$$\|A\| = \sup_{u\in S^{n-1}} \|Au\| = \sup_{u\in S^{d-1}} \langle u, Au\rangle = \max\{|\lambda_1(A)|, |\lambda_n(A)|\}.$$

De forma similar si $A \in \mathbb{R}^{n\times m}$

$$\|A\| = \sup_{u\in S^{m-1}} \|Au\| = \sup_{\substack{u\in S^{m-1}\\ v\in S^{n-1}}} \langle v, Au\rangle = \sigma_1(A).$$

Hoy acotaremos la norma espectral.

Mallas, numeros de recubrimiento y empaquetamiento

Estudiaremos un método sencillo, pero poderoso que se basa en ε-mallas.

Def: Sea (T, d) un espacio métrico. Considere un conjunto K y fije un $\varepsilon>0$. Un subconjunto $N \subseteq K$ es una ε-malla si $\forall x \in K \ \exists y \in N$ tq

$$d(x, y) \leq \varepsilon.$$

Def: La cardinalidad más pequeña de una

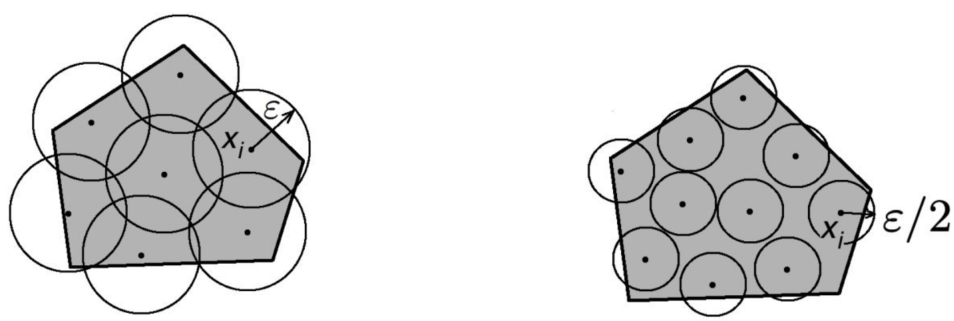
ε-malla de K se llama **número de recubrimiento**, lo denotamos $\mathcal{N}(K, d, \varepsilon)$.

Recorderis: Un conjunto es precompacto si y solo si $\mathcal{N}(K,d,\varepsilon)$ $\forall \varepsilon > 0$.

Def: Un subconjunto $\mathcal{N}$ de (T,d) está ε-separado si $\forall x, y \in \mathcal{N}$ distintos,

$$d(x,y) > \varepsilon.$$

La cardinalidad del conjunto ε-separado más grande contenido en K se llama número de empaquetamiento y se denota $\mathcal{P}(K,d,\varepsilon)$.



Lemma:
$$\mathcal{P}(K,d,2\varepsilon) \leq \mathcal{N}(K,d,\varepsilon) \leq \mathcal{P}(K,d,\varepsilon).$$

Prueba: Probamos solo la cota superior. Suponga que $\mathcal{N}$ es un conjunto ε-separado maximal separado. Veamos que también es una ε-malla. Suponga que no es una ε-malla

$$\Rightarrow \exists y \in K \text{ tq } d(x,y) > \varepsilon \ \forall x \in \mathcal{N}$$

$\Rightarrow \mathcal{N} \cup \{y\}$ está ε-separado $\lightning$.

La cota inferior queda como Ejercicio $\square$

Números de recubrimiento y volumen

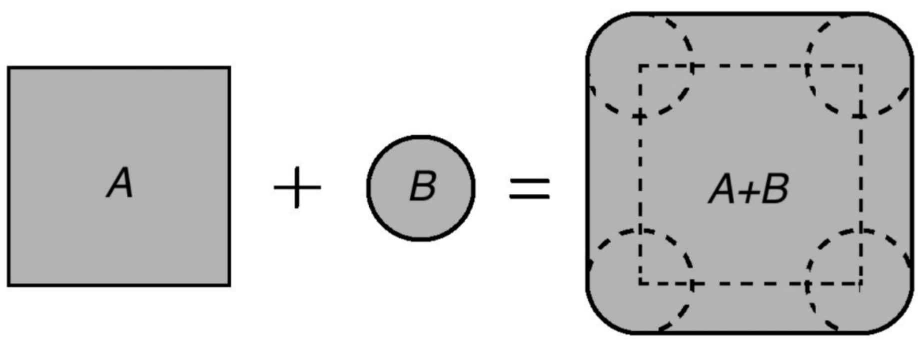
Suponga que $T = \mathbb{R}^n$ y $d(x,y) = \|x-y\|_2$.

Usamos la notación

$$\mathcal{N}(K,\varepsilon) := \mathcal{N}(K,d,\varepsilon)$$
$$\mathcal{P}(K,\varepsilon) := \mathcal{P}(K,d,\varepsilon).$$

Recuerden que la suma de Minkowski de dos conjuntos $A, B \subseteq \mathbb{R}^n$ se define como

$$A+B = \{a+b \mid a \in A,\ b \in B\}.$$



Proposición: Sea $K \subseteq \mathbb{R}^n$ y $\varepsilon > 0$. Entonces

$$\frac{\mathrm{Vol}(K)}{\mathrm{Vol}(\varepsilon B_2^n)} \leq \mathcal{N}(K,\varepsilon) \leq \mathcal{P}(K,\varepsilon) \leq \frac{\mathrm{Vol}(K + \frac{\varepsilon}{2} B_2^n)}{\mathrm{Vol}(\frac{\varepsilon}{2} B_2^n)}.$$

Bola euclidiana

Volumen en $\mathbb{R}^n$.

Prueba:

Cota inferior: Sea $N = \mathcal{N}(K,\varepsilon)$. Entonces K

se puede cubrir con N ε-bolas

$$\Rightarrow \quad \mathrm{Vol}(K) \leq N \, \mathrm{Vol}(\varepsilon B_2^n).$$

Cota superior: Sea $N = \mathcal{P}(K, \varepsilon)$. Podemos construir N bolas disjuntas de radio $\varepsilon/2$ que están contenidas en $K + \frac{\varepsilon}{2} B_2^n$. Entonces

$$N \cdot \mathrm{Vol}(\tfrac{\varepsilon}{2} B_2^n) \leq \mathrm{Vol}(K + \tfrac{\varepsilon}{2} B_2^n).$$

$\square$

Corolario: Para cualquier $\varepsilon > 0$

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}(B_2^n, \varepsilon) \leq \left(\frac{2}{\varepsilon} + 1\right)^n.$$

La cota superior aplica para S^{n-1} también.

Prueba: Note que $\mathrm{Vol}(r B_2^n) = r^n \mathrm{Vol}(B_2^n)$. Entonces

$$\left(\frac{1}{\varepsilon}\right)^n = \frac{\mathrm{Vol}(B_2^n)}{\mathrm{Vol}(\varepsilon B_2^n)} \leq \mathcal{N}(B_2^n, \varepsilon)$$

$$\leq \frac{\mathrm{Vol}((1 + \varepsilon/2) B_2^n)}{\mathrm{Vol}(\varepsilon/2 \, B_2^n)}$$

$$= \left(\frac{2}{\varepsilon} + 1\right)^n.$$

El argumento para S^{n-1} es analogo. $\square$

Cotas para matrices sub-gaussianas

Lemma: Sea $A \in \mathbb{R}^{n\times m}$ y $\varepsilon \in [0,1)$. Entonces para cualquier ε-malla $\mathcal{N} \subseteq \mathcal{S}^{m-1}$ tenemos

$$(-.-) \quad \sup_{x\in\mathcal{N}} \|Ax\| \leq \|A\| \leq \frac{1}{1-\varepsilon} \sup_{x\in\mathcal{N}} \|Ax\|.$$

Más aún si $\mathcal{M} \subseteq \mathcal{S}^{n-1}$ es una ε-malla,

$$(\cdot\_\cdot) \quad \sup_{\substack{x\in\mathcal{N}\\ y\in\mathcal{M}}} \langle y, Ax\rangle \leq \|A\| \leq \frac{1}{1-2\varepsilon} \sup_{\substack{x\in\mathcal{N}\\ y\in\mathcal{M}}} \langle y, Ax\rangle$$

$\varepsilon \in [0, 1/2)$

$\dashv$

Prueba: Probamos $(-.-)$ y dejamos $(\cdot\_\cdot)$ como ejercicio. La cota inferior es trivial. Sea $x \in \mathcal{S}^{m-1}$ t.q. $\|A\| = \|Ax\| \Rightarrow \exists\, x_0 \in \mathcal{N}$ t.q $\|x - x_0\| \leq \varepsilon$. Luego

$\|Au\| \leq \|A\|\|u\|$.

$$\|Ax_0\| \geq \|Ax\| - \|A(x-x_0)\| \geq \|A\| - \varepsilon\|A\| = (1-\varepsilon)\|A\|.$$

Desigualdad triangular

$\square$

Teorema: Sea A una matriz $n\times m$ con entradas independientes sub-gaussianas con $\mathbb{E}A_{ij} = 0$. Entonces, para todo $t > 0$ tenemos

$$:= \max \|A_{ij}\|_{\psi_2}$$

$$\|A\| \leq cK(\sqrt{m} + \sqrt{n} + t)$$

con probabilidad por lo menos $1 - 2\exp(-t^2)$.

Prueba:

Paso 1: Aproximación. Escoja $\varepsilon = 1/4$. Usando el Corolario ♡ obtenemos ε-mallas $\mathcal{N} \subseteq S^{m-1}$ y $\mathcal{M} \in S^{n-1}$ tales que

$$|\mathcal{N}| \leq 9^n \quad \text{y} \quad |\mathcal{M}| \leq 9^n$$

Cardinalidad

Gracias al Lema de arriba

$$\|A\| \leq 2 \max_{\substack{x \in \mathcal{N} \\ y \in \mathcal{M}}} \langle y, Ax \rangle.$$

Paso 2: Concentración. Note que para un par $(x, y) \in \mathcal{N} \times \mathcal{M}$ fijo

$$\|\langle y, Ax\rangle\|_{\psi_2}^2 = \Big\| \sum_{i,j} A_{ij} x_i y_j \Big\|_{\psi_2}^2$$

Proposición (++)

$$\leq C \sum \|A_{ij} x_i y_j\|_{\psi_2}^2$$

$$\leq C K^2 \sum_{ij} x_i^2 y_j^2$$

$$= C K^2 \Big(\sum_i x_i^2\Big)\Big(\sum y_i^2\Big)$$

$$= C K^2.$$

Entonces, $\forall\, u \geq 0$

$$\mathbb{P}(\langle y, Ax\rangle \geq u) \leq 2\exp(-cu^2/K^2).$$

Paso 3: Uniendo todo. Usamos la desigualdad de Boole:

$$\mathbb{P}(\|A\| \geq u) \overset{\text{Paso 1}}{\leq} \mathbb{P}\left\{ \max_{\substack{x \in \mathcal{N} \\ y \in \mathcal{M}}} \langle y, Ax \rangle \geq u \right\}$$

$$\overset{\text{Boole}}{\leq} \sum_{\substack{x \in \mathcal{N} \\ y \in \mathcal{M}}} \mathbb{P}(\langle y, Ax \rangle \geq u)$$

$$\overset{\text{Paso 2}}{\leq} 2 \cdot 9^{n+m} \cdot \exp(-cu^2/K^2).$$

Ahora tomemos $u = CK(\sqrt{n} + \sqrt{m} + t)$.
$\quad$ Por definir

Note que $u^2 \geq C^2K^2(n+m+t^2)$ y si tomamos $C = 3/\sqrt{c} \Rightarrow cu^2/K^2 \geq 3(n+m) + t^2$.

Entonces

$$\mathbb{P}(\|A\| \geq u) \leq 2 \cdot 9^{n+m} \exp(-(3(n+m)+t^2))$$
$$\leq 2\exp(-t^2). \qquad \square$$

Corolario ϟ: Sea $A \in \mathbb{R}^{n \times n}$ una matriz simétrica con entradas sub-gaussianas independientes con esperanza cero en su parte triangular superior. Entonces

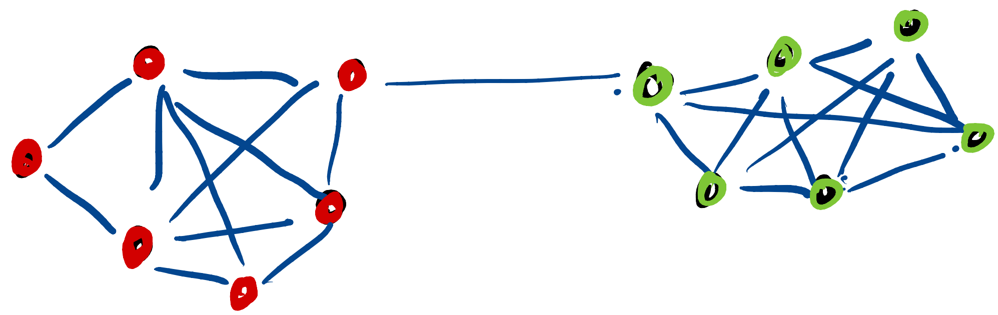
$$\|A\| \leq CK(\sqrt{n} + t)$$

$K := \max \|A_{ij}\|_{\psi_2}$

con probabilidad por lo menos $1 - 4e^{-t^2}$.

Aplicación: Detección de comunidades

Recordemos que teníamos un grafo aleatorio $G(n, p, q)$ con dos comunidades.



Algoritmo

- Calcule el segundo vector propio u_2 de la matriz de adjacencia A_G.
- Returne $\text{sign}(u_2)$ como etiquetas.

Teorema (☺): Suponga $p > q$ y $\min\{q, p-q\} = \mu > 0$. Entonces con probabilidad $1-4e^{-n}$, el algoritmo solo se equivoca en a lo sumo C/μ^2 etiquetas.

Para probar este teorema necesitamos hacer un pequeño recorderis de teoría de perturbación.

Teorema (Davis-Kahan) Sean S y T matrices simétricas $n \times n$. Fije un i tal que

$$\min_{j : j \neq i} |\lambda_i(S) - \lambda_j(S)| = \delta > 0.$$

Entonces

$$\sin \underbrace{\angle}_{\text{ángulo entre } 0 \text{ y } \pi/2} (\underbrace{v_i(S), v_i(T)}_{\text{vector propio}}) \leq \frac{2\|S-T\|}{\delta}$$

Intuición Si S y T están cerca, sus vectores propios también lo están:

$$\exists\, \theta \in \{\pm 1\} \text{ t.q. } \|v_i(S) - \theta v_i(T)\|_2 \leq \frac{2^{3/2}\|S-T\|}{\delta}.$$

¡Chequear!

Prueba del Teorema (:):

Considere $\underbrace{A_G}_{T} = \underbrace{\mathbb{E} A_G}_{S} + \underbrace{(A_G - \mathbb{E} A_G)}_{R}$.

Recuerde que

$$\lambda_1(S) = \frac{(p+q)}{2} n \qquad \lambda_2(S) = \frac{(p-q)}{2} n$$

y $\lambda_i(S) = 0 \quad \forall i > 2$. Entonces

$$\delta = \underbrace{\min\left(\frac{p-q}{2}, q\right)}_{\mu :=} n$$

Las entradas de $A_G - \mathbb{E} A_G$ son v.a.

sub-gaussianas ind con esperanza cero. Entonces por Corolario 4 tenemos

$$\| A_G - \mathbb{E} A_G \| \leq C\sqrt{n}$$

con probabilidad $1-4e^{-n}$. Luego por Davis-Kahan tenemos que

$$\min_{\theta \in \{\pm 1\}} \| v_2(A_G) - v_2(\mathbb{E} A_G) \|_2 \leq \frac{C\sqrt{n}}{\mu n} = \frac{C}{\mu \sqrt{n}}.$$

norma 1

Note que si multiplicamos por $\sqrt{n}$ $\Rightarrow$ con $u_2 := \sqrt{n}\, v_2$ tenemos que $v_2(\mathbb{E} A_G)$ tiene entradas en $\{\pm 1\}$. Entonces

$$\begin{aligned} |\{ i : u_2(A_G) \neq u_2(\mathbb{E} A_G) \}| &\leq \sum_{i=1}^{n} (u_2^{(i)}(A_G) - u_2^{(i)}(\mathbb{E} A_G))^2 \\ &\leq C^2/\mu^2. \end{aligned}$$

$\square$