מבוא למערכות לומדות תרגיל קצר 1

אורי מינץ - 314616897

תרגיל 1

(1)

$$\mathbb{E}\left[\overline{X}\right]=\mathbb{E}\left[\frac{1}{m}\sum_{i}X_{i}\right]=\frac{1}{m}\sum_{i}\mathbb{E}\left[X_{i}\right]=\frac{1}{m}\sum_{i}\mu=\mu$$

(2)

$$\text{Var}\left[\overline{X}\right]=\mathbb{E}\left[\left(\overline{X}\right)^{2}\right]-\left(\mathbb{E}\left[\overline{X}\right]\right)^{2}$$

$$\mathbb{E}\left[\overline{X}\right]=\mu\Rightarrow\left(\mathbb{E}\left[\overline{X}\right]\right)^{2}=\mu^{2}$$

$$\begin{aligned}\mathbb{E}\left[\left(\overline{X}\right)^{2}\right]&=\mathbb{E}\left[\left(\frac{1}{m}\sum_{i}X_{i}\right)^{2}\right]=\frac{1}{m^{2}}\mathbb{E}\left[\left(\sum_{i=1}^{m}X_{i}\right)^{2}\right]=\frac{1}{m^{2}}\mathbb{E}\left[\sum_{i=1}^{m}\sum_{j=1}^{n}X_{i}X_{j}\right]\\&=\sum_{i=1}^{m}\sum_{j=1}^{n}\mathbb{E}\left[X_{i}X_{j}\right]=\frac{1}{m^{2}}\left(\sum_{i=1}^{m}\mathbb{E}\left[X_{i}^{2}\right]+\sum_{i\neq j}\mathbb{E}\left[X_{i}X_{j}\right]\right)\\&\overset{iid}{=}\frac{1}{m^{2}}\left(\sum_{i=1}^{m}\mathbb{E}\left[X_{i}^{2}\right]+\sum_{i\neq j}\mathbb{E}\left[X_{i}\right]\mathbb{E}\left[X_{j}\right]\right)\end{aligned}$$

$$\mathbb{E}\left[X_{i}^{2}\right]=\text{Var}\left[X_{i}\right]+\left(\mathbb{E}\left[X_{i}\right]\right)^{2}=\sigma^{2}+\mu^{2}$$

$$\sum_{i\neq j}\mathbb{E}\left[X_{i}\right]\mathbb{E}\left[X_{j}\right]=\left(m^{2}-m\right)\mu^{2}$$

$$\begin{aligned}\mathbb{E}\left[\left(\overline{X}\right)^{2}\right]&=\frac{1}{m^{2}}\left(m\left(\sigma^{2}+\mu^{2}\right)+\left(m^{2}-m\right)\mu^{2}\right)\\&=\frac{\sigma^{2}}{m}+\mu^{2}\end{aligned}$$

$$\text{Var}\left[\overline{X}\right] = \frac{\sigma^2}{m} + \mu^2 - \mu^2 = \frac{\sigma^2}{m}$$

תרגיל 2

(1)

$$\theta_i \sim \text{Bin}\left(50, p\right)$$

(2)

$$\mathbb{E}\left[\theta_i\right] = 50p$$

(3)

$$\mathbb{P}\left[\left|\overline{\theta}\left(m\right) - \mu\right| > 1\right] \leq 0.01$$

$$\begin{aligned} 2\exp\left\{-\frac{2m}{\left(50-0\right)^2}\right\} &\leq 0.01 \\ \Rightarrow \exp\left\{-\frac{2m}{50^2}\right\} &\leq \frac{0.01}{2} \\ \Rightarrow -\frac{2m}{50^2} &\leq \ln\frac{0.01}{2} \\ \Rightarrow m &\geq -1250 \cdot \ln\frac{0.01}{2} = 6622.9 \\ \Rightarrow m &= 6623 \end{aligned}$$

תרגיל 3

(1) נוכיח

הוכחה. יהי λ ע"ע, ויהי v ו"ע של A מתאים. כיוון ש-$A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ ו-$v \in \mathbb{R}^n$ (ואז $\|v\|_2^2 > 0$).

$$0 < v^T A v = v^T \lambda v = \lambda v^T v = \lambda \|v\|_2^2$$

כיוון ש-$\|v\|_2^2 > 0$, $\lambda > 0$. ■

(2) הפרכה.

נבחר $A = Id$ ו-$B = 3Id$, כיוון ש- Id היא PD וסימטרית, מתקיים כי Id היא PSD.

לכן $(2A - B) = -Id$,

נבחר $v=\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}$, ואז:

$$v^T\left(2A-B\right)v=v^T\left(-Id\right)v=-\left(v^Tv\right)=-1<0$$

לכן $\left(2A-B\right)$ לא PSD.

תרגיל 4

(1)

$$\nabla_w f=\left[\frac{\partial f}{\partial w_1},\ldots,\frac{\partial f}{\partial w_d}\right]^T$$

$$\frac{\partial f}{\partial w_i}=\frac{\partial}{\partial w_i}w^Tx+b=\frac{\partial}{\partial w_i}\left(w_1x_1+\ldots+w_ix_i+\ldots+w_dx_d+b\right)=x_i$$

לכן:

$$\nabla_w f=\left[x_1,\ldots,x_d\right]^T=\begin{pmatrix}x_1\\\vdots\\x_d\end{pmatrix}=x$$

(2)

$$\nabla_w^2 f=\begin{pmatrix}\frac{\partial^2 f}{\partial w_1^2} & \frac{\partial^2 f}{\partial w_1\partial w_2} & \cdots & \frac{\partial^2 f}{\partial w_1\partial w_n}\\ \frac{\partial^2 f}{\partial w_2\partial w_1} & \frac{\partial^2 f}{\partial w_2} & \cdots & \frac{\partial^2 f}{\partial w_2\partial w_n}\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial w_n\partial w_1} & \frac{\partial^2 f}{\partial w_n\partial w_2} & \cdots & \frac{\partial^2 f}{\partial w_d^2}\end{pmatrix}$$

$$\left(\nabla_w^2 f\right)_{i,j}=\frac{\partial^2 f}{\partial w_j\partial w_i}=\frac{\partial x_j}{\partial w_i}=0_{d\times d}$$

לכן:

$$\nabla_w^2 f=0$$

(3) כן, כיוון שלכל $0_n\neq v\in\mathbb{R}^n$:

$$v^T\nabla_w^2 fv=v^T0v=0\geq 0$$

לכן היא PSD לפי הגדרה.

(4)

$$g\left(w\right)=\lambda\left\|w\right\|^{2}=\lambda\sum_{i=1}^{n}w_{i}^{2}$$

$$\nabla_{w}g=\left[\frac{\partial g}{\partial w_{1}},\ldots,\frac{\partial g}{\partial w_{n}}\right]^{T}=\left[2\lambda w_{1},\ldots,2\lambda w_{n}\right]^{T}=2\lambda w$$

(5)

$$\left(\nabla_{w}^{2}g\right)_{i,j}=\frac{\partial^{2}g}{\partial w_{j}\partial w_{i}}=\frac{\partial2\lambda w_{i}}{\partial w_{j}}=\begin{cases}2\lambda & i=j\\0 & i\neq j\end{cases}$$

לכן:

$$\nabla_{w}^{2}g=2\lambda Id$$

(6) המטריצה היא PD

יהי $0_{n}\neq v\in\mathbb{R}^{n}$:

$$v^{T}\left(\nabla_{w}^{2}g\right)v=2\underbrace{\lambda}_{>0}\underbrace{\left\|v\right\|_{2}^{2}}_{>0}>0$$

לכן חיובית מוגדרת לפי הגדרה.