تمرین ۰ یادگیری ماشین

نویسنده:

سید احسان حسن بیگی - ۴۰۲۲۱۱۷۲۳

آمار و احتمال

پرسش ۱

$X_i \sim \text{Uniform}(0,1)$

⊛ داریم $Y_n = \max\{X_1, \dots, X_n\}$ پس احتمال اینکه Y_n کمتر از $\overbrace{\text{مقداری}}^{0 \leq x \leq 1}$ باشد برابر است با احتمال اینکه تمام X_i ها کمتر از آن مقدار باشند

$$\underbrace{P(Y_n \leq x)}_{\text{تعریف CDF است}} = P(X_1 \leq x \cap X_2 \leq x \cap \dots \cap X_n \leq x) \overset{X_i \text{ ها مستقل اند}}{=\!=\!=} \prod_{i=1}^{n} \underbrace{P(X_i \leq x)}_{\frac{x}{1-0}} = x^n$$

$$\longrightarrow F_Y(x) = x^n \underset{0 \leq x \leq 1}{} \xrightarrow{\text{مشتق}} f_Y(x) = n x^{n-1} \underset{0 \leq x \leq 1}{}$$

$$\longrightarrow E[Y] = \int_0^1 x f(x)\, dx = \int_0^1 n x^n \, dx = n\left(\frac{1}{n+1} x^{n+1}\right)\Big|_0^1 = \boxed{\frac{n}{n+1}}$$

پرسش ۲

⊛ در توزیع نمایی داشتیم $f_X(x)=\lambda e^{-\lambda x}$ پس $a=\lambda=\frac{1}{2}$. همچنین داشتیم $E[X]=\frac{1}{\lambda}$ پس $E[X]=2$. همچنین داشتیم $Var[X]=\frac{1}{\lambda^2}$
پس $Var[X]=4$. اما اگر بخواهیم تشریحی بنویسیم داریم:

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \longrightarrow \int_0^{\infty} a e^{-\frac{x}{2}}dx = 1 \longrightarrow -2ae^{-\frac{x}{2}}\Big|_0^{\infty} = 1 \longrightarrow \underbrace{\cancel{e^{-\infty}}}_{0} - \underbrace{\cancel{e^{0}}}_{1} = \frac{-1}{2a}$$

$$\longrightarrow 2a = 1 \longrightarrow \boxed{a=\frac{1}{2}}$$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^{\infty} \underbrace{\frac{1}{2}x}_{u}\underbrace{e^{-\frac{x}{2}}dx}_{dv} \xlongequal{\text{انتگرال جزء به جزء}} \frac{1}{2}\left[\overbrace{(x)(-2e^{-\frac{x}{2}})\Big|_0^{\infty}}^{0} - \overbrace{4e^{-\frac{x}{2}}\Big|_0^{\infty}}^{-4}\right] = \frac{1}{2}\times 4 = \boxed{2}$$

$$E[X^2] = \int_0^{\infty} \underbrace{\frac{1}{2}x^2}_{u}\underbrace{e^{-\frac{x}{2}}dx}_{dv} \xlongequal{\text{انتگرال جزء به جزء}} \frac{1}{2}\left[\underbrace{(x^2)(-2e^{-\frac{x}{2}})\Big|_0^{\infty}}_{0} - \underbrace{\int_0^{\infty} -2e^{-\frac{x}{2}}\,2x\,dx}_{\underbrace{-8E[X]}_{-16}}\right] = \frac{1}{2}\times 16 = 8$$

$$Var[X] = E[X^2] - E^2[X] = 8 - 4 = \boxed{4}$$

پرسش ۳

1.3.1 (law of total expectation)

⊛ اگر X و Y پیوسته باشند داریم:

$$E_X[X|Y] = \int_{\mathrm{Range}_X} x\, f_{X|Y}(x|y)\, dx \longrightarrow E_Y\left[E_X[X|Y]\right] = \int_{R_Y}\int_{R_X} x\, f_{X|Y}(x|y)\, dx\; f_Y(y)\, dy$$

$$= \int_{R_Y}\int_{R_X} x\, f_{X|Y}(x|y)\, f_Y(y)\, dx\, dy = \int_{R_X} x \underbrace{\int_{R_Y} f_{X|Y}(x|y)\, f_Y(y)\, dy}_{f_X(x)}\, dx$$

$$= \int_{R_X} x\, f_X(x)\, dx = E_X[x]$$

احتمال joint ای که y را sum out کنیم

⊛ به همین صورت اگر به جای $\int$ از $\sum$ استفاده کنیم برای متغیرهای تصادفی گسسته نیز اثبات می‌شود

1.3.2 (law of total variance)

$$E[X^2] \overset{\text{سوال قبل}}{=} E\left[E[X^2|Y]\right]$$

$$\overset{⊛}{=} E\left[\mathrm{var}[X|Y] + E^2[X|Y]\right]$$

$$= E\left[\mathrm{var}[X|Y]\right] + E\left[E^2[X|Y]\right]$$

⊛ از درس آمار، داریم که:

$$\mathrm{var}[X] = E[X^2] - E^2[X]$$

$$\rightarrow E[X^2] = \mathrm{var}[X] + E^2[X] \quad ⊛$$

$$\xrightarrow[-E^2[X]]{\text{طرفین}} \underbrace{E[X^2] - E^2[X]}_{\mathrm{var}[X]} = E\left[\mathrm{var}[X|Y]\right] + E\left[E^2[X|Y]\right] - \underbrace{E^2[X]}_{E^2[E[X|Y]]}$$

طبق سوال قبل

$$\longrightarrow \mathrm{var}[X] = E\left[\mathrm{var}[X|Y]\right] + \mathrm{var}\left[E[X|Y]\right]$$

پرسش ۴

X, Y are iid geometric(p) , $q=1-p$

$W = X - Y$

$Z = \min(X, Y)$

$P_{Z,W}(Z,W) = P_Z(Z) \cdot P_W(W)$ می خواهیم نشان دهیم ⊛

ابتدا $P_Z(Z)$, $P_W(W)$ را به دست می آوریم

$$P_W(w) = P(W=w) = P(X-Y=w) \xlongequal[\text{formula}]{\text{convolution}} \sum_{x=\max(0,w)}^{\infty} P(X=x, Y=x-w)$$

چون X ، Y هر دو geometric هستند باید داشته باشیم
- $x \geq 0$
- $x-w \geq 0 \rightarrow x \geq w$

$$\xrightarrow[X,Y]{\text{استقلال}} P_W(w) = \sum_{x=\max(0,w)}^{\infty} P(X=x)\,P(Y=x-w) = \sum_{x=\max(0,w)}^{\infty} (q^x p)(q^{x-w} p)$$

$$= p^2 q^w \sum_{x=\max(0,w)}^{\infty} q^{2(x-w)} = p^2 q^w \sum_{x=\max(-w,0)}^{\infty} q^{2x}$$

$$\text{حالت } 1 \xrightarrow{w \geq 0} p^2 q^w \underbrace{\cancel{\sum_{x=0}^{\infty} q^{2x}}}_{} \overset{\frac{1}{1-q^2}}{} = \frac{p^2 q^w}{1-q^2}$$

$$\text{حالت } 2 \xrightarrow{w < 0} p^2 q^w \underbrace{\cancel{\sum_{x=-w}^{\infty} q^{2x}}}_{\frac{q^{-2w}}{1-q^2}} = \frac{p^2 q^{-w}}{1-q^2}$$

$$\Bigg\} \longrightarrow \boxed{P_W(w) = \frac{p^2 q^{|w|}}{1-q^2} \quad w \in \text{اعداد صحیح}}$$

توزیع هندسی یعنی تعداد (k) سکه هایی که نیاز است بیاندازیم تا برنده شویم (با احتمال برد p برای هر پرتاب)

$Z=\min(X,Y)$ یعنی این آزمایش را 2 بار همزمان و مستقل انجام می دهیم و در هر مرحله هرکدام از X یا Y که برنده شود، برنده شده ایم

این معادل آن است که یک توزیع هندسی با احتمال باخت q^2 داشته باشیم زیرا در هر مرحله برای باختن باید هم X ببازد و هم Y $\left(q_Z = q^2\right)$

$Z \sim$ geometric $(1-q^2)$

$$P_Z(z) = q_Z^{\,z}\, p_Z = \left(q^{2z}\right) \underbrace{\left(1-(1-p)^2\right)}_{(2-p)p} = p q^{2z} (2-p) = \boxed{p q^{2z}(1+q) \quad z \in \{0,1,2,\dots\}}$$

$$\Longrightarrow P_W(w)\, P_Z(z) = \frac{p^2 q^{|w|}}{\cancel{(1-q)}\cancel{(1+q)}} \times \cancel{p} q^{2z} \cancel{(1+q)} = p^2 q^{2z+|w|} \quad \begin{matrix} w \in \text{اعداد صحیح} \\ z \in \{0,1,\dots\} \end{matrix}$$

⊛ حال نشان می دهیم که احتمال joint نیز به همین مقدار می رسد

$$P_{Z,W}(z,w) = P\left(\min(X,Y)=z,\ X-Y=w\right)$$

حالت 1 $\xrightarrow{X \geq Y}$ $P(Y=z\ ,\ X=w+z) \underset{X,Y}{\overset{\text{استقلال}}{=\!=}} P(Y=z)\,P(X=w+z)$

شروط $\rightarrow$ $w \geq 0$ ⊛ $,\ z \geq 0\ ,\ w+z \geq 0$

حالت 2 $\xrightarrow{X < Y}$ ⊛⊛ $P(X=z\ ,\ Y=z-w) \underset{X,Y}{\overset{\text{استقلال}}{=\!=}} P(X=z)\,P(Y=z-w)$

شروط $\rightarrow$ ⊛⊛ $w < 0\ ,\ z \geq 0\ ,\ z-w \geq 0$

$$\rightarrow P_{Z,W}(z,w) = \underbrace{(q^{z}p)(q^{w+z}p)}_{w \geq 0,\ z \geq 0}^{\text{حالت 1}} + \underbrace{(q^{z}p)(q^{z-w}p)}_{w<0,\ z \geq 0}^{\text{حالت 2}} = p^{2}q^{2z+|w|}$$

$w \in$ اعداد صحیح

$z \in \{0,1,\dots\}$

پرسش ۵

$$W \sim N(\overset{\mu}{22.8}, \overset{\sigma}{0.5})$$

1.5.1

$$P(W \leqslant 25) = \int_{-\infty}^{25} f_W(w)\,dw = F_W(25) \xlongequal{W \to Z} \Phi\left(\frac{25-22.8}{0.5}\right) = \Phi(4.4)$$

در جداول برای مقادیر بالای 4 تابع Φ را ذکر نمی‌کردند چون

احتمال آن بسیار بالا می‌شود ← از calculator استفاده کردم $\Phi(4.4) = 0.99999458$

1.5.2

$$P(25 \leqslant W \leqslant 26.5) = \int_{25}^{26.5} f_W(w)\,dw = F_W(26.5) - F_W(25) \xlongequal{W \to Z} \Phi(7.4) - \Phi(4.4)$$

با استفاده از calculator ← $\Phi(7.4) - \Phi(4.4) = 0.00000541$

1.5.3

W' ای می‌خواهیم که

$$P(W \geqslant W') = 0.75 \longrightarrow F_W(W') = 0.25 \xrightarrow{W \to Z} \Phi\left(\frac{W'-22.8}{0.5}\right) = 0.25$$

$$\Phi(-0.68) = 0.25$$ از روی جدول می‌دانیم

$$\left.\right\} \frac{W'-22.8}{0.5} = -0.68$$

$$\boxed{W' = 22.46\ \text{kg}}$$

جبر خطی

پرسش ۱

2.1.1

$$a^T x = [a_1, a_2, \dots, a_n]\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} = \sum_{i=1}^{n} a_i x_i \longrightarrow \begin{matrix}\frac{\partial a^T x}{\partial x_1}\\ \vdots \\ \frac{\partial a^T x}{\partial x_n}\end{matrix} = \begin{bmatrix}a_1\\\vdots\\a_n\end{bmatrix} = \frac{d a^T x}{d x}$$

2.1.2

$$x^T A x = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{ni} x_i \end{bmatrix}$$

$$\begin{aligned} &= x_1 (a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n) \\ &\quad + x_2 (a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n) \\ &\quad \vdots \\ &\quad + x_n (a_{n1} x_1 + a_{n2} x_2 + \cdots + a_{nn} x_n) \end{aligned}$$

$$\begin{matrix} \dfrac{\partial x^T A x}{\partial x_1} = \\ \dfrac{\partial x^T A x}{\partial x_2} = \\ \vdots \\ \dfrac{\partial x^T A x}{\partial x_n} = \end{matrix} \begin{bmatrix} 2a_{11} x_1 + (a_{12}+a_{21}) x_2 + (a_{13}+a_{31}) x_3 + \cdots + (a_{1n}+a_{n1}) x_n \\ (a_{21}+a_{12}) x_1 + 2a_{22} x_2 + (a_{23}+a_{32}) x_3 + \cdots + (a_{2n}+a_{n2}) x_n \\ \vdots \\ (a_{n1}+a_{1n}) x_1 + (a_{n2}+a_{2n}) x_2 + (a_{n3}+a_{3n}) x_3 + \cdots + 2a_{nn} x_n \end{bmatrix} = \frac{d x^T A x}{d x}$$

$$= \begin{bmatrix} 2a_{11} & a_{12}+a_{21} & \cdots & a_{1n}+a_{n1} \\ a_{21}+a_{12} & 2a_{22} & \cdots & a_{2n}+a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}+a_{1n} & a_{n2}+a_{2n} & \cdots & 2a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = (A + A^T) x$$

2.1.3

$$x^T A = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} \overset{f_1}{\sum_{i=1}^{n} a_{i1} x_i} & \cdots & \overset{f_n}{\sum_{i=1}^{n} a_{in} x_i} \end{bmatrix}$$

$$\frac{d\, x^T A}{d x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} = A^T$$

پرسش ۲

2.2.1

$$Av = \lambda v \longrightarrow (A - \lambda I)v = 0 \longrightarrow \det(A - \lambda I) = (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i) = 0$$

$$\xrightarrow[\lambda = 0]{\text{جاگذاری}} \det(A - 0) = \underbrace{(-1)^n \prod_{i=1}^{n} -\lambda_i}_{(-1)^n \prod_{i=1}^{n} \lambda_i} \longrightarrow \det(A) = \prod_{i=1}^{n} \lambda_i$$

2.2.2

می دانیم که $\det(\lambda I - A) = \prod_{i=1}^{n} (\lambda - \lambda_i)$. توسط دو روش ، ابتدا توسط سمت راست تساوی و سپس توسط سمت چپ ، ضریب λ^{n-1} که آن را c_{n-1} می نامیم را محاسبه می کنیم و آن مقادیر را برابر قرار می دهیم

$\boxed{\text{سمت راست تساوی } (p(\lambda))}$ برای اینکه λ به توان $n-1$ برسد ، n بار ، هر دفعه از یکی از $(\lambda - \lambda_i)$ ها باید λ_i اش انتخاب شده و از سایر عبارت ها λ باید انتخاب شود

$$(-\lambda_1 \lambda^{n-1}) + (-\lambda_2 \lambda^{n-1}) + \cdots + (-\lambda_n \lambda^{n-1}) = \left(-\sum_{i=1}^{n} \lambda_i\right) \lambda^{n-1}$$

$$\longrightarrow c_{n-1} = \boxed{-\sum_{i=1}^{n} \lambda_i \; *}$$

سمت چپ تساوی (ماتریس)

$$\begin{array}{c}\begin{bmatrix} \lambda - a_{11} & \cdots & -a_{1n} \\ \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & \lambda - a_{nn} \end{bmatrix} \\ \hspace{5em} q_{nn}\end{array}$$

یکی از روش های محاسبه دترمینان ماتریس Q آن است که به ازای تمام جایگشت های ممکن اعداد 1 تا n که آنها را $j_1, \dots, j_n$ می نامیم، المان های ماتریس در موقعیت های $q_{1j_1}, q_{2j_2}, \dots, q_{nj_n}$ را در هم ضرب کنیم و تمام $n!$ مقدار به دست آمده از هر جایگشت را با هم جمع/تفریق کنیم (اگر جایگشت فرد باشد تفریق و اگر زوج باشد جمع)

با توجه به اینکه λ ها تنها روی قطر اصلی ماتریس $\lambda I - A$ ظاهر شده اند، تنها یکی از جایگشت ها توانایی ایجاد λ^{n-1} را دارد و آن هم جایگشت $q_{11}, q_{22}, \dots, q_{nn}$ (قطر اصلی) است. زیرا اگر حتی یک موقعیت از جایی به غیر قطر اصلی انتخاب شود q_{ii} موجود در آن سطر و آن ستون غیرقابل انتخاب می شوند و درجه λ حداکثر λ^{n-2} خواهد بود.

بنابراین c_{n-1} تنها در $\prod_{i=1}^{n} q_{ii}$ خواهد بود

مشابه استدلال قبل داریم:

$$\left(-a_{11}\lambda^{n-1}\right) + \left(-a_{22}\lambda^{n-1}\right) + \dots + \left(-a_{nn}\lambda^{n-1}\right) = \left(-\sum_{i=1}^{n} a_{nn}\right)\lambda^{n-1}$$

$$\longrightarrow c_{n-1} = -\sum_{i=1}^{n} a_{nn} = \boxed{-\mathrm{trace}(A) \; (**)}$$

$$\xrightarrow{(*)\;(**)} \sum_{i=1}^{n} \lambda_i = \mathrm{trace}(A)$$

2.2.3

- می دانیم که برای هر ماتریس Q، $\det(Q) = \det(Q^T)$
- می دانیم که $\underbrace{(A-\lambda I)}_{Q}^{T} = A^T - \lambda I^T = \underbrace{A^T - \lambda I}_{Q^T}$

$$\det(A - \lambda I) = \det(A^T - \lambda I)$$

A و A^T معادله مشخصه های یکسان و eigen value های یکسان دارند

2.2.4

تعریف eigen vector آن است که بردار هایی (v) وجود دارند که اعمال ماتریس A بر روی آن ها مانند اسکیل کردنشان به اندازه λ است $(Av = \lambda v)$

بنابراین A^k به معنی k بار اعمال این اسکیل است $\left(A^k v = A^{k-1}\overset{\lambda v}{\cancel{(Av)}} = \dots = \lambda^k v\right)$

پس داریم که λ^k مقدار ویژه ماتریس A^k است

پرسش ۳

2.3.1

با استقرا ثابت می‌کنیم

حالت پایه ← $k=2$ برهان خلف ← فرض می‌کنیم $\lambda_1 \neq \lambda_2$ اما q_1 و q_2 وابسته خطی هستند

$$\alpha_1 q_1 + \alpha_2 q_2 = 0 \xrightarrow{\text{اعمال } A} \alpha_1 A q_1 + \alpha_2 A q_2 = A0 \longrightarrow \alpha_1 \lambda_1 q_1 + \alpha_2 \lambda_2 q_2 = 0 \quad \mathrm{I}$$

$$\alpha_1 q_1 + \alpha_2 q_2 = 0 \xrightarrow{\text{ضرب } \lambda_1 \text{ در طرفین}} \alpha_1 \lambda_1 q_1 + \alpha_2 \lambda_1 q_2 = 0 \quad \mathrm{II}$$

$$\xrightarrow{\mathrm{I}-\mathrm{II}} (\lambda_2 - \lambda_1)\alpha_2 q_2 = 0 \xrightarrow[q_2 \neq 0]{\lambda_1 \neq \lambda_2} \alpha_2 = 0 \xrightarrow[q_1 \neq 0]{} \alpha_1 = 0$$

بنابراین به تناقض رسیدیم و q_1 و q_2 مستقل خطی اند

$k \to k+1$ فرض می‌کنیم $q_1, \dots, q_k$ طبق شرایط ذکرشده مستقل خطی باشند. می‌خواهیم نشان دهیم q_{k+1} نیز از بقیه مستقل است

برهان خلف ← فرض می‌کنیم q_{k+1} به سایر q_i ها وابسته خطی است

$$\sum_{i=1}^{k} \alpha_i q_i = q_{k+1}$$

$$\xrightarrow{\text{اعمال } A} \sum_{i=1}^{k} \alpha_i A q_i = A q_{k+1} \longrightarrow \sum_{i=1}^{k} \alpha_i \lambda_i q_i = \lambda_{k+1} \underbrace{q_{k+1}}_{\sum_{i=1}^{k} \alpha_i q_i} = \sum_{i=1}^{k} \alpha_i \lambda_{k+1} q_i$$

$$\longrightarrow \sum_{i=1}^{k} (\lambda_i - \lambda_{k+1}) \alpha_i q_i = 0 \xrightarrow[q_i \neq 0]{\lambda_i \neq \lambda_{k+1}} \alpha_i = 0, \quad 1 \leq i \leq k$$

بنابراین به تناقض رسیدیم و

q_{k+1} نیز نسبت به بقیه بردارها مستقل خطی است

2.3.2

برای هر ماتریس حقیقی A ، بردار های q_i و q_j می‌دانیم $\quad \circledast \; Aq_i \cdot q_j = q_i \cdot A^T q_j$

$$\lambda_i(q_i \cdot q_j) = \lambda_i q_i \cdot q_j = Aq_i \cdot q_j \overset{\circledast}{=} q_i \cdot \overset{A}{A^T} q_j = q_i \cdot \lambda_j q_j = \lambda_j (q_i \cdot q_j)$$

$$\rightarrow \lambda_i (q_i \cdot q_j) = \lambda_j (q_i \cdot q_j) \rightarrow \left.\begin{array}{l} (\lambda_i - \lambda_j)(q_i \cdot q_j) = 0 \\ \lambda_i \neq \lambda_j \end{array}\right\} \rightarrow q_i \cdot q_j = 0$$

بردار های q_i و q_j بر هم عمود هستند

پرسش ۴

2.4.1

$$Av = \lambda v \to (A-\lambda I)v = 0 \to \det\left(\begin{bmatrix}-\lambda & 1\\ \frac{1}{2} & \frac{1}{2}-\lambda\end{bmatrix}\right) = 0 \to \lambda^2 - \frac{\lambda}{2} - \frac{1}{2} = 0$$

$$\to (\lambda-1)(\lambda+\tfrac{1}{2}) = 0 \to \lambda_1 = -\frac{1}{2} \;,\; \lambda_2 = 1$$

$$\lambda_1 = -\frac{1}{2} \to \begin{bmatrix}\frac{1}{2} & 1\\ \frac{1}{2} & 1\end{bmatrix} V_1 = 0 \xrightarrow{\text{elimination}} \begin{bmatrix}\frac{1}{2} & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1\\ V_2\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \to$$

تمامی بردار هایی که $\frac{1}{2}V_1 + V_2 = 0$ جواب هستند $\xrightarrow{\text{نماینده}} V_1 = \begin{bmatrix}-2\\ 1\end{bmatrix}$

$$\lambda_2 = 1 \to \begin{bmatrix}-1 & 1\\ \frac{1}{2} & -\frac{1}{2}\end{bmatrix} V_2 = 0 \xrightarrow{\text{elimination}} \begin{bmatrix}-1 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1'\\ V_2'\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix} \to$$

تمامی بردار هایی که $-V_1' + V_2' = 0$ جواب هستند $\xrightarrow{\text{نماینده}} V_2 = \begin{bmatrix}1\\ 1\end{bmatrix}$

2.4.2

$$A = \underbrace{P}_{\text{eigenvector matrix}} \; \underbrace{\Lambda}_{\text{diag}(\lambda_1,\lambda_2)} \; P^{-1} = \begin{bmatrix}-2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}-\frac{1}{2} & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix} \qquad P^{-1} \neq P^T$$

2.4.3

$$A^k = (P\Lambda \cancel{P^{-1}})(\cancel{P}\Lambda \cancel{P^{-1}}) \cdots (\cancel{P}\Lambda P^{-1}) = P\Lambda^k P^{-1} \xrightarrow[\text{است}]{\Lambda\ \text{قطری}} \begin{bmatrix}-2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}(-\frac{1}{2})^k & 0\\ 0 & 1^k\end{bmatrix}\begin{bmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix}$$

$$\lim_{k\to\infty} A^k = \begin{bmatrix}-2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}-\frac{1}{3} & \frac{1}{3}\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix} = \begin{bmatrix}-2 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}0 & 0\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix} = \begin{bmatrix}\frac{1}{3} & \frac{2}{3}\\ \frac{1}{3} & \frac{2}{3}\end{bmatrix}$$

پرسش ۵

$$A_{n\times m} = U\Sigma V^T$$

⊛ در singular value decomposition برای ماتریس حقیقی A داریم

- U,V ماتریس های orthogonal هستند ← $V^T=V^{-1},\ U^T=U^{-1}$ $\left(U_{n\times n}\ ,\ V_{m\times m}\right)$
- Σ ماتریس nxm قطری با اعداد نامنفی است

2.5.1

$$(A^TA)^{-1} = \left(\left(V\Sigma^T\cancel{U^T}\right)\left(\cancel{U}\Sigma V^T\right)\right)^{-1} = (V^T)^{-1}\left(\Sigma^T\Sigma\right)^{-1}V^{-1} = V\left(\Sigma^T\Sigma\right)^{-1}V^T$$

2.5.2

$$(A^TA)^{-1}A^T = \left(V\left(\Sigma^T\Sigma\right)^{-1}\cancel{V^T}\right)\left(\cancel{V}\Sigma^TU^T\right) = V\underbrace{\left(\Sigma^T\Sigma\right)^{-1}\Sigma^T}_{\text{pseudo inverse of }\Sigma}U^T$$

2.5.3

$$A\left(A^TA\right)^{-1} = \left(U\Sigma\cancel{V^T}\right)\left(\cancel{V}\left(\Sigma^T\Sigma\right)^{-1}V^T\right) = U\Sigma\left(\Sigma^T\Sigma\right)^{-1}V^T$$

2.5.4

$$A\left(A^TA\right)^{-1}A^T = \left(U\Sigma\left(\Sigma^T\Sigma\right)^{-1}\cancel{V^T}\right)\left(\cancel{V}\Sigma^TU^T\right) = U\Sigma\underbrace{\left(\Sigma^T\Sigma\right)^{-1}\Sigma^T}_{\text{pseudo inverse of }\Sigma}U^T$$
