שאלה 1

צ״ל להוכיח כי עבור מטריצה סימטרית A, $A = xx^t \iff A$ היא PSD.

$\Leftarrow$ נניח כי $A = xx^t$. אזי לכל וקטור v,

$$v^t A v = v^t x x^t v = (x^t v)^t (x^t v) = \| x^t v \|^2 \geq 0$$

ומכאן ש- A היא PSD ע״פ הגדרה.

$\Rightarrow$ נניח כי $v^t A v \geq 0$, $\forall v$. נסמן את הוקטורים העצמיים של A ב- $\{v_i\}_{i=1}^n$ ואת הע״ע המתאימים $\{\lambda_i\}_{i=1}^n$.

$$\forall 1 \leq i \leq n, \quad 0 \underset{\text{הנחה, עבור הוקטור } v_i}{\leq} v_i^t A v_i = v_i^t \lambda_i v_i = \lambda_i \cdot \| v_i \|^2$$

$$\Rightarrow \lambda_i \geq 0$$

* במקרה שהנורמה הינה 0, לא ניתן לחלק את הביטוי. אם כך, במקרה זה הו״ע הוא וקטור האפס (שכן הנורמה היא סכימה של ריבועים אי-שליליים) ומכאן גם הע״ע הינו 0, ובפרט ≥ 0.

$$\Rightarrow D = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix} =$$

$$= \hat{D} \cdot \hat{D} = \hat{D}^t \cdot \hat{D}$$

$$\Rightarrow A = QDQ^t = Q\hat{D}^t \hat{D} Q^t = (\hat{D}Q^t)^t (\hat{D}Q^t) \underset{x = \hat{D}Q^t}{=} x^t x$$

* לצורך הדיוק ניתן היה לסמן $y = x^t$ ולהגיע לאותה מסקנה עם נוסחה - $A = yy^t$.

שאלה 1

נתחיל מפישוט הביטוי ואז נוכל לגזור

$$x^t A x = x^t \begin{pmatrix} \sum_{j=1}^{n} A_{1j} x_j \\ \vdots \\ \sum_{j=1}^{n} A_{nj} x_j \end{pmatrix} = \sum_{i=1}^{n} x_i \cdot \left(\sum_{j=1}^{n} A_{ij} x_j \right) =$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} A_{ij} x_i x_j$$

ולסיום, נקבל את הנגזרת הדרושה:

$$\frac{\partial x^t A x}{\partial x_i} = \underbrace{1 \cdot \sum_{j=1}^{n} A_{ij} \cdot x_j}_{\text{הגורם היחיד שעבר מהסכימה הראשונה}} + \underbrace{1 \cdot \sum_{m=1}^{n} A_{mi} x_m}_{\text{הגורם היחיד שמתאים הגדרה כיוון שהוא גזור להיפוך סדר סכימה (ואותו מכן נגזרה)}} = \left((A + A^t) x \right)_i$$

וקיבלנו את הדרוש.

שאלה 2

א) $R(\lambda) = \ln(E[e^{\lambda x}])$. נגזור פעם ראשונה:

$$R(\lambda)' = \frac{\frac{\partial E[e^{\lambda x}]}{\partial \lambda}}{E[e^{\lambda x}]}$$

מתוך ההגדרה: $E[e^{\lambda x}] = \int_{-\infty}^{\infty} e^{\lambda x} f_x(x)\,dx$. הפונק' בתוך האינטגרל רציפה, וגזירה

ומאחר והיא חסומה בכל ת-ה, נוכל את גזירה והאינטגרציה להחליף (כמו שראינו בהרצאה) :

$$\frac{\partial E[e^{\lambda x}]}{\partial \lambda} = \frac{\partial}{\partial \lambda}\int_{-\infty}^{\infty} e^{\lambda x} f_x(x)\,dx =$$

$$= \int_{-\infty}^{\infty} \frac{\partial}{\partial \lambda} e^{\lambda x} f_x(x)\,dx =$$

$$= \int_{-\infty}^{\infty} x e^{\lambda x} f_x(x)\,dx = E[xe^{\lambda x}]$$

אם כן מהנגזרת הראשונה קיבלנו

$$R(\lambda)' = \frac{E[xe^{\lambda x}]}{E[e^{\lambda x}]}$$

נגזור את נגזרת שניה:

$$R(\lambda)'' = \frac{E[xe^{\lambda x}]' \cdot E[e^{\lambda x}] - E[e^{\lambda x}]' \cdot E[xe^{\lambda x}]}{E[e^{\lambda x}]^2}$$

הנגזרת של $E[xe^{\lambda x}]$ מחושבת – באותה דרך של $E[e^{\lambda x}]$ שחישבנו, רק הפעם מתקבל $E[xe^{\lambda x}]' = E[x^2 e^{\lambda x}]$.

$$\Rightarrow R(\lambda)'' = \frac{E[x^2 e^{\lambda x}]\, E[e^{\lambda x}] - (E[xe^{\lambda x}])^2}{E[e^{\lambda x}]^2} =$$

$$= \frac{E[x^2 e^{\lambda x}]}{E[e^{\lambda x}]} - \left(\frac{E[xe^{\lambda x}]}{E[e^{\lambda x}]}\right)^2$$

כפי שרצינו.

ב) לצורך חישוב השונות של ההתפלגות החדשה, נרצה לחשב תחילה את המומנט הראשון והשני. נעשה זאת:

$$E_g[X] = \int_{-\infty}^{\infty} x \cdot \frac{e^{\lambda x}}{E_f[e^{\lambda x}]} f_X(x)\,dx =$$

$$= \frac{1}{E_f[e^{\lambda x}]} \cdot \int_{-\infty}^{\infty} x\, e^{\lambda x} f_X(x)\,dx = \frac{E[xe^{\lambda x}]}{E[e^{\lambda x}]}$$

$$E_g[X^2] = \int_{-\infty}^{\infty} x^2 \cdot \frac{e^{\lambda x}}{E_f[e^{\lambda x}]} f_X(x)\,dx =$$

$$= \frac{1}{E_f[e^{\lambda x}]} \cdot \int_{-\infty}^{\infty} x^2 \cdot e^{\lambda x} \cdot f_X(x)\,dx = \frac{E[x^2e^{\lambda x}]}{E[e^{\lambda x}]}$$

ומנוסחת שונות, כפי שכתבנו כי $Var_g(X) = E_g[X^2] - E_g[X]^2 =$

$$= R(\lambda)''$$

נותר להוכיח את החסם הבא: $Var_g(X) \le B^2$

כיוון שהתומך של X הוא $[-B, B]$, ניתן להסיק ש $E_g[X^2] \le B^2$

$$\Rightarrow Var_g(X) = E_g[X^2] - E_g[X]^2 \le E_g[X^2] \le B^2$$

כנדרש.

ג)

$* \ R(\lambda)'' \leq B^2$

$* \ R(0) = \ln(\mathbb{E}[e^{0\cdot x}]) = \ln 1 = 0$

$* \ R'(0) = \dfrac{\mathbb{E}[xe^{0x}]}{\mathbb{E}[e^{0\cdot x}]} = \mathbb{E}[x] = 0$

לפי משפט היסודי של החדו"א ; $\int_a^b f(x)\,dx = F(b) - F(a)$.

$$\Rightarrow \int_0^\lambda \int_0^t R''(s)\,ds\,dt = \int_0^\lambda R'(t) - R'(0)\,dt = R(\lambda) - R(0) = R(\lambda)$$

$$; \int_0^\lambda \int_0^t R''(s)\,ds\,dt \leq \int_0^\lambda \int_0^t B^2\,ds\,dt = \int_0^\lambda t \cdot B^2\,dt = B^2 \frac{\lambda^2}{2}$$

$$\Rightarrow \ln(\mathbb{E}[e^{\lambda x}]) \leq B^2 \frac{\lambda^2}{2} \Rightarrow \mathbb{E}[e^{\lambda x}] \leq e^{B^2 \cdot \frac{\lambda^2}{2}}, \ \forall \lambda \in \mathbb{R}$$

כנדרש.

שאלה 1

הגדרנו את $h = \underset{f: X \to Y}{\arg\min}\ \mathbb{E}[l_{0-1}(f(X), Y)]$.

נרצה להוכיח כי $h(x) = \underset{i \in Y = \{1, \dots, L\}}{\arg\max}\ \mathbb{P}(Y = i \mid X = x)$

נקודה וסגורה כלשהי, נעבוד בהתניה כי $X = \bar{x}$

$$\Rightarrow L(h) = \mathbb{E}[l_{0-1}(h(\bar{x}), Y)] = \sum_{i=1}^{L} l_{0-1}(h(\bar{x}), i) \cdot \mathbb{P}(X = \bar{x}, Y = i)$$

מעתה h קבוע ב Y.

- $\mathbb{P}(X = \bar{x}, Y = i) = \mathbb{P}(X = \bar{x}) \cdot \mathbb{P}(Y = i \mid X = \bar{x})$

$$\Rightarrow L(h) = \mathbb{P}(X = \bar{x}) \cdot \sum_{i=1}^{L} l_{0-1}(h(\bar{x}), i) \cdot \mathbb{P}(Y = i \mid X = \bar{x}) =$$

$$= \left\{ \mathbb{P}(X = \bar{x}) \cdot \sum_{\substack{i=1 \\ i \neq k}}^{L} \mathbb{P}(Y = i \mid X = \bar{x}) \qquad h(\bar{x}) = k,\ k \in Y = \{1, \dots, L\} \right.$$

$$l_{0-1}(h(\bar{x}), i) = \begin{cases} 1 & h(\bar{x}) \neq i \\ 0 & h(\bar{x}) = i \end{cases}$$

$$= \left\{ \mathbb{P}(X = \bar{x}) \cdot (1 - \mathbb{P}(Y = k \mid X = \bar{x})) \qquad h(\bar{x}) = k \right.$$

לכן הערך המינימלי של $L(h)$ יתקבל כאשר

$$h(\bar{x}) = i = \underset{i \in Y}{\arg\max}\ \mathbb{P}(Y = i \mid X = \bar{x})$$

כפי שרצינו.

שאלה 2

א) ״לא להגשה״..

ב) נחשב את הפונק' :

$$f_0(x,\mu_0,\Sigma) = \frac{1}{2\pi\cdot\sqrt{2}}\cdot e^{-\frac{1}{2}x^t\Sigma^{-1}x} =$$

$$= \frac{1}{2\pi\cdot\sqrt{2}}\cdot e^{-\frac{1}{2}(x_1,x_2)\cdot\begin{pmatrix}1 & 0\\ 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}} =$$

$$= \frac{1}{2\pi\cdot\sqrt{2}}\cdot e^{-\frac{1}{2}(x_1^2+\frac{1}{2}x_2^2)}$$

$$f_1(x,\mu_1,\Sigma) = \frac{1}{2\pi\cdot\sqrt{2}}\cdot e^{-\frac{1}{2}(x_1-1,x_2-1)\begin{pmatrix}1 & 0\\ 0 & \frac{1}{2}\end{pmatrix}\begin{pmatrix}x_1-1\\ x_2-1\end{pmatrix}} =$$

$$= \frac{1}{2\pi\cdot\sqrt{2}}\cdot e^{-\frac{1}{2}((x_1-1)^2+\frac{1}{2}(x_2-1)^2)}$$

לסיום, הכנסתי את הפונק' הזו לקוד היוצר את הגרף הבא.

ג)

$$\mathbb{P}(y=1 \mid x) > \mathbb{P}(y=0 \mid x)$$

נוסחת בייס $\iff \dfrac{f_{x|y=1}(x)\,\mathbb{P}(y=1)}{f_x(x)} > \dfrac{f_{x|y=0}(x)\cdot\mathbb{P}(y=0)}{f_x(x)}$

$\mathbb{P}(y=1)=p$, $y \in \{0,1\}$ $\iff p \cdot f_1(x, \mu_1, \underset{\Sigma}{\overset{}{\Sigma_1}}) > (1-p)\cdot f_0(x, \mu_0, \Sigma_0)$, $\Sigma_1 = \Sigma$, $\Sigma_0 = \Sigma$

$$\iff \frac{1-p}{p} < \frac{e^{-\frac{1}{2}(x-\mu_1)^t \Sigma^{-1}(x-\mu_1)}}{e^{-\frac{1}{2}(x-\mu_0)^t \Sigma^{-1}(x-\mu_0)}}$$

$$\iff \frac{1-p}{p} < e^{-\frac{1}{2}(x-\mu_1)^t \Sigma^{-1}(x-\mu_1) + \frac{1}{2}(x-\mu_0)^t \Sigma^{-1}(x-\mu_0)}$$

$$\iff 2\ln\left(\frac{1-p}{p}\right) < (x-\mu_0)^t \Sigma^{-1}(x-\mu_0) - (x-\mu_1)^t \Sigma^{-1}(x-\mu_1)$$

ננסה לפתח את הביטוי שבאגף ימין עוד יותר. את התוצאה נוכל לסמן, אם נקבל:

$$\iff 2\ln\left(\frac{1-p}{p}\right) < (\mu_1^t - \mu_0^t)\Sigma^{-1}(2x - \mu_0 - \mu_1)$$

וקיבלנו בייזני מסווג ליניארי.

ב)

ראינו בשאלה כי עבור $d=1$ נקבל נקודה. ניתן אינטואיטיבית — להבין כי עבור $d=2$ נקבל ישר.
פורמלית, נדרוש שוויון משוו- השוויון מהסעיף הקודם :

$$2\ln\left(\frac{1-p}{p}\right) = (\mu_1^t - \mu_0^t)\,\Sigma^{-1}\,(2x - \mu_0 - \mu_1)$$

$$\Leftrightarrow \underbrace{2(\mu_1^t - \mu_0^t)\,\Sigma^{-1}}_{\text{וקטור}}\,x = \underbrace{2\ln\left(\frac{1-p}{p}\right) + (\mu_1^t - \mu_0^t)\,\Sigma^{-1}\cdot(\mu_0 + \mu_1)}_{\text{סקלר}}$$

כלומר הפתרון הוא מישור d - ממדי.

ה) מסעיף הקודם קיבלנו משוואה. נציב את הנתונים הנתונים Σ, μ_1, μ_0 :

$$(2,2)\cdot\begin{pmatrix}1 & 0\\ 0 & 1/2\end{pmatrix} x = 2\ln\left(\frac{1-p}{p}\right) + (1,1)\begin{pmatrix}1 & 0\\ 0 & 1/2\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}$$

$$\Rightarrow (2,1)\,x = 2\ln\left(\frac{1-p}{p}\right) = \left(1,\tfrac{1}{2}\right)\begin{pmatrix}1\\1\end{pmatrix}$$

$$\Rightarrow 2x_1 + x_2 = 2\ln\left(\frac{1-p}{p}\right) + \frac{3}{2}$$

נוסיף את הפתרון לגרף עבור הערכים $p=0.25, 0.5, 0.99$.

$$\frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma_1^2}}}{\frac{1}{\sqrt{2\pi\sigma_0^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma_0^2}}} = \frac{1-p}{p}$$ (1

$$\Leftrightarrow \quad \frac{1-p}{p} \cdot \frac{\sigma_0}{\sigma_1} = e^{\frac{(x-\mu)^2}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)}$$

$$\Leftrightarrow \quad \ln\left(\frac{1-p}{p} \cdot \frac{\sigma_0}{\sigma_1}\right) = \frac{(x-\mu)^2}{2}\left(\frac{1}{\sigma_0^2} - \frac{1}{\sigma_1^2}\right)$$

$$\vdots$$

$$\Leftrightarrow \quad x = \mu \pm \sigma_0 \sigma_1 \sqrt{\frac{\ln\left(\frac{1-p}{p} \cdot \frac{\sigma_0}{\sigma_1}\right)}{\sigma_1^2 - \sigma_0^2}}$$

נקודות של כתובי.

שאלה 3

נתון $S = X_1 + \dots + X_n$ ואנו רוצים למצוא $N \in \mathbb{N}$ כך שלכל $n \geq N$,

$$\mathbb{P}(S > n^2 + 0.2n) < 0.1$$

$$* \; \mathbb{E}[X_i] = \frac{-3+5}{2} = 1 \;, \quad \forall\, 1 \leq i \leq n$$

כפי שהגדרנו בכיתה, נגדיר $\bar{X} = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

$$* \; \mathbb{E}[\bar{X}] = \frac{1}{n} \cdot n = 1$$

$$\Rightarrow \underset{\text{הופדינג}}{\uparrow} \quad \mathbb{P}(|\bar{X} - 1| \geq \varepsilon) \leq 2e^{-\frac{\varepsilon^2}{2(5+3)^2} \cdot n}$$

כעת,

$$\mathbb{P}(S > n^2 + 0.2n) = \mathbb{P}\left(\frac{S}{n} > n + 0.2\right) =$$

$$= \mathbb{P}(\bar{X} - 1 > n - 0.8) \leq$$

מונוטוניות פונק' ההסתברות $\rightarrow$ $\leq \mathbb{P}(|\bar{X} - 1| > n - 0.8) \leq$

הופדינג $\rightarrow$ $\leq 2e^{-\frac{n(n-0.8)^2}{2 \cdot 8^2}} := \alpha$

נרצה לדרוש $\alpha < 0.1$.

$$\Rightarrow -\frac{n(n-0.8)^2}{2 \cdot 8^2} < \ln(0.05)$$

$$\Rightarrow n(n-0.8)^2 > -2 \cdot 8^2 \cdot \ln(0.05) \approx 383.45$$

אז לפי משוואה זאת, מבדיקה ידנית – פשוטה נמצא כי ה-N הרצוי הוא

$$N = 8$$

וקיבלנו את הרצוי.