פרק 9. שערוך אופטימאלי של המצב – מסנן קלמן

בפרקים קודמים טיפלנו בבעיה של שחזור וקטור המצב (x_k) מתוך מדידות (y_k), עבור המערכת הליניארית

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned}$$

לצורך זה הצענו משחזר מצב (state observer) במבנה הבא:

$$\hat{x}_{k+1} = A\hat{x}_k + Bu_k + L(y_k - C\hat{x}_k)$$

ראינו כי התנהגות שגיאת השערוך נקבעת על ידי המטריצה $A_L = A - LC$. בתנאים אידאליים, ובהנחה כי המערכת אובזרווביליות, ניתן להבטיח כי שגיאת השערוך תדעך לאפס במהירות כרצוננו, על ידי בחירה מתאימה של הערכים העצמיים של מטריצה זו.

בפועל, ניתוח זה מתעלם משתי בעיות מעשיות חשובות:

1. המדידה y_k אינה מדויקת, ועלולה לכלול **רעש מדידה** משמעותי.
2. מודל המערכת (משוואות המצב) הוא רק קירוב למערכת האמיתית.

לשחזור נאמן של המצב (שגיאה קטנה), יש להתחשב בנקודות אלו.

בפרק זה ניגש לבעיית שחזור המצב מנקודת ראות סטטיסטית. אי הדיוקים הנ"ל במדידה ובמודל ייוצגו על ידי **רעשים** מתאימים: רעש המדידה, ורעש המצב. בעיית שחזור המצב תטופל בהתאם בכלים של שערוך סטטיסטי אופטימאלי (שערוך = estimation, אמידה סטטיסטית).

הפתרון שנקבל נקרא **מסנן קלמן** (Kalman Filter). מבנה המסנן דומה למשחזר המצב שלעיל – כאשר ההגבר L נבחר באופן אופטימאלי, והוא תלוי בזמן באופן כללי.

נציין כי השימוש במודלים סטוכסטיים לצורך סינון אותות זמניים ראשיתו בשנות ה-1940`, שם פותחה תורת הסינון האופטימאלי של וינר-קולמוגורוב, ובפרט מסנן וינר לסינון אותות סטציונריים טבולים ברעש. תורה זו הינה מוגבלת מעיקרה לאותות סטציונריים, דבר שהגביל מאוד את השימוש המעשי במסנן.

ב-1960, הציע קלמן (Rudolf Kalman, 1930-2016) את השימוש <u>במודל מצב</u> לניסוח בעיית הסינון האופטימאלי, ופיתח את המסנן הקרוי על שמו. פריצת דרך זו אפשרה יישום של תורת הסינון האופטימאלי לאותות ומערכות שאינם סטציונריים, ובהמשך גם למערכות לא לינאריות. מסנן קלמן הינו פשוט יחסית לחישוב ומימוש, והוא אחד מהכלים ההנדסיים המוכרים והשימושיים ביותר בתחומי בקרה, ניווט, עקיבת מטרות, מיזוג חיישנים, עיבוד אותות, ועוד.

9.1 וקטורים אקראיים גאוסיים (חזרה)

הערה: בסעיף זה נסמן משתנים ווקטורים אקראיים באותיות גדולות (X), וערכיהם באותיות קטנות (x). בסעיפים הבאים לא נבדיל ביניהם ונסמן את שניהם באותיות קטנות.

וקטור אקראי $X=(X_1,\ldots,X_n)^T$ הוא גאוסי (או נורמלי) אם מתקיים $X=AZ+m$, כאשר Z הוא וקטור של משתנים מקריים גאוסיים מנורמלים ובלתי תלויים. במקרה הלא-מנוון ($\Sigma=AA^T>0$), X בעל פונקציית צפיפות הסתברות:

$$f_X(x)=\frac{1}{(2\pi)^{n/2}\sqrt{\det\Sigma}}\exp\{-\tfrac{1}{2}(x-m)^T\Sigma^{-1}(x-m)\},\quad x\in\mathbb{R}^n$$

$m=E(X)$ הוא וקטור הממוצעים, ואילו $\Sigma=E\{(X-m)(X-m)^T\}\triangleq\operatorname{cov}(X)$ הינה מטריצת הקווריאנס. נסמן $X\sim N(m,\Sigma)$. מעתה נניח שהצפיפות קיימת.

וקטורים אקראיים (ו"א) X,Y הם <u>גאוסיים במשותף</u> אם הוקטור $(X^T,Y^T)^T$ הוא גאוסי. נזכור כי זו דרישה חזקה יותר מכך שכל אחד משניהם הוא ו"א גאוסי.

אי-תלות: כזכור ו"א X,Y הם בלתי-תלויים (סטטיסטית) אם $f_{XY}(x,y)=f_X(x)f_Y(y)$. אם X,Y גאוסיים <u>במשותף</u>, אזי הם בלתי-תלויים אם ורק אם הם חסרי קורלציה, כלומר $\Sigma_{XY}=0$, כאשר

$$\Sigma_{XY}\triangleq E\{(X-m_X)(Y-m_Y)^T\}\equiv\operatorname{cov}(X,Y)$$

התמרה לינארית של ו"א גאוסיים: יהי $X\sim N(m_X,\Sigma_X)$ ו"א גאוסי. נגדיר

$$Y=AX+B$$

אזי Y ו"א גאוסי, כלומר $Y\sim N(m_Y,\Sigma_Y)$, כאשר $m_Y=Am_X+B$, $\Sigma_Y=A\Sigma_XA^T$.
מכאן נובע גם כי X,Y גאוסיים במשותף (מדוע?).

פילוג מותנה: נזכיר כי בהינתן שני ו"א בעלי פונקציית צפיפות משותפת $f_{XY}(x,y)$, הצפיפות המותנית של X בהינתן Y נתונה על ידי

$$f_{X|Y}(x\,|\,y)=f_{XY}(x,y)\,/\,f_Y(y)\,.$$

ניתן לראות בפילוג זה כמייצג וקטור אקראי מותנה, $(X\,|\,Y=y)$. הממוצע והקווריאנס של משתנה מותנה זה יחושבו בהתאם על ידי

$$m_{X|Y=y}\triangleq E(X\,|\,Y=y)=\int x\,f_{X|Y}(x\,|\,y)dx$$

$$\Sigma_{X|Y=y}\triangleq\operatorname{cov}(X\,|\,Y=y)=E((X-E(X\,|\,Y))(X-E(X\,|\,Y))^T\,|\,Y=y)$$

פילוג מותנה גאוסי : יהיו X,Y ו"א גאוסיים במשותף. על ידי חישוב ישיר ניתן לוודא את התכונה החשובה הבאה :

- לכל y , $f_{X|Y}(x\,|\,y)$ הוא פילוג גאוסי, בעל ממוצע וקווריאנס להלן :

$$m_{X|Y=y} = m_X + \Sigma_{XY}\Sigma_{YY}^{-1}(y - m_Y)$$
$$\Sigma_{X|Y=y} = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}$$

נשים לב כי $\Sigma_{X|Y=y}$ המתקבל במקרה הגאוסי אינו תלוי בערך ה"נמדד" y של Y. ניתן לפיכך לסמנו בקיצור $\Sigma_{X|Y}$. תכונה זו היא תכונה ייחודית וחשובה של הפילוג הגאוסי.

9.2 משערך MMSE (חזרה)

בעיית השערוך הסטטיסטי (ניסוח בייסיאני): נתונים ו״א X, Y בעלי פילוג משותף ידוע F_{XY}. נדרש להעריך את ערכו של X, כאשר נמדד הערך y של Y בלבד.

משערך (estimator) של X על סמך Y הוא פונקציה (דטרמיניסטית) $\hat{x}(y)$ של הערך הנמדד y. כאשר מציבים במקום הערך y את הו״א Y מתקבל ו״א $\hat{X} = \hat{x}(Y)$.

ההפרש $e = X - \hat{X}$ הוא שגיאת השערוך.

קריטריון הטיב המקובל ביותר למדידת טיב השערוך הינו השגיאה הריבועית הממוצעת (MSE):

$$\text{MSE}(\hat{X} \mid y) \triangleq E(\| X - \hat{X} \|^2 \mid Y = y) = E(\| X - \hat{x}(y) \|^2 \mid Y = y)$$

משפט: המשערך $\hat{x}(y)$ המביא למינימום את השגיאה הריבועית הממוצעת נתון על ידי התוחלת המותנית:

$$\hat{x}_{opt}(y) = E(X \mid Y = y) \equiv \int x\, f_{X|Y}(x \mid y) dx$$

או בקיצור $\hat{X}_{opt} = E(X \mid Y)$.

תכונות נוספות של המשערך האופטימלי:

1. המשערך אינו מוטה (unbiased), כלומר התוחלת (המותנית) שלו שווה לזו של X:

$$E(X - \hat{X}_{opt} \mid Y = y) = 0$$

2. מטריצת הקווריאנס של שגיאת השערוך האופטימאלית נתונה על ידי

$$\Sigma_{e|y} \triangleq \text{cov}(e \mid Y = y) = E\{(X - E(X \mid Y)(X - E(X \mid Y)^T \mid Y = y\}$$

מטריצה זו מהווה מדד מרכזי לטיב השערוך, ולכן רצוי לחשבה יחד עם המשערך עצמו.

3. מהשוואה לביטוי בסעיף הקודם, ניתן לראות כי

$$\Sigma_{e|y} = \Sigma_{X|Y=y}$$

כלומר, מטריצת הקווריאנס של השגיאה האופטימלית זהה למטריצת הקווריאנס $\Sigma_{X|Y=y}$ של המשתנה המותנה $(X \mid Y = y)$. בר מקל על חישובה.

נציין כי איברי האלכסון של מטריצת הקוווריאנס $\Sigma_{e|y}$ מבטאים את שגיאת השערוך הריבועית עבור כל רכיב של הווקטור X. נורמת השגיאה הריבועית מתקבלת על ידי סיכומם:

$$\text{MSE}(\hat{X} \mid y) = trace\{\Sigma_{e|y}\} \quad \Rightarrow \quad \text{MSE}(\hat{X}_{opt} \mid y) = trace\{\Sigma_{X|Y=y}\}$$

המקרה הגאוסי: כאשר X, Y הם ו"א גאוסיים במשותף, נקבל:

$$\hat{x}_{opt}(y) = E(X \mid Y = y) \equiv m_{X|Y=y}$$
$$= m_X + \Sigma_{XY}\Sigma_{YY}^{-1}(y - m_Y)$$

$$\Sigma_{e|y} = \Sigma_{X|Y=y} = \Sigma_X - \Sigma_{XY}\Sigma_Y^{-1}\Sigma_{YX}$$

ו"א גאוסי טבול ברעש: נתבונן עתה במקרה המיוחד שבו המדידה Y נתונה על ידי הסכום

$$Y = CX + V,$$

כאשר C מטריצה נתונה, ואילו X, V ו"א גאוסיים בלתי תלויים, עם פרמטרים

$$X \sim N(m_X, \Sigma_X), \quad V \sim N(0, \Sigma_V)$$

נובע כי X, V גאוסיים במשותף (עקב אי-תלות), ולכן גם X, Y גאוסיים במשותף.

לחישוב המשערך האופטימאלי בעזרת הנוסחאות הקודמות נשים לב כי:

$$m_Y = Cm_X$$
$$\Sigma_{XY} = E(X - m_X)(Y - m_Y)^T = \Sigma_X C^T$$
$$\Sigma_{YY} = E(Y - m_Y)(Y - m_Y)^T = C\Sigma_X C^T + \Sigma_V$$

ולפיכך

$$\hat{x}_{opt}(y) = m_X + \Sigma_X C^T (C\Sigma_X C^T + \Sigma_V)^{-1}(y - Cm_X)$$
$$\Sigma_{X|Y=y} = \Sigma_X - \Sigma_X C(C\Sigma_X C^T + \Sigma_V)^{-1} C\Sigma_X$$

ניתן לבטא זאת בקיצור:

$$\hat{x}_{opt}(y) = m_X + K(y - Cm_X)$$

$$\text{cov}(X \mid Y = y) = \Sigma_X - KC\Sigma_X = (I - KC)\Sigma_X$$

כאשר:

$$K = \Sigma_X C^T (C\Sigma_X C^T + \Sigma_V)^{-1}$$

9.3 מודל המערכת ובעיית שיערוך המצב

נתבונן במערכת המצב הלינארית (לאו דווקא קבועה בזמן) הבאה, בזמן בדיד:

$$x_{k+1} = A_k x_k + B_k u_k + w_k \,, \quad k \ge 0$$
$$y_k = C_k x_k + v_k$$

כבעבר, x_k הוא וקטור מצב n-מימדי, u_k אות בקרה (ידוע ודטרמיניסטי), ו- y_k אות המדידה, שהינו במימד $n_y \ge 1$.

לאותות אלה נוספו אותות הרעש הבאים:

- $w_k \in \mathbb{R}^n$: רעש המצב
- $v_k \in \mathbb{R}^{n_y}$: רעש המדידה

ההנחות הסטטיסטיות הבסיסיות לגבי המודל הינן כלהלן:

1. רעש המצב (w_k) הינו תהליך אקראי גאוסי, לבן (כלומר: הו"א w_k ו- w_j הינם בת"ס עבור כל $k \neq j$), בעל ממוצע 0, ומטריצת קווריאנס $W_k = E(w_k w_k^T)$. נסמן, כרגיל, $w_k \sim N(0, W_k)$.
2. רעש המדידה (v_k) הינו תהליך אקראי גאוסי, לבן, בעל ממוצע 0, ומטריצת קווריאנס $V_k = E(v_k v_k^T)$. נסמן $v_k \sim N(0, V_k)$.
3. המצב ההתחלתי $x_0 \sim N(\bar{x}_0, P_0)$ הוא וקטור אקראי גאוסי, בעל ממוצע $\bar{x}_0$ ומטריצת קווריאנס $P_0 = E((x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T)$.
4. התהליכים (w_k), (v_k), והמצב ההתחלתי x_0 הינם בת"ס (בלתי תלויים סטטיסטית).
5. מתכונות מטריצת קווריאנס נובע כמובן כי $P_0, W_k, V_k \ge 0$. אנו נוסיף את הדרישה כי $V_k > 0$ (משמעותה: כל רכיבי המדידה הינם רועשים).

נדגיש כי בשלב זה אנו מגבילים את הדיון לרעשים גאוסיים, ונפתח את מסנן קלמן כמסנן אופטימאלי תחת הנחה זו. בהמשך נתייחס גם למקרה בו הרעשים אינם בהכרח גאוסיים: במקרה כללי זה מסנן קלמן יהיה המסנן הליניארי האופטימאלי.

<u>בעיית השיערוך</u>: נניח כי בזמן k נתונות המדידות $y_{0:k} \triangleq \{y_0,\ldots,y_k\}$. נדרש להעריך את המצב x_k על סמך וקטור מדידות זה. אנו נתעניין במשערך $\hat{x}_k = \hat{x}_k(y_{0:k})$ אשר ממזער את תוחלת השגיאה הריבועית:

$$E(\| x_k - \hat{x}_k \|^2) \to \min$$

משערך זה הינו כזכור המשערך האופטימאלי במובן MMSE (Minimal Mean Square Error).

כפי שידוע לנו מתורת השיערוך של משתנים מקריים, המשערך האופטימאלי במובן MMSE נתון על ידי התוחלת המותנית:

$$\hat{x}_k = E(x_k \mid y_{0:k})$$

נותר לנו אם כן למצוא נוסחה מפורשת עבור התוחלת המותנית.

<u>הערה</u>: מההנחות הנ"ל על המערכת נובע כי המצב x_k הינו ו"א גאוסי (לכל $k \geq 0$), דהיינו $x_k \sim N(m_k, \Sigma_k)$. הממוצע והקוואריאנס של וקטור זה (<u>ללא</u> התניה במדידות) ניתנים לחישוב באופן רקורסיבי כלהלן:

$$\begin{aligned} m_{k+1} &= A_k m_k + B_k u_k \\ \Sigma_{k+1} &= A_k^T \Sigma_k A_k + W_k \end{aligned} \quad , \qquad k \geq 0$$

עם תנאי התחלה $m_0 = \bar{x}_0$, $\Sigma_0 = P_0$.

המשוואה שקיבלנו לעדכון קווריאנס המצב נקראת משוואת ליאפונוב הדיסקרטית, בצורת שערוך (השוו למשוואת ליאפונוב שהתקבלה בפרק על יציבות).

אנו מצפים כמובן כי הקווריאנס של המשערך המתקבל על סמך המדידות יהיה קטן יותר מהקווריאנס Σ_k שאינו מתחשב במדידות.

9.4 פיתוח מסנן קלמן (רעש גאוסי)

<u>סימונים</u> : נגדיר את שני המשערכים הבאים של x_k , באמצעות התוחלות המותנות :

$$\hat{x}_k^+ = E(x_k \mid y_{0:k})$$
$$\hat{x}_k^- = E(x_k \mid y_{0:k-1})$$

המשערך $\hat{x}_k^-$ נקרא גם <u>חזאי-צעד-אחד-קדימה</u> של x_k .

המשערך $\hat{x}_k^+$ הוא $\hat{x}_k$ שהוגדר קודם, כאשר הסימן + נוסף כדי להבדילו מ- $\hat{x}_k^-$.

<u>סימון חליפי</u> הנהוג הספרות הוא $\hat{x}_k^+ \equiv \hat{x}_{k|k}$, $\hat{x}_k^- \equiv \hat{x}_{k|k-1}$.

מטריצות הקווריאנס המתאימות יסומנו

$$P_k^+ = \operatorname{cov}(x_k \mid y_{0:k}) \equiv \operatorname{cov}(x_k - \hat{x}_k^+ \mid y_{0:k})$$
$$P_k^- = \operatorname{cov}(x_k \mid y_{0:k-1}) \equiv \operatorname{cov}(x_k - \hat{x}_k^- \mid y_{0:k-1})$$

<u>**פיתוח המסנן**</u> : חישוב המשערך האופטימאלי יבוצע באופן רקורסיבי. נפתח את נוסחאות המשערך באינדוקציה.

<u>איתחול</u> : אנו מתחילים בזמן $k=0$ ללא מדידות. לפיכך :

$$\hat{x}_0^- = x_0 \sim N(\bar{x}_0, P_0)$$

<u>הנחת האינדוקציה</u> : $x_k \mid y_{0:k-1} \sim N(\hat{x}_k^-, P_k^-)$

(כלומר : $f(x_k \mid y_{0:k-1})$ הינו פילוג גאוסי, עם פרמטרים $\hat{x}_k^-, P_k^-$ שכבר חושבו). נראה כי זה מתקיים גם עבור $k+1$, ונחשב את $\hat{x}_{k+1}^-, P_{k+1}^-$.

<u>שלב 1</u> ("עדכון המדידה") : נחשב את הפילוג של $(x_k \mid y_{0:k})$. נשיב לב כי $y_{0:k} = \{y_{0:k-1}, y_k\}$. לאחר מדידת $y_{0:k-1}$ ניתן להתייחס למצב המותנה $(x_k \mid y_{0:k-1})$ כמשתנה אקראי גאוסי z_k בעל פילוג $z_k \sim N(\hat{x}_k^-, P_k^-)$ (לפי הנחת האינדוקציה). עתה, מתוך $y_k = C_k x_k + v_k$ מתקיים השיוויון הבא (במובן פילוגים) :

$$y_k = C_k z_k + v_k$$

ניתן להראות כי v_k בלתי תלוי ב- z_k (עקב הנחות האי-תלות הבסיסיות שלנו, הגוררות אי תלות של v_k ב- x_k וב- $y_{0:k-1}$). לפיכך, לפי נוסחאות השיערוך שפיתחנו קודם עבור ו"א גאוסיים:

$$\hat{x}_k^+ = \hat{x}_k^- + K_k(y_k - C_k\hat{x}_k^-)$$
$$P_k^+ = (I - K_kC_k)P_k^-$$

כאשר $K_k = P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1}$.

<u>שלב 2</u> ("עדכון הזמן"): נחשב את הפילוג של $(x_{k+1} \mid y_{0:k})$. מתוך משוואת המצב $x_{k+1} = A_k x_k + B_k u_k + w_k$ והנחת אי התלות של w_k ברעשים קודמים, נובע כי:

$$\hat{x}_{k+1}^- = A_k \hat{x}_k^+ + B_k u_k$$
$$P_{k+1}^- = A_k P_k^+ A_k^T + W_k$$

בכך סיימנו למעשה את פיתוח מסנן קלמן למקרה הגאוסי.

9.5 משוואות המסנן

א. עדכון דו-שלבי : נסכם ראשית את המשוואות שפיתחנו בסעיף הקודם.

- תנאי התחלה : $\hat{x}_0^- = \bar{x}_0,\ P_0^- = P_0$.
- "עדכון המדידה" ($\hat{x}_k^- \to \hat{x}_k^+$) :

$$K_k \triangleq P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1}$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - C_k \hat{x}_k^-)$$

$$P_k^+ = (I - K_k C_k) P_k^-$$

- "עדכון הזמן" ($\hat{x}_{k+1}^- \leftarrow \hat{x}_k^+$) :

$$\hat{x}_{k+1}^- = A_k \hat{x}_k^+ + B_k u_k$$

$$P_{k+1}^- = A_k P_k^+ A_k^T + W_k$$

ב. חישוב מאוחד : את שני השלבים שלעיל (עדכון הזמן והמדידה) ניתן לאחד למשוואה אחת. על ידי הצבת נוסחאות עדכון המדידה באלו של עדכון הזמן נקבל :

$$\hat{x}_{k+1}^- = A_k \hat{x}_k^- + B_k u_k + L_k (y_k - C_k \hat{x}_k^-)$$

כאשר

$$L_k = A_k K_k = A_k P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1}$$

באופן דומה, עבור P_k^- נקבל :

$$P_{k+1}^- = A_k (I - K_k C_k) P_k^- A_k^T + W_k$$
$$= (A_k - L_k C_k) P_k^- A_k^T + W_k$$

ובצורה מפורשת,

$$P_{k+1}^- = A_k P_k^- A_k^T - A_k P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1} C_k P_k^- A_k^T + W_k$$

אנו רואים כי נוסחת העדכון של $\hat{x}_k^-$ היא בצורת משחזר מצב, כאשר ההגבר L_k מחושב באופן אופטימאלי.

המשוואה האחרונה לעדכון P_k^- נקראת משוואת ריקאטי הדיסקרטית (בגרסת שערוך). ניתן לראות כי היא דואלית למשוואת ריקאטי שקיבלנו עבור בעיית הבקר האופטימאלי.

הערות לגבי מטריצות הקוואריאנס P_k:

- ניתן לראות כי מטריצות הקוואריאנס P_k^+ , P_k^- נדרשות כחלק מחישוב ההגבר האופטימאלי.
- בנוסף לכך, יש להן תפקיד חשוב בהערכת טיב השיערוך המתקבל. נזכור כי מתוך הגדרת $\hat{x}_k^+ = E(x_k \mid y_{0:k})$ נובע

$$P_k^+ \triangleq \mathrm{cov}(x_k \mid y_{0:k}) = \mathrm{cov}(x_k - \hat{x}_k^+ \mid y_{0:k}) = E((x_k - \hat{x}_k^+)(x_k - \hat{x}_k^+)^T \mid y_{0:k})$$

כאשר $x_k - \hat{x}_k^+$ הוא וקטור שגיאת השיערוך. לפיכך P_k^+ נותנת ישירות את קוואריאנס שגיאת השיערוך (ובאופן דומה עבור P_k^-).

שאלה: למה שווה $E(\| x_k - \hat{x}_k^+ \|^2 \mid y_{0:k})$? $E(\| x_k - \hat{x}_k^+ \|^2)$?

- נציין כי מטריצות הקוואריאנס המתקבלות אינן תלויות במדידות, כך שניתן (תיאורטית לפחות) לחשבן מראש, ולפיכך לחשב מראש את הגבר המסנן האופטימאלי. אולם במרבית היישומים מקובל לחשב את ההגבר בזמן פעולת המסנן (פרט למקרה של הגבר קבוע, שיידון בהמשך).

לסיכום, יישום של מסנן קלמן לבעייה ספציפית יתבצע עקרונית באופן הבא:

1. נקבע מודל לדינמיקת המצב, דהיינו: $x_{k+1} = Ax_k + Bu_k + w_k$, נעריך את גודל רעש המצב w_k ובהתאם נקבע את מטריצת הקוואריאנס W_k עבור רעש זה.
2. נקבע את מודל מדידה $y_k = Cx_k + v_k$, ונעריך את גודל הקוואריאנס V_k של רעש המדידה.
3. נעריך את תנאי ההתחלה x_0 (ערך ראשוני סביר $\bar{x}_0$, ואי ודאות P_0).
4. נחשב (מראש, או תוך פעולה) את מטריצת הקוואריאנס P_k , ומכאן את ההגבר L_k .
5. נציב במשוואת המסנן (האובזרוור) לקבלת המשערך האופטימאלי.

9.5 המסנן הסטציונרי

נתבונן עתה במקרה החשוב שבו מטריצות המערכת A,B,C ופרמטרי הרעש W,V הינם קבועים בזמן. באופן כללי, הגבר המסנן L_k יהיה עדיין תלוי בזמן, עקב השפעת תנאי ההתחלה, $P_0=\mathrm{cov}(x_0)$. בפרט, P_k^- עדיין מתקבל מתוך משוואת ההפרש של ריקאטי:

$$P_{k+1}^- = AP_k^-A^T - AP_k^-C^T(CP_k^-C^T+V)^{-1}CP_k^-A^T + W$$

נניח כי תנאי ההתחלה $P_0=P$ מקיימים את משוואת ריקאטי האלגברית (ARE), המתקבלת על ידי הצבת אותו P בשני צידי המשוואה:

$$P = APA^T - APC^T(CPC^T+V)^{-1}CPA^T + W \qquad (*)$$

קל לראות כי במקרה זה $P_k^- = P_{k-1}^- = \ldots = P_0^- = P$, ולכן $L_k \equiv L \triangleq APC^T(CPC^T+V)^{-1}$. המסנן המתקבל

$$\hat{x}_{k+1}^- = A\hat{x}_k^- + Bu_k + L(y_k - C\hat{x}_k^-)$$

הוא בצורת משחזר מצב, עם הגבר קבוע בזמן. מסנן זה נקרא **מסנן קלמן הסטציונרי**.

גם כאשר תנאי ההתחלה P_0 שונים מ-P, ניתן לצפות כי ככל שנתקדם בזמן "נשכח" את תנאי ההתחלה ונתקרב למצב היציב, בו $P_k^- \simeq P$, ולכן $L_k \simeq L$. למעשה, מתקיים המשפט הבא:

משפט (ללא הוכחה): נניח כי הצמד (A,C) דטקטיבילי, וכן כי $W=B_1B_1^T$ כאשר הצמד (A,B_1) סטביליזבילי. אזי,

(1) למשוואת ריקאטי (*) קיים פתרון אחד בלבד שהוא אי-שלילי מוגדר ($P\ge 0$).

(2) לכל תנאי התחלה $P_0^-=P_0$ מתקיים $\lim_{k\to\infty} P_k^- = P$, כאשר $P\ge 0$ הינו פתרון זה.

במקרים רבים נוח להשתמש במסנן הסטציונרי גם כאשר תנאי ההתחלה אינם P_0 אינם בהכרח זהים ל- P. מקרה אופייני הוא כאשר המערכת פועלת באופן מתמשך לפרק זמן ארוך (כדוגמת מערכת בקרה). מסנן זה הינו תת-אופטימלי בזמן סופי, אולם הוא אופטימאלי ביחס לשגיאת השערוך האסימפטוטית :

$$\lim_{k\to\infty} E(\| x_k - \hat{x}_k^- \|^2 \mid y_{0:k-1})$$

כלומר, שגיאת השערוך האסימפטוטית של המסנן הסטציונרי זהה לזו של המסנן האופטימלי.

9.6 מסנן קלמן למקרה הלא-גאוסי

במקרה שבו רעשי המדידה ו\או תנאי ההתחלה אינם גאוסיים, מסנן קלמן אינו המסנן האופטימלי במובן MMSE. אך, עדיין, ניתן להראות כי הוא המסנן <u>הליניארי</u> האופטימאלי. כזכור משערך לינארי מוגדר כפונקציה לינארית של המדידות, וקריטריון האופטימאליות פה הוא השגיאה הריבועית (הבלתי מותנית!) $E(\| x_k - \hat{x}_k \|^2)$.

<u>ההנחות הסטטיסטיות</u> הנדרשות במקרה הנוכחי הן אך ורק לגבי המומנטים הראשון והשני של הרעשים. בפרט, נניח כי:

1. רעש המצב (w_k) הינו תהליך אקראי בעל ממוצע 0, חסר קורלציה בזמן ($E(w_k w_l^T) = 0$ עבור $k \neq l$), עם קווריאנס $W_k = E(w_k w_k^T)$.
2. רעש המדידה (v_k) הינו תהליך אקראי בעל ממוצע 0, חסר קורלציה בזמן, ובעל קווריאנס $V_k = E(v_k v_k^T)$.
3. המצב ההתחלתי $x_0 \sim N(\bar{x}_0, P_0)$ הוא ו"א בעל ממוצע $\bar{x}_0$ וקווריאנס P_0.
4. התהליכים (w_k), (v_k), והמצב ההתחלתי x_0 הינם חסרי קורלציה הדדית.

בהנחות אלו, משוואות המסנן המתקבל (כמסנן הליניארי האופטימאלי) הינן ללא שינוי. הזהות בין משוואות המסננים נובעת מהתיאוריה הכללית של שיערוך לינארי אופטימאלי, ולא נפרט מעבר לכך.

9.7 מסנן קלמן בזמן רציף

נתאר בקצרה את מסנן קלמן המתקבל בזמן רציף; מסנן זה נקרא גם מסנן Kalman-Bucy. הטיפול הטכני פה מסובך יותר, אולם ניתן לקבל את המשוואות באופן פורמלי ממשוואות המסנן הבדיד על ידי דיסקרטיזציה. לא נפרט זאת פה.

מודל המערכת בזמן רציף:

$$\dot{x}_t = Ax_t + Bu_t + w_t; \quad x_0$$
$$y_t = Cx_t + v_t$$

כאשר $\{w_t\}$ ו- $\{v_t\}$ הינם רעשים לבנים, גאוסיים, בעלי ממוצע אפס, ובלתי תלויים זה בזה ובמצב ההתחלתי x_0. נסמן

$$x_0 \sim N(\bar{x}_0, P_0)$$
$$E(w_s w_t^T) = W_t \delta(t-s)$$
$$E(v_s v_t^T) = V_t \delta(t-s) \qquad (V_t > 0)$$

המשערך האופטימאלי (במובן MMSE) של x_t על סמך המדידות $y_{0:t} = \{y_s, 0 \le s < t\}$ הינו

$$\hat{x}_t \triangleq E(x_t \mid y_{0:t})$$

עם קווריאנס שערוך

$$P_t \triangleq \operatorname{cov}(x_t \mid y_{0:t}) = \operatorname{cov}(x_t - \hat{x}_t \mid y_{0:t})$$

המסנן האופטימאלי הוא

$$\tfrac{d}{dt}\hat{x}_t = A\hat{x}_t + Bu_t + L_t(y_t - C\hat{x}_t); \quad \hat{x}_0 = \bar{x}_0$$

כאשר $L_t = P_t C^T V_t^{-1}$, ואילו הקווריאנס P_t מקיים את משוואת ריקאטי:

$$\dot{P}_t = AP_t + P_t A^T - P_t C^T V_t^{-1} C P_t + W_t$$

זו משוואת ריקאטי הדיפרנציאלית (בגרסת שערוך). שתי המשוואות האחרונות מהוות את מסנן קלמן בזמן הרציף. נשים לב שהמשוואות פה קצרות יותר: בפרט, לא קיימת ההבחנה בין P_t^- ו- P_t^+.

<u>המסנן הסטציונרי</u>: נניח עתה כי (A,C,W,V) קבועים בזמן. משוואת ריקאטי האלגברית לזמן הרציף מתקבלת עתה על ידי הצבת $\dot{P}_t = 0$, דהיינו:

$$AP + PA^T - PC^T V^{-1} CP + W = 0 \qquad \text{(ARE-CT)}$$

המסנן הסטציונרי בזמן רציף הינו בהתאם:

$$\tfrac{d}{dt}\hat{x}_t = A\hat{x}_t + Bu_t + L(y_t - C\hat{x}_t); \quad \hat{x}_0 = \bar{x}_0$$

כאשר $L = PC^T V^{-1}$, ו- $P \geq 0$ הינו פתרון משוואת ריקאטי הנ"ל.

התוצאות שתיארנו בזמן הבדיד (לגבי קיום ויחידות הפתרון למשוואת ריקאטי האלגברית וההתכנסות למסנן הסטציונרי) תקפות גם פה.

9.8 הערות נוספות לגבי השימוש במסנן קלמן

1. **מדידות בלתי-סדירות** : מסנן קלמן אינו מחייב קבלת מדידות במרווחי זמן סדירים. אם בזמן k כלשהו לא מקבלת מדידה, ניתן פשוט לדלג על שלב עדכון המדידה בזמן זה.

2. **מיזוג חיישנים** : מסנן קלמן מהווה מסגרת נוחה וגמישה למיזוג חיישנים (sensor fusion), כלומר שימוש במדידות המתקבלות ממספר חיישנים שונים לצורך שיערוך אחוד ומשופר של הגודל המבוקש. דוגמא לכך היא שילוב של מספר מכשירי ניווט (GPS, מדידים אינרציאליים, מדי מהירות וכו') לקבלת הערכה מאוחדת של המיקום והמהירות.
גם פה, הדבר מתאפשר עקב ההפרדה בין עדכון המדידה ועדכון הזמן. למשל, אם מתקבלות מספר מדידות (ממדידים שונים) באותו רגע זמן, ניתן לבצע עדכוני מדידה עוקבים עבורם (כאשר הסדר אינו משנה את התוצאה).

3. **מסנן קלמן המורחב עבור מערכות לא-לינאריות** : ביישומים רבים המערכת עצמה (דינמיקת המצב ו/או משוואת המדידה) אינה לינארית. הפתרון הנפוץ במקרה זה הינו שימוש במסנן תת-אופטימאלי, לרוב מסנן קלמן המורחב המתקבל על ידי לינאריזציה של המערכת סביב הערך הנוכחי של המצב המשוערך ($\hat{x}_k$) .

9.9 * מסנן קלמן המורחב (Extended Kalman Filter)

כאשר המערכת אינה לינארית, פתרון מדויק למשערך אופטימלי אינו מעשי חישובית. לפיכך נעזרים בפתרונות מקורבים. הקרוב הבסיסי והשימושי ביותר הוא מסנן קלמן המורחב, אשר מבוסס על לינאריזציה מקומית של דינמיקת המערכת. נתאר בקצרה את הגרסה הבסיסית ביותר של מסנן זה, שלו חשיבות מעשית רבה.

נתבונן במערכת המצב הלא-לינארית, בזמן בדיד:

$$\begin{aligned} x_{k+1} &= f(x_k, u_k) + w_k, \quad k \ge 0 \\ y_k &= h(x_k) + v_k \end{aligned}$$

ההנחות לגבי סטטיסטיקת הרעשים ותנאי ההתחלה כמו קודם.

הרעיון הוא לקרב בכל שלב k את הפונקציות הלא-לינאריות של המצב על ידי משוואות לינאריות עם פרמטרים מתאימים A_k, C_k, ולהציב מטריצות אלו היכן שנחוץ במשוואות המסנן הרגילות.

לצורך זה נבצע לינאריזציה של המשוואות סביב וקטור מצב מתאים $\hat{x}$. אידאלית היינו רוצים לבחור $\hat{x} = x_k$, אך כיוון שערך זה אינו ידוע נבחר בשערוך הטוב ביותר שלו לאותו שלב.

המסנן המתקבל (השינויים מצוינים בכחול):

- "עדכון המדידה" ($\hat{x}_k^- \rightarrow \hat{x}_k^+$):

$$C_k = \left.\frac{\partial h(x)}{\partial x}\right|_{x=\hat{x}_k^-}$$

$$K_k = P_k^- C_k^T (C_k P_k^- C_k^T + V_k)^{-1}$$

$$\hat{x}_k^+ = \hat{x}_k^- + K_k (y_k - f(\hat{x}_k^-))$$

$$P_k^+ = (I - K_k C_k) P_k^-$$

- "עדכון הזמן" ($\hat{x}_{k+1}^- \leftarrow \hat{x}_k^+$):

$$\hat{x}_{k+1}^- = f(\hat{x}_k^+, u_k)$$

$$A_k = \left.\frac{\partial f(x, u_k)}{\partial x}\right|_{x=\hat{x}_k^+}$$

$$P_{k+1}^- = A_k P_k^+ A_k^T + W_k$$