תרגיל 1

שאלה 1

סעיף A

נניח את המקרה הרציף והמקרה הבדיד זהה פשוט צריך להחליף את האינטגרל בסכום. בנוסף בעזרת אנליזה מודרנית ניתן להתייחס לסכום כאינטגרל עם פונקציית מידה מתאימה.

$$
\begin{aligned}
&\mathbb{E}[XY] \\
&= \int_{\mathbb{R}} xP(XY = x)\, dx \\
&= \int_{\mathbb{R}} x \int_{\mathbb{R}} P(X = \frac{x}{y}, Y = y)\, dydx \\
&= \int_{\mathbb{R}} x \int_{\mathbb{R}} P(X = \frac{x}{y})P(Y = y)\, dydx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} xP(X = \frac{x}{y})P(Y = y)\, dydx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} xyP(X = x)P(Y = y)\, dydx \\
&= \int_{\mathbb{R}} xP(X = x)\, dx \int_{\mathbb{R}} yP(Y = y)\, dy \\
&= \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}
$$

כדרוש

סעיף B

נגדיר $X \sim U(-1,1)$ ו$Y = X^2$, ברור שהם תלויים אבל,

$$\mathbb{E}[XY] = \mathbb{E}[X^3] = 0 = 0 \cdot \mathbb{E}[Y] = \mathbb{E}[X]\mathbb{E}[Y]$$

כדרוש.

סעיף C

נגדיר $X, Y \sim U\{0,1\}$ בת"ל ו$Z = 1_{X \neq Y}$ כלומר 0 אם הם שווים ו1 אחרת. ברור שכל שתיים נותנים את השלישי (לדוגמה בעזרת חיבור מודולו 2). ובנוסף

$$P(X = 0, Z = 0) = P(X = 0, Y = 0) = P(X = 0)P(Y = 0) = \frac{1}{4}$$

$$P(X = 1, Z = 0) = P(X = 1, Y = 1) = P(X = 1)P(Y = 1) = \frac{1}{4}$$

$$P(X = 0, Z = 1) = P(X = 0, Y = 1) = P(X = 0)P(Y = 1) = \frac{1}{4}$$

$$P(X = 1, Z = 1) = P(X = 1, Y = 0) = P(X = 1)P(Y = 0) = \frac{1}{4}$$

נשים לב שגם $Z \sim U\{0,1\}$ ולכן

$$P(X = x)P(Z = z) = \frac{1}{4}$$

כלומר X, Z בת"ל ולכן בה"כ גם Y, Z בת"ל (שכן X, Y סימטריים בהגדרה)
לכן קיבלנו את הדרוש.

סעיף D

עבור, x_i כמתואר ו$z \in \mathbb{R}^d$

$$z^TCz = z^T\left(\sum_{i=1}^{n} x_i x_i^T\right)z = \sum_{i=1}^{n}(z^Tx_i)(z^Tx_i)^T = \sum_{i=1}^{n}||z^Tx_i||^2 \geq 0$$

שכן סכום של אי-שליליים הוא גם אי-שלילי.

שאלה 2

סעיף A

$$P(-1 \leq X \leq 1) = 1 - 2P(X > 1) = 1 - \Phi(1) = 1 - 2 \cdot 0.159 = 0.683$$

(חושב בעזרת טבלת Z)

סעיף B

$$P(||X|| \leq 1) \leq P(\forall i : |x_i| \leq 1) = \prod_{i=1}^{d} P(|x_i| \leq 1) = 0.683^d$$

שכמובן שואף לאפס כש$d \to \infty$ ולכן גם

$$\lim_{d\to\infty} P(||X|| \leq 1) = 0$$

שאלה 3

$$Var(y) = \mathbb{E}[(y - \mathbb{E}[y])^2] = \mathbb{E}[y^2] - \mathbb{E}[y]^2$$

נחשב את ערכים אלו בנפרד ונחבר

$$\mathbb{E}[y^2] = \mathbb{E}[(w^TX)(w^TX)^T] = \mathbb{E}[w^TXX^Tw] = w^T\mathbb{E}[XX^T]w$$
$$\mathbb{E}[y]^2 = \mathbb{E}[w^TX]^2 = (w^T\mathbb{E}[X])^2 = (w^T\mathbb{E}[X])(w^T\mathbb{E}[X])^T = w^T\mathbb{E}[X]\mathbb{E}[X]^Tw$$

קיבלנו ש

$$Var(y) = \mathbb{E}[y^2] - \mathbb{E}[y]^2 = w^T(\mathbb{E}[XX^T] - \mathbb{E}[X]\mathbb{E}[X]^T)w = w^TCw$$

שכן

$$C = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T] = \mathbb{E}[XX^T] - \mathbb{E}[X]\mathbb{E}[X]^T$$

שאלה 4

סעיף A

ראשית, נוכיח כי $specA \subseteq specA^T$.
יהי λ ע"ע של A, אזי $\det(A - \lambda I) = 0$, שכן קיים $v \neq 0$ עבורו $(A - \lambda I)v = 0$ כלומר הגרעין לא טריוויאלי ולכן הדטרמיננטה מתאפסת.
אבל

$$\det(A^T - \lambda I) = \det((A - \lambda I)^T) = \det(A - \lambda I) = 0$$

ולכן הגרעין של $A^T - \lambda I$ לא טריוויאלי, כלומר קיים $v \neq 0$ כך ש

$$(A^T - \lambda I)v = 0 \implies A^Tv = \lambda v$$

כלומר v ו"ע של A^T עם ע"ע λ.
לכן אכן קיבלנו ש $specA \subseteq specA^T$ ובגלל ש$A = (A^T)^T$ נקבל ש $specA^T \subseteq specA$ כלומר הם שווים.

סעיף B

נשים לב ש

$$\det(A) = \det(P\Lambda P^{-1}) = \det(P)\det(\Lambda)\det(P^{-1}) = \det(P)\det(\Lambda)\det(P)^{-1} = \det(\Lambda) = \prod_{i=1}^{n}\lambda_i$$

כאשר המעבר האחרון נובע מדטרמיננטה של מטריצה אלכסונית.

שאלה 5

סעיף A

$$f(x) = \begin{cases} 0 & w^T x \geq 1 \\ 1 - w^T x & o.w. \end{cases}$$

נשים לב שזו הרכבה של שתי פונקציות

$$f = (x \mapsto \max(1-x, 0)) \circ (x \mapsto w^T x)$$

לכן אם f גזירה בx אזי בפרט

$$\nabla f(x) = g(w^T x) w$$

כשg מוגדר על $x \neq 1$ ע"י

$$g(x) = \begin{cases} 0 & x > 1 \\ -1 & x < 1 \end{cases}$$

ולכן $\nabla f(x)$ מוגדר עבור $w^T x \neq 1$ כלומר f גזיר ב

$$\{x \in \mathbb{R}^d \mid w^T x \neq 1\}$$

ושם

$$\nabla f(x) = \begin{cases} 0 & w^T x > 1 \\ -w & w^T x < 1 \end{cases}$$

סעיף B

$$f(x) = \frac{1}{1 + e^{-w^T x}}$$

נשים לב שזו הרכבה של שתי פונקציות

$$f = \left(x \mapsto \frac{1}{1 + e^{-x}}\right) \circ (x \mapsto w^T x)$$

לכן אם f גזירה בx אזי בפרט

$$\nabla f(x) = \frac{e^{-w^T x}}{\left(1 + e^{-w^T x}\right)^2} w$$

ונשים לב ש

$$\frac{e^{-w^T x}}{\left(1 + e^{-w^T x}\right)^2} = f(x)(1 - f(x))$$

ולכן

$$\nabla f(x) = f(x)(1 - f(x)) w$$