פתרון מבחן מסכם 67653, מועד א'

שיטות סטטיסטיות בהנדסה ומדעי המחשב
מרצה - עמי ויזל

7.7.23

שאלה 1 - 34 נקודות

יהיו Y, X משתנים מקריים יוניפורמיים. Y מתפלג יוניפורמית בין אפס ל-A דטרמיניסטי ידוע $Y \sim Uni(0, A)$, ואילו X מתפלג בין $-Y$ לבין Y - כלומר $X|Y \sim Unif(-Y, Y)$.

1. חשבו משערך MMSE ל-X בהנתן $Y = y$
2. חשבו משערך LMMSE ל-X בהנתן $Y = y$
3. חשבו משערך MMSE ל-Y בהנתן $X = x$
4. חשבו משערך LMMSE ל-Y בהנתן $X = x$

לשימושכם:
משתנה מקרי $Z \sim Uni(a, b)$:

$$\mathbb{E}[Z] = \frac{a+b}{2}, \ \mathbb{V}ar(Z) = \frac{(b-a)^2}{12}$$

משפט השונות השלמה: $\mathbb{V}ar(X) = \mathbb{E}_Y[\mathbb{V}ar(X|Y)] + \mathbb{V}ar(\mathbb{E}_X[X|Y])$

פתרון

1. על מנת לחשב משערך MMSE ל-X בהנתן $Y = y$, נתחיל מלבדוק את ההתפלגות המותנית, ונבחר את התוחלת שלה כמשערך:

$$X|Y = y \sim Uni(-y, y)$$

$$\hat{X}_{MMSE}(y) = \mathbb{E}[X|Y = y] = \frac{-y+y}{2} = 0$$

מכיוון ש-$X|Y = y$ מתפלג יוניפורמי בתחום סימטרי סביב האפס התוחלת שלו היא אפס.

2. המשערך האופטימלי הוא לינארי, ולכן הוא גם המשערך הלינארי-אופטימלי

$$\hat{X}_{LMMSE}(y) = \hat{X}_{MMSE}(y) = 0$$

3. על מנת לחשב משערך MMSE ל-Y בהנתן $X = x$ נתחיל מלבדוק את ההתפלגות המותנית. בבדיקת תחומים, מתחייב ש $|x| \leq Y$ אחרת אי אפשר היה לקבל $X = x$. כמו כן הערך של Y חסום ע"י A, ולכן ההסתברות המותנית שונה מאפס רק עבור $|x| \leq Y \leq A$:

$$f(Y|X=x) = \frac{f(X=x|Y)f(Y)}{f(X=x)} = \frac{\frac{1}{2Y}\frac{1}{A}}{\int_{|x|}^{A}\frac{1}{2\hat{y}}\frac{1}{A}d\hat{y}}$$

$$\int_{|x|}^{A}\frac{1}{2\hat{y}}\frac{1}{A}d\hat{y} = \frac{1}{2A}ln(\hat{y})|_{|x|}^{A} = \frac{ln(A) - ln(|x|)}{2A}$$

$$f(Y|X=x) = \frac{2A}{2Y \cdot A \cdot (ln(A) - ln(|x|))} = \frac{1}{Y(ln(A) - ln(|x|))}$$

כאשר הנסוחא נכונה רק בתחום האמור. נעבור לחישוב התוחלת:

$$\begin{aligned}\hat{Y}_{MMSE}(x) = \mathbb{E}\left[Y|X=x\right] &= \int_{-\infty}^{\infty} yf(y|X=x)dy \\ &= \int_{|x|}^{A} y\frac{1}{y(ln(A) - ln(|x|))}dy \\ &= \frac{1}{ln(A) - ln(|x|)}y|_{|x|}^{A} = \frac{A - |x|}{ln(A) - ln(|x|)}\end{aligned}$$

4. על מנת לחשב משערך LMMSE ל-Y בהנתן $X = x$ נשתמש בנוסחא לתוחלת מותנית גאוסית שהיא המשערך הלינארי האופטימלי. נתחיל מלחשב את מרכיבי הנוסחא, ונשתמש בנוסחא לשונות של משתנה מקרי יוניפורמי:

$$\mu_Y = \frac{0 + A}{2} = \frac{A}{2}$$

$$\mu_X = \mathbb{E}_Y\left[\mathbb{E}\left[X|Y\right]\right] = \mathbb{E}_Y\left[0\right] = 0$$

$$\mathbb{V}ar\left(X|Y\right) = \frac{(b-a)^2}{12} = \frac{(Y - (-Y))^2}{12} = \frac{4Y^2}{12} = \frac{Y^2}{3}$$

$$\begin{aligned}\mathbb{V}ar\left(X\right) &= \mathbb{E}_Y\left[\mathbb{V}ar\left(X|Y\right)\right] + \mathbb{V}ar\left(\mathbb{E}_X\left[X|Y\right]\right) \\ &= \mathbb{E}_Y\left[\frac{Y^2}{3}\right] + \mathbb{V}ar\left(0\right) \\ &= \int_0^A \frac{y^2}{3}\frac{1}{A}dy = \frac{y^3}{9A}|_0^A = \frac{A^2}{9}\end{aligned}$$

$$\begin{aligned}\mathbb{C}ov\left(X,Y\right) &= \mathbb{E}\left[XY\right] - \mu_X\mu_Y \\ &= \mathbb{E}_Y\left[\mathbb{E}\left[XY|Y\right]\right] - 0 = \mathbb{E}_Y\left[0\right] = 0\end{aligned}$$

$$\hat{Y}_{LMMSE}(x) = \mu_Y - \frac{\mathbb{C}ov\left(X,Y\right)}{\mathbb{V}ar\left(X\right)}(x - \mu_x) = \frac{A}{2} - 0$$

אתם (ואנחנו) היינו יכולים לחסוך את העבודה המפורטת אם היינו מתחילים מחישוב הקורלציה.

שאלה 2 - 31 נקודות

יהי $w[n] \sim_{i.i.d.} N(0,1)$, ונתון תהליך אקראי ממוצע נע $x[n] = \alpha w[n-1] + w[n]$ כאשר α דטרמיניסטי לא ידוע.

1. האם התהליך $x[n]$ סטציונארי במובן הצר או הרחב? הוכיחו.
2. מהי הנראות של α בהנתן זוג דגימות סמוכות של התהליך $x[1], x[2]$? כתבו ביטוי מפורט
3. בהנתן זוג דוגמאות סמוכות $\{x[1], x[2]\}$, כתבו מבחן יחס נראות לאפשרות $H_0 : \alpha = 0$ לעומת $H_1 : \alpha = 1$, וכתבו את תוחלת המבחן תחת שתי ההיפותזות.
4. בהנתן דוגמאות מזמנים זוגיים בתהליך $x[0], x[2], ..., x[2k]$, מה היא פונקציית הנראות של α?
5. בהנתן אוסף דוגמאות מזמנים זוגיים בתהליך, מצאו משערך ML ל-α^2. שימו לב כי המשערך צריך להיות אי-שלילי, אך בשאלה זו ניתן להתעלם מכך.
6. חשבו את הטיית המשערך

פתרון

1. התהליך $x[n]$ בנוי מקומבינציות לינאריות דטרמיניסטיות של תהליך גאוסי, ולכן הוא בעצמו תהליך גאוסי. נבדוק את התוחלת והאוטוקורלציה של התהליך:

$$\begin{aligned} m_x[n] &= \mathbb{E}\left[x[n]\right] = \mathbb{E}\left[\alpha w[n-1] + w[n]\right] \\ &= \alpha \mathbb{E}\left[w[n-1]\right] + \mathbb{E}\left[w[n]\right] = 0 \end{aligned}$$

$$\begin{aligned} R_x(n, n+k) &= \mathbb{E}\left[x[n]x[n+k]\right] \\ &= \mathbb{E}\left[(\alpha w[n-1] + w[n])(\alpha w[n+k-1] + w[n+k])\right] \\ &= \alpha^2 \mathbb{E}\left[w[n-1]w[n+k-1]\right] + \alpha \mathbb{E}\left[w[n]w[n+k-1]\right] \\ &\quad + \alpha \mathbb{E}\left[w[n-1]w[n+k]\right] + \mathbb{E}\left[w[n]w[n+k]\right] \\ &= \alpha^2 R_w(k) + \alpha R_w(k+1) + \alpha R_w(k-1) + R_w(k) \end{aligned}$$

$$R_x(k) = \begin{cases} 1 + \alpha^2 & k = 0 \\ \alpha & k = 1 \\ 0 & else \end{cases}$$

התוחלת קבועה בזמן והאוטוקורלציה תלויה רק בהפרש, ולכן התהליך סמ"ר. מכיוון שזהו תהליך גאוסי וסמ"ר, הוא גם סמ"צ.

2. התהליך גאוסי, ומכיוון שהוא ממורכז האוטוקורלציה שווה לאוטוקוואריאנס, ונותר

לנו רק להציב:

$$\vec{x} = \begin{pmatrix} x[i] \\ x[i-1] \end{pmatrix} \sim N\left(\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1+\alpha^2 & \alpha \\ \alpha & 1+\alpha^2 \end{pmatrix}\right)$$

$$L(\alpha; \vec{x}) = \frac{exp(-\frac{1}{2}\vec{x}^T\Sigma^{-1}\vec{x})}{\sqrt{(2\pi)^d|\Sigma|}}$$

נפתח את הלייקליהוד המלא:

$$det(\Sigma) = (1+\alpha^2)^2 - \alpha^2 = 1 + \alpha^2 + \alpha^4$$

$$\Sigma^{-1} = \begin{pmatrix} 1+\alpha^2 & \alpha \\ \alpha & 1+\alpha^2 \end{pmatrix}^{-1} = \frac{1}{det(\Sigma)} \begin{pmatrix} 1+\alpha^2 & -\alpha \\ -\alpha & 1+\alpha^2 \end{pmatrix}$$

$$\begin{aligned} \vec{x}^T\Sigma^{-1}\vec{x} &= \begin{pmatrix} x[i] & x[i-1] \end{pmatrix} \frac{1}{det(\Sigma)} \begin{pmatrix} 1+\alpha^2 & -\alpha \\ -\alpha & 1+\alpha^2 \end{pmatrix} \begin{pmatrix} x[i] \\ x[i-1] \end{pmatrix} \\ &= \frac{(1+\alpha^2)(x[i]^2 + x[i-1]^2) - 2\alpha(x[i]x[i-1])}{det(\Sigma)} \end{aligned}$$

$$L(\alpha; \vec{x}) = \frac{1}{\sqrt{(2\pi)^2(1+\alpha^2+\alpha^4)}} exp\left(-\frac{(1+\alpha^2)(x[i]^2 + x[i-1]^2) - 2\alpha(x[i]x[i-1])}{2(1+\alpha^2+\alpha^4)}\right)$$

3. נכתוב את מבחן יחס הנראות:

$$T(\vec{x}) := \frac{L(1; \vec{x})}{L(0; \vec{x})} \gtrless \gamma$$

$$L(1; \vec{x}) = \frac{1}{\sqrt{(2\pi)^2(1+1+1)}} exp\left(-\frac{(1+1)(x[i]^2 + x[i-1]^2) - 2(x[i]x[i-1])}{2(1+1+1)}\right)$$

$$L(0; \vec{x}) = \frac{1}{\sqrt{(2\pi)^2(1+0)}} exp\left(-\frac{(1+0)(x[i]^2 + x[i-1]^2) + 0x[i]x[i-1]}{2(1+0)}\right)$$

$$T(\vec{x}) := \frac{\sqrt{(2\pi)^2}}{\sqrt{(2\pi)^2 3}} exp\left(-\frac{2(x[i]^2 + x[i-1]^2) - 2x[i]x[i-1]}{6} + \frac{(x[i]^2 + x[i-1]^2)}{2}\right) \gtrless \gamma$$

נחלק בקבועים, נוציא לוג ונכפיל בשש:

$$-2x[i]^2 - 2x[i-1]^2 + 2x[i]x[i-1] + 3x[i]^2 + 3x[i-1]^2) = x[i]^2 + 2x[i]x[i-1] + x[i-1]^2$$
$$T(\vec{x}) = (x[i] + x[i-1])^2 \gtrless \gamma'$$

תוחלת המבחן תחת ההשערות השונות:

$$\mathbb{E}\left[T(\vec{x})\right] = \mathbb{E}\left[(x[i] + x[i-1])^2\right] = R_x(0) + 2R_x(1) + R_x(0) = \begin{cases} 1+0+1 = 2 & H_0 \\ 2 + 2\cdot 1 + 2 = 6 & H_1 \end{cases}$$

4. מכיוון שהתהליך הוא גאוסי ושדוגמאות במרחק 2 הן חסרות קורלציה, אנחנו מבינים שהן גם בלתי תלויות סטטיסטית. הנראות של הדוגמאות היא לכן:

$$L(\alpha;\vec{x}) = f_x(x[0],x[2],...,x[2k]) = \prod_{i=0}^{k} f_x(x[2i])$$
$$= \frac{1}{\sqrt{2\pi(1+\alpha^2)}^{k+1}} exp\left(-\frac{\sum_{i=0}^{k} x[i]^2}{2(1+\alpha^2)}\right)$$

5. כדי למצוא משערך ML ל-α^2 נוציא לוג ללייקליהוד, נגזור ונשווה לאפס:

$$LL = -\frac{k+1}{2}log(2\pi) - \frac{k+1}{2}log(1+\alpha^2) - \frac{\sum_{i=0}^{k} x[i]^2}{2(1+\alpha^2)}$$

$$0 = \frac{\partial LL}{\partial \alpha^2} = -\frac{k+1}{2(1+\alpha^2)} + \frac{1}{2(1+\alpha^2)^2}\sum_{i=0}^{k} x[i]^2$$

$$1+\alpha^2 = \frac{1}{k+1}\sum_{i=0}^{k} x[i]^2$$

$$\hat{\alpha^2}_{ML} = -1 + \frac{1}{k+1}\sum_{i=0}^{k} x[i]^2$$

$\alpha^2 \geq 0$ הוא אי-שלילי, ולכן אם השונות האמפירית תהיה קטנה מאחד נצטרך לקחת את אפס בתור הפתרון.

6. כדי לחשב את הטיית המשערך נתחיל מחישוב התוחלת שלו:

$$\mathbb{E}\left[-1 + \frac{1}{k+1}\sum_{i=0}^{k} x[i]^2\right] = -1 + \frac{1}{k+1}\sum_{i=0}^{k}\mathbb{E}\left[x[i]^2\right] = -1 - \frac{k+1}{k+1}(1+\alpha^2) = \alpha^2$$
$$bias = \mathbb{E}\left[\hat{\alpha^2}_{ML} - \alpha^2\right] = 0$$

המשערך חסר הטיה

שאלה 3 - 35 נקודות

יהי $y = h\theta + n$ כאשר θ הינו סקלאר דטרמיניסטי לא ידוע ו- $\mathbb{E}[n] = \vec{0},\ \mathbb{C}ov(n) = I$. $h \in \mathbb{R}^M$ דטרמיניסטית.

1. השתמשו במשפט קושי שוורץ כדי להוכיח שהמינימום של $\|a\|^2$ תחת התנאי $a^T b = 1$ מתקבל בנקודה $\frac{b}{\|b\|^2}$
2. נבחר משערך לינארי $\hat{\theta}(y) = g^T y$. מהו תנאי נדרש על מנת שהמשערך יהיה חסר הטיה?
3. מהי השונות של $\hat{\theta}(y)$
4. מהו פתרון ה-Least Squares לבעיה?
5. בסעיף זה הניחו כי הרעש n מתפלג גאוסית. מהו פתרון ה-Maximum Likelihood לבעיה?
6. מהי השונות של משערך ה-Least Squares?
7. האם ניתן לפתח משערך לינארי חסר הטיה בעל שונות קטנה מזאת של Least Squares?

להזכרכם $a^T a = \|a\|^2$, ומשפט קושי שוורץ לנורמות הוא $|a^T b| \leq \|a\| \|b\|$

פתרון

1. ראשית נראה שהוקטור $\frac{y}{\|y\|^2}$ עומד בתנאי הנדרש:

$$1 = \left(\frac{y}{\|y\|^2}\right)^T y = \frac{y^T y}{\|y\|^2} = 1$$

נציב את התנאי הנדרש במשפט קושי שוורץ ומכיוון שנורמה היא אי-שלילית אפשר יהיה להעלות בריבוע:

$$\begin{aligned} |x^T y| &\leq \|x\| \|y\| \\ |1| &\leq \|x\| \|y\| \\ \frac{1}{\|y\|} &\leq \|x\| \\ \frac{1}{\|y\|^2} &\leq \|x\|^2 \end{aligned}$$

נציב את הוקטור $\frac{y}{\|y\|^2}$ ונראה שמתקבל שוויון, ולכן כל וקטור x אחר שעומד בתנאי המשפט יהיה עם נורמה ריבועית גדולה יותר:

$$\frac{1}{\|y\|^2} \leq \left\| \frac{y}{\|y\|^2} \right\|^2 = \frac{\|y\|^2}{\|y\|^4} = \frac{1}{\|y\|^2}$$

2. אם המשערך הוא חסר הטיה:

$$\begin{aligned}
0 = \mathbb{E}\left[\hat{\theta} - \theta\right] &= \mathbb{E}\left[g^T y - \theta\right] \\
&= \mathbb{E}\left[g^T (h\theta + n) - \theta\right] \\
&= (g^T h - 1)\theta + \mathbb{E}\left[n\right] = (g^T h - 1)\theta \\
\theta = 0 \quad or \quad g^T h &= 1
\end{aligned}$$

3. נחשב את השונות של y ובעזרתה את השונות של $\hat{\theta}(y)$:

$$\mathbb{C}ov\left(y\right) = \mathbb{C}ov\left(h\theta + n\right) = \mathbb{C}ov\left(n\right) = I$$

$$\begin{aligned}
\mathbb{V}ar\left(g^T y\right) &= g^T \mathbb{C}ov\left(y, y\right) g \\
&= g^T \mathbb{C}ov\left(y, y\right) g \\
&= g^T I g = g^T g = \|g\|^2
\end{aligned}$$

4. פתרון ה-Least Squares לבעיה הוא

$$(h^T h)^{-1} h^T y$$

5. פתרון ה-Least Squares הוא גם פתרון ML כאשר הרעש הוא גאוסיאן ממורכז

6. השונות של משערך ה-Least Squares היא:

$$\begin{aligned}
\mathbb{V}ar\left((h^T h)^{-1} h^T y\right) &= (h^T h)^{-1} h^T \mathbb{C}ov\left(y, y\right) h (h^T h)^{-1} \\
&= (h^T h)^{-1} h^T I h (h^T h)^{-1} = \frac{1}{h^T h}
\end{aligned}$$

7. נוכיח שמשערך ה-Least Squares הוא מינימלי מבחינת השונות ביחס לכל המשערכים הלינארים חסרי ההטיה נשאר לחבר את הסעיפים:

כל משערך לינארי חסר הטיה נדרש לעמוד בתנאי $g^T h = 1$. מסעיף שלוש אנחנו רואים שהשונות של משערכים לינאריים במקרה שלנו היא הנורמה בריבוע, כלומר המזעור של השונות מבין המשערכים הלינארים חסרי ההטיה זו בעיה שקולה למה שהוכחנו בסעיף הראשון, ולכן המשערך Least Squares הוא האופטימלי!