archi gupta
21111014

Q 1

(a) $$\min_{x \in \mathbb{R}^d} \|Ax-b\|^2 + \|Cx-d\|^2 \qquad A, C \in \mathbb{R}^{n\times d}, \quad b, d \in \mathbb{R}^n.$$

Let $S \sim N(0, \frac{1}{k})$ be a random matrix of dim $k \times n$.

① It preserves the column space of $[A|b]$ approx.

② It preserves the column space of $[C|d]$ approx.

$$① \Rightarrow P\left[\forall x \in \mathbb{R}^{d+1} \ \|S[A\ b]x\| \in (1 \pm \varepsilon)\|[A\ b]x\|\right] > 1-\delta$$

$$② \Rightarrow P\left[\forall x \in \mathbb{R}^{d+1} \ \|S[C|d]x\| \in (1 \pm \varepsilon)\|[C\ d]x\|\right] > 1-\delta$$

$k = O\left(\frac{d}{\varepsilon^2}\log\frac{1}{\delta}\right)$ for both ① & ②

$$\|SAx-Sb\|^2 + \|SCx-Sd\|^2 \in \underset{(1 \pm O(\varepsilon))}{(1\pm\varepsilon)^2}\left(\|Ax-b\|^2 + \|Cx-d\|^2\right)$$

with prob $1-\delta$.

21111014

Q1

(b) $\min_{x \in \mathbb{R}^d} \|Ax-b\|^2 + \|Cx-d\|^2$. $A, C \in \mathbb{R}^{n\times d}$, $b, d \in \mathbb{R}^n$

(1) Using differentiation :

$$\|Ax-b\|^2 + \|Cx-d\|^2$$
$$= (Ax-b)^T(Ax-b) + (Cx-d)^T(Cx-d)$$
$$= (x^TA^T-b^T)(Ax-b) + (x^TC^T-d^T)(Cx-d)$$
$$f = x^TA^TAx + b^Tb - x^TA^Tb - b^TAx + x^TC^TCx + d^Td - x^TC^Td - d^TCx$$

Taking gradient w.r.t x.

$$\nabla f = 2A^TAx - \underbrace{A^Tb - b^TA}_{\text{same}} + 2C^TCx - \underbrace{C^Td - d^TC}_{\text{same}}$$

$$\nabla f = 2(A^TA + C^TC)x - 2(A^Tb - C^Td).$$

for minima $\nabla f = 0$ is a necessary condition [let minima $= x^*$]

$$\nabla f = 0 \Rightarrow (A^TA + C^TC)x^* = A^Tb + C^Td.$$

and $\nabla^2 f$ at x^* should be positive semi definite

$\nabla^2 f = 2(A^TA + C^TC)$ is positive semi definite ~~iff $(A^TA + C^TC)$ is positive semi definite~~

(c) $x^* = (A^TA + C^TC)^-(A^Tb + C^Td)$ [general solution]

where $(A^TA+C^TC)^-$ is pseudo inverse of $(A^TA + C^TC)$

archi gupta
21111014

(d) Also algebraic solution :

$$\min_{x \in \mathbb{R}^d} \underbrace{\|Ax-b\|_2^2 + \|Cx-d\|_2^2}.$$

$$\Downarrow$$

$$\left\| \begin{bmatrix} Ax-b \\ Cx-d \end{bmatrix} \right\|_2^2 .$$

$$\Downarrow$$

$$\min_{x \in \mathbb{R}^d} \left\| \underbrace{\begin{bmatrix} A \\ C \end{bmatrix}}_{2n \times d} x - \underbrace{\begin{bmatrix} b \\ d \end{bmatrix}}_{2n \times 1} \right\|_2^2 .$$

$$x^* = \begin{bmatrix} A \\ C \end{bmatrix}^{-} \begin{bmatrix} b \\ d \end{bmatrix}.$$

archi gupta
21111014

Q2

$$\min_{\text{rank}(X) \le k} \|AX - B\|_F \qquad A \in \mathbb{R}^{m\times n}, \quad B \in \mathbb{R}^{m\times p}, \quad X \in \mathbb{R}^{n\times p}$$

(a) consider $\|AX - B\|_F^2$

$$= \|AX - AA^{-}B - (I - AA^{-})B\|^2$$

$$= \underbrace{\|AX - AA^{-}B\|_F^2}_{\text{contains } X} + \underbrace{\|(I - AA^{-})B\|_F^2}_{\text{constant term}}$$

$$\arg\min_{r(X)\le k} \|AX - B\| = \arg\min_{r(X)\le k} \|AX - B\|^2 = \arg\min_{r(X)\le k} \|AX - AA^{-}B\|_F^2$$

$\min_{r(X)\le k} \|AX - AA^{-}B\|_F^2$ by Eckart's young's theorm. best approximation to $\|X_k - AA^{-}B\|_F^2$ is $X_k = [AA^{-}B]_k$

$$\therefore AX^* = X_k = [AA^{-}B]_k$$

$$X^* = A^{-}[AA^{-}B]_k \text{ ——— } (1)$$

$$AX^* = \underbrace{AA^{-}}_{\substack{\text{full col space} \\ \text{of } A}} \underbrace{[AA^{-}B]_k}_{\substack{\text{col sp of } A \\ \& \text{ rank } k}} = [AA^{-}B]_k$$

(b) $S \Rightarrow$ affine embedding $\in \mathbb{R}^{r\times m}$ s.t.

$$\|SAA^{-}[AA^{-}B]_k X - SB\|_F \in (1 \pm \varepsilon)\|AA^{-}[AA^{-}B]_k X - B\|_F$$

it is sufficient to col space of $A_{m\times n}$ & norm of $B_{m\times p}$.

$$r = O\Big[\underbrace{\frac{n}{\varepsilon^2}\log\frac{1}{\delta}}_{\substack{\text{for col space} \\ \text{of } A \Rightarrow m\times n}} + \underbrace{\frac{1}{\varepsilon^2}\log\frac{p}{\delta}}_{\text{for norm of } B \approx p \text{ col vectors in } B}\Big] \text{ ——— } (2)$$

(c) $\arg\min_X \| SAA^-[AA^-B]_k X - SAA^-B\|_F = \hat{X}$

let $AA^-[AA^-B]_k = A X_k$ (optimal solution to $\|AX-B\|$ $\min_{r(X)\le k}$), we have:

$$\|SAX_kX - SAA^-B\| \in (1\pm\varepsilon)\|AX_kX - AA^-B\|$$

(let S also preserve norm of $\underbrace{\|AA^-B\|}_{(m\times p)}$.) ↳ (2) is still sufficient

$$(1-\varepsilon)\|AX_kX - AA^-B\| \le \underbrace{\|SAX_kX - SAA^-B\|}_{\text{let optimal sol } \hat{X}} \le (1+\varepsilon)\underbrace{\|AX_kX - AA^-B\|}_{\text{optimal sol } = X^*}$$

$$(1-\varepsilon)\|AX_k\hat{X} - AA^-B\| \le \|SAX_k\hat{X} - SAA^-B\| \le \|SAX_kX^* - SAA^-B\|$$

$$\le (1+\varepsilon)\|\underbrace{AX_kX^*}_{\text{max rank} = k} - AA^-B\|$$

$AX_k = [AA^-B]_k$.

$(AA^-[AA^-B]_k)$

∴ optimal solution is I.

$$\boxed{(1-\varepsilon)^2\|AX_k\hat{X} - AA^-B\|^2 \le (1+\varepsilon)^2\|AX_k - AA^-B\|^2.}$$

$$+(1-\varepsilon)^2\|(I-AA^-)B\|^2 \qquad + (1+\varepsilon)^2\|(I-AA^-)B\|^2. \quad (3)$$

$$(1-\varepsilon)^2\|AX_k\hat{X} - B\|_F^2 \le (1+\varepsilon)^2\|AX_k - B\|_F^2$$

$$\|AX_k\hat{X} - B\|_F^2 \le (1+O(\varepsilon))\underbrace{\|AX_k - B\|_F^2}_{\min_{r(X)\le k}\|AX-B\|_F^2}$$

$\hat{X} = \arg\min \|SAX_k\hat{X} - SAA^-B\|$

$\hat{X} = (SAX_k)^-(SAA^-B)$.

$= (\underbrace{SAA^-[AA^-B]_k}_{\text{max rank} = k \text{ (say } U_k\Sigma_kV_k^T)})^-(SAA^-B)$ } $\hat{X}$ has rank at max K

∴ pseudo inverse's rank is also $\le K$

[cut off] gupta
21/10/14

(d) $$\min_{\mathrm{rank}(Y)\le k} \| Y(SAA^{-}B) - AA^{-}B\|_F .$$

equation (3) can be written as:

$$(1-\varepsilon)\, \| \underbrace{AA^{-}[AA^{-}B]_k\overbrace{(SAA^{-}[AA^{-}B]_k)^{-}}^{*}}_{\text{has rank} \le k}(SAA^{-}B) - AA^{-}B\| \le (1+\varepsilon)\,\|AX_k - AA^{-}B\| .$$

is one possible solution for Y

$$\le (1+\varepsilon)\,\|AX_k - B\| .$$

$\therefore$

$$\min_{r(Y)\le k} \|Y(SAA^{-}B) - AA^{-}B\| \le \|AX_k(SAX_k)^{-}SAA^{-}B - AA^{-}B\| .$$

$$\le \|AX_k - AA^{-}B\|\,(1+\varepsilon) .$$

$$\le \|AX_k - B\|\,(1+O(\varepsilon))$$

$$\Rightarrow \min_{r(X)\le k}\|AX - B\|\cdot(1+O(\varepsilon)) .$$

(e) $\|YSAA^{-}BR - BR\|_F^2$

consider $SAA^{-}BR = T$

$$\|YT - BR\|_F^2 = \|\underbrace{YT - BR(T^{-}T)}_{} - \underbrace{BR(I - T^{-}T)}_{}\|_F^2$$

(orthogonal).

Projection matrix of row space of T is $T^{-}T$

$$= \|YT - BRT^{-}T\|_F^2 + \|BR - BRT^{-}T\|_F^2$$

$\therefore$

$$\|YSAA^{-}BR - BR\|_F^2 = \|YSAA^{-}BR - BR(SAA^{-}BR)^{-}(SAA^{-}BR)\|_F^2 + \|BR - BR(SAA^{-}BR)^{-}(SAA^{-}BR)\|_F^2$$

21111014
archi gupta

Q3

(a) $S_{ij} = b_{ij}\varepsilon_{ij}$

$$(Sx)_i^2 = \left(\sum_{j=1}^{n} b_{ij}\varepsilon_{ij}x_j\right)^2$$

$$= \underbrace{\sum_{j=1}^{n}(b_{ij}\varepsilon_{ij}x_j)^2}_{\Downarrow} + 2\sum_{j<j'} b_{ij}b_{ij'}\varepsilon_{ij}\varepsilon_{ij'}x_jx_{j'}$$

$(b_{ij})^2 = b_{ij}$

$(\varepsilon_{ij})^2 = 1$.

$$(Sx)_i^2 = \sum_{j=1}^{n} b_{ij}x_j^2 + 2\sum_{1\le j<j'\le n} b_{ij}b_{ij'}\varepsilon_{ij}\varepsilon_{ij'}x_jx_{j'}$$

(b)

$$\|Sx\|^2 = \sum_{i=1}^{n}(Sx)_i^2$$

$$= \underbrace{\sum_{i=1}^{k}\sum_{j=1}^{n} b_{ij}x_j^2}_{\Downarrow} + \sum_{i=1}^{k}\sum_{j<j'} b_{ij}b_{ij'}\varepsilon_{ij}\varepsilon_{ij'}x_jx_{j'}$$

$$\sum_{j=1}^{n}\underbrace{\sum_{i=1}^{k} b_{ij}}x_j^2$$

$$= \sum_{j=1}^{n} x_j^2 = \|x\|^2 = 1.$$

$$\therefore \|Sx\|^2 - 1 = \sum_{i=1}^{k}\sum_{j<j'} b_{ij}b_{ij'}\varepsilon_{ij}\varepsilon_{ij'}x_jx_{j'}$$

5

Scanned By Scanner Go

archi gupta.

(c)

$$E[\|Sx\|^2 - 1] = \sum_{i=1}^{n} \sum_{j<j'} 2E[b_{ij} b_{ij'} \varepsilon_{ij} \varepsilon_{ij'} x_j x_{j'}]$$

(by linearity of expectation)

now b_{ij}'s & ε_{ij}'s & x_j's are independent from each other

$$E[\|Sx\|^2 - 1] = \sum_{i=1}^{n} \sum_{j<j'} 2E[b_{ij} b_{ij'}] \underbrace{E[\varepsilon_{ij} \varepsilon_{ij'}]} E[x_j x_{j'}]$$

all radamachar variables are independent

$$E[\varepsilon_{ij} \varepsilon_{ij'}] = \underset{0}{E[\varepsilon_{ij}]} \cdot \underset{0}{E[\varepsilon_{ij'}]} = 0$$

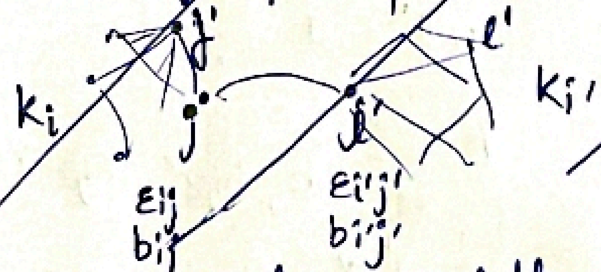
$$E[\|Sx\|^2 - 1] = 0.$$

(4)

(a) $E[\varepsilon_{ij} \varepsilon_{ij'} \varepsilon_{i'l} \varepsilon_{i'l'}] = 0$ irrespective of (j, j') & (l, l')

$i \neq i'$

for each row index i, we have a complete graph K_i (n vertices) and for a pair (j, j') there is an edge corresponding to $T^i_{jj'}$.



since all radamacher variables are independent

$$E[\varepsilon_{ij} \varepsilon_{ij'} \varepsilon_{i'l} \varepsilon_{i'l'}] = E[\varepsilon_{ij}] E[\varepsilon_{ij'}] E[\varepsilon_{i'l}] E[\varepsilon_{i'l'}]$$

$$= 0 \cdot 0 \cdot 0 \cdot 0$$

$$= 0.$$

Scanned By Scanner Go

21111014

4 (b) $E[T^i_{jj'}, T^{i'}_{ll'}]$

$$= E[4 \cdot b_{ij} b_{ij'} b_{i'l} b_{i'l'} \varepsilon_{ij}\varepsilon_{ij'}\varepsilon_{i'l}\varepsilon_{i'l'} x_j x_{j'} x_l x_{l'}].$$

$$= 4E[b_{ij} \cdot b_{ij'} \cdot b_{i'l} \cdot b_{i'l'}] \, \underbrace{E[\varepsilon_{ij} \varepsilon_{ij'} \varepsilon_{i'l} \varepsilon_{i'l'}]}_{=0} \, E[x_j x_{j'} x_l x_{l'}]$$

if $i \neq i'$ $\Downarrow$ $=0$.

∴ i has to be equal to i' for non zero expectation

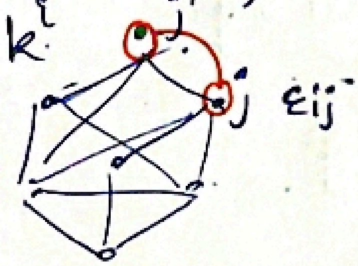
(c) $E[T^i_{jj'} \cdot T^i_{ll'}]$.

$$= 4E[b_{ij} \cdot b_{ij'} b_{il} b_{il'}] \, E[\underbrace{\varepsilon_{ij} \varepsilon_{ij'} \cdot \varepsilon_{il} \varepsilon_{il'}}_{\text{these are still independent}}] \, E[x_j x_{j'} x_l x_{l'}]$$

if $j \neq j'$ or $l \neq l'$ then ε_{ij} & $\varepsilon_{ij'}$ or $\varepsilon_{il}, \varepsilon_{il'}$ will be independent & expectation will be zero

Hence both $j = j'$ & $l = l'$ should hold.

(d)



$E[T^i_{jj'} \times T^i_{jj'}] \neq 0$.

Both ε_{ij} & $\varepsilon_{ij'}$ should be paired with themselves for non zero expectation

21111014

(e) (i) $T^i_{jj'} T^i_{jj'} = 4(b_{ij} b_{ij'})^2 (\varepsilon_{ij} \varepsilon_{ij'})^2 (x_j x_{j'})^2$

$= 4 b_{ij} b_{ij'} \cdot 1 \cdot x_j x_{j'}$ $\quad (b_{ij}^2 = b_{ij})$.

(f) $E[(T^i_{jj'})^2] = 4 E[b_{ij} b_{ij'} x_j^2 x_{j'}^2]$.

$= 4 E[b_{ij} b_{ij'}] E[x_j^2 x_{j'}^2]$.

column wise they are independent.

$= 4 \cdot E[b_{ij}] E[b_{ij'}] \cdot x_j^2 x_{j'}^2$

$= 4 \cdot \frac{1}{k} \cdot \frac{1}{k} \cdot x_j^2 x_{j'}^2$ ——— (6)

$\left(E[b_{ij}] = \frac{1}{k} \cdot 1 + \frac{k-1}{k} \cdot 0 = \frac{1}{k} \right)$

(g) $E[(\|Sx\|^2 - 1)^2]$

$= E\left[\left(\sum_i \sum_{j<j'} T^i_{jj'} \right)^2 \right]$. by equation (5)

$= E\left[\left(\sum_i \sum_{j<j'} T^i_{jj'} \right) \left(\sum_i \sum_{j<j'} T^i_{jj'} \right) \right]$.

$= E\left[\sum_i \sum_{i'} \sum_{j<j'} \sum_{\ell<\ell'} T^i_{jj'} T^{i'}_{\ell\ell'} \right]$.

$= \sum_i \sum_{i'} \sum_{j<j'} \sum_{\ell<\ell'} E[T^i_{jj'} T^{i'}_{\ell\ell'}]$.

$= \sum_i \sum_{j<j'} E[(T^i_{jj'})^2]$

21111014
archi gupta.

(4) (g) using equation (6)

$$\Rightarrow \frac{4}{k^2} \sum_{i} \sum_{j<j'} x_j^2 x_{j'}^2$$

$$\Rightarrow \frac{4}{k^2} \cdot k \sum_{j<j'} x_j^2 x_{j'}^2$$

$$\leq \frac{4}{k} \left[\left(\sum x_j^2 \right)^2 - \sum_{l} x_l^4 \right]$$

$$\Rightarrow (x_1^2 + x_2^2 + \dots x_n^2)^2 - (x_1^4 + x_2^4 + x_2^4 \dots x_n^4)$$

$$\Rightarrow 2 \cdot \sum_{j<j'} x_j^2 x_{j'}^2 \geq \sum_{j<j'} x_j^2 x_{j'}^2$$

$$E\left[(\|Sx\|^2 - 1)^2\right] \leq \frac{4}{k}\left[1 - (+ve)\right] \leq \frac{4}{k}.$$

(h) $$P\left[|\|Sx\|^2 - 1| > \varepsilon\right] \leq \frac{4}{k\varepsilon^2}.$$

Using chebychev's inequality :

$$P\left[|\|Sx\|^2 - 1| > \varepsilon\right] \leq \frac{\overbrace{E\left[(\|Sx\|^2 - 1)^2\right] - \left(E\left[\|Sx\| - 1\right]\right)^2}^{\sigma^2}}{\varepsilon^2}.$$

$$P\left[|\|Sx\|^2 - 1| > \varepsilon\right] \leq \frac{4}{k\varepsilon^2}.$$

(i) $$P\left[|\|Sx\|^2 - 1| > \varepsilon\right] \leq \underbrace{\frac{4}{k\varepsilon^2} \leq \delta}.$$

$$\frac{4}{\delta\varepsilon^2} \leq k$$