הטכניון - מכון טכנולוגי לישראל
הפקולטה להנדסת חשמל

אותות אקראיים 044202

משה זכאי, משה סידי, אדם שורץ, עפר זיתוני

תשס"ג 2003

תודות

מהדורת 1989:
ברצוננו להודות לגברת אנט ברג ז"ל על עזרתה בהדפסת ובהכנת השרטוטים, לגברת יפה לוי על עזרתה בהדפסה ולגברת חנה ביסמוט על עזרתה בהכנת השרטוטים.

כמו כן ברצוננו להודות למר דורון שקד על עזרתו בעדכון החוברת.

מהדורת 2003:
תודות לגברת לזלי פרייס על הדפסת החוברת ולגברת חנה ביסמוט על הכנת השרטוטים.

תוכן ענינים

0 תכנית הקורס

מספר שעות ההרצאה המוקדשות לכל נושא מתואר בסוגריים () ליד כל סעיף.

סדר הנושאים, ובמידה מסויימת משך הזמן המוקדש לכל נושא תלויים במורה המקצוע.

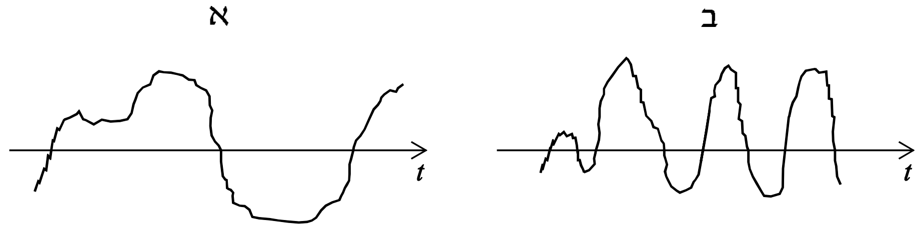
1. (1) הערות על הקורס. מבוא כללי ומוטיבציה.
2. (1) גישות לתורת ההסתברות: גישה סטטיסטית ודטרמיניסטית. חזרה על מושגים בהסתברות: משתנה אקראי, ווקטור אקראי, פילוג וצפיפות. תוחלת ותכונותיה, מומנטים. קווריאנס וקורלציה. פונקציה אפיינית.
3. (2) ווקטורים אקראיים: הגדרה, תוחלת, מומנטים, ומטריצת קווריאנס. פונקציה אפיינית. השפעת טרנספורמציות לינאריות על תוחלת ומטריצת קווריאנס. ווקטורים גאוסיים והשפעת טרנספורמציות לינאריות על פילוגם.
4. (2) שערוך. הגדרת הבעיה-קריטריון שגיאה ריבועית ממוצעת. חזרה: תוחלת מותנית ומשפט ההחלקה, פילוג גאוסי מותנה, שערוך לינארי סקלרי. שערוך מווקטור, עקרון הניצבות. השוואת משערך אופטימלי ומשערך לינארי אופטימלי.
5. (2) תהליך אקראי בזמן בדיד. דוגמאות לתהליכים בזמן בדיד כולל רעש לבן, הילוך שיכור ותהליך יציאה ממערכת לינארית (זמן בדיד). איפיון סטטיסטי של תהליכים בזמן בדיד: פילוגים, תוחלת, אוטוקורלציה וקווריאנס. סטציונריות וארגודיות.
6. (5) חזרה: הסתברות מותנית. שרשרות מרקוב. הגדרות, שרשרות הומוגניות, סיווג מצבים ופרוק מרחב המצב. פילוג סטציונרי וארגודיות.
7. (3) תהליכים אקראיים בזמן רציף: דוגמאות. הגדרה ואפיון סטטיסטי: פילוגים, תוחלת, אוטוקורלציה וקווריאנס. תהליך פואסון. סטציונריות וארגודיות. תהליכים מרקוביים בזמן רציף: הגדרה. תור $M/M/1$.
8. (3) מעבר תהליכים דרך מערכות לינאריות: תיאור, מעבר הממוצע, מעבר אוטוקורלציה וקרוס קורלציה. מערכות קבועות בזמן ותהליכים סטציונריים במובן הרחב. ניתוח במרחב התדר: צפיפות הספק ספקטרלית, משפט גבול לצפיפות הספק ספקטרלית.
9. (2) סינון לינארי (מסננת וינר). תיאור, מסננת לא סיבתית, פיתרון בעית הסינון במישור הזמן ובמישור התדר. חישוב השגיאה הריבועית הממוצעת, דוגמאות.
10. (3) רעשים פיזיקליים במערכות. רעש לבן, רעש תרמי. מקורות מרובי נגדים. הספק מצוי, טמפרטורת רעש אפקטיבית או סיפרת רעש, חיבור בשרשרת של מגברים. רעש דיודה.
11. נושאים נוספים: משפט הדגימה. תור $M/G/1$. תהליכים דו ממדיים. הילוך אקראי על גרפים ורשתות נגדים, סימולציות מונטה קרלו.

1 מבוא וחזרה על הסתברות

1.1 מבוא

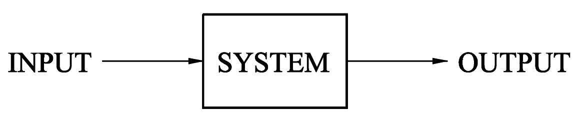
בתורת ההסתברות למדנו איך לאפיין אפיון הסתברותי ו"לתמצת" גדלים אקראיים שקראנו להם משתנים אקראיים או וקטורים אקראיים (חוק ההסתברות, מומנטים ועוד). המאפיין משתנים אלה הוא קיום פרמטר של "מזל". דוגמה: תוצאת זריקת קוביה או מספר קוביות. בקורס זה (אותות אקראיים) נעסוק באפיון ותמצות תהליכים אקראיים. בתהליך אקראי (שמות נוספים: אות אקראי, פונקציה אקראית, תהליך סטוכסטי, סידרה אקראית) בנוסף לפרמטר ה"מזל", קיים גם פרמטר זמן (דיסקרטי או רציף). דוגמאות: רעש מגבר, שיחות טלפון, גלי הים, מספר מכוניות העוברות בצומת, רעידות מכונית וכו'.

עבור צורת גל דטרמיניסטית אנו יכולים לדבר על רוחב סרט ולהגיד (לפחות במקרים מסויימים) איזו משתי צורות גל "מהירה" יותר ע"י השואת התמרת פוריה שלהן. במקרה ההסתברותי, נעיין בשתי צורות גל טיפוסיות



ציור 1.1:

כאשר א' צורת גל טיפוסית של מכונה עם בולם זעזועים מסוג א', ו-ב' צורת גל טיפוסית עבור אותה מכונה עם בולם זעזועים מסוג ב'. הזעזועים אקראיים ואינם חוזרים על עצמם (תהליכים אקראיים). בהרגשה, תהליך ב' מהיר יותר ובעתיד נראה איך לתת לזה מובן. נראה בעיה זו כחלק מבעיה כללית ועקרונית יותר כדלקמן:

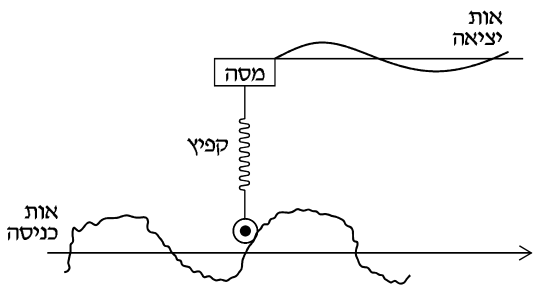


ציור 1.2:

כאשר עסקנו בצורות גל דטרמיניסטיות עסקנו במודל המתואר בציור 1.2 ולמדנו (לפחות במקרה של מערכת לינארית, ובמיוחד במקרה של מערכת לינארית שאינה משתנה בזמן), איך לאפיין את המערכת ואיך לחשב את היציאה עבור כניסה שרירותית. במיוחד, עבור מערכת לינארית שאינה משתנה בזמן למדנו לאפיין את המערכת בשתי צורות: אפיון במרחב הזמן (תגובת המערכת להלם - פונקצית דירק), ואפיון במרחב התדר (ע"י תגובה לערור הרמוני $e^{i\omega t}$). בעזרת אפיונים אלה ובעזרת עקרון הסופרפוזיציה, קבלנו את התגובה של המערכת לכל אות כניסה בשתי הצורות - במרחב התדר ובמרחב הזמן. בקורס זה נעסוק באפיון ההסתברותי של היציאה כאשר נתון האפיון ההסתברותי של הכניסה ואפיון המערכת. אנו נעסוק תמיד במערכת דטרמיניסטית, רק הכניסה (ולכן גם היציאה) יהיו תהליכים אקראיים.

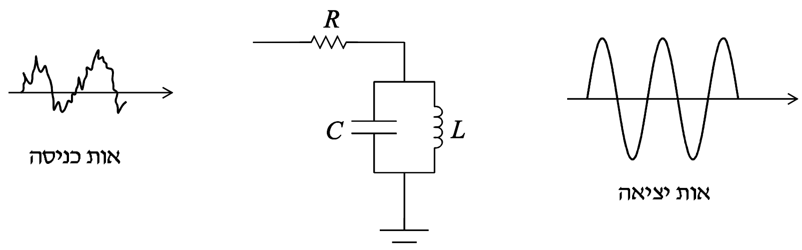
<u>דוגמאות:</u>

א. בולם זעזועים



ציור 1.3:

ב. מעגל RLC



ציור 1.4:

<u>הערה:</u> בשני המקרים ציירנו את היציאה "פחות עצבנית" מהכניסה (אם כי לא בהכרח חלשה יותר, במגבר היא תהיה חזקה יותר). בהמשך ההרצאות נראה מדוע.

נחזור ונדגיש שאנו עוסקים <u>במשפחות</u> של פונקציות מבלי שיהיה לנו עניין בדגם מיוחד. הבעיה שלנו תהיה איך לאפיין אותות (משפחות) כאלה ואיך מערכות לינאריות מגיבות למשפחות כאלה.

<u>מוטיבציה</u>

מנקודת מבטו של מחשב המשמש כשרת תקשורת, משתמשים חדשים מתחברים בזמנים אקראיים (ראה דוגמה 4.1). כיצד נתאר את מספר המשתמשים שהגיעו למחשב עד רגע t נתון? מצד אחד זהו תהליך, שכן יש תלות בזמן—אם נבדוק את מספר המשתמשים זמן מאוחר יותר, יתכן שהמספר ישתנה. מצד שני, בכל רגע נתון מספר המשתמשים שהגיעו עד רגע זה הוא משתנה אקראי (הגדרה 8.9) , משום שמשך הזמן בין הגעה להגעה הוא אקראי. ל"יצור" שכזה אנו קוראים תהליך אקראי (הגדרה 3.2).

דוגמאות נוספות:

- רעש אנלוגי במקלט רדיו (אות אנלוגי). מקור האקראיות: תהליכים פיסיקליים שונים היוצרים רעש, כגון תנועות אקראיות של אלקטרונים. ניתן לבנות מודלים לרעשים מסוגים שונים.
- אות דיבור. מקור האקראיות: מודל מתמטי של אות עבורו לא קיים אפיון דטרמיניסטי.
- מכונית בתנועה. מקור האקראיות: מודל של כביש בתנאים מציאותיים (למרות שהכביש דטרמיניסטי!).
- מערכת עקיבה. מקור האקראיות: מודל לתנועת הגוף (מטוס, אניה וכו') אשר אינה ידועה מראש, או מודל של שגיאות מדידה שונות.

אנו רואים שתהליך אקראי מאופיין ע"י תלות בזמן (ולכן הוא "תהליך"), ואקראיות.

מודלים

כדי לראות כיצד נראה מודל מתמטי של תהליך אקראי, נזכר במושגים משתנה אקראי (הגדרה 8.9) פונקציה ואות. נתאר לנו אין-סוף מקלטים, שכולם הופעלו בזמן t_0. כולם מאותו סוג, וכולם מכוונים לאותו תדר. נניח שאין שידור בתדר זה, ולכן במוצא המקלט יהיה רק הרעש שנוצר בו עצמו. אפשר, למשל, לחשוב על מרחב המדגם (הגדרה 8.1) Ω כעל אוסף כל המקלטים, כך שכל $\omega\in\Omega$ יהיה מקלט מסויים. נסמן ב- $X(t,\omega)$ את היציאה ממקלט מספר ω ברגע t. כך למשל, אם נקבע זמן מסויים, למשל $t=2$, אזי בתלות ב"פרמטר המזל" ω, גודל זה $X(2,\omega)$ הוא משתנה אקראי. מצד שני, אם נקבע משדר מסויים, כלומר נקבע את $\omega=\omega_0$, אזי לפי המודל שלנו הפונקציה $X(t,\omega_0)$, היא פונקציה של המשתנה t בלבד: לפונקציה זו של משתנה הזמן (כאשר פרמטר המזל קבוע) קוראים פונקציית מדגם (הגדרה 3.2).

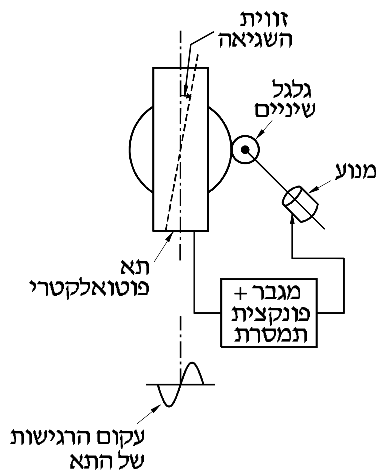
דוגמה 1.1 לפעמים אפשר לתת ביטוי מפורש לתהליך אקראי (להמחשה ראה 7.1).

- יהיו Y ו-Z משתנים אקראיים. נגדיר $X(t,\omega)=Y(\omega)\cdot t+Z(\omega)$. זהו תהליך אקראי, ופונקציות המדגם (ω קבוע, ומתבוננים בפונקציה של המשתנה t) הן קווים ישרים.
- יהיו A ו-ϕ מ"א. נגדיר $X(t,\omega)=A(\omega)\sin(2\pi ft+\phi(\omega))$. צורות הגל (פונקציות המדגם) במקרה זה הן תנודות הרמוניות בעלות אמפליטודה (אקראית) A ופאזה (אקראית) ϕ.
- יהיה N שלם חיובי קבוע או אקראי, ויהיו X_n משתנים אקראיים. נגדיר

$$X(t,\omega)=\sum_{n=1}^{N}X_n(\omega)\sin nt$$

אזי פונקציות המדגם הן סכום משוקלל של תנודות הרמוניות. אם $N=N(\omega)$ הוא אקראי, אזי מספר האברים בסכום תלוי ב-ω.

<u>דוגמא נוספת:</u> מערכת לעקיבה אחרי כוכב (או רובוט העוקב אחרי משהו) בנויה כמצויר:



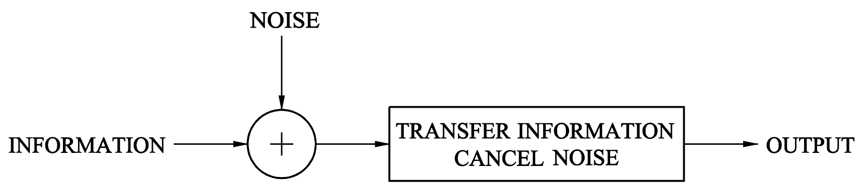
ציור 1.5:

אור הכוכב נופל על התא הפוטואלקטרי. באם אור הכוכב אינו נופל על מרכז התא (שגיאה אפס), נוצר אות שגיאה חיובי או שלילי בהתאם לכוון השגיאה. אות השגיאה מוגבר ומפעיל את המנוע בכוון לאפוס השגיאה. מערכת שבאופן עקרוני דומה למערכת כזו מופיעה במערכות עקיבה של מכ"ם בטווח ובזוית וכן במעגל הנקרא "חוג נעול פאזה" המשמש למטרות שונות וחשובות ביותר בקומוניקציה. נחזור למערכת העקיבה המצוירת ונניח שהיא נמצאת על פלטפורמה מתנדנדת ולכן יש שגיאות עקיבה אחרי הכוכב בשל תנודות הפלטפורמה. כמו כן יש רעש שנוצר בתא הפוטואלקטרי שאף הוא גורם שגיאות עקיבה. המטרה היא עקיבה נאמנה אחרי הכוכב. נניח שיש לנו בקרת מהירות, כלומר, מהירות המנוע פרופורציונלית לזוית השגיאה בין ציר הטלסקופ והכוכב (בכוון אפוס השגיאה). נניח שהמערכת יציבה: האם רצוי הגבר גדול או קטן? יש כאן דרישות מנוגדות בין הדרישה לעקיבה נאמנה (אפוס מהיר של שגיאות העקיבה בגלל תנודות הפלטפורמה), הדורשת הגבר גבוה, לבין הקטנת שגיאות העקיבה בשל רעש התא הפוטואלקטרי, הדורשת הגבר נמוך. איך נאפיין את הסיגנל? במקרה שאנו עוסקים בו הסיגנל הוא תנועת הכוכב, או תנועת הפלטפורמה, או תנועת הגוף שעליו נעול הרובוט. זה נראה די טבעי ומובן מאליו שרעש טעון אפיון הסתברותי. מה שאולי פחות מובן מאליו הוא שגם הסיגנל הרצוי (האינפורמציה) הוא בעצם גודל אקראי.

בעבר הרחוק יותר השתמשו במודל מאד נאיבי לסיגנל ורעש בקומוניקציה. במודל זה, סינוס מתאר את האות הרצוי (האינפורמציה) וסינוס אחר (או תהליך אקראי) מתאר את הרעש. מודל זה אינו מספיק וכיום הן בקומוניקציה והן בבקרה רואים הן את הסיגנל הרצוי והן את הרעש, כל אחד כתהליך אקראי. בקומוניקציה יהיה, לכן, המודל הטבעי:

1.2 מהלך ההרצאות

<u>חלק א':</u> מבוא וחזרה על הסתברות.



ציור 1.6:

חלק ב': בחלק זה נעסוק בבעיות כגון הבעיה הבאה: נתונות המדידות Y_1 ו- Y_2,

$$Y_1 = X + n_1 \qquad ; \qquad Y_2 = X + n_2$$

X מ"א שרוצים לדעת את ערכו, n_1, n_2 רעשים; אין לנו גישה ישירה ל-X ואנו יודעים את Y_1, Y_2 בלבד. המטרה היא לתת "ניחוש חכם" ל-X על סמך ידיעת Y_1, Y_2 (וחוקי ההסתברות של X, n_2, n_1). בעיות מסוג זה נקראות שערוך.

חלק ג': בחלק השלישי (והעיקרי) של הקורס נעסוק בתהליכים אקראיים, אפיונם ומעברם דרך מערכות לינאריות.

חלק ד': בחלק זה נעסוק באפיון הרעש הפיזיקלי הבסיסי - רעש הנגד, רעש הדיודה, אפיון רעש מגברים ותכונות הרעש של שרשרת מגברים.

1.3 חזרה על הסתברות

מרחב הסתברות

* מרחב המדגם $\Omega = \{\omega\}$: אוסף כל התוצאות האפשריות של ניסוי. דוגמא: במקרה של קוביה מרחב המדגם הוא מרחב עם 6 אלמנטים. במקרה של רולטה: אוסף הנקודות על היקף מעגל היחידה. במקרה של רעש היציאה ממגבר בקטע הזמן $[0,1]$, מרחב מדגם יכול להיות אוסף כל הפונקציות הרציפות בקטע הזמן $[0,1]$.

* מרחב המאורעות $F = \{A\}$: אוסף של תת קבוצות של Ω. יהיה $\omega_1 \in \Omega$, כלומר, ω_1 מדגם מסוים. בדוגמאות לעיל, עבור המקרה של קוביה אנו יכולים ליחס ל-ω_1 הסתברות שונה מאפס. בשתי הדוגמאות האחרות ההסתברות להופעת ω_1 מסוים היא אפס ולכן את ההסתברות עלינו ליחס למאורעות ולא לדגמים. לדוגמא נוכל לשאול עבור הרעש מהמגבר מה ההסתברות ש - $\int_0^1 n^2(t)dt < 3$. על מנת שאפשר יהיה לבנות תורה מבוססת של הסתברות יש לדרוש שאוסף תת הקבוצות של Ω, $F = \{A\}$ יקיים את התנאים הבאים:

1. $\Omega \in F$, כלומר תת הקבוצה Ω של Ω היא מאורע.
2. אם A_1, A_2 מאורעות (ז.א. $A_i \in F$) אזי גם $A_1 \cup A_2$ מאורע ($A_1 \cup A_2 \in F$).
3. אם A_i שייך ל-F אזי גם A_i^c שייך ל-F.
4. איחוד ניתן להמנות של מאורעות אף הוא מאורע, כלומר, אם $A_i \in F$ $i = 1, 2, \cdots$ גם $\cup A_i \in F$.

הערה: מאורעות A, B $(A \in F, B \in F)$ המקיימים $A \cap B = \emptyset$ נקראים מאורעות זרים.

* <u>פונקצית הסתברות</u> $P(A)$: עבור כל מאורע A, כאשר $A \in F$, $P(A)$ מקיים: $0 \le P(A) \le 1$; $P(A^c) + P(A) = 1$
אם נתון $A_1, A_2, \cdots$ $A_i \cap A_j = \emptyset$ עבור כל $i \neq j$ אזי $P(\cup A_i) = \sum_i P(A_i)$; כמו כן $P(\Omega) = 1$.

* השלשה $\{\Omega, F, P\}$ נקראת מרחב הסתברות.

<u>משתנה אקראי</u> משתנה אקראי (מ"א) הוא פונקציה $\omega \in \Omega, X(\omega)$ על מרחב הדגמים כך שעבור כל a ממשי $\{\omega : X(\omega \le a\}$ הוא מאורע.

דוגמאות של מ"א על רעש המגבר (נניח שרעש היציאה מהמגבר, $n(t)$ הוא צורת גל רציפה)

1. $n(0.25) = n(t)|_{t=0.25}$ (הרעש ברגע $t = 0.25$).
2. $X = \int_0^1 n^2(t)dt$ (האנרגיה של אות הרעש בתחום הזמן $[0,1]$).
3. $Y = \max_{t\in[0,1]} n(t)$ (הערך המכסימלי של אות הרעש בתחום הזמן $[0,1]$).

הטענות בדוגמאות הן אינטואיטיביות והן טעונות הוכחה (ההוכחה מתבססת על רציפות הדגמים של הרעש).

פונקצית פילוג

פונקצית הפילוג של המשתנה האקראי $X(\omega)$ מגדירה את חוק ההסתברות של $X(\omega)$:

$$F_X(a) = \mathrm{Prob}\{X(\omega) \le a\}$$

מוגדרת תמיד כפונקציה של a (X הוא רק אינדקס), היא מונוטונית (לא יורדת), וניתן להוכיח שהיא רציפה מימין. בגבולות מקבלת פונקצית הפילוג את הערכים הבאים: $F_X(-\infty) = 0, F_X(+\infty) = 1$. עבור מ"א בדיד $F_X(a)$ עולה בקפיצות (צייר את $F_X(\alpha)$ עבור קובית משחק). אם קיים $f_X(\alpha)$ כך ש- $F_X(\alpha) = \int_{-\infty}^{\alpha} f_X(\theta)d\theta$ אזי $f_X(\alpha)$ נקראת פונקצית הצפיפות או הפילוג הסגולי (פ"ס) של X. מהגדרת פונקצית הפילוג נובע מיד שעבור $a_1 < a_2$, מתקיים $\mathrm{Prob}\{a_1 < X \le a_2\} = F_X(a_2) - F_X(a_1)$, (כי המאורע $\{X \le a_2\}$ הוא איחוד של המאורעות הזרים $\{a_1 < X \le a_2\}$ ו- $\{X \le a_1\}$).

תוחלת

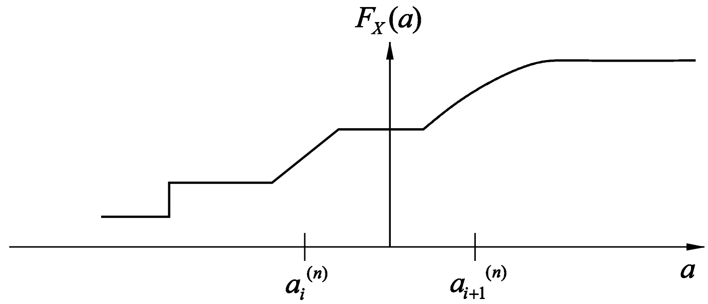
עבור מ"א בדיד: $m_x = \overline{X} = E[X] = \sum_i \alpha_i \mathrm{Prob}\{X_i = \alpha\}$.

עבור מ"א עם פ"ס: $m_X = \overline{X} = E[X] = \int_{-\infty}^{\infty} \alpha f_X(\alpha)d\alpha$, וזאת בהנחה ש- $\int_{-\infty}^{\infty} |\alpha| f_X(\alpha)d\alpha < \infty$. באופן כללי עבור כל n, תהיה $\alpha_i^{(n)}$; $i = 1, 2, \cdots, m$ (כאשר m תלוי ב- n) חלוקה סופית של הקטע $(-n, n)$; $\alpha_i^{(n)} < \alpha_{i+1}^{(n)}$.
נניח שמתקיים $\max_i |\alpha_{i+1}^{(n)} - \alpha_i^{(n)}| \underset{n\to\infty}{\longrightarrow} 0$ (ז.א. ככל ש- n גדל החלוקה נעשית עדינה יותר).

נגדיר:

$$E[X]=\int_{-\infty}^{\infty}\alpha dF_X(\alpha)=\lim_{n\to\infty}\sum_i \alpha_i^{(n)}\left[F_X(\alpha_{i+1}^{(n)})-F_X(\alpha_i^{(n)})\right]$$

בדוק שעבור המקרים הקודמים (בדיד, פ"ס) ההגדרה האחרונה אכן נותנת את התוצאה הנכונה.



ציור 1.7:

יהיה $Y=g(X)$: אזי לפי ההגדרה $E[Y]=E[g(X)]=\int_{-\infty}^{\infty}\alpha dF_Y(\alpha)$. טענה חשובה ללא הוכחה: $E[Y]=\int_{-\infty}^{\infty}g(\alpha)dF_X(\alpha)$ דהיינו, במקרה זה, על מנת לחשב את התוחלת של Y - אין צורך לחשב קודם את פונקצית הפלוג של Y ומתוך פילוג זה לחשב את EY. אפשר לחשב את EY ישירות מתוך ידיעת פונקצית הפילוג של X (הדרך השניה היא כמעט תמיד יותר קצרה).

<u>הערה:</u> לא לכל מ"א יש תוחלת. לדוגמא, עבור מ"א עם פ"ס

$$f_X(\alpha)=\begin{cases}0, & \alpha\geq 0\\ \dfrac{2/\pi}{1+\alpha^2}, & \alpha<0\end{cases}.$$

מתקיים: $\int_{-\infty}^{\infty}\alpha f_X(\alpha)d\alpha=\infty$.

<u>תכונות פשוטות של התוחלת</u>

א. יהיה A מאורע ונסמן ב- I_A מ"א המקבל ערכים 0 או 1 בלבד כדלקמן:

$$I_A=\begin{cases}1 & \omega\in A\\ 0 & \omega\in A^c\end{cases}$$

I_A נקרא פונקצית האינדיקטור של המאורע A ומתקיים:

$$EI_A=\int\alpha dF(\alpha)=0\cdot\mathrm{Prob}(A^c)+1\cdot\mathrm{Prob}(A)=\mathrm{Prob}(A)$$

ב. אם $X=\text{const}$, כלומר X דטרמיניסטי או במילים אחרות משתנה אקראי מנוון, אזי מתקיים:
$E[\text{constant}]=\text{constant}$.

ג. לינאריות: X,Y מ"א; b,a קבועים. אזי: $E[aX+bY]=aE[X]+bE[Y]$ (ולא חשוב אם X,Y תלויים אם לאו).

ד. מונוטוניות: אם $X(\omega)\geq Y(\omega)$ לכל ω אזי $EX\geq EY$.

<u>מומנטים מסדר גבוה יותר</u>

מומנט שני: $E[X^2]$

מומנט מסדר n: $E[X^n]$

מומנט מוחלט מסדר n: $E[|X|^n]$

מומנט מרכזי מסדר n: $E[(X-\overline{X})^n]$ (במיוחד $E[(X-\overline{X}=0$)

שונות, וריאנס: $E[(X-\overline{X})^2]=\text{Var}(X)$

סטית התקן: $\sigma_X=\sqrt{\text{Var}(X)}$

הגדרה: X מ"א מנורמל אם $E[X^2]=1, E[X]=0$ ולכן אם $Z=\frac{X-E[X]}{\sigma_X}$ אזי Z מ"א מנורמל.

עבור משתנה אקראי בודד, חוק ההסתברות $(F_X(a),\ a\in\mathbb{R})$, הוא "כל מה שאפשר להגיד על משתנה אקראי זה", בעקבות זאת אפשר לדבר על $EX, EX^2\cdots$. בכוון ההפוך, אם נתונים $EX, EX^2, EX^3,\cdots$ ואם $E^{1/n}|X|^n$ אינו עולה מהר מדי אזי אפשר להראות שהמומנטים מגדירים את חוק ההסתברות של X. לצורך הרבה בעיות טכניות חוק ההסתברות לא ידוע ולא מעניין; מספיק לדעת את EX ואת EX^2 (שני המומנטים הראשונים שבדרך כלל אינם מגדירים את חוק ההסתברות) היכולים לתת תמונה כללית ולאפשר פתרון הבעיות. למשל, X מ"א, Y מ"א המתאר מדידה של X. אזי $E(X-Y)^2/EX^2$ נותן הערכה טובה על השגיאה.

<u>אי השוויון של צ'ביצ'ב</u>

לפי תכונות א' וד' לעיל,

$$\begin{aligned}E[X^2] &\geq E\left[X^2\,1_{|X|\geq\varepsilon}\right]\\ &\geq \varepsilon^2 E\,1_{|X|\geq\varepsilon}\\ &= \varepsilon^2 P(|X|\geq\varepsilon)\end{aligned}$$

ומכאן:

$$P(|X|\geq\varepsilon)\leq\frac{EX^2}{\varepsilon^2}\,.$$

אם נקח $X=Y-EY$ אזי $EX^2=\text{Var}\,Y$ ומכאן:

$$P(|Y-EY|\geq\varepsilon)\leq\frac{\text{Var}\,Y}{\varepsilon^2}$$

הפונקציה האופינית של מ"א

הפונקציה האופינית של מ"א X, $\Phi_X(\nu)$, מוגדרת ע"י:

$$\Phi_X(\nu) = E[e^{i\nu X}] = E[\cos X + i\sin\nu X]$$

כאשר למשתנה האקראי X יש פילוג סגולי $f_X(\alpha)$ אזי

$$\Phi_X(\nu) = \int_{-\infty}^{\infty} e^{i\nu b} f_X(\alpha) d\alpha$$

דהיינו $\phi_X(\nu)$ הוא הצמוד הקומפלקסי של התמרת פוריה של $f_X(\alpha)$. אפשר להראות שבכל מקרה, הפונקציה האופינית $\phi_X(\cdot)$ מגדירה חד משמעית את פונקצית הפילוג $F_X(\cdot)$.

יתכן ש-$E[X^n]$ לא קיים, אולם $\Phi_X(\nu)$ תמיד קיים כי $E[|\cos\nu X|]$ ו-$E[|\sin\nu X|]$ תמיד קיים (חסימות הסינוס והקוסינוס). בהנחות מתאימות נוכל לפרק את $e^{i\nu X}$ לטור חזקות ולהחליף את סדר הסיכום עם התוחלת ונקבל

$$\Phi_X(\nu) = \sum_k \frac{(i\nu)^k}{k!} E[X^k]$$

ושוב, בלי לבדוק תנאי קיום וכו', נגזור את הביטוי האחרון עבור $\Phi_X(\nu)$ לפי ν, ונקבל:

$$\phi_X(0) = 1; \qquad \left.\frac{\partial\phi_X(\nu)}{\partial\nu}\right|_{\nu=0} = i\,EX; \qquad \left(\frac{\partial^n\phi_X(\nu)}{\partial\nu^n}\right)_{\nu=0} = i^n EX^n$$

<u>טענה:</u> אם $Y = aX + b$ אזי $\phi_Y(\nu) = e^{i\nu b}\phi_X(a\nu)$

<u>הוכחה:</u>

$$\phi_Y(\nu) = E[e^{i\nu(aX+b)}] = e^{i\nu b}E[e^{i\nu aX}] = e^{i\nu b}\phi_X(\nu a)$$

שני משתנים אקראיים

אם Y, X מ"א

$$F_{X,Y}(a,b) = \text{Prob}\{X \le a, Y \le b\}$$

$$f_{X,Y}(a,b) = \frac{\partial^2 F}{\partial a \partial b}$$

$$E[g(X,Y)] = \iint_{-\infty}^{\infty} g(\alpha,\beta) f_{X,Y}(\alpha,\beta) d\alpha d\beta$$

ובמיוחד עבור $g(X,Y) = X^m Y^n$ נקבל את המומנטים המשולבים של X עם Y.

שני מ"א Y, X נקראים בלתי תלויים (ב"ת) סטטיסטית אם עבור כל a, b ממשיים מתקיים

$$\text{Prob}\{X \le a, Y \le b\} = \text{Prob}\{X \le a\} \cdot \text{Prob}\{Y \le b\}$$

או

$$F_{X,Y}(a,b) = F_X(a)F_Y(b)\,.$$

ללא הוכחה: Y, X ב"ת אמ"ם לכל שתי פונקציות חסומות g.h מתקיים $Eg(X)h(Y) = Eg(X)Eh(Y)$ מכאן נובע: עבור Y, X בלתי תלויים ו- $Z = X + Y$ מתקיים

$$\phi_Z(\nu) = \phi_X(\nu) \cdot \phi_Y(\nu)$$

<u>הקווריאנס</u> של זוג מ"א Y, X מוגדר ע"י:

$$\mathrm{Cov}(X,Y) = E(X - \overline{X})(Y - \overline{Y})$$

אם $\mathrm{Cov}(X,Y) = 0$ אומרים שהמ"א Y ו-X <u>חסרי קורלציה</u> או <u>בלתי תלויים לינארית</u> (בת"ל).

<u>תרגיל:</u> הוכח שאי תלות סטטיסטית מחייבת אי תלות לינארית אבל ההיפך אינו בהכרח נכון.

<u>מקדם קורלציה</u>

עבור שני משתנים אקראיים X_2 ו-X_1 נניח למען הפשטות $EX_1 = EX_2 = 0$. עבור איזה λ הביטוי $E(X_1 - \lambda X_2)^2$ מינימלי? ע"י גזירה לפי λ נקבל ש - $\lambda^* = \frac{\mathrm{Cov}(X_1,X_2)}{\mathrm{Var}(X_2)}$ הוא האופטימלי. אם נציג את λ^* לביטוי עליו עשינו את המינימיזציה נקבל

$$E(X_1 - \lambda^* X_2)^2 = EX_1^2 - 2\lambda^*\,\mathrm{Cov}(X_1,X_2) + (\lambda^*)^2 EX_2^2 = EX_1^2 - \frac{(\mathrm{Cov}(X_1,X_2))^2}{\mathrm{Var}(X_2)} \geq 0$$

ומכאן

$$\mathrm{Var}(X_1)\,\mathrm{Var}(X_2) \geq (\mathrm{Cov}(X_1,X_2))^2$$

למקדם

$$\rho = \frac{\mathrm{Cov}(X_1,X_2)}{\sqrt{\mathrm{Var}(X_1)\,\mathrm{Var}(X_2)}}$$

קוראים מקדם הקורילציה ו- $|\rho| \leq 1$ לאור התוצאה האחרונה. אם X_1, X_2 בלתי תלויים לינארית, אזי $\rho = 0$. בהמשך נראה שביטוי זה מופיע בפתרון בעיות שנעסוק בהן.

1.4 וקטורים אקראיים

<u>סימונים:</u> $\underline{X}, \underline{a}$ יסמנו וקטורים. $\underline{\underline{B}}, \underline{\underline{A}}$ יסמנו מטריצות. יהיו $\underline{a}, \underline{b}$ וקטורים n מימדיים. אזי $\underline{b}^T\underline{a} = \sum a_i b_i$ הוא סקלר. לעומת זאת $\underline{a}\underline{b}^T$ היא מטריצה $n \times n$. מכפלת מטריצות: $\underline{\underline{A}}, \underline{\underline{B}}$ מתאפשר להכפיל: $\underline{\underline{A}} \cdot \underline{\underline{B}}$ כאשר $\underline{\underline{A}}$ היא מטריצה $m \times n$ (m שורות n עמודות) ו-$\underline{\underline{B}}$ היא מטריצה $n \times k$. המטריצה המתקבלת היא $m \times k$. לדוגמא:

$$\begin{pmatrix} a_1 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} b_1 & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_1 b_1 + a_{12} b_{21} + a_{13} b_{31} & a_1 b_{12} + a_{12} b_{22} + a_{13} b_{32} \\ a_{21} b_1 + a_{22} b_{21} + a_{23} b_{31} & a_{21} b_{12} + a_{22} b_{22} + a_{23} b_{32} \end{pmatrix}$$

יהי $\underline{X}$ וקטור אקראי בעל n רכיבים

$$\underline{X}=\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}\quad ;\quad \underline{X}^T=(X_1,X_2,\cdots,X_n)$$

* $i=1\cdots n, EX_i$ נקראים המומנטים מסדר ראשון.
* $i=1,2,\cdots,n, E|X_i|$ נקראים המומנטים המוחלטים מסדר ראשון.
* $1\le i\le n, 1\le j\le n, EX_iX_j$ נקראים המומנטים מסדר שני.
* $E(X_i-EX_i)(X_j-EX_j)$ עבור $1\le i\le n, 1\le j\le n$ נקראים המומנטים המרכזיים מסדר שני.

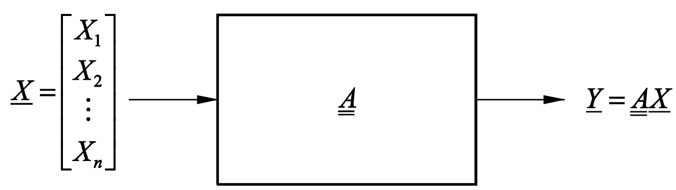
עבור וקטורים אקראיים, כאשר n גדול, חוק ההסתברות של וקטור $\underline{X}$

$$F_{X_1,X_2}\cdots,X_n(a_1,a_2,\cdots,a_n)=\text{Prob}\{X_1\le a_1,X_2\le a_2,\cdots,X_n\le a_n\}$$

(או בכתיב וקטורי $F_{\underline{X}}(\underline{\alpha})$) יכול להיות מורכב ביותר ולא ידוע. אולם לבעיות רבות מספיק לדעת את המומנטים מסדר ראשון ושני. אחת הבעיות שנעסוק בהן בהמשך היא הבעיה הבאה: נתונים המומנטים מסדר ראשון ומסדר שני של $\underline{X}$ והוקטור $\underline{Y}$ מוגדר כדלקמן:

$$\underline{Y}=\underline{\underline{A}}\,\underline{X}+\underline{b}$$

כאשר $\underline{\underline{A}}$ מטריצה קבועה, לא אקראית מספר העמודות של $\underline{\underline{A}}$ צריך כמובן להיות שווה למספר השורות (אלמנטים) של $\underline{X}$.



ציור 1.8:

האם מתוך הנתון אפשר למצוא את המומנטים מסדר ראשון ומסדר שני של $\underline{Y}$?

<u>תוחלות</u>

תוחלת וקטור אקראי מוגדרת כוקטור התוחלות, לאמר:

$$E\underline{X}=\begin{pmatrix}EX_1\\EX_2\\ \vdots\\EX_n\end{pmatrix}$$

טענה: אם $\underline{Y} = \underline{\underline{A}}\,\underline{X} + \underline{B}$ אזי $E\underline{Y} = E[\underline{\underline{A}}\,\underline{X} + \underline{b}] = \underline{\underline{A}}E\underline{X} + \underline{b}$.

הוכחה: התוצאה נובעת מהחשבון הישיר הבא:

$$E\left\{\begin{pmatrix} a_1 & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \\ \vdots & & & \vdots \\ a_{m1} & & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} X_1 \\ \\ \\ X_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \\ \\ b_n \end{pmatrix}\right\} = E\left\{\begin{pmatrix} a_1X_1 + a_{12}X_2 + & \cdots & +a_{1n}X_n + b_1 \\ & \vdots & \\ a_{m1}X_1 + a_{m2}X_2 + & \cdots & +a_{mn}X_n + b_m \end{pmatrix}\right\}$$

$$= \left\{\begin{pmatrix} a_1EX_1 + a_{12}EX_2 + & \cdots & +a_{1n}EX_n + b_1 \\ & \vdots & \\ a_{m1}EX_1 + a_{m2}EX_2 + & \cdots & +a_{mn}EX_n + b_m \end{pmatrix}\right\} = \underline{\underline{A}}\begin{pmatrix} EX_1 \\ \vdots \\ EX_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

הרחבה של הטענה האחרונה: תהיה $\underline{\underline{Y}}$ מטריצה אקראית(m שורות, n עמודות) עם אלמנטים אקראיים:

$$\underline{\underline{Y}} = \begin{pmatrix} Y_1 & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & & & \\ \vdots & & & \vdots \\ Y_{m1} & & \cdots & Y_{mn} \end{pmatrix} = [Y_{ij}]$$

נגדיר - $E\underline{\underline{Y}} = [EY_{ij}]$. תהיה $\underline{\underline{A}}$ מטריצה קבועה (לא אקראית). אזי קל לראות ש-$E\underline{\underline{A}}\,\underline{\underline{Y}} = \underline{\underline{A}}E\underline{\underline{Y}}$ וכן $E\underline{\underline{Y}}\,\underline{\underline{B}} = (E\underline{\underline{Y}}) \cdot \underline{\underline{B}}$.

מומנטים מסדר שני

נעבור כעת למומנטים מסדר שני וכאן נעסוק בוקטורים אקראיים בלבד. נעיין ב-$E[\underline{X} \cdot \underline{X}^T]$

$$E[\underline{X} \cdot \underline{X}^T] = E\begin{pmatrix} X_1 & (X_1, X_2, \cdots, X_n) \\ X_2 & \\ \vdots & \\ X_n & \end{pmatrix} = \begin{pmatrix} EX_1X_1 & EX_1X_2 & \cdots & EX_1X_n \\ EX_2X_1 & & & \\ \vdots & & & \\ EX_nX_1 & & & EX_nX_n \end{pmatrix}$$

שים לב, מתקבלת מטריצה $n \times n$ סימטרית! מטריצה זו נקראת מטריצת המומנטים מסדר שני. בצורה דומה נגדיר את מטריצת הקווריאנס:

$$\left\{E(\underline{X} - E\underline{X})(\underline{X} - E\underline{X})^T\right\} = \left\{\mathrm{Cov}(X_i, X_j)\right\} = \left\{E(X_i - \overline{X}_i)(X_j - \overline{X}_j)\right\}$$

הערה: עבור $n>1$, $\det(\underline{X}\cdot\underline{X}^T)$ תמיד אפס, אבל זה לא בהכרח נכון ש-$\det(E\underline{X}\cdot\underline{X}^T)=0$ לדוגמא, $EX_1=EX_2=0, \quad EX_1X_2=0, \quad i=1,2 \quad EX_i^2=1$.

נעיין כעת בבעיה הבאה: נניח שנתון ו"א $\underline{X}$ וידוע $E\underline{X}$ וכן $E[\underline{X}\cdot\underline{X}^T]$. יהיה $\underline{Y}$ מוגדר ע"י $\underline{Y}=\underline{\underline{A}}\underline{X}$, אזי כבר הראינו בסעיף הקודם ש-$E\underline{Y}=\underline{\underline{A}}E\underline{X}$. נשאלת השאלה מהי מטריצת המומנטים מסדר שני של $\underline{Y}$, כלומר,

$$E[\underline{Y}\cdot\underline{Y}^T]=?$$

נעיין בדוגמא:

$$\underline{Y}=\begin{pmatrix}Y_1\\ Y_2\end{pmatrix} \quad ; \qquad \underline{\underline{A}}=\begin{pmatrix}a_1 & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

במקרה זה:

$$\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}=\underline{\underline{A}}\underline{X}=\begin{pmatrix}a_1X_1+a_{12}X_2\\ a_{21}X_1+a_{22}X_2\end{pmatrix}$$

וחשבון ישיר נותן:

$$\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\end{pmatrix}^T=\begin{pmatrix}(a_1X_1+a_{12}X_2)^2 & (a_1X_1+a_{12}X_2)(a_{21}X_1+a_{22}X_2)\\ (a_1X_1+a_{12}X_2)(a_{21}X_1+a_{22}X_2) & (a_{21}X_1+a_{22}X_2)^2\end{pmatrix}$$

מהסתכלות בתוצאה האחרונה ברור שאפשר לחשב את המטריצה $E[\underline{Y}\cdot\underline{Y}^T]$ מתוך ידיעת $E[\underline{X}\cdot\underline{X}^T]$ אולם החשבון נראה די מיגע. מושג המטריצות והכפל של מטריצות מאפשר לנו לסכם את החשבון בצורה מאוד פשוטה:

$$E[\underline{Y}\underline{Y}]^T=E[\underline{\underline{A}}\underline{X}(\underline{\underline{A}}\underline{X})^T]=E[\underline{\underline{A}}\underline{X}\underline{X}^T\underline{\underline{A}}^T]=\underline{\underline{A}}E[\underline{X}\underline{X}^T]\underline{\underline{A}}^T$$

ובצורה דומה:

$$\Big(\mathrm{Cov}(Y_i,Y_j)\Big)=E\Big[(\underline{Y}-E\underline{Y})(\underline{Y}-E\underline{Y})^T\Big]=E\Big[(\underline{\underline{A}}\underline{X}-\underline{\underline{A}}E\underline{X})(\underline{\underline{A}}\underline{X}-\underline{\underline{A}}E\underline{X})^T\Big]$$
$$=\underline{\underline{A}}E\Big[(\underline{X}-E\underline{X})(\underline{X}-E\underline{X})^T\Big]\underline{\underline{A}}^T=\underline{\underline{A}}\Big[\mathrm{Cov}(X_i,X_j)\Big]\underline{\underline{A}}^T$$

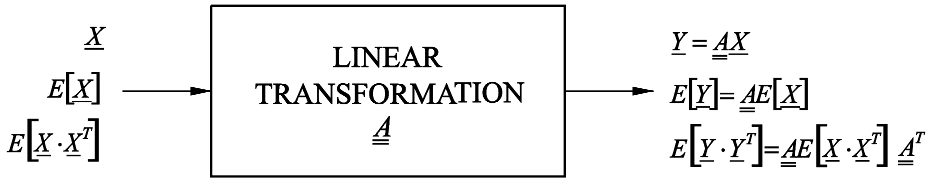
את התוצאות על המומנטים מסדר ראשון ושני כאשר מבצעים טרנספורמציה ליניארית על וקטור אקראי נסכם בציור הבא:

תכונה בסיסית של מטריצת המומנטים מסדר שני

מטריצה ריבועית סימטרית C, $(n\times n)$, נקראת לא שלילית אם עבור כל וקטור n-מימדי $\underline{a}$, מתקיים:

$$\underline{a}^T\underline{\underline{C}}\underline{a}\geq 0$$

(שים לב ש-$\underline{a}^T\underline{\underline{C}}\underline{A}$ הוא סקלר) ואם, בנוסף, $\underline{a}^T\underline{\underline{C}}\underline{a}=0$ אך ורק כאשר כל רכיבי $\underline{a}$ מתאפסים אזי $\underline{\underline{C}}$ נקראת חיובית או חיובית מוגדרת.



ציור 1.9:

טענה: מטריצת המומנטים מסדר שני תמיד לא שלילית.

הוכחה: נקח $\alpha = \underline{a}^T\underline{X}$. אזי α הוא משתנה אקראי ולכן:

$$0 \le E\alpha^2 = E[\underline{a}^T\underline{X}\underline{a}^T\underline{X}] = E[\underline{a}^T\underline{X}\,\underline{X}^T\underline{a}] = \underline{a}^T E[\underline{X}\,\underline{X}^T]\underline{a}$$

הערה: כפי שכבר הערנו, $\underline{X}\,\underline{X}^T$ הינה תמיד מטריצה סינגולרית אבל $E[\underline{X}\,\underline{X}^T]$ אינה בהכרח סינגולרית. על משמעות המקרה שבו $E[\underline{X}\,\underline{X}^T]$ סינגולרית נעמוד בהמשך. אפשר להראות שאם $E[\underline{X}\,\underline{X}^T]$ אינה סינגולרית אזי היא חיובית מוגדרת.

פונקציה אופינית של וקטור אקראי

יהי $\underline{X}$ וקטור אקראי ו-$\underline{\nu}$ וקטור דטרמיניסטי.

$$\underline{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad ; \quad \underline{\nu} = \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix}$$

נגדיר

$$\phi_{\underline{X}}(\underline{\nu}) = E\, e^{i\sum_{j=1}^n X_j\nu_j} = E\, e^{i\underline{X}^T\underline{\nu}} = E\, e^{i\underline{\nu}^T\underline{X}} = E\, e^{i(\underline{X},\underline{\nu})}$$

כאשר $(\underline{X},\underline{\nu})$ מסמן את המכפלה הסקלרית של שני הוקטורים. חוק ההסתברות של $\underline{X}$ מגדיר את $\phi_{\underline{X}}(\nu)$. בכוון ההפוך, ללא הוכחה, $\phi_{\underline{X}}(\nu)$ עבור כל $\underline{\nu}\in\mathbb{R}^n$, מגדיר את חוק ההסתברות של הוקטור האקראי $\underline{X}$. קל לראות שאם כל הרכיבים של $\underline{X}$ בלתי תלויים אזי

$$\phi_{\underline{X}}(\underline{\nu}) = \prod_{i=1}^n \phi_{X_i}(\nu_i)$$

ולהפך, אפשר להראות שאם הפונקציה האופינית נתנת לייצוג כמכפלה כנ"ל אזי רכיבי הוקטור $\underline{X}$ הינם בלתי תלויים.

שאלה: ידוע $\phi_{\underline{X}}(\underline{u})$, $\underline{u}\in\mathbb{R}^n$. נבצע את הטרנספורמציה המוכרת $\underline{Y}=\underline{\underline{A}}\underline{X}$. מה נוכל לומר על $\phi_{\underline{Y}}(\underline{\nu})$? תשובה:

$$\phi_{\underline{Y}}(\underline{\nu}) = E\, e^{i\underline{Y}^T\underline{\nu}} = E\, e^{i(\underline{\underline{A}}\underline{X})^T\underline{\nu}} = E\, e^{i\underline{X}^T\underline{\underline{A}}^T\underline{\nu}} = \phi_{\underline{X}}(\underline{\underline{A}}^T\underline{\nu})$$

לדוגמא נראה טכניקה להעלמת משתנה: נתון הפלוג של X_1, X_2, X_3 , $F_{X_1,X_2,X_3}(a_1,a_2,a_3)$. מעונינים רק ב-$X_1, X_2$, אזי

$$F_{X_1,X_2}(a_1,a_2) = F_{X_1,X_2,X_3}(a_1,a_2,\infty)$$

עבור הפונקציה האופינית נקבל:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ \\ 0 & 1 & 0 \end{pmatrix} \quad ; \quad n = 3$$

$$\phi_{\underline{Y}}(\nu_1, \nu_2) = \phi_{\underline{X}}\left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \nu_1 \\ \\ \nu_2 \end{pmatrix} \right) = \phi_{\underline{X}}(\nu_1, \nu_2, 0)$$

ולכן אם רוצים להעלים משתנה כגון X_i אזי לוקחים את $\phi_{\underline{X}}(\underline{\nu})$ ומציבים בו $\nu_i \equiv 0$.

דוגמא נוספת:

$$Y = \underline{a}^T \underline{X} = a_1 X_1 + \cdots a_n X_n = (a_1, \ldots, a_n) \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$$

במקרה זה Y ו-ν הם חד מימדיים

$$\phi_Y(\nu) = \phi_{\underline{X}}\left(\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \nu \right) = \phi_{\underline{X}} \begin{pmatrix} a_1 \nu \\ \vdots \\ a_n \nu \end{pmatrix}$$

ובמיוחד, אם הרכיבים של $\underline{X}$ ב"ת ו-$a_1 = a_2 = \cdots = a_n = 1$ אזי

$$\phi_Y(\nu) = \phi_{X_1}(\nu) \phi_{X_2}(\nu) \cdots \phi_{X_n}(\nu)$$

<u>מסקנה:</u> הפונקציה האופינית של סכום מ"א ב"ת היא מכפלת הפונקציות האופיניות.

<u>הערה:</u> יהיו Y, X מ"א.

$$E\left[e^{iX\nu_1 + iY\nu_2}\right] = E\left\{ \sum_{m=0}^{\infty} \frac{(i\nu_1)^m}{m!} X^m \cdot \sum_{k=0}^{\infty} \frac{(i\nu_2)^k}{k!} Y^k \right\}$$

ולכן אם נפתח את $\phi_{X,Y}(\nu_1, \nu_2)$ לטור חזקות ב-ν_1, ν_2 נוכל לקבל את המומנטים $EX^m Y^k$ (בתנאי שהם קיימים).

1.5 הפילוג הגאוסי

<u>המקרה הסקלרי</u>

יהא X מ"א גאוסי. כזכור, הפילוג הגאוסי החד מימדי מוגדר ע"י:

$$f_X(\alpha) = \frac{1}{\sqrt{2\pi\sigma^2}} \, e^{-\frac{(\alpha - m)^2}{2\sigma^2}} \tag{1.1}$$

$$E[X]m, \quad \mathrm{Var}(X) = \sigma^2$$

והפונקציה האופינית (ללא הוכחה) של מ"א גאוסי נתונה ע"י:

$$\phi_X(\nu) = e^{im\nu - \frac{1}{2}\sigma^2\nu^2} \tag{1.2}$$

(מה בנוגע למקרה המנוון $\sigma^2 = 0$, כלומר $X \equiv m$?). עבור $Y = aX + b$

$$\phi_Y(u) = E\, e^{iauX+ibu} = e^{ibu} e^{iaum - \frac{1}{2}a^2u^2\sigma^2} = e^{iu(b+am) - \frac{1}{2}a^2u^2\sigma^2}$$

ואם נסכים להרחיב את משפחת המשתנים האקראים הגאוסים ע"י הכללת כל המ"א המנוונים ($\sigma^2 = 0$) לתוך המשפחה אזי נקבל את התוצאה שטרנספורמציה לינארית של מ"א גאוסי היא תמיד מ"א גאוסי (ללא צורך בהוספת דרישות מיוחדות). כפי שנראה בהמשך נוח וכדאי לעשות הכללה זו ולכן נגדיר את המ"א לפי (1.1) בצרוף המקרה המנוון, או ישירות לפי (1.2).

המקרה הוקטורי

הגדרה: וקטור אקראי (ו"א) $\underline{X}$ n-מ ימדי נקרא גאוסי אם עבור כל וקטור לא אקראי n מימדי $\underline{a}$, המשתנה האקראי $\underline{a}^T\underline{X} = (\underline{a}, \underline{X})$ הוא מ"א גאוסי (אפשר גם לדרוש $|\underline{a}| = 1$; זה לא ישנה דבר. אז נוכל להגיד שו"א הוא גאוסי אם כל השלכה שלו לכל כוון הוא מ"א גאוסי).

אם ו"א $\underline{X}$ הוא גאוסי אזי כל רכיב שלו הוא השלכה ולכן כל רכיב הוא מ"א גאוסי. מה עם ההפך? אפשר להביא דוגמא של וקטור אקראי שבמערכת קואורדינטות מסוימת כל הרכיבים הם גאוסיים אולם הוקטור האקראי איננו וקטור אקראי גאוסי! ולכן, אם נתון וקטור אקראי שכל רכיביו גאוסיים זה עדיין לא מחייב שהוקטור הוא וקטור אקראי גאוסי.

תכונות של וקטורים אקראיים גאוסיים:

א. טענה: אם $\underline{X}$ ו"א גאוסי, אזי (1) גם $\underline{X} + \underline{b}$ ו"א גאוסי; (2) גם $\underline{Y} = \underline{\underline{A}}\underline{X}$ ו"א גאוסי.

הוכחת (2):

$$(\underline{a}, \underline{Y}) = (\underline{a}, \underline{\underline{A}}\underline{X}) = \underline{a}^T\underline{\underline{A}}\underline{X} = (\underline{\underline{A}}^T\underline{a})^T\underline{X}$$

ולכן כל השלכה של $\underline{Y}$ בכיוון מסוים הוא השלכה של $\underline{X}$ (לא בהכרח באותו כיוון כמובן), ולכן $\underline{Y}$ ו"א גאוסי.

ב. טענה: $\underline{X}$ ו"א גאוסי n-מימדי אם ורק אם קיים וקטור n-מימדי $\underline{m}$ ומטריצה $\underline{\underline{\Lambda}}$ סימטרית ואי שלילית $n \times n$ כך שהפונקציה האופינית של $\underline{X}$ נתונה ע"י:

$$\begin{aligned}\phi_{\underline{X}}(\underline{\nu}) &= e^{i(\underline{\nu}^T\underline{m}) - \frac{1}{2}\underline{\nu}^T\underline{\underline{\Lambda}}\underline{\nu}} \\ &= e^{i\sum \nu_i m_i - \frac{1}{2}\sum_i\sum_j \nu_i\nu_j\lambda_{ij}}\end{aligned} \tag{***}$$

ואז מתקיים

$$E\underline{X} = \underline{m}$$

$$\underline{\underline{\Lambda}} = E(\underline{X} - E\underline{X})(\underline{X} - E\underline{X})^T = [\lambda_{ij}] = \Big[\mathrm{Cov}(X_i, X_j)\Big]$$

<u>הוכחה:</u> נניח ש-$\underline{X}$ הוא ווקטור גאוסי. נסמן את ווקטור הממוצעים ואת מטריצת הקווריאנס ב-

$$\underline{m} \doteq \mathbb{E}[\underline{X}]$$
$$\underline{\underline{\Lambda}} \doteq \mathbb{E}\left[(\underline{X}-\underline{m})(\underline{X}-\underline{m})^T\right]$$

נסמן את אברי המטריצה ב-$\underline{\underline{\Lambda}} \doteq \{\lambda_{ij}\}$. נבחר ווקטור $\underline{\nu}$ ונגדיר מ"א חדש $Y = \underline{\nu}^T\underline{X}$. מהגדרת ווקטור גאוסי נובע כי Y מ"א גאוסי. מלינאריות התוחלת נובע כי הממוצע שלו הוא

$$m_Y = \underline{\nu}^T\underline{m}$$

ומחישוב קודם של קווריאנסים נקבל שהווריאנס של Y הוא

$$\sigma_Y^2 = \underline{\nu}^T\underline{\underline{\Lambda}}\underline{\nu} = \sum_{i,j=1}^{n} \nu_i\nu_j\lambda_{ij}$$

כוון ש-Y גאוסי, הפונקציה האפיינית שלו בנקודה 1 היא

$$\phi_Y(1) = \mathbb{E}\left[e^{iY}\right] = e^{i\underline{\nu}^T\underline{m}-\frac{1}{2}\underline{\nu}^T\underline{\underline{\Lambda}}\underline{\nu}}$$

אולם מהגדרת Y,

$$\begin{aligned}\mathbb{E}\left[e^{iY}\right] &= \mathbb{E}\left[e^{\sum_{i=1}^{n}\nu_i X_i}\right]\\ &= \mathbb{E}\left[e^{\underline{\nu}^T\underline{X}}\right]\\ &\doteq \phi_X(\underline{\nu})\end{aligned}$$

כיוון שהחישוב נכון לכל $\underline{\nu}$ הוכחנו כי לווקטור גאוסי יש פונקציה אפיינית כדרוש, כאשר הווקטור $\underline{m}$ והמטריצה $\underline{\underline{\Lambda}}$ המופיעים בפונקציה האפיינית הם בדיוק ווקטור הממוצעים ומטריצת הקווריאנס.
להוכחת הכיוון השני, נתונים הווקטור $\underline{m}$ והמטריצה $\underline{\underline{\Lambda}}$. נבחר ווקטור $\underline{a}$ ומגדיר שוב

$$Y = \underline{a}^T\underline{X}.$$

מתכונות הפונקציה האפיינית,

$$\begin{aligned}\phi_Y(u) &= \phi_X(\underline{a}\cdot u)\\ &= \phi_X\begin{pmatrix}a_1\cdot u\\ a_2\cdot u\\ \vdots\\ a_n\cdot u\end{pmatrix}\\ &= e^{iu(\underline{a}^T\underline{m})-\frac{1}{2}u^2(\underline{a}^T\underline{\underline{\Lambda}}\underline{a})}\end{aligned}$$

כאשר השוויון האחרון נובע מההנחה על $\underline{X}$. מכאן נובע כי Y הוא מ"א גאוסי, וכיוון שהדבר נכון לכל $\underline{a}$, הרי מההגדרה $\underline{X}$ הוא וקטור אקראי גאוסי. מהוכחת הכוון הראשון אנו יודעים כי עבור ווקטור גאוסי, הווקטור $\underline{m}$ והמטריצה $\underline{\underline{\Lambda}}$ המופיעים בפונקציה האפיינית הם בדיוק ווקטור הממוצעים ומטריצת הקווריאנס.

ג. נניח ש-$\underline{\underline{\Lambda}}$ מטריצה אלכסונית, $\lambda_{ij} = \sigma_i^2 \delta_{ij}$ (כאשר $\delta_{ij} = 0$ עבור $j \neq i$ ו $\delta_{ij} = 1$ עבור כל i). אזי משמע שהמשתנים האקראיים $X_1, X_2, \cdots$ בלתי תלויים לינארית בזוגות, כלומר $i \neq j$, $\text{Cov}(X_i, X_j) = 0$ אם בנוסף $\underline{X}$ ו"א גאוסי, אזי

$$\phi_{\underline{X}}(\underline{\nu}) = e^{i \sum \nu_i m_i - \frac{1}{2} \sum \lambda_{jj} \nu_j^2} = \prod_{j=1}^{n} e^{i \nu_i m_i - \frac{1}{2} \lambda_{jj} \nu_j^2}$$

ולכן, תמיד נכון ש-$\underline{\underline{\Lambda}}$ אלכסונית אם ורק אם קיימת אי תלות לינארית בזוגות אולם במקרה הגאוסי מתקיימת גם אי תלות סטטיסטית, ולכן עבור וקטור אקראי גאוסי אי תלות ליניארית גוררת אי תלות סטטיסטית.

הערה: אם $\underline{X}$ הוא ו"א, שרכיביו הם בת"ס וכל אחד מהם הוא מ"א גאוסי, אזי $\underline{X}$ הוא ו"א גאוסי.

ד. שאלה: מתי m הרכיבים הראשונים של ווקטור אקראי גאוסי n-מימדי בלתי תלויים ב-$(n-m)$ הרכיבים הנותרים?

תשובה (ללא הוכחה): כאשר למטריצת הקווריאנס הצורה:

$$\underline{\underline{\Lambda}} = \begin{pmatrix} \frac{m \times m}{0} & \frac{0}{(n-m) \times (n-m)} \end{pmatrix}$$

ה. שאלה: אם $\underline{X}$ ו"א גאוסי, רכיבים ב"ת, $\underline{Y} = \underline{\underline{A}}\underline{X}$ אזי בדרך כלל הרכיבים של $\underline{Y}$ הם תלויים. מה בכוון ההפוך? התשובה נתונה ע"י:

טענה: אם $\underline{X}$ ו"א גאוסי n-מימדי (לשם פשטות נדון כאן במקרה של תוחלת אפס אבל זה כלל לא חשוב), אזי קיימת מטריצה לא סינגולרית $n \times n$, נקרא לה $\underline{\underline{D}}$ כך ש-$\underline{Y} = \underline{\underline{D}}\underline{X}$ והרכיבים של $\underline{Y}$ בלתי תלויים!

הערה: יתכן שחלק מהרכיבים של $\underline{Y}$ יהיו מנוונים (כלומר קבועים). למשל, $\underline{X} = \binom{X_1}{X_1}$ כאשר X_1 מ"א גאוסי ו $\underline{Y}$ מוגדר ע"י:

$$\underline{Y} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \sqrt{2} & \sqrt{2} \\ -\sqrt{2} & \sqrt{2} \end{pmatrix} \begin{pmatrix} X_1 \\ X_1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} X_1 \\ 0 \end{pmatrix}$$

אזי הרכיבים של $\underline{Y}$ בלתי תלויים. נסמן

$$\underline{\underline{D}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \quad ; \quad \det \underline{\underline{D}} = 1$$

שים לב ש-$\underline{\underline{D}}$ מקיימת

$$\underline{\underline{D}}^T \underline{\underline{D}} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ולכן $\underline{\underline{D}}^T = \underline{\underline{D}}^{-1}$. לצורך הוכחת הטענה, נקבל ללא הוכחה את המשפט הבא: אם $\underline{\underline{\Lambda}}$ מטריצה סימטרית לא שלילית אזי קיימת מטריצת אוניטרית ($\underline{\underline{D}}^T = \underline{\underline{D}}^{-1}$ ולכן כמובן לא סינגולרית) כך שהמטריצה $\underline{\underline{D\Lambda D}}^T$ היא אלכסונית, כלומר

$$\underline{\underline{D\Lambda D}}^T = \begin{pmatrix} c_1 & \dots & 0 \\ \vdots & c_2 & \vdots \\ 0 & \dots & c_n \end{pmatrix} \tag{1.3}$$

מיד נחזור ונעיין ב-(1.3) אולם לפני זה נוכיח שלוקטור האקראי $\underline{Y}$ המתקבל מתוך $\underline{Y} = \underline{\underline{D}}\underline{X}$ רכיבים ב"ת לינארית (זכור $E\underline{X} = 0$):

$$E[\underline{Y}\underline{Y}^T] = E[\underline{\underline{D}}\underline{X}\,\underline{X}^T\,\underline{\underline{D}}^T] = \underline{\underline{D\Lambda D}}^T = \begin{pmatrix} c_1 & \dots & 0 \\ \vdots & c_2 & \vdots \\ 0 & \dots & c_n \end{pmatrix}$$

כעת נוסיף את הנחת הגאוסיות ולכן אי תלות לינארית גוררת אי תלות סטטיסטית, ולכן הרכיבים של $\underline{Y}$ בלתי תלויים. (באופן כללי יותר:

$$E\, e^{i\underline{\nu}^T\underline{Y}} = E\, e^{i\underline{\nu}^T\underline{\underline{D}}\underline{X}} = e^{i\underline{\nu}^T\underline{\underline{D}}m_z - \frac{1}{2}\underline{\nu}^T\underline{\underline{D\Lambda D}}^T\underline{\nu}} = e^{i\underline{\nu}^T\underline{\underline{D}}m_z - \frac{1}{2}\sum_i \nu_i^2 e_i}$$

ולכן הרכיבים של $\underline{Y}$ בלתי תלויים גם אם התוחלת של $\underline{X}$ שונה מאפס.)

נחזור ל-(1.3): איך מוצאים את המטריצה $\underline{\underline{D}}$? נסתפק בהערה הבאה: נסמן $\underline{\underline{D}}^T = (\underline{d_1}, \underline{d_2}, \dots, \underline{d_n})$ כאשר $\underline{d_i}$ וקטור n-מימדי. אזי נוכל לרשום:

$$\underline{\underline{\Lambda D}}^T = (\underline{\underline{\Lambda}}\underline{d_1}, \underline{\underline{\Lambda}}\underline{d_2}, \dots, \underline{\underline{\Lambda}}\underline{d_n})$$

$$\underline{\underline{D}}^T \begin{pmatrix} c_1 & \dots & 0 \\ \vdots & c_2 & \vdots \\ 0 & \dots & c_n \end{pmatrix} = (c_1\underline{d_1}, c_2\underline{d_2}, \dots, c_n\underline{d_n})$$

לכן נובע מתוך (1.3), היות ו-$\underline{\underline{D}}$ אוניטרית:

$$\underline{\underline{\Lambda D}}^T = \underline{\underline{D}}^T \begin{pmatrix} c_1 & \dots & 0 \\ \vdots & c_2 & \vdots \\ 0 & \dots & c_n \end{pmatrix}$$

וע"י השוואת עמודות נקבל $\underline{\underline{\Lambda}}\underline{d_i} = c_i\underline{d_i}$. לכן, הוקטורים $\underline{d_i}$ פותרים את המשואה $\underline{\underline{\Lambda}}\underline{d} = c\underline{d}$, במילים אחרות c_i הם הערכים העצמיים של $\underline{\underline{\Lambda}}$ ו-$\underline{d_i}$ הם הוקטורים העצמיים של $\underline{\underline{\Lambda}}$ (ולכן גם מתקיים $\underline{\underline{D}}^T = \underline{\underline{D}}^{-1}$).

ו. ללא הוכחה: אם $\underline{\underline{\Lambda}}$ לא סינגולרית אזי הפילוג הסגולי ה-n- מימדי הגאוסי נתון ע"י:

$$f_{\underline{X}}(\underline{a}) = \frac{1}{(2\pi)^{n/2}(\det\underline{\underline{\Lambda}})^{1/2}}\, e^{\frac{1}{2}(\underline{a}-\underline{m})^T\underline{\underline{\Lambda}}^{-1}(\underline{a}-\underline{m})} = \frac{1}{(2\pi)^{n/2}(\det\underline{\underline{\Lambda}})^{1/2}}\, e^{-\frac{1}{2}\sum\sum(a_i-m_i)(a_j-m_j)\theta_{ij}}$$

כאשר: $\underline{\underline{\Lambda}}^{-1} = [\theta_{ij}]$

ז. <u>ללא הוכחה:</u> אם $\underline{Y} = \begin{pmatrix} \frac{X_1}{X_2} \end{pmatrix}$ ו"א גאוסי אזי חוק ההסתברות של $\underline{X_1}$ כאשר נתון $\underline{X_2}$ הוא גם כן גאוסי. (ואז, בנוסחת הפילוג הסגולי המותנה של $\underline{X_1}$ נתון $\underline{X_2}$ או בנוסחת הפונקציה האופינית המותנית, תופיע התוחלת המותנית של $\underline{X_1}$ נתון $\underline{X_2}$ והפיזור המותנה. מתברר שבמקרה זה הפיזור המותנה איננו תלוי בצורה מפורשת בהתניה).

<u>סיכום</u>

א. תכונת הגאוסיות אינוריאנטית לטרנספורמציות לינאריות.

ב. הפילוג של ו"א גאוסי $\underline{X}$ נקבע חד משמעית ע"י $\underline{\underline{\Lambda}}, \underline{m}$ (שני המומנטים הראשונים).

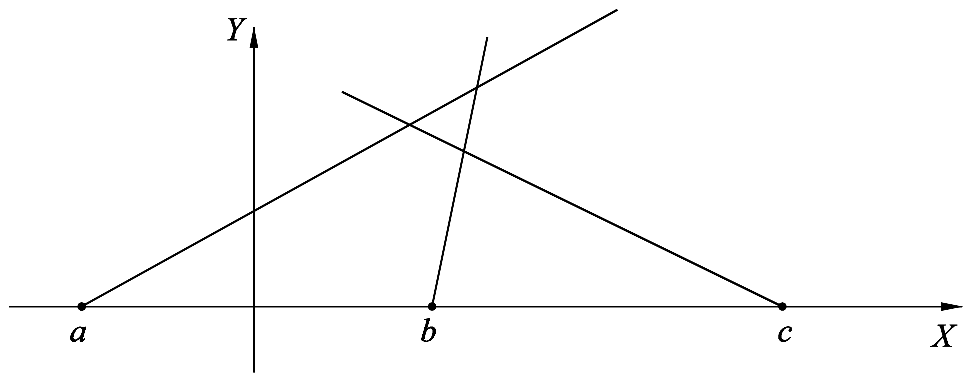
ג. אי תלות לינארית גוררת אי תלות סטטיסטית.

ד. ניתן לעבור למ"א ב"ת על ידי סבוב מערכת הצירים.

2 שערוך

2.1 מבוא

נעיין בבעיה של אתור קורן: משלוש הנקודות a, b ו- c נקבע כוון לקורן (שמיקומו אינו ידוע ורוצים לקבוע אותו).



ציור 2.1:

אם היו לנו רק שתי מדידות למשל, אחת מ-a ואחת מ-c , היינו קובעים את המיקום כנקודת החיתוך (זו אינה הקביעה הטובה ביותר תמיד אבל כרגע נתעלם מזה), אולם כאשר שלושת המדידות יוצרות שלוש נקודות חיתוך נשאלת השאלה איך לקבוע את המיקום המשוער של הקורן (השערוך), ובנוסף, מה יהיה סדר הגודל של שגיאת השערוך. בתרגיל המסכם (ראה סעיף 1.6) היו רק שתי מדידות ושם קבענו את החיתוך כמיקום המשוער, לכן הבעיה שנותרה לנו במקרה הנ"ל היתה הערכת השגיאה בלבד.

ניסוח כללי של בעית השערוך:

א. מצוי הוקטור האקראי $\underline{X}$, רוצים לשערך את המשתנה האקראי $\underline{Y}$, ו-$(Y\underline{X})$ הוא ו"א. משערך כל שהוא (טוב או רע) הוא פונקציה של המדידה $\phi(\underline{X})$.

ב. הגדרת קריטריון טיב $g(\cdot,\cdot)$: עבור כל משערך $\phi(\underline{X})$, נגדיר $E[g(Y,\phi(\underline{X}))] \triangleq$ השגיאה הממוצעת (לפי הקריטריון g). לדוגמא, עבור $g(a,b)=(a-b)^2$ השגיאה הממוצעת היא $E[(Y-\phi(\underline{X}))^2]$. דוגמא אחרת (תוחלת הערך המוחלט של השגיאה) $\delta(a,b)=|a-b|$ השגיאה הממוצעת היא $E[|Y-\phi(\underline{X})|]$.

ג. הבעיה: לאחר שהחלטת על קריטריון טיב מצא $\phi(\cdot)$ (פונקציה של n משתנים) כך שהשגיאה הממוצעת תהיה מינימלית.

כדי שניתן יהיה להפעיל את הגישה הזו יש להקפיד על שתי הנקודות הבאות: (1) קיים פילוג הסתברות מלכתחילה (a–priori) למיקום של הקורן; (2) קובעים קריטריון שגיאה ואז שואפים למצוא משערך אופטימלי על פי קריטריון

זה. לאחר מכן מחשבים את השגיאה הממוצעת (שאינה תלויה במיקום הקורן אלא ממוצעת ע"פ כל מיקומי הקורן לפי הפילוג מלכתחילה) של המשערך שנמצא.

לגבי (2) קיים העניין של בחירת קריטריון וקריטריון שונה יכול לתת משערך אופטימלי שונה. לגבי (1) קיימת נקודה עקרונית: בהרבה מקרים הגיוני שנוכל ליחס ל-Y פילוג מלכתחילה אולם אין הדבר כך בכל מקרה. למשל, התגלה כוכב לכת חדש ורוצים למדוד את המרחק אליו. איזה מובן ניתן ליחס לפילוג מלכתחילה של המרחק מאתנו לכוכב זה? מאידך, בבעיות אחרות (כגון בעיות קומוניקציה) יש מובן לפילוג האפריורי. כאשר אנו מניחים שיש מובן לקיום הסתברות מלכתחילה הבעיה נקראת בעיה בייסיאנית (ע"ש חוק Bayes).

לפני שנמשיך, נחזור על מספר מושגים שישמשו אותנו בהמשך.

<u>הסתברות מותנית</u>

Y, X מ"א; נניח X מ"א בדיד אזי הפילוג המותנה של Y כאשר נתון X היא

$$F_{Y|X}(\alpha|\beta) = \mathrm{Prob}\{Y \le \alpha | X = \beta\} = \frac{\mathrm{Prob}\{Y \le \alpha, X = \beta\}}{\mathrm{Prob}\{X = \beta\}}$$

ואם X מ"א רצוף

$$F_{Y|X}(\alpha|\beta) = \lim_{\varepsilon \to 0} \frac{\mathrm{Prob}\{Y \le \alpha\, ;\, \beta < X \le \beta + \varepsilon\}}{\mathrm{Prob}\{\beta < X \le \beta + \varepsilon\}}$$

<u>פילוג סגולי מותנה</u>

$$f_{Y|X}(\alpha|\beta) = \frac{\partial F_{Y|X}(\alpha|\beta)}{\partial \alpha} = \frac{f(\alpha, \beta)}{f(\beta)}$$

ובמיוחד, כאשר Y, X בלתי תלויים אזי

$$f_{Y|X}(\alpha|\beta) = f_Y(\alpha)$$

<u>תוחלת מותנית</u>

$$E[Y|X = \beta] = \int_{-\infty}^{\infty} \alpha f_{Y|X}(\alpha|\beta) \cdot d\alpha$$

ובאופן כללי יותר

$$E[Y|X = \beta] = \int_{-\infty}^{\infty} \alpha F_{Y|X}(d\alpha|\beta)$$

שיב לב שהתוחלת של Y מותנה ב-X היא פונקציה של β. אפשר להראות שגם במקרה זה קיים הקשר:

$$E[g(Y)|X = \beta] = \int_{-\infty}^{\infty} g(\alpha) f_{Y|X}(\alpha|\beta) d\alpha$$

ובאופן כללי יותר

$$E[g(Y)|X = \beta] = \int_{-\infty}^{\infty} g(\alpha) F_{Y|X}(d\alpha|\beta)$$

ושוב זו פונקציה של ההתניה.

אם $Y = Z_1 + Z_2$ אזי לא קשה להראות ש:

$$E[Y|X=\beta] = E[Z_1|X=\beta] + E[Z_2|X=\beta]$$

כמו כן, כאשר X ו Y בלתי תלויים, מתקיים: $E[Y|X=\beta] = E[Y]$.

מכאן שאם $Y = X + N$ כאשר X ו-N ב"ת אזי

$$E[Y|X=\beta] = E[X|X=\beta] + E[N|X=\beta] = \beta + E[N]$$

כאמור $E[Y|X=\beta]$ היא פונקציה של β, נסמן פונקציה זאת ב $\Psi(\beta)$. אם כעת נציב בפונקציה זאת את המשתנה האקראי X עצמו, תהיה התוצאה $\Psi(X)$ משתנה אקראי חדש. נהוג לסמן משתנה אקראי זה ב-$E[Y|X]$ דהיינו: $E[Y|X] = \Psi(X)$ כאשר $\Psi(\beta) = E[Y|X=\beta]$.

שים לב: כאשר X ו Y בלתי תלויים, $E[Y|X] = E[Y]$.

תוצאה חשובה ללא הוכחה: אם $(Y, \underline{X})$ ו"א אקראי גאוסי ($n+1$ מימדי) $EY = 0, E\underline{X} = 0$ אזי קיים וקטור n מימדי $\underline{a}$ כך ש:

$$E[Y|\underline{X}] = \underline{a}^T \underline{X}$$

הערה: שכנע את עצמך שמתוצאה זאת נובע, שאם $(Y, \underline{X})$ הוא וקטור אקראי גאוסי, ללא הדרישה של תוחלת אפסית, אזי קיים $\underline{a}$ כך ש:

$$E[Y|\underline{X} = E[Y] + \underline{a}^T(\underline{X} - E[\underline{X}])$$

(במידה וקשה לך להשתכנע, נסה תחילה לחשוב על המקרה בו $\underline{X}$ וקטור חד-מימדי).

2.2 שערוך אופטימלי

נחזור לבעית השערוך שהוצגה במבוא (נתרכז במקרה בו X הוא מ"א ולא ו"א); Y רצוי, X מצוי, קיים פילוג הסתברות משותף ל-Y ו-X. המשערך ל-Y כאשר נמדד X מסומן ב-$\hat{Y} = \phi(X)$. קובעים סיפרת טיב $g(\cdot,\cdot)$ ובעזרתה מגדירים את השגיאה הממוצעת $E[g(Y, \phi(X))]$. עבור קריטריון השגיאה הריבועית הממוצעת $g(a,b) = (a-b)^2$ הגודל ε המוגדר כ-$\varepsilon \triangleq Y - \phi(X)$ נקרא שגיאת השערוך, ε^2 הוא השגיאה הריבועית ו-$E[\varepsilon^2]$ הוא השגיאה הריבועית הממוצעת של המשערך. בכל מקרה מחפשים את $\phi_0(\cdot)$ שעושה מינימציה לשגיאה הממוצעת. מעתה והלאה נבחר בקריטריון השגיאה הריבועית הממוצעת ובכל מקום שלא נתיחס לקריטריון במפורש יהיה זה קריטריון השגיאה הריבועית הממוצעת.

טענה: המשערך $\phi_0(\cdot)$ שעבורו $E[(Y - \phi_0(X))^2]$ הוא מינימלי (על פני כל המשערכים $\phi(\cdot)$) הוא:

$$\hat{Y}_{\text{opt}} = \phi_0(\beta) = E[Y|X=\beta]$$

להוכחת טענה זו ניגש מיד לאחר הצגת שלוש דוגמאות שבהן מחושב המשערך האופטימלי במפורש.

דוגמא א' (סתם דוגמא טכנית):

$$f_{Y,X}(\alpha,\beta)=\alpha+\beta \quad ; \quad 0\le\alpha,\beta\le 1$$

$$f_X(\beta)=\int_0^1(\alpha+\beta)d\alpha=\frac{1}{2}+\beta$$

$$E(Y|X=\beta)=\int_0^1\alpha\frac{\alpha+\beta}{\frac{1}{2}+\beta}d\alpha=\frac{\frac{1}{3}+\frac{\beta}{2}}{\frac{1}{2}+\beta}$$

דוגמא ב': X,Y משתנים אקראיים גאוסים במשולב. נחשב את $E[Y|X]$.

נסמן:

$$X_c=X-E[X] \quad ; \quad Y_c=Y-E[Y]$$

ונבנה משתנה אקראי כדלקמן:

$$Z=Y_c-\frac{\text{Cov}(X,Y)}{\text{Var}(X)}X_c$$

מתוך הגדרת Z ברור ש-$E[Z]=0$ ו-$E[ZX_c]=0$. קל לראות ש-Z ו-X_c גאוסיים במשותף ולכן בלתי תלויים. מכאן ש $E[Z|X]=E[Z]=0$, ולכן:

$$E\left[Y_c-\frac{\text{Cov}(X,Y)}{\text{Var}(X)}X_c|X\right]=0$$

$$EZ\,X_c=E\,Y_c\,X_c-\frac{\text{Cov}(X,Y)}{\text{Var}(X)}EX_c\,X_c$$

$$E[Y|X]-E[Y]-\frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X-E[X])=0$$

מכאן:

$$E[Y|X]=E[Y]+\frac{\text{Cov}(X,Y)}{\text{Var}(X)}(X-E[X])$$

ולכן המשערך האופטימלי של Y מתוך X נתון על ידי:

$$\phi_0(\beta)=E[Y]+\frac{\text{Cov}(X,Y)}{\text{Var}(X)}(\beta-E[X])$$

שים לב לכך שהמשערך האופטימלי במקרה זה, הוא פונקציה לינארית של המדידות. (השווה עם השורות האחרונות של סעיף 2.1).

$\underline{X} = (X_1, \cdots, X_n)^T$, $(\underline{X}, Y)$ ו"א גאוסי.
נחשב את $E(Y|\underline{X})$.
נגדיר

$$\underline{X}_c = \underline{X} - E\,\underline{X}$$
$$Y_c = Y - E\,Y$$
$$Z = Y_c - \underline{X}_c^T \Lambda_X^{-1} E\,Y_c\,\underline{X}_c\,.$$

אזי $E\,Z = 0$ וכן $E\underline{X}_c Z = 0$. לכן $Z_1\underline{X}_c$ גאוסיים במשותף וב"ת.
מכאן $E(Z|\underline{X}) = EZ = 0$.
כלומר:

$$E(Y_c|\underline{X}) = E\Big[X_c^T \Lambda_X^{-1} E(Y_c \underline{X}_c)|\underline{X}\Big]$$
$$= X_c^T \Lambda_X^{-1}\, EY_c\,\underline{X}_c\,.$$

דוגמא ג': דוגמא פשוטה אולם חשובה וחשוב לזכור אותה כולל התוצאות המסומנות ב-*:

$$X = Y + n$$
$$E[Y] = E[n] = 0 \tag{2.1}$$
$$E[Y^2] = 1, \quad E[n^2] = N \tag{2.2}$$

Y, n גאוסיים ב"ת (Y הוא הסיגנל, n הרעש, מודדים את X כלומר את הסיגנל טבול ברעש ורוצים לשערך את הסיגנל הנקי Y).

הנחנו ש-Y, n גאוסיים ב"ת, ולכן הם גאוסיים במשותף (תוחלת אפס). לצורך דוגמא זאת לא נשתמש בתוצאות דוגמא ב', אלא במשפט שהופיע ללא הוכחה בסוף סעיף 2.1. לכן נוכל לרשום:

$$E[Y|X] = c_0 X$$

עבור קבוע לא ידוע c_0. על מנת למצוא את c_0, נבצע מינימיזציה על הביטוי

$$E\Big[(Y - cX)^2\Big] = E\Big[(Y - cY - cn)^2\Big] = (1-c)^2 + c^2 N$$

ומינימיזציה נותנת (ע"י גזירה והשוואה ל-0),

$$c_0 = \frac{1}{1+N} \tag{*}$$

הערה: במבט ראשון על (1) נראה שהשערוך הטוב ביותר של Y הוא X אבל (*) נותן תוצאה אחרת ובמבט שני הגיונית יותר; כאשר עוצמת הרעש נמוכה (N קטן), המשערך הטוב ביותר ל-Y הוא באמת המדידה; אך כאשר עוצמת הרעש גבוהה (N גדול), המדידה חסרת משמעות למעשה ולכן המשערך הטוב ביותר ל-Y הוא הממוצע שלו.

השגיאה הנותרת היא:

$$E\left[(Y-c_0X)^2\right]=\left(\frac{N}{1+N}\right)^2+\frac{N}{(1+N)^2}=\frac{N}{1+N}=\frac{1}{1+\frac{1}{N}} \qquad (*)$$

נחזור להוכחת הטענה לעיל.

<u>הוכחה:</u> נתחיל בהערות:

א) $E[X|X=\beta]=\beta$

ב) $E[h(Y)g(X)|X=\beta]=g(\beta)E[h(Y)|X=\beta]$

ובאופן כללי יותר: $E[h(Y,X)g(X)|X=\beta]=g(\beta)E[h(\beta,y)|X=\beta]$

ג) $E[U]=E[E[U|X]]$

<u>הערה:</u> תכונות (ב) ו-(ג) נקראות תכונות ההחלקה. מהן גם נובע הנסוח הבא של תכונת ההחלקה:

$$E\Big[h(X,Y)g(X)\Big]=E\Big[g(X)E[h(X,Y)|X]\Big] \qquad (+)$$

"מעין הוכחה" של (+):

$$\begin{aligned}E\Big[g(X)E[h(X,Y)|X]\Big]&=\int_{-\infty}^{\infty}g(\beta)\Big[\int_{-\infty}^{\infty}h(\beta,\alpha)f_{Y|X}(\alpha|\beta)d\alpha\Big]f_X(\beta)d\beta\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(\beta)h(\beta,\alpha)f_{Y|X}(\alpha|\beta)f_X(\beta)d\alpha d\beta\\&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(\beta)h(\beta,\alpha)f_{Y|X}(\alpha,\beta)d\alpha d\beta=E\Big[g(X)h(X,Y)\Big]\end{aligned}$$

כעת, להוכחת הטענה: נדרש להראות שלמשערך כל שהוא ϕ מתקיים: $E(\phi(X)-Y)^2\geq E(\phi_0(X)-Y)$. עבור משערך כל שהוא $\phi(\cdot)$ (כאשר $\phi_0(X)$ היא התוחלת המותנית של Y כאשר נתון X), מתקיים:

$$\begin{aligned}E\Big[(Y-\phi(X))^2\Big]&=E\Big[(Y-\phi_0(X)+\phi_0(X)-\phi(X))^2\Big]\\&E\Big[(Y-\phi_0(X))^2\Big]+E\Big[(\phi_0(X)-\phi(X))^2\Big]+2E\Big[(Y-\phi_0(X))(\phi_0(X)-\phi(X))\Big]\end{aligned}$$

נעיין באיבר האחרון $E[(Y-\phi_0(X))(\phi_0(X)-\phi(X))]$; נבצע קודם תוחלת מותנית (מותנה ב-X) ואח"כ תוחלת על X. לפי (ב) ו (ג) נקבל

$$E\Big[(Y-\phi_0(X))(\phi_0(X)-\phi(X))\Big]=E\Big[(\phi_o(X)-\phi(X))E[Y-\phi_0(X)|X]\Big]$$

ומתקיים

$$E\Big[Y-\phi_0(X)|X\Big]=E[Y|X]-\phi_0(X)=\phi_0(X)-\phi_0(X)=0$$

כלומר האיבר השלישי מתאפס. בנוסף, $E[(\phi_0(X)-\phi(X))^2]\geq 0$ ולכן:

$$E\Big[(Y-\phi(X))^2\Big]\geq E\Big[(Y-\phi_0(X))^2\Big]$$

ומכאן ברור ש-$\phi_0(X)$ הוא המשערך הטוב ביותר על פי קריטריון השגיאה הריבועית הממוצעת (אולי יש טוב כמוהו אולם אין טוב ממנו).

בכך השלמנו את הוכחת הטענה. שים לב שהתוצאה נשארת ללא שינוי גם אם במקום מ"א X יהיה ו"א $\underline{X}$.

נעבור כעת לדוגמא נוספת לחישוב מפורש של המשערך האופטימלי. דוגמא זו עוסקת בשערוך של תוצאת זריקת קוביה "הוגנת" $p_i=1/6$; $i=1,\cdots,6$ לפי קריטריון השגיאה הרבועית המינימלית.

א ללא מדידות המשערך האופטימלי הוא הממוצע, כלומר: $E[Y]=3.5$, והשגיאה הנותרת

$$E\Big[(Y-EY)^2\Big]=E[Y^2]-\Big(E[Y]\Big)^2=\frac{1}{6}(1+4+9+16+25+36)-\left(3\frac{1}{2}\right)^2=\frac{91}{6}-\left(3\frac{1}{2}\right)^2=2.72$$

ב אנו מקבלים אינפורמציה: מישהו מספר לנו אם התוצאה Y היא אחד מהשלשה 4, 5, 6 או ש-Y אחד מהשלשה 1, 2, 3. נגדיר מ"א X כדלקמן:

$$X=0\quad\text{if}\quad Y=4\quad\text{or}\quad 5\quad\text{or}\quad 6$$
$$X=1\quad\text{if}\quad Y=1\quad\text{or}\quad 2\quad\text{or}\quad 3$$

מתקיים:

$$\text{Prob}\{Y=1|X=0\}=0$$
$$\text{Prob}\{Y=4|X=0\}=1/3$$

וכנ"ל עבור $Y=2$ וכו'. לכן:

$$E[Y|X=0]=\frac{1}{3}(4+5+6)=5$$
$$E[Y|X=1]=2$$

לכן אם $X=1$ אזי $\hat{Y}_{\text{opt}}=2$ ואם $X=0$ אזי $\hat{Y}_{\text{opt}}=5$. השגיאה הנותרת תהיה:

$$E[\varepsilon^2]=E\Big[(Y-E[Y|])^2\Big]=E[Y^2]-E\Big[(E[Y|X])^2\Big]=\frac{91}{6}-\frac{1}{2}(5^2+2^2)=\frac{2}{3}$$

2.3 שערוך לינארי

בהרבה מקרים חישוב התוחלת המותנית קשה ביותר. בכל מקרה דורש חישוב זה ידע של חוק ההסתברות המשותף של $\underline{X}$ ו-Y. כאשר חוק זה אינו ידוע או כאשר חישוב התוחלת המותנית אינו אפשרי, לא ניתן למצוא את המשערך האופטימלי של Y בהינתן $\underline{X}$. במקרים כאלה מסתפקים בדרך כלל במשערך טוב פחות, אך ניתן לחישוב. להלן נתרכז במשפחת משערכים שהם פונקציה לינארית של הגדלים המתאימים.

המקרה הסקלרי

נתחיל מן המקרה הפשוט שבו X הוא מ"א ולא ו"א, כלומר השערוך מתבצע על סמך מדידה אחת. במקרה זה למשערך שהוא פונקציה לינארית של X יש באופן כללי את הצורה הבאה:

$$\hat{Y}^l = aX + b$$

כאשר a ו-b הם קבועים כלשהם.

את בעית מציאת המשערך הלינארי האופטימלי נגדיר כדלקמן:

בעיה: מצא משערך מהצורה $\hat{Y}^l = aX + b$ כך ששגיאת השערוך הריבועית הממוצעת תהיה מינימלית. במילים אחרות, מצא קבועים a ו-b כך שהביטוי

$$E[\varepsilon^2] = E\Big[(Y - \hat{Y}^l)^2\Big] = E\Big[(Y - aX - b)^2\Big]$$

יהיה מינימלי. שים לב שזו בעיה הרבה יותר פשוטה ממציאת המשערך האופטימלי כי כאן יש למצוא שני קבועים בלבד ולא פונקציה שלמה $\hat{Y} = (X)$.

פתרון: נרשום פעם נוספת את הביטוי לשגיאת השערוך הריבועית הממוצעת:

$$E[\varepsilon^2] = E\Big[(Y - aX - b)^2\Big] = E[Y^2] - 2aE[YX] - 2bE[Y] + a^2e[X^2] + 2abE[X] + b^2$$

אם נגזור את הביטוי לעיל לפי b ונשווה ל-0 את הנגזרת נקבל ש-b האופטימלי, b^*, צריך לקיים:

$$b^* = E[Y] - aE[X]$$

בהצבת b^* לתוך הביטוי לשגיאת השערוך הריבועית הממוצעת נקבל:

$$E[\varepsilon^2] = E\Big[(Y_c - aX_c)^2\Big] = E[Y_c^2] - 2aE[Y_cX_c] + a^2E[X_c^2]$$

כאשר $X_c = X - E[X]$ ו-$Y_c = Y - E[Y]$. שוב, ע"י גזירת הביטוי לעיל לפי a והשוואה ל-0 נקבל ש-a האופטימלי, a^*, צריך לקיים:

$$a^* = \frac{E[Y_cX_c]}{E[X_c^2]} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}$$

ולכן המשערך הלינארי האופטימלי הוא:

$$\hat{Y}^l_{\mathrm{opt}} = E[Y] + \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}\Big[X - E[X]\Big]$$

(מה הקשר בין תוצאה זאת לדוגמה ב' בסעיף 2.2? ראה הערה מס' 4 בהמשך). השגיאה הריבועית הממוצעת המתקבלת במקרה זה היא:

$$E[\varepsilon^2_{\min} = \mathrm{Var}(Y) - \frac{[\mathrm{Cov}(X,Y)]^2}{\mathrm{Var}(X)} = \mathrm{Var}(Y)(1-\rho^2)$$

כאשר ρ מקדם הקורליליציה; $|\rho| \leq 1$.

הערות:

1. שים לב שלצורך חישוב המשערך הלינארי האופטימלי יש צורך לדעת רק מומנטים מסדר ראשון ושני.
2. משערך אופטימלי לעולם אינו גרוע יותר ממשערך לינארי אופטימלי; משערך לינארי אופטימלי לעולם אינו גרוע יותר ממשערך ללא מדידה כלל.
3. כאשר $\mathrm{Cov}(X,Y) = 0$ המשערך הלינארי האופטימלי עם מדידה X הוא $E[Y]$, כלומר, אותו מספר כאילו ולא נמדד דבר. מכאן שכאשר X ו-Y חסרי קורילציה, X אינו עוזר בשערוך לינארי של Y.
4. במקרים מסוימים המשערך הליניארי האופטימלי הוא המשערך האופטימלי. לדוגמא, כאשר X ו-Y בלתי תלויים. דוגמא אחרת היא המקרה הגאוסי (וזאת ומאחר וכבר צוין, ללא הוכחה שבמקרה הגאוסי $E[Y|X] = c_0X$ עבור קבוע c_0 מסוים). דוגמאות נוספות יובאו בתרגילים.

המקרה הוקטורי

נעבור למקרה בו השערוך מתבצע על סמך מספר מדידות $X_1, X_2, \cdots X_n$ לשם פשטות נניח ש-$E[Y] = 0$ וכן $E[\underline{X}] = 0$ כאשר $\underline{X}$ הוא הוקטור האקראי שרכיביו הם $X_1, X_2, \ldots, X_n$ במקרה זה למשערך שהוא פונקציה לינארית של $\underline{X}$ יש באופן כללי את הצורה הבאה:

$$\hat{Y}^l = \underline{a}^T\underline{X} = \sum a_iX_i$$

כאשר $\underline{a}^T = (a_1, a_2, \ldots, a_n)$ הם קבועים כלשהם. בעית מציאת המשערך הלינארי האופטימלי במקרה זה היא:

בעיה: מצא משערך מהצורה $\hat{Y}^l = \underline{a}^T\underline{X}$ כך ששגיאת השערוך הריבועית הממוצעת תהיה מינימלית. במילים אחרות, מצא קבועים $(a_1, a_2, \ldots, a_n)$ כך שהביטוי

$$E[\varepsilon^2] = E\Big[(Y - \hat{Y}^l)^2\Big] = E\Big[(Y - \underline{a}^T\underline{X}^2\Big]$$

יהיה מינימלי (על פני כל הוקטורים הקבועים $\underline{a}$).

פתרון: נרשום פעם נוספת את הביטוי לשגיאת השערוך הריבועית הממוצעת:

$$E[\varepsilon^2] = E\Big[(Y - \sum a_iX_i)^2\Big] = E[Y^2] - 2\sum a_iE[YX_i] + \sum\sum a_ia_jE[X_iX_j] \qquad (2.3)$$

אם נגזור את הביטוי לעיל לפי a_i ונשווה ל-0 את הנגזרת נקבל שערכי a_i האופטימליים, a_i^* צריכים לקיים:

$$0 = \frac{\partial E[\varepsilon^2]}{\partial a_i} = -2E[YX_i] + 2\sum_{j=1}^{n} a_j^*E[X_iX_j] \qquad 1 \leq i \leq n$$

בכתיב וקטורי:

$$\begin{pmatrix} E[YX_1] \\ E[YX_2] \\ \vdots \\ E[YX_n] \end{pmatrix} = \begin{pmatrix} E[X_1^2] & E[X_1X_2] & E[X_1X_n] \\ E[X_1X_2] & E[X_2^2] & \\ & & \\ E[X_1X_n] & & E[X_n^2] \end{pmatrix} \begin{pmatrix} a_1^* \\ \\ \\ a_n^* \end{pmatrix}$$

ובצורה מקוצרת:

$$E[Y\underline{X}] = E[\underline{X}\underline{X}^T] \cdot \underline{a}^*$$

כאשר באגף השמאלי מופיע וקטור n מימדי ובאגף הימני מכפלה של מטריצה $n \times n$ בוקטור n מימדי. לצורך ההמשך נרשום תוצאה זאת בצורה הבאה:

$$\mathrm{Cov}(Y, \underline{X}) = \mathrm{Var}\,\underline{X} \cdot \underline{a}^* \qquad (2.4)$$

בהנחה שמטריצת הקווריאנס $E[\underline{X}\underline{X}^T]$ אינה סינגולרית, יש פתרון (והוא יחיד) עבור $\underline{a}^*$ (במקרה הסינגולרי נטפל אח"כ). פתרון זה נתון ע"י

$$\underline{a}^* = \left(E[\underline{X}\underline{X}^T]\right)^{-1} E[Y\underline{X}]$$

והמשערך הלינארי האופטימלי הוא

$$\hat{Y}^l_{\text{opt}} = (\underline{a}^*)^T \underline{X} = \sum_{i=1}^{n} a_i X_i \qquad (2.5)$$

מה השגיאה הנותרת $E[(Y - (\underline{a}^*)^T \cdot \underline{X}^2] = ?$. על מנת לחשב את השגיאה הנותרת נרשום את (2.4) בצורת $\underline{b} = \underline{\underline{C}} \cdot \underline{a}^*$ כאשר $\underline{b} = E[Y\underline{X}]$ ו-$\underline{\underline{C}} = E[\underline{X} \cdot \underline{X}^T]$ ואז נציב לתוך (2.3) ונקבל

$$(E[\varepsilon^2])_{\min} = E[Y^2] - 2\underline{b}^T \underline{a}^* + (\underline{a}^*)^T \underline{\underline{C}} \underline{a}^* = E[Y^2] - (\underline{a}^*)^T \underline{\underline{C}} \underline{a}^* \qquad (2.6)$$

או

$$\left(E[\varepsilon^2]\right)_{\min} = E[Y^2] - E\left[((\underline{a}^*)^T \underline{X})^2\right]$$

(2), (3) ו-(4) נותנים את התשובה המלאה לבעית השערוך הלינארי במקרה הוקטורי כאשר המומנטים מסדר ראשון מתאפסים (לפחות כאשר $\underline{\underline{C}}$ הפיכה).

נעיין כעת במקרה ש-$E[\underline{X}] \neq 0, E[Y] \neq 0$. במקרה זה המשערך הלינארי האופטימלי הוא:

$$\hat{Y}^l_{\text{opt}} = E[Y] + \sum_{i=1}^{n} a_i^* (X_i - E[X_i])$$

כאשר ה-a_i^* נתונים ע"י פתרון משואה (2.3) והשגיאה הנותרת תהיה:

$$\left(E[\varepsilon^2]\right)_{\min} = E\left[(Y - E[Y])^2\right] - E\left[((\underline{a}^*)^T (\underline{X} - E[\underline{X}]))^2\right]$$

(הוכח תוצאות אלה).

ההערות שניתנו במקרה הסקלרי תופסות גם כאן.

<u>דוגמא:</u> נתונים

$$E[Y]=E[X_1]=E[X_2]=0 \qquad n=2$$

$$E[X_1,X_2]=0; \quad E[X_1^2]=1; \quad E[X_2^2]=1; \quad E[YX_1]=5; \quad E[YX_2]=7$$

אזי

$$\begin{pmatrix}5\\ \\ 7\end{pmatrix}=\begin{pmatrix}1&0\\ \\ 0&1\end{pmatrix}\begin{pmatrix}a_1^*\\ \\ a_2^*\end{pmatrix}$$

ומכאן

$$a_1^*=5, \quad a_2^*=7$$

$$\hat{Y}^l_{\text{opt}}=5X_1+7X_2$$

$$(E[\varepsilon^2])_{\min}=E\Big[(Y-\underline{X}^T\underline{a}^*)^2\Big]=E[Y^2]-E\Big[(5X_1+7X_2)^2\Big]=E[Y^2]-25-49$$

מדוע לא ותכן שהביטוי האחרון יהיה שלילי? (זכור: $E[X^2]E[Y^2]\geq(E[XY])^2$, מטריצת הקווריאנס אינה שלילית).

<u>מטריצת קווריאס סינגולרית</u>

כמובטח, נעבור כעת לטפל במקרה שבו מטריצת הקווריאס $\underline{\underline{C}}=[\text{Cov}(X_i,X_j)]$ היא סינגולרית. במקרה זה קיים וקטור קבוע n מימדי, $\underline{h}$, כך ש $\underline{\underline{C}}\underline{h}=0$ ולכן:

$$E\Big[(\underline{h}^T\underline{X})^2\Big]=\underline{h}^T\underline{\underline{C}}\underline{h}=0$$

נעיין בוקטור $\underline{h}$ ונניח (ללא הגבלת הכלליות) ש-$h_n\neq 0$ ולכן נוכללשנות את קנה המדה כך ש-$h_n=1$ לפיכך

$$E\Big[(X_n+h_{n-1}X_{n-1}+\cdots+h_1X_1)^2\Big]=0$$

ולכן

$$X_n=-h_{n-1}X_{n-1}-h_{n-2}X_{n-2}-\cdots-h_1X_1$$

דהיינו, אם נתונים $X_1,\ldots,X_{n-1}$ אזי נוכל לנחש את X_n ללא שגיאה. במלים אחרות, באם ידועים $X_1,\ldots,X_{n-1}$ אזי אין צורך בידיעת X_n. לכן, במקום לשערך את Y נתון $X_1,\ldots,X_n$ מספיק לשערך את Y כשנתון $X_1,\ldots,X_{n-1}$. אם $h_n=0$ אבל $h_i\neq 0$ עבור i כלשהוא יהיה החשבון בדיוק אותו דבר. מוסר השכל: אם $\underline{\underline{C}}$ סינגולרית אזי קיים i_0 (לפחות אחד) כך שאפשר לותר על X_{i_0} מבלי לקלקל את השערוך. אם המטריצה החדשה $(n-1)\times(n-1)$ שמתקבלת ע"י מחיקת העמודה והשורה ה-i_0 היא מטריצה לא סינגולרית - נמשיך בשערוך לפי התוצאות שבידינו עבור הבעיה המוקטנת. אם המטריצה החדשה עדיין סינגולרית, נמשיך בתהליך הויתור עד שנגיע למטריצה לא סינגולרית.

2.4 עקרון ההשלכה

בחלק זה נתרכז במקרה בו $EY = 0; E\underline{X} = 0$. עבור מקרה זה מצאנו שהמשערך הלינארי האופטימלי של Y כשנמדד $\underline{X}$ הוא $\hat{Y}^l_{\text{opt}} = (\underline{a}^*)^T$ כאשר הקבועים $\underline{a}^*$ מקיימים:

$$E[Y \cdot \underline{X}] = E[\underline{X} \cdot \underline{X}^T] \cdot \underline{a}^* \tag{2.7}$$

אנו נראה דרך נוספת למציאת תוצאה זו, הפעם דרך נימוקים "גיאומטריים".

הערה א': שים לב שאם $\underline{a}^*$ הוא פתרון משואת השערוך הלינארי אזי

$$E\Big[(Y - (\underline{a}^*)^T\underline{X}) \cdot X_i\Big] = 0 \quad \text{עבור כל } i \tag{2.8}$$

הוכחה: מתוך (2.7) נרשום:

$$\Big(E[Y\underline{X}]\Big)^T = (\underline{a})^T E[\underline{X} \cdot \underline{X}^T]$$

ולכן

$$E[YX_i] = (\underline{a}^*)^T E[\underline{X} \cdot X_i]$$

הערה ב': יהיה $\underline{Y}$ וקטור דטרמיניסטי במרחב האוקלידי ה-m ממדי E_m. יהיו $\underline{X}_1, \ldots, \underline{X}_n$ וקטורים דטרמיניסטיים באותו מרחב m ממדי. נעיין בתת המרחב $M(\underline{X}_1, \ldots, \underline{X}_n)$ הנפרש ע"י $\sum_{i=1}^n \theta_i \underline{X}_i$ ונניח שזה תת מרחב אמיתי של E_m (ז.א. לא E_m עצמו). נסמן ב-$\hat{\underline{Y}}$ את ההשלכה של $\underline{Y}$ על M.

טענה: $\hat{\underline{Y}}$ ההשלכה של $\underline{Y}$ על M אם ורק אם עבור כל i מתקיים: $(\underline{Y} - \hat{\underline{Y}}) | \underline{X}_i$.

שאלה ב- E_m: נתונים $\underline{Y}, \underline{X}_1, \ldots, \underline{X}_n$ וקטורים ב-E_m. מצא ההשלכה $\hat{\underline{Y}}$ של $\underline{Y}$ על תת המרחב הנפרש על ידי $\underline{X}_1, \cdots, \underline{X}_n$: אזי $\hat{\underline{Y}}$ הוא צרוף לינארי של הוקטורים $\underline{X}_i$:

$$\hat{\underline{Y}} = h_1^* \underline{X}_1 + h_2^* \underline{X}_2 + \cdots + h_n^* \underline{X}_n$$

ועקרון ההשלכה קובע שעבור כל i:

$$\Big((\underline{Y} - \hat{\underline{Y}}), \underline{X}_i\Big) = 0$$

לכן $\underline{h}^*$ מוגדר ע"י:

$$(\underline{Y}, \underline{X}_i) = \sum_{j=1}^n h_j^* (\underline{X}_i, \underline{X}_j)$$

את התוצאה ניסחנו במרחב m ממדי E_m אבל זה ניתן להרחבה למרחבים "אין סוף ממדיים" בתנאי שנגדיר כהלכה

מרחבים כאלה ויהיה לרשותנו מושג של מכפלה סקלרית במרחבים אלה. ננסה כעת לבדוק את האנלוגיה הבאה:

$$\underline{X}_1, \underline{X}_2 \text{ מ"א סקלריים } E[X_i]=0 \longleftrightarrow \underline{X}_1, \underline{X}_2 \text{ וקטורים דטרמיניסטיים ב- } E_m$$

$$\mathrm{Cov}(X_1,X_2) \longleftrightarrow \text{ מכפלה סקלרית } (\underline{X}_1, \underline{X}_2)$$

$$\sqrt{\mathrm{Var}(X)} \longleftrightarrow \text{ אורך וקטור } = ||\underline{X}|| = (\underline{X}, \underline{X})^{\frac{1}{2}}$$

$$\mathrm{Cov}(X_1,X_2)=0 \longleftrightarrow \text{ ניצבות } (\underline{X}_1, \underline{X}_2)=0$$

נעבור לבעית השערוך הלינארי של המ"א Y מתוך $X_1,\ldots,X_n$; נפעיל את טבלת האנלוגיה לתוצאה ב-E_m ונקבל לכל i:

$$\mathrm{Cov}(Y,X_i)=\sum_{j=1}^{n} h_j^* \,\mathrm{Cov}(X_i,X_j)$$

וזה בדיוק (2.8)! ולכן האנלוגיה ועקרון ההשלכה נותנים תוצאות נכונות.

<u>נסכם:</u> השגיאה של המשערך הלינארי האופטימלי $Y-\hat{Y}=Y-\sum a_i^* X_i$ ניצבת לתת המרחב שעליו אנו משליכים, דהיינו:

$$E\left[\left(Y-\sum_{i=1}^{n} a_i^* X_i\right)\cdot X_j\right]=0; \quad j$$

לגבי השגיאה הנותרת: ב-E_m אנו יודעים בגלל הנצבות שמתקיים

$$||\underline{Y}||^2 = ||\underline{\hat{Y}}||^2 + ||\underline{Y}-\underline{\hat{Y}}||^2$$

בדומה, בבעית השערוך

$$E[\varepsilon^2]=E\left[(Y-\hat{Y})^2\right]-E\left[(Y-\hat{Y})\cdot\hat{Y}\right]$$
$$=E\left[(Y-\hat{Y})Y\right]=E[Y^2]-E[\hat{Y}^2]$$

<u>שים לב:</u> $E[(Y-\sum a_i^* X_i)\cdot X_j]=0,\ j=1,\ldots,N$ מאפיין את המשערך הלינארי האופטימלי ואילו $E[(Y-\hat{Y}(\underline{X}))\cdot g(X_1,\ldots,X_n)]=0$ מאפיין את המשערך האופטימלי (התוחלת המותנית). אפשר לכן לנסח עקרון השלכה גם עבור השערוך האופטימלי.

3 תהליכים אקראיים בזמן בדיד

"וקטור אקראי" (הגדרה 8.14) הוא אוסף של מספר משתנים אקראיים: זוהי הרחבה של מושג ה"משתנה אקראי" למספר מימדים. המושג "תהליך אקראי" הוא הרחבה נוספת, כאשר מספר המשתנים יכול להיות אין סופי. בנוסף, אנו מתייחסים לאוסף המשתנים כאילו הופיעו בנקודות זמן עוקבות.

דוגמה 3.1 נניח שמשדרים אות סיפרתי: כלומר, בכל יחידת זמן k משודרת הסיפרה 0 או הסיפרה 1. נניח שלאות זה מתווסף רעש: כלומר בזמן k אנו קולטים את $X(k)+n(k)$, המורכב מהגודל המשודר $X(k)$ ומרעש $n(k)$. האות שאנו קולטים לכן מורכב ממשתנים אקראיים: משתנה אחד בכל יחידת זמן. לאות כזה קוראים תהליך אקראי.

הגדרה 3.2 תהליך אקראי בזמן בדידהוא סידרה של משתנים אקראיים $\{X(t),\ T_1 \leq t \leq T_2\}$. פונקציית מדגםשל התהליך האקראי היא כל פונקציה של משתנה הזמן t עבור $\omega=\omega_0$ קבוע. כלומר, זוהי כל פונקציה מהצורה

$$X(t,\omega_0),\ T_1 \leq t \leq T_2, \qquad \omega_0 \text{ קבוע}$$

בדרך כלל t מקבל ערכים שלמים, T_2 יהיה סופי או $+\infty$, ו-T_1 יקבל את הערך 0 או $(-\infty)$. מקובל לסמן שלמים באותיות i,j,k,l,m,n, וכמו כן נשתמש בשני הסימונים השקולים $X(t)$ או X_t וכו'.

3.1 פילוג של תהליך אקראי

עבור וקטור אקראי, ניתן להגדיר פילוג (הגדרה 8.15). אולם בבואנו להרחיב הגדרה זו לתהליך אקראי, אנו נתקלים בקשיים: כוון שתהליך אקראי יכול להיות מוגדר עבור אין-סוף נקודות זמן, לא ברור כיצד נחשב הסתברויות של מאורעות, המוגדרים על ידי אין סוף משתנים. אחד הבעיות מתוארת בתרגיל הבא.

תרגיל 3.3 נתונה סדרת משתנים אקראיים $\{X_i(\omega),\ i=1,2,\dots\}$ המתארים זריקות מטבע בלתי תלויות (1 מתאר עץ, 0 מתאר פלי). הראה כי נדע לחשב את ה"פילוג האין-סופי"

$$\mathbb{P}\{X_1 \leq a_1, X_2 \leq a_2, \dots\} \tag{3.1}$$

לכל סידרה אין-סופית $a_1,a_2,\dots$ אם ורק אם מתקיים התנאי הבא. לכל N, לכל סידרת זמנים (אינדקסים) $t_1 < t_2 < \cdots < t_N$ ולכל סידרה $a_1,a_2,\dots,a_N$ באורך N אנו יודעים לחשב את הפילוג של הווקטור

$$F_{X(t_1),X(t_2),\dots,X(t_N)}(a_1,a_2,\dots,a_N) = \mathbb{P}\{X(t_1)\leq a_1, X(t_2)\leq a_2,\dots,X(t_N)\leq a_N\}$$

רמז: בדוק בניפרד את המקרה בו מספר הפעמים ש-$a_i < 1$ הוא אין-סופי (ואז ההיסתברות ב-(3.1) היא 0), ואת המקרה בו $a_i = 1$ פרט למספר סופי של פעמים (ואז ניתן לרשום את (3.1) על ידי אוסף סופי של מ"א).

הקושי העיקרי במקרה זה הוא כי בבואנו לחשב הסתברות של מאורע המוגדר ע"י מספר אין סופי של תנאים, בדרך כלל הסתברות זו תהייה שווה 0. לכן נוח להעזר בהכללה הבאה.

הגדרה 3.4 הפילוגאו חוק הפילוגשל תהליך אקראי בזמן בדיד הוא אוסף כל פונקציות הפילוג המשותפות, המוגדרות עבור סידרת זמנים סופיים. כלומר, זהו האוסף של הפילוגים הסופיים

$$F_{X(t_1),X(t_2),\ldots,X(t_N)}(a_1,a_2,\ldots,a_N)=\mathbb{P}\left\{X(t_1)\le a_1, X(t_2)\le a_2,\ldots,X(t_N)\le a_N\right\}$$

עבור כל N, כל סידרת זמנים (אינדקסים) $t_1<t_2<\cdots<t_N$ וכל סידרה $a_1,a_2,\ldots,a_N$. הגודל $F_{X(t)}$, שהוא הפילוג של המשתנה האקראי $X(t)$ נקרא הפילוג החד-מימדי, או הפילוג השולי של התהליך.

כמובן שלא כל אוסף פילוגים סופיים מתאר פילוג של תהליך. לדוגמה, נתבונן בתהליך $\{X(1),X(2)\}$. האם יתכן שמתקיימים שני השוויונים

$$\mathbb{P}\left\{X(1)\le a\right\}=0.5$$

$$\mathbb{P}\left\{x(1)\le a,\ X(2)\le b\right\}=1$$

בו זמנית? זה ודאי לא יתכן, שכן למאורע הראשון בוודאי הסתברות גדולה יותר מאשר לשני. לכן ברור כי חייב להיות תנאי עיקביות כלשהוא.

אוסף פילוגים מתאר פילוג של תהליך אם ורק אם הוא מקיים את דרישת העיקביות (קונסיסטנטיות) הבאה.

הגדרה 3.5 דרישת עיקביות: לכל N ו- $k<N$, לכל סידרת זמנים (אינדקסים) $t_1<t_2<\cdots<t_N$ ולכל סידרה $a_1,a_2,\ldots,a_N$ מתקיים התנאי

$$\begin{aligned}&\mathbb{P}\left\{X(t_1)\le a_1,\ldots,X(t_{k-1})\le a_{k-1},X(t_{k+1})\le a_{k+1},\ldots,X(t_N)\le a_N\right\}\\&\quad=\mathbb{P}\left\{X(t_1)\le a_1,\ldots,X(t_{k-1})\le a_{k-1},X(t_k)<\infty,X(t_{k+1})\le a_{k+1},\ldots,X(t_N)\le a_N\right\}\end{aligned}$$

או, בסימון של פונקציות פילוג,

$$\begin{aligned}&F_{X(t_1),\ldots,X(t_{k-1}),X(t_{k+1}),\ldots,X(t_N)}(a_1,\ldots,a_{k-1},a_{k+1},\ldots,a_N)\\&\quad=F_{X(t_1),\ldots,X(t_N)}(a_1,\ldots,a_{k-1},\infty,a_{k+1},\ldots,a_N)\end{aligned}$$

תרגיל 3.6 הראה כי בהגדרה 3.4, מספיק להשתמש בסדרות זמן מהצורה $T_1,T_1+1,\ldots,T_1+N$.

מספר דוגמאות של אותות אקראיים בזמן בדיד.

דוגמה 3.7 יהיה $\{X(n),\ n=1,2,\ldots\}$ אוסף של משתנים בלתי תלויים סטטיסטית ושווי פילוג. (בקיצרה i.i.d) אפשר להתייחס לאוסף המסודר כאל תהליך אקראי בזמן בדיד, כאשר הזמנים הם שלמים $t=1,2,\ldots$. תהליך כזה נקרא רעש לבן.להמחשה ראה דוגמה 1.1.

תהליך כזה מופיע למשל כסידרת הזכיות בהפעלות חוזרות של מכונת הימורים. בנוסף, במקרים רבים מודל הרעש במדידות הוא כזה. קל לחשב את הפילוג של תהליך בלתי תלוי ושווה פילוג. נסמן ב-F_X את חוק הפילוג של המ"א $X(1)$ (שהוא גם חוק הפילוג של כל אחד מהמשתנים האחרים). יהיו $\{a_i,\ i=1,2,\dots\}$ מספרים ממשיים כלשהם. אזי בגלל אי התלות,

$$F_{X(t_1),X(t_2),\dots,X(t_k)}(a_1,a_2,\dots,a_k) = F_{X(t_1)}(a_1)\cdot F_{X(t_2)}(a_2)\cdots F_{X(t_k)}(a_k) = \prod_{i=1}^{k} F_X(a_i)$$

דוגמה 3.8 הילוך אקראי,או הילוך שיכורהוא תהליך הפרשים בת"ס ושווי פילוג. נגדיר תהליך אקראי $\{Y(n),\ n=0,1,\dots\}$ על ידי

$$Y(0)=0, \qquad Y(n)=Y(n-1)+X(n)$$

כאשר המשתנים $\{X(n),\ n=1,2,\dots\}$ הם בת"ס. התהליך $Y(n)$ יקרא תהליך הפרשים בת"ס. אם לא יאמר אחרת, אנו נניח שהמשתנים $\{X(n),\ n=1,2,\dots\}$ הם שווי פילוג. להמחשה, ראה 7.2.

תהליך הפרשים בת"ס ושווי פילוג נקרא גם הילוך אקראי, או הילוך שיכור (מטעמים מובנים). אפשר (ומקובל) להגדיר תהליך כזה דרך ההפרשים:

$$Y(n)=\sum_{i=1}^{n} X(i) \tag{3.2}$$

דוגמאות מעשיות לתהליך כזה הוא הרווח המצטבר בסידרת הימורים במכונת מזל.

עבור תהליך כזה, אם הוא מקבל למשל ערכים שלמים, ניתן לרשום ביטוי פשוט עבור פונקצית ההיסתברות. בסימונים של דוגמה 3.7,

$$\begin{aligned}
&\mathbb{P}\{Y(1)=a_1,\dots Y(k)=a_k\}\\
&\qquad = \mathbb{P}\{Y(1)=a_1,\dots,Y(k-1)=a_{k-1},Y(k-1)+X(k)=a_k\}\\
&\qquad = \mathbb{P}\{Y(1)=a_1,\dots,Y(k-1)=a_{k-1},X(k)=a_k-a_{k-1}\}\\
&\qquad = \mathbb{P}\{Y(1)=a_1,\dots,Y(k-1)=a_{k-1}\}\cdot\mathbb{P}\{X(k)=a_k-a_{k-1}\}
\end{aligned}$$

בגלל אי התלות. בצורה דומה נמשיך ונפתח

$$\begin{aligned}
&\mathbb{P}\{Y(1)=a_1,\dots Y(k)=a_k\}\\
&\quad = \mathbb{P}\{Y(1)=a_1,\dots,Y(k-2)=a_{k-2}\}\cdot\mathbb{P}\{X(k-1)=a_{k-1}-a_{k-2}\}\cdot\mathbb{P}\{X(k)=a_k-a_{k-1}\}\\
&\quad = \mathbb{P}\{Y(1)=a_1\}\cdot\mathbb{P}\{X(1)=a_1-a_0\}\cdots\mathbb{P}\{X(k-1)=a_{k-1}-a_{k-2}\}\cdot\mathbb{P}\{X(k)=a_k-a_{k-1}\}\\
&\quad = \prod_{n=1}^{k}\mathbb{P}\{X(n)=a_n-a_{n-1}\}
\end{aligned}$$

3.2 תוחלת ומומנטים של תהליך אקראי

הגדלים הבסיסיים ביותר המתארים תהליך אקראי הם פונקציית התוחלת, הווריאנס ופונקצית אוטוקורלציה. עבור משתנים אקראיים, התוחלת (הגדרה 8.17), המומנטים, השונות (הגדרה 8.22), והקווריאנס (הגדרה 8.23), הם מספר-ים. בעזרתם נוכל להגדיר תכונות מקבילות של תהליך אקראי, הכוללות מומנטים עד (כולל) סדר שני, ע"י התבוננות במשתנה האקראי $X(t_1)$ או בזוג המשתנים $(X(t_1), X(t_2))$. הגדלים יהיו כמובן תלויים בזמנים $t_1,\ t_2$.

הגדרה 3.9 עבור תהליך אקראי $\{X(t),\ -\infty < t < \infty\}$ (כאשר t שלם) נגדיר את

1. פונקציית התוחלת

$$\mu_X(t) = \mathbb{E}[X(t)]$$

2. פונקצית האוטוקורלציה

$$\mathbf{R}_X(t_1, t_2) = \mathbb{E}\left[X(t_1) \cdot X(t_2)\right]$$

3. פונקצית הקווריאנס

$$\begin{aligned}\mathbf{K}_X(t_1, t_2) &= \mathbb{E}\left[(X(t_1) - \mu_X(t_1)) \cdot (X(t_2) - \mu_X(t_2))\right] \\ &= \mathbf{R}_X(t_1, t_2) - \mu_X(t_1) \cdot \mu_X(t_2)\end{aligned}$$

פונקציות התוחלת והאוטוקורלציה מספקות הערכות גסות על התהליך: לא רק הממוצע (קירוב דטרמיניסטי) והפיזור (ווריאנס), אלא גם הערכה של התלות הסטטיסטית (הלינארית) של ערכי התהליך בנקודות זמן שונות.

הבה נחשב פוקציות אילו עבור רעש לבן (דוגמה 3.7) והילוך שיכור (דוגמה 3.8).

דוגמה 3.10 עבור התהליך $X(n)$,

$$\begin{aligned}\mu_X(n) &= \mathbb{E}[X(n)] \\ &= \mathbb{E}[X(1)]\end{aligned}$$

כיון שהפילוג אינו תלוי ב-n, ולכן $\mu_X(n) = \mu_X$ אינו תלוי בזמן n. פונקצית האוטוקורלציה היא

$$\begin{aligned}\mathbf{R}_X(t_1, t_2) &= \mathbb{E}[X(t_1)X(t_2)] \\ &= \begin{cases}\mathbb{E}[(X(1))^2] & t_1 = t_2 \text{ אם} \\ (\mu_X)^2 & t_1 \neq t_2 \text{ אם.}\end{cases}\end{aligned}$$

בגלל אי התלות בזמן, ובגלל אי התלות הסטטיסטית בין $X(t_1)$ לבין $X(t_2)$ כאשר הזמנים הם שונים.

עבור התהליך $Y(n)$, נעזר ביצוג 3.2 ונקבל בעזרת הלינאריות של התוחלת (טענה 8.20),

$$\begin{aligned}\mu_Y(n) &= \mathbb{E}[Y(n)] \\ &= \left[\sum_{i=1}^{n} \mathbb{E}[X(i)]\right] \\ &= \sum_{i=1}^{n} \mu_X(i)\end{aligned}$$

כוון שההפרשים הם שווי פילוג, נקבל את הנוסחה הפשוטה

$$\mu_Y(n) = n\mu_X$$

כעת, עבור $k \leq n$,

$$\begin{aligned}\mathbf{R}_Y(k,n) =& \mathbb{E}[Y(n) \cdot Y(k)] \\ =& \mathbb{E}\left[\sum_{i=1}^{n} X(i)\right] \cdot \left[\sum_{j=1}^{k} X(j)\right] \\ =& \mathbb{E}\left[\sum_{j=1}^{k}\sum_{i=1}^{n} X(i)X(j)\right] \\ =& \mathbb{E}\left[\sum_{j=1}^{k}\sum_{i=1,\ i\neq j}^{n} X(i)X(j) + \sum_{i=1}^{k}\big(X(i)\big)^2\right]\end{aligned}$$

נסמן $\mu = \mathbb{E}[X(1)]$ ו-$\sigma^2 = \mathbb{E}[(X(i)-\mu)^2]$, ונשתמש באי התלות, כך ש-$\mathbb{E}[X(i)X(j)] = \mu^2$ עבור $i \neq j$. כוון ש-

$$\begin{aligned}\mathbb{E}[(X(i))^2] &= \mathbb{E}[(X(1)-\mu+\mu)^2] \\ &= \mathbb{E}[(X(1)-\mu)^2] + \mu^2 + 2\mu\,\mathbb{E}[X(1)-\mu] \\ &= \sigma^2 + \mu^2\end{aligned}$$

ומכוון שלכל $k \leq j$ ישנו בדיוק ערך אחד של $i \leq n$ כך ש-$j = i$, נקבל

$$\begin{aligned}\mathbb{E}[Y(n)\cdot Y(k)] =& k(n-1)\mu^2 + k(\sigma^2+\mu^2) \\ =& kn\mu^2 + k\sigma^2\end{aligned}$$

לסיכום, עבור הילוך שיכור

$$\boxed{\text{כאשר } k \leq n: \quad \mathbb{E}[Y(n)\cdot Y(k)] = k\sigma^2 + nk\mu^2}$$

דוגמה כללית יותר מתקבלת אם בוחנים אלגוריתמים רקורסיביים. מקובל לרשום אלגוריתמים כאילו בצורה הבאה.

דוגמה 3.11 יהיה $\{X(n),\ n=0,1,\dots\}$ רעש לבן (תהליך אקראי המורכב ממ"א בלתי תלויים סטטיסטית). תהי g פונקציה נתונה, ונקבע את $Y(0)=y_0$. נגדיר תהליך $\{Y(n),\ n=0,1,\dots\}$ ע"י

$$Y(n+1)=Y(n)+g(Y(n),X(n)), \qquad n=0,1,\dots, \qquad Y(0)=y_0$$

זהו ניסוח כללי של אלגוריתם רקורסיבי "בנוכחות רעש". נוסחה זו היא כללית מדי מכדי שנוכל לנתח אותה. אולם ניתן לנתח מיקרים פשוטים, למשל:

תרגיל 3.12 יהי $\{X(n),\ n=1,2,\dots\}$ אוסף של מ"א בת"ס ושווי פילוג,

$$\mathbb{P}\{X(1)=1\}=\mathbb{P}\{X(1)=-1\}=\frac{1}{2}$$

יהיה $Y(0)=0$ ונגדיר פונקציה $g(y,x)=x-\frac{y}{2}$. עבור דוגמאות 3.7, 3.8 ו-3.11, חשב את (או רשום נוסחאות עבור פונקציות התוחלת μ_X ו-$\mu_Y(t)$, ופונקציות האוטוקורלציה $\mathbf{R}_X(t_1,t_2)$ ו-$\mathbf{R}_Y(t_1,t_2)$. שרטט (רצוי בעזרת MATLAB) את פונקציות התוחלת, מספר פונקציות מדגם של התהליך, ואת הממוצע של פונקציות המדגם (אם פונקציות המדגם הן $X(t,\omega_1),X(t,\omega_2),\dots,X(t,\omega_k)$ אזי הממוצע של פונקציות המדגם הוא

$$\frac{1}{k}\sum_{n=1}^{k}X(t,\omega_n)$$

שהיא פונקציה של משתנה הזמן t). שרטט את פונקציות האוטוקורלציה $\mathbf{R}_X(1,t)$ ו-$\mathbf{R}_Y(1,t)$ וכן את הפונקציות

$$C_X(1,t;\omega)=X(1,\omega)\cdot X(t,\omega)$$

עבור מספר ערכים של ω ואת הממוצע של פונקציות אילו. הסק מסקנות.

תרגיל 3.13 בידינו רכיב אלקטרוני לינארי אשר בתחום העבודה שלו מקיים את המשוואה $v=i+a$ כאשר הפרמטר a הוא קבוע אך ערכו אינו ידוע. נקבע את i על ערך קבוע (וידוע). את v אננו יכולים למדוד במדויק, שכן המדידה ה-n נותנת את $v(n)=v+X(n)=i+a+X(n)$, כלומר מדידה ועוד רעש. אנו יודעים כי הרעש $\{X(n),\ n=1,2,\dots\}$ מורכב ממשתנים בת"ס ושווי פילוג, עם ממוצע 0. אנו מפעילים את האלגוריתם הבא לחיפוש הערך של a:

$$\begin{aligned}a(n+1)&=a(n)+\frac{1}{n}\left[v(n)-(i+a(n))\right]\\&=a(n)+\frac{1}{n}\left[(i+a+X(n))-(i+a(n))\right]\end{aligned}$$

שם לב כי האלגוריתם ניתן לביצוע (הוא משתמש רק בגדלים ידועים או נמדדים).

פשט אלגוריתם זה על ידי ההצבה $b(n)=a(n)-a$. רשום נוסחה מפורשת עבור התהליך $b(n)$, ותאר את התנהגות האלגוריתם לאחר מספר איטרציות רב (n גדול).

3.3 סטציונריות וארגודיות

כאשר מעבירים אות קבוע בזמן, או אות מחזורי, דרך מערכת לינארית יציבה, מקבלים תגובה המורכבת מתופעת מעבר וממצב יציב. תופעת מעבר קשורה לתנאי ההתחלה, והתאמתם לאות הכניסה. אפשר לראות תופעה דומה בתהליכים אקראיים. תהליך קבוע (אולי פרט לגודל אקראי) הוא המקביל לאות קבוע, אך הוא אינו מעניין... משפחה מעניינת יותר של אותות בעלי תכונה חלשה יותר של "קביעות בזמן" הם האותות הסטציונריים.

הגדרה 3.14 סטציונריות.תהליך אקראי בזמן בדיד $\{X(n),\ -\infty < n < \infty\}$ נקרא סטציונרי אם לכל קבוע τ ולכל t_1, t_2 הפילוג של $\{X(t_1), X(t_1+1), \ldots, X(t_2)\}$ זהה לפילוג של $\{X(t_1+\tau), X(t_1+1+\tau), \ldots, X(t_2+\tau)\}$. כלומר, הפילוג אינו משתנה תחת הזזות זמן.

דוגמה 3.15 יהיה Y מ"א כלשהוא, ויהיה Z מ"א המתאר זריקת קוביה, כלומר Z מקבל את הערכים $\{1,2,3,4,5,6\}$ בהסתברות שווה.

1. נגדיר $X(n) = Y$ לכל n. זהו כמובן תהליך סטציונרי (אם כי לא מעניין).
2. נגדיר כעת $X(n) = Y[1+\sin(\pi\alpha n)]$. אם α שלם, אזי אנו במקרה הראשון והתהליך סטציונרי. אם α אינו שלם, אזי התהליך אינו סטציונרי: כדי לראות זאת מספיק להתבונן בפילוג החד-מימדי $F_{X(n)}$.
3. נגדיר $X(n) = \sin[\alpha\pi(n+Z)]$. אם $\alpha = 1/3$ אזי התהליך הוא סטציונרי. כדי לראות זאת, נבחר τ כלשהוא ונגדיר $\tilde{Z} = [Z+\tau] \bmod 6 + 1$. אזי גם $\tilde{Z}$ מתאר זריקת קוביה, ולכן הפילוג של $\tilde{X}(n) = \sin[\alpha\pi(n+\tilde{Z})]$ זהה לפילוג של התהליך $X(n)$. מכוון ש-

$$\begin{aligned}(1/3)\pi m =& (1/3)\pi([m \bmod 6] + 6k)\\ =& 2k\pi + (1/3)\pi[m \bmod 6]\end{aligned}$$

עבור שלם k כלשהוא, קיבלנו שהפילוג של $X(n+\tau)$ זהה לפילוג של $\tilde{Z}(n)$ ולכן הפילוג נשמר תחת הזזת זמן. בצורה דומה ניתן להראות שגם הפילוגים הרב-מימדיים אינם משתנים תחת הזזת זמן.

שים לב שהתוצאה תלויה בערך של α, ובאופן כללי (לערכים אחרים של α) התהליך אינו סטציונרי.

4. התהליך $Z\sin(\alpha\pi n)$ הוא סטציונרי עבור $\alpha = 1/3$, אך לא בהכרח עבור ערכים אחרים (לאילו ערכים נקבל סטציונריות?)

עד כה הנחנו שהפילוגים של המ"א או התהליכים ידועים לנו. כאשר אנו מנסים לתאר תופעה טיבעית (או טכנולוגית) ע"י מודל של משתנים אקראיים, אין לנו ידע שכזה. אפשר כמובן להניח הנחות, אולם כיצד ניבדוק את ההנחות? אם נחשוב על המדידות שבידינו לבצע, הרי שנוכל למדוד רק פונקציית מדגם אחת של התהליך (ω נבחר ע"י הטבע, ואינו בידינו!).

אך יש משפחה של תהליכים עבורם ניתן להתבונן בפונקציית מדגם אחת (לאורך זמן ארוך מספיק) וללמוד מכך על הפילוגים של התהליך. זוהי משפחת התהליכים הארגודיים.

הגדרה 3.16 ארגודיות.תהליך אקראי סטציונרי $\{X(n),\ n=\dots,-2,-1,0,1,2,\dots\}$ נקרא ארגודי אם מתקיים לכל k, ולכל פונקציה (חסומה) g של k משתנים:

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}g(X_n,\dots,X_{n+k-1})=Eg(X_0,\dots,X_{k-1})\quad\forall\omega \tag{3.3}$$

(למען הדיוק, נדרוש שהשוויון יתקיים בהסתברות 1, כלומר שלמאורע שיש שוויון תהייה הסתברות 1).

אם בידינו תהליך ארגודי, נוכל למשל לבחור מספר a ולהגדיר פונקציה מציינת

$$\mathbf{1}_{(-\infty,a]}(x)=\begin{cases}1 & \text{אם } x\le a,\\ 0 & \text{אחרת}\end{cases}$$

כיוון ש-

$$\mathbb{E}[\mathbf{1}_{(-\infty,a]}(X(0))]=1\cdot\mathbb{P}\{X(0)\le a\}=F_{X(0)}(a)$$

קיבלנו דרך לקרב את הפילוג של המשתנים האקראיים, ע"י מדידה של התהליך לאורך זמן וחישוב הממוצע (האר-יתמטי, לא הסטטיסטי, שאינו ידוע). בצורה דומה ניתן לקרב פילוגים רב ממדיים של תהליך ארגודי. המסקנה היא שעבור תהליך ארגודי, ניתן ללמוד את הפילוג מתוך התבוננות בפונקציית מדגם אחת!

דוגמה 3.17 רעש לבן (דוגמה 3.7) הוא ארגודי. עובדה זו נובעת מחוק המספרים הגדולים. לעומת זאת, התהליך $\{X(n)=X,\ n=\dots,-1,0,1,\dots\}$ (אינו תלוי בזמן) אינו ארגודי, שכן הממוצע האריתמטי תמיד שווה ל-X.

המודלים המקובלים בהנדסה הם בדרך כלל של אותות על פרק זמן ארוך, אולם עם נקודת התחלה ונקודת סיום. עבור תהליך כזה, אנו מניחים שפרק הזמן הוא ארוך מספיק כדי שתכונות כמו 3.3 יתקיימו בקרוב עבור T סופי. אנו גם מניחים ש"תופעות מעבר", אם ישנן, כבר דעכו.

דוגמה 3.18 יהי $\{X(n),n=0,1,\dots\}$ רעש לבן. נגדיר משתנה אקראי נוסף $Y(0)$ ונדרוש שיהיה בת"ס ב-$\{X(n),n=0,1,\dots\}$. נגדיר כעת תהליך $\{Y(n),\ n=0,1,\dots\}$ ע"י

$$Y(n+1)=\frac{1}{2}Y(n)+X(n),\quad n=0,1,\dots$$

תהליך כזה נקרא תהליך AR (תהליך Auto Regressive). נגדיר משתנה אקראי Z על ידי

$$Z=\sum_{k=0}^{\infty}\left(\frac{1}{2}\right)^k X(k)$$

התהליך $Y(n)$ בדרך כלל אינו סטציונרי (אלא אם נתחיל אותו בתנאי התחלה מתאימים: אילו?). אולם עבור n גדול, הפילוג של $Y(n)$ קרוב לפילוג של Z. מסיבה זו, התהליך $Y(n)$ מקיים את הגירסה הבאה של תכונת הארגודיות 3.3 :

$$\lim_{T\to\infty}\frac{1}{T}\sum_{n=0}^{T}g[Y(n)]=\mathbb{E}[g(Z)]$$

בקירוב הנדסי, אם נתבונן בתהליך החל מזמן N_0 גדול מספיק, התהליך יהיה בקירוב סטציונרי וארגודי.

4 שרשרות מרקוב

להילוך שיכור (דוגמה 3.8) מבנה פשוט למדי. בפרט, אם ברגע t השיכור נמצא בנקודה x, קל לחשב היכן ימצא ברגע $t+1$: זאת משום שפילוג הצעד ידוע לנו. לצורך חישוב זה, אין זה משנה איזה מסלול עבר השיכור עד לנקודה x. תכונה זו נקראת מרקוביות.

4.1 דוגמאות והגדרות

אחד המודלים הבסיסיים של רשתות מחשבים הוא תור עם שרת בודד.

דוגמה 4.1 תור עם שרת בודד. אל תור יחיד מגיעים משתמשים חדשים לפי תהליך רעש לבן בינארי: כלומר, בכל רגע מגיע משתמש נוסף באופן בלתי תלוי ובהסתברות קבועה λ. שרת יחיד משרת את כל המשתמשים, ובכל רגע זמן אחד מהם מסיים את עבודתו באופן בלתי תלוי, בהסתברות μ. נניח שברגע 0 היו $X(0)$ משתמשים, ונסמן ב-$X(n)$ את מספר המשתמשים אשר עדיין מחוברים, ברגע מספר n.

התור עם שרת בודד קרוב מאד להילוך שיכור: כל עוד התור אינו ריק, פילוג הצעד הבא הוא קבוע: צעד של $+1$ אם מגיע משתמש ולא מסתיים שרות, כלומר בהסתברות $\lambda(1-\mu)$, וצעד של (-1) אם לא מגיע משתמש והסתיים שרות, כלומר בהסתברות $\mu(1-\lambda)$. בהסתברות $(1-\lambda-\mu)+2\cdot\lambda\cdot\mu$ התור נשאר ללא שינוי: אם בגלל שלא קרה דבר (בהסתברות $1-(\lambda+\mu-\lambda\cdot\mu)$), או משום שהיו גם הגעה וגם עזיבה (בהסתברות $\lambda\cdot\mu$). כאשר התור ריק אין כמובן עזיבות, ולכן פילוג הצעד הבא שונה ($+1$ בהסתברות λ, ו-0 בהסתברות $1-\lambda$). לכן תהליך זה אינו הילוך שיכור. למרות זאת, התהליך שומר על תכונת המרקוביות: בהנתן גודל תור של $X(n)$ ברגע n, פילוג גודל התור ברגע $(n+1)$ אינו תלוי באורך התור בעבר הרחוק יותר. בניגוד להילוך שיכור, החישוב תלוי במצב הנוכחי: החישוב כשהתור ריק שונה מהחישוב כאשר אינו ריק.

מודל כללי יותר עם תכונת מרקוביות הוא האלגוריתם הרקורסיבי המתואר בדוגמה 3.11. גם עבור אלגוריתם זה, בהנתן שהערך הוא $Y(n)$ ברגע n, אפשר לחשב את הפילוג של $Y(n+1)$. אם למשל ידוע ש-$Y(n)=y$, אזי ההסתברות של המאורע $\{Y(n+1)\le\alpha\}$ היא ההיסתברות ש-$\{y+g(y,X(n))\le\alpha\}$. ברור כי הסתברות זו אינה תלוייה כלל בעבר הרחוק: ברגע שידוע הערך הנוכחי, נקבע פילוג הצעד הבא.

בפרק זה נעסוק בשרשרות מרקוביות, כלומר בתהליכים מרקוביים המקבלים ערכים בדידים. הדבר ייקל מאד על ההבנה. נזכר בהגדרה 8.28 של הסתברות מותנית.

הגדרה 4.2 שרשרת מרקוב. יהי $\{X(n),\ n=1,2,\dots\}$ תהליך אקראי בזמן בדיד המקבל ערכים שלמים. התהליך יקרא שרשרת מרקוב אם לכל $k<n$ ו-$\{i,j_{n-k},\dots,j_n\}$,

$$\begin{aligned}&\mathbb{P}\left\{X(n+1)=i \mid X(n)=j_n, X(n-1)=j_{n-1},\dots,X(n-k)=j_{n-k}\right\}\\&\qquad=\mathbb{P}\left\{X(n+1)=i\mid X(n)=j_n\right\}\end{aligned}$$

כלומר, בהנתן הערך בהווה, העתיד אינו תלוי בעבר.

הביטוי $\mathbb{P}\left\{X(n+1)=i \mid X(n)=j_n\right\}$ נקרא הסתברות המעבר של השרשרת. הגודל $X(n)$ נקרא המצב של התהליך ברגע n. אוסף המצבים האפשריים של השרשרת נקרא מרחב המצב, ונסמן אותו ב-S. ההגדרה תקיפה כמובן גם לתהליך בזמן בדיד שאינו שלם, או לזמנים שלמים ולאו דוקא חיוביים, וגם אם ערכי התהליך אינם שלמים-ובלבד שיהיו בדידים.

שם לב שגם עבור שרשרת מרקובית,

$$\mathbb{P}\left\{X(n+1)=i \mid X(n)=j_n, X(n+2)=l\right\} \neq \mathbb{P}\left\{X(n+1)=i \mid X(n)=j_n\right\} \tag{4.1}$$

תרגיל 4.3 יהי $\{Z(n)\}$ רעש לבן המקבל ערכים ± 1 כאשר $\mathbb{P}\{Z(1)=1\}=p$. יהי $\{Y(n)\}$ הילוך שיכור (דוגמה 3.8) עם הפרשים $\{Z(n)\}$. הראה כי שני התהליכים הם מרקוביים. חשב את הסתברויות המעבר והראה כי אינן תלויות בזמן (מדוע?). הראה כי עבור רעש לבן מתקיים שוויון ב- (4.1), אולם הילוך שיכור אינו מקיים שוויון כזה.

תרגיל 4.4 יהי $\{Z(n)\}$ רעש לבן המקבל ערכים $\{-1,-2,\ldots,-K\}$ בהסתברות $\mathbb{P}\{Z(n)=-j\}=p_j$. תהי $g=g(k,l)$ פונקציה המקבלת ערכים $\{1,2,\ldots,K\}$. הראה כי האלגוריתם הרקורסיבי (דוגמה 3.11) מייצר תהליך מרקובי $\{Y(n), n\geq 0\}$. חשב את הסתברויות המעבר.

המבנה שקיבלנו בתרגיל האחרון הוא כללי למדי, ובעזרתו ניתן לבנות מודלים לתופעות ותהליכים רבים כולל רשתות תקשורת ומחשבים, אותות דיבור, התנהגות הבורסה, מערכות שרות, תהליכי ייצור ועוד. מסתבר (אם כי לא נראה זאת כאן) שהמודל של אלגוריתם רקורסיבי עם רעש לבן שקול למודל של תהליך מרקובי (שרשרת מרקובית אך עם ערכים לאו דוקא שלמים), במובן שכל אלגוריתם מתאר תהליך מרקובי, וכל תהליך מרקובי אפשר לתאר על ידי אלגוריתם רקורסיבי, כאשר הרעש הוא לבן.

הסיבה לחשיבות של שרשרות מרקוב היא כפולה. מצד אחד כפי שראינו, מודל זה עשיר מספיק כדי לכסות מגוון רחב של תהליכים ומערכות. מצד שני כפי שנראה לשרשרת מרקוב יש מבנה מוגדר מספיק כדי לאפשר חישובים מפורשים, בצורה פשוטה למדי (לפחות מבחינה נומרית).

4.2 שרשרות הומוגניות

נתרכז מעתה במקרה בו הסתברויות המעבר אינן תלויות במפורש בזמן. להגבלה זו מספר סיבות: כפי שנראה מייד, החישובים נעשים פשוטים בהרבה, ומצד שני רוב המודלים המעניינים אכן מקיימים תכונה זו. בנוסף, ניתן לייצג שרשרת מרקובית כזו על ידי רשת-או דיאגרמה. ייצוג זה תורם רבות להבנת התהליך.

ניראה כי חישובים של פילוגים במקרה זה אפשר לבצע על ידי פעולות אלגבריות פשוטות. לצורך החישובים והפיתוחים נזדקק למושגים של ווקטור, מטריצה, וכפל מטריצות.

הגדרה 4.5 שרשרת מרקוב נקראת שרשרת הומוגניתאם לכל n

$$\mathbb{P}\left\{X(n+1)=j \mid X(n)=i\right\}=\mathbb{P}\left\{X(n)=j \mid X(n-1)=i\right\}$$

כלומר אם הסתברויות המעבר אינן תלויות בזמן.

במקרה זה נשתמש באחד מהסימונים המקובלים הבאים:

$$\boxed{\mathbb{P}\left\{X(n+1)=j \mid X(n)=i\right\}=\mathrm{p}_{ij}=\mathrm{p}(j \mid i)}$$

ונכנה אותן "הסתברות המעבר מ-i ל-j. מטריצת הסתברויות המעבר P מוגדרת כ-

$$\mathrm{P}=\{\mathrm{p}_{ij}\}_{i,j}=\begin{bmatrix} \mathrm{p}_{11} & \mathrm{p}_{12} & \mathrm{p}_{13} & \cdots & \mathrm{p}_{1N} \\ \mathrm{p}_{21} & \mathrm{p}_{22} & \mathrm{p}_{23} & \cdots & \mathrm{p}_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathrm{p}_{N1} & \mathrm{p}_{N2} & \mathrm{p}_{N3} & \cdots & \mathrm{p}_{NN} \end{bmatrix}$$

כאשר מספר המצבים כאן הוא N ובחרנו לקרוא להם $\{1,2,\ldots,N\}$. זוהי מטריצה בה כל שורה מייצגת את המצב הנוכחי, והעמודה את המצב אליו עוברים. המינוח "שרשרת הומוגנית" ("homogeneous Markov chain") אינו לגמרי סטנדרטי. יש הקוראים לשרשרת כזו "שרשרת סטציונרית" ("stationary Markov chain"). מינוח זה מטעה, שכן שרשרת הומוגנית אינה בהכרח תהליך אקראי סטציונרי! מסיבה זו אנו נקפיד להשתמש במינוח "שרשרת הומוגנית".

שרשרת מרקובית יכולה להיות מוגדרת אם כן דרך אוסף הסתברויות המעבר, או בצורה שקולה דרך מטריצת המעבר. ייצוג נוסף הוא דרך דיאגרמת המעברים.

דוגמה 4.6 בהמשך לדוגמה 4.1, נניח שבתור יש מספר סופי K של מקומות. במקרה זה, אם מגיע משתמש לתור כאשר התור מלא, הוא נחסם (ולכן אינו משפיע על אורך התור). לכן, הסתברויות המעבר הן:

$$\mathrm{p}_{i(i+1)}=\begin{cases}\lambda(1-\mu) & i<K \text{ אם} \\ 0 & i=K \text{ אם}\end{cases}$$

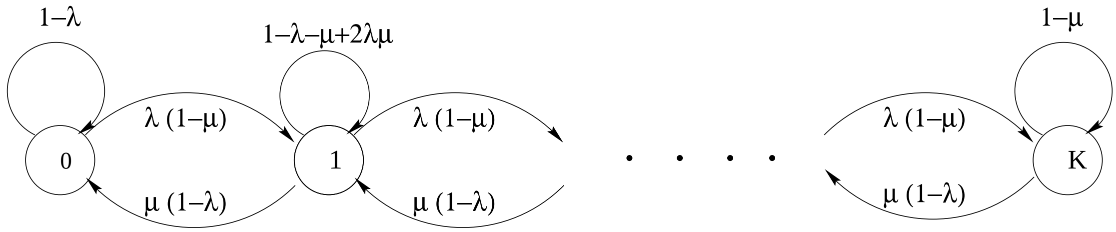
$$\mathrm{p}_{ii}=\begin{cases}1-\lambda & i=0 \text{ אם} \\ 1-\mu & i=K \text{ אם} \\ 1-\lambda-\mu+2\cdot\lambda\cdot\mu & \text{אחרת}\end{cases}$$

$$\mathrm{p}_{i(i-1)}=\begin{cases}\mu(1-\lambda) & i>0 \text{ אם} \\ 0 & i=0 \text{ אם}\end{cases}$$

$$\mathrm{p}_{ij}=0 \qquad .|i-j|>1 \text{ אם}$$

(בהסתברויות אילו טמונה הנחה לגבי הסדר שבין הגעות ועזיבות. מהי?) ניתן לייצג הסתברויות אילו בצורה של דיאגרמת מעברים.

בדוגמה זו, מצבי השרשרת הם $\{0,1,\ldots,K\}$. כל מצב מיוצג על ידי עיגול, שבתוכו שם המצב. כל חץ מתאר מעבר אפשרי בין מצבים, ולידו מספר, המייצג את הסתברות המעבר.



ציור 4.1: שרשרת מרקוב: תור סופי

בצורה דומה ניתן לצייר דיאגרמת מעברים לכל שרשרת מרקוב סופית והומוגנית. בדיאגרמה כזו מספר הצמתים הוא כמספר המצבים של השרשרת. מספר הקשתות הוא כמספר המעברים שהסתברותם אינה 0. השרשרת מדוגמה 4.6 נקראת גם "תהליך לידה מיתה".

טענה 4.7 הסתברויות המעבר של שרשרת מרקוב הומוגנית מקיימות את התנאים הבאים:

$$0 \leq \mathrm{p}_{ij} \leq 1$$

$$\sum_j \mathrm{p}_{ij} = 1$$

התנאי הראשון נובע מהגדרת הסתברות מותנית. התנאי השני טוען כי בהסתברות 1, לאחר צעד אחד נמצא את עצמינו במצב כלשהוא. בסימון מטריצי אפשר לכתוב את התנאי השני כ-

$$\mathrm{P} \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

כלומר ווקטור העמודה $(1, 1, \ldots, 1)^T$ הוא ווקטור עצמי (ימני) של מטריצת הסתברויות המעבר P, עם ערך עצמי 1.

4.3 חישוב הפילוג והסתברויות המעבר

מהו הסיכוי ששרשרת מרקובית תמצא במצב j לאחר n צעדים? ומהו הסיכוי ששרשרת מרקובית תעבור ממצב i למצב j ב-n צעדים? הבה נראה כי זאת ניתן לחשב בקלות. לשם הנוחיות, נגדיר סימון לפונקציית ההסתברות ולהסתברות המעבר במספר צעדים.

הגדרה 4.8 נקבע שרשרת מרקוב מסויימת $\{X(n),\ n = 0, 1, \ldots\}$, ונסמן את מצבי השרשרת ב- $\{1, 2, \ldots, K\}$.

נסמן את ההסתברות שהשרשרת נמצאת בזמן n במצב k ב-$\nu_k(n)$. כעת נגדיר וקטור שורה

$$\begin{aligned}\underline{\nu}(n) &= (\nu_1(n), \nu_2(n), \ldots, \nu_K(n)) \\ &= (\mathbb{P}\{X(n)=1\}, \mathbb{P}\{X(n)=2\}, \ldots, \mathbb{P}\{X(n)=K\})\end{aligned}$$

נסמן את הסתברות המעבר ממצב i למצב j ב-n צעדים ב-

$$\mathrm{p}_{ij}^{(n)} = \mathbb{P}\{X(n)=j \mid X(0)=i\}$$

תרגיל 4.9 ראה המחשה של פונקציות המדגם ושל חישוב הפילוג של שרשרת מרקוב פשוטה (תרגיל 7.3).

טענה 4.10 חישוב הפילוג והסתברויות המעבר של שרשרת הומוגנית וסופית.

1. ההסתברות למעבר ב-n צעדים מקיימת את המשוואה

$$\mathrm{p}_{ij}^{(n)} = \sum_k \mathrm{p}_{ik}^{(n-1)}\,\mathrm{p}_{kj}\ , \qquad \mathrm{p}_{ij}^{(1)} = \mathrm{p}_{ij}$$

2. המטריצה של הסתברויות המעבר ב-n צעדים היא $\{\mathrm{p}_{ij}^{(n)}\}_{ij} = \mathrm{P}^n$, כלומר החזקה ה-$n$ של מטריצת המעברים.
קיבלנו לכן את נוסחת צ'פמן-קולמוגורוב (Chapman-Kolmogorov):

$$\begin{aligned}\mathrm{p}_{ij}^{(n+m)} &= \sum_k \mathrm{p}_{ik}^{(n)} \cdot \mathrm{p}_{kj}^{(m)} \\ \mathrm{P}^{(n+m)} &= \mathrm{P}^n \cdot \mathrm{P}^m\end{aligned}$$

3. את פונקציית ההסתברות של התהליך ברגע n ניתן לחשב על ידי

$$\begin{aligned}\nu_j(n) &= \sum_k \mathbb{P}\{X(0)=k\} \cdot \mathbb{P}\{X(n)=j \mid X(0)=k\} \\ &= \sum_k \underline{\nu}_k(0) \cdot \mathrm{p}_{kj}^{(n)} \\ \underline{\nu}(n) &= \underline{\nu}(0) \cdot \mathrm{P}^n\end{aligned}$$

כאשר הביטוי האחרון הוא בייצוג ווקטורי.

הוכחת הטענה. נרשום

$$\mathbb{P}\left\{X(n)=j \mid X(0)=i\right\}=\sum_k \mathbb{P}\left\{X(n)=j,\ X(n-1)=k \mid X(0)=i\right\}$$

מהגדרת הסתברות מותנית

$$\begin{aligned}\mathbb{P}&\left\{X(n)=j,\ X(n-1)=k \mid X(0)=i\right\}\\ &=\mathbb{P}\left\{X(n)=j \mid X(n-1)=k,\ X(0)=i\right\}\cdot\mathbb{P}\left\{X(n-1)=k \mid X(0)=i\right\}\end{aligned}$$

ובגלל המרקוביות,

$$\mathbb{P}\left\{X(n)=j \mid X(n-1)=k,\ X(0)=i\right\}=\mathbb{P}\left\{X(n)=j \mid X(n-1)=k\right\}$$

לכן נקבל

$$\begin{aligned}\mathrm{p}_{ij}^{(n)}&=\mathbb{P}\left\{X(n)=j \mid X(0)=i\right\}\\ &=\sum_k \mathbb{P}\left\{X(n)=j,\ X(n-1)=k \mid X(0)=i\right\}\\ &=\sum_k \mathbb{P}\left\{X(n)=j \mid X(n-1)=k,\ X(0)=i\right\}\cdot\mathbb{P}\left\{X(n-1)=k \mid X(0)=i\right\}\\ &=\sum_k \mathbb{P}\left\{X(n)=j \mid X(n-1)=k\right\}\cdot\mathbb{P}\left\{X(n-1)=k \mid X(0)=i\right\}\\ &=\sum_k \mathrm{p}_{ik}^{(n-1)}\,\mathrm{p}_{kj}\end{aligned}$$

ובכך סיימנו להוכיח את טענה 1. כיוון שלפי ההגדרה

$$\mathrm{P}=\{\mathrm{p}_{ij}\}_{ij}$$

נובע מסעיף 1 כי מטריצת המעבר בשני צעדים היא P^2, ובאינדוקציה נקבל שהטענה נכונה לכל n. כיוון שכך, משוואת צ'פמן-קולמוגורוב נובעת מיידית מתכונות של מטריצות. כדי להוכיח את סעיף 3 נשתמש בהגדת הסתברות מותנית ונרשום

$$\begin{aligned}\nu_j(n)&=\mathbb{P}\left\{X(n)=j\right\}\\ &=\sum_k \mathbb{P}\left\{X(n)=j \mid X(0)=k\right\}\cdot\mathbb{P}\left\{X(0)=k\right\}\\ &=\sum_k \nu_k(0)\cdot\mathrm{p}_{kj}^{(n)}\end{aligned}$$

או בסימון ווקטורי

$$\underline{\nu}(n)=\underline{\nu}(0)\,\mathrm{P}^n$$

ובכך סיימנו את ההוכחה.

אם כך, החישוב של פונקציית ההסתברות מסתכם בכפל מטריצות. כדי לדעת מהוא הפילוג של התהליך, עלינו להיות מסוגלים לחשב את כל הפילוגים הרב-מימדיים. האם גם זאת ניתן לעשות בשיטות אלגבריות (כפל מטריצות)? נבדוק לדוגמה:

$$
\begin{aligned}
&\mathbb{P}\{X(5)=i,\ X(100)=j\} \\
&\qquad = \sum_k \mathbb{P}\{X(100)=j,\ X(5)=i,\ X(0)=k\} \\
&\qquad = \sum_k \mathbb{P}\{X(100)=j \mid X(5)=i,\ X(0)=k\} \cdot \mathbb{P}\{X(5)=i,\ X(0)=k\} \\
&\qquad = \sum_k \mathbb{P}\{X(100)=j \mid X(5)=i\} \cdot \mathbb{P}\{X(5)=i \mid X(0)=k\} \cdot \mathbb{P}\{X(0)=k\} \\
&\qquad = \mathrm{p}_{ij}^{(95)} \cdot \sum_k \mathrm{p}_{ki}^{(5)}\, \nu_k(0)
\end{aligned}
$$

בצורה דומה ניתן לחשב כל פילוג רב ממדי. מכאן נובעת המסקנה החשובה הבאה:

טענה 4.11 חוק הפילוג של שרשרת מרקובית הומוגנית $\{X(n),\ n \geq 0\}$ נקבע באופן חד משמעי על ידי הפילוג בזמן 0, ומטריצת הסתברויות המעבר.

הוכחה: נבחר סדרת זמנים $\{t_1 < t_2 < \ldots < t_k\}$ וסדרת מצבים $\{j_1, j_2, \ldots, j_k\}$ כלשהם. נחשב את הפילוג המשותף:

$$
\begin{aligned}
&\mathbb{P}\{X(t_k)=j_k, X(t_{k-1})=j_{k-1}, \ldots, X(t_1)=j_1\} \\
&\qquad = \mathbb{P}\{X(t_k)=j_k \mid X(t_{k-1})=j_{k-1}, \ldots, X(t_1)=j_1\} \\
&\qquad\quad \times \mathbb{P}\{X(t_{k-1})=j_{k-1}, \ldots, X(t_1)=j_1\} \\
&\qquad = \mathbb{P}\{X(t_k)=j_k \mid X(t_{k-1})\} \cdot \mathbb{P}\{X(t_{k-1})=j_{k-1}, \ldots, X(t_1)=j_1\}
\end{aligned}
$$

בגלל המרקוביות. נמשיך בשיטה זהה ונקבל

$$
\begin{aligned}
&= \mathbb{P}\{X(t_k)=j_k \mid X(t_{k-1}=j_{k-1})\} \cdot \mathbb{P}\{X(t_{k-1})=j_{k-1} \mid X(t_{k-2})=j_{k-2}\} \\
&\quad \times \cdots\ \mathbb{P}\{X(t_2)=j_2 \mid X(t_1)=j_1\} \cdot \mathbb{P}\{X(t_1)=j_1\} \\
&= \left(\prod_{l=2}^{k} \mathbb{P}\left\{X(t_l)=j_l \mid X(t_{l-1}=j_{l-1})\right\}\right) \cdot \mathbb{P}\{X(t_1)=j_1\} \\
&= \left(\prod_{l=2}^{k} \mathrm{p}_{j_{l-1}\,j_l}^{(t_l-t_{l-1})}\right) \cdot \nu_{j_1}(t_1)
\end{aligned}
$$

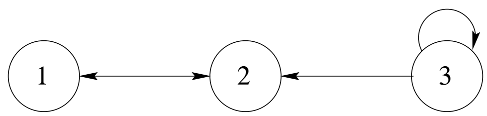
כלומר, אפשר לחשב את הפילוג הרב-מימדי אם יודעים את הפילוג החד-מימדי (אותו ראינו שניתן לחשב מתוך הפילוג ההתחלתי והסתברויות המעבר), ואת הסתברויות המעבר.

הערה: את החישובים שעשינו ניתן להרחיב למקרה של שרשרות שאינן סופיות. הקושי היחידי הוא בסימון המטריצי. אולם נחשוב על "מטריצות אין-סופיות", אשר מקיימות את כללי החיבור והכפל הרגילים בין מטריצות. אזי כל הפיתוחים שעשינו תקפים, ולכן המסקנות נכונות גם עבור שרשרות לא סופיות.

4.4 מצבים נשנים וחולפים

ניתן לסווג את המצבים של שרשרת מרקוב, לפי הקריטריון: האים נחזור למצב שוב ושוב, או שמא נבקר בו רק מספר סופי של פעמים?

דוגמה 4.12 בדוגמה שלפנינו לא רשומות הסתברויות המעבר. אולם כל קשת מתארת הסתברות חיובית.



ציור 4.2: מצבים חולפים ונשנים

אם נתחיל במצב 3, אזי בכל צעד יש סיכוי לעבור למצב 2, ולכן מעבר זה מובטח (אם כי לא ברור מתי). מרגע שעברנו, לא ניתן לחזור: לכן מספר הפעמים שנהיה במצב זה הוא סופי (אם כי אקראי). לעומת זאת, למצבים 1 ו-2 נחזור שוב ושוב.

דוגמה 4.13 סופו של מהמר, או ה-Gambler's ruin. נניח שבכיסינו $X(0)$ שקלים, ואנו מהמרים בכל רגע על שקל אחד. בהסתברות $1/2$ נרוויח שקל, ובהסתברות זהה נפסיד שקל. יהי $\{X(n),\ n \geq 0\}$ סכום הכסף שבכיסינו לפני ההימור מספר n. זוהי שרשרת מרקובית, דומה לתור עם שרת יחיד (דוגמה 4.1), אך עם הבדל קטן: כאשר יגמר הכסף שבכיסינו, יגמר הבילוי, ולא נוכל עוד להמר. כלומר, מצב "0" הוא מצב מיוחד: כאשר הגענו אליו, שם נשאר.

באשר לשאר המצבים, המצב אינו לגמרי ברור. האים יתכן שהמשחק ימשך לנצח? אם לא, פירוש הדבר כי מספר הפעמים שנהיה בכל מצב שאינו 0 הוא סופי (אקראי, כמובן).

נגדיר

$$\begin{aligned}
\rho_{ji} &= \mathbb{P}\left\{X(n) = i, \quad \text{כלשהוא } n > 0 \text{ עבור} \mid X(0) = j\right\} \\
&= \sum_{n=1}^{\infty} \mathbb{P}\left\{X(n) = i, X(n-1) \neq i, \ldots, X(2) \neq i, X(1) \neq i \mid X(0) = j\right\}
\end{aligned}$$

הגדרה 4.14 מצב x_0 נקרא <u>מצב נשנה</u> (Recurrent) אם, כאשר השרשרת מתחילה במצב x_0, מובטח לנו שנשוב אליו לפחות פעם אחת בעתיד, כלומר אם $\rho_{ii} = 1$. מצב שאינו נשנה נקרא <u>מצב חולף</u> (Transient).

דוגמה 4.15 אם נתחיל במצב 1, לא נוכל לעבור למצב 2, ולהיפך. לפי ההגדרה, שני המצבים נשנים.

ציור 4.3: מצבים נשנים

הגדרה 4.16 תהליך מרקובי נקרא <u>תהליך חולף</u> אם כל מצביו חולפים.

תהליך המקיים את המשוואה

$$X(n+1) = X(n) + 1$$

הוא חולף, שכן אם נתחיל במצב כלשהוא $X(0) = i$, לא נחזור אליו לעולם. כלומר, כל המצבים חולפים.

לכאורה, לא הגבלנו את מספר הביקורים במצב חולף אים התחלנו במצב אחר. אולם בגלל המרקוביות, הדברים קשורים.
נסמן ב-N_i את מספר הביקורים של התהליך במצב i (זהו משתנה אקראי).

טענה 4.17 מצב i הוא חולף אם ורק אם לכל j,

$$\mathbb{E}\left[N_i \mid X(0) = j\right] < \infty \tag{4.2}$$

הוכחה: לצורך חישוב התוחלת נשתמש בנוסחה שבטענה 8.21. נחשב תחילה את

$$\mathbb{P}\left\{N_i \geq 1 \mid X(0) = j\right\} = \rho_{ji} \tag{4.3}$$

$$\mathbb{P}\left\{N_i \geq 2 \mid X(0) = j\right\}$$

$$= \mathbb{P}\left\{X(n) = i \text{ כלשהוא } n > 0 \text{ עבור}, X(m) = i \text{ כלשהוא } m > n \text{ עבור} \mid X(0) = j\right\}$$

אבל אם $X(n)=i$ אזי, בגלל המרקוביות וההומוגניות,

$$\begin{aligned}
&\mathbb{P}\left\{X(n)=i,\ X(m)=i \text{ עבור } m>n \text{ כלשהוא} \mid X(0)=j\right\}\\
&\quad=\mathbb{P}\left\{X(m)=i \text{ עבור } m>n \text{ כלשהוא} \mid X(n)=i,\ X(0)=j\right\}\cdot\mathbb{P}\left\{X(n)=i\mid X(0)=j\right\}\\
&\quad=\mathbb{P}\left\{X(m)=i \text{ עבור } m>n \text{ כלשהוא} \mid X(n)=i\right\}\cdot \mathrm{p}_{ji}^{(n)}\\
&\quad=\mathbb{P}\left\{X(n)=i \text{ עבור } n>0 \text{ כלשהוא} \mid X(0)=i\right\}\cdot \mathrm{p}_{ji}^{(n)}\\
&\quad=\rho_{ii}\cdot \mathrm{p}_{ji}^{(n)}
\end{aligned}$$

ולכן גם

$$\mathbb{P}\left\{X(n)=i \text{ עבור } n>0 \text{ כלשהוא}, X(m)=i \text{ עבור } m>n \text{ כלשהוא} \mid X(0)=j\right\}=\rho_{ji}\cdot\rho_{ii}$$

באותה צורה נקבל גם

$$\mathbb{P}\left\{N_i\geq k\mid X(0)=j\right\}=\rho_{ji}\cdot\rho_{ii}^{k-1}$$

עבור מצב חולף, לפי ההגדרה $\rho_{ii}<1$. נחשב כעת את התוחלת בעזרת טענה 8.21:

$$\begin{aligned}
\mathbb{E}[N_i\mid X(0)=j]&=\sum_{k=1}^{\infty}\mathbb{P}\left\{N_i\geq k\mid X(0)=j\right\}\\
&=\sum_{k=1}^{\infty}\rho_{ji}\cdot\rho_{ii}^{k-1}\\
&=\rho_{ji}\frac{1}{1-\rho_{ii}}<\infty
\end{aligned}\tag{4.4}$$

שכן כאמור $\rho_{ii}<1$. הוכחנו כי עבור מצב חולף, $\mathbb{E}[N_i\mid X(0)=j]<\infty$. לעומת זאת, אם המצב נשנה, נבחר $j=i$ ואז $\rho_{ii}=1$ ואז הסכום במשוואה (4.4) יתן ∞. בכך הוכחנו את הטענה.

<u>הערה:</u> מספיק לבדוק את סופיות הביטוי במשוואה (4.2) עבור $j=i$.

כעת נשתמש בתוצאה זו כדי להראות את התוצאה (האינטואיטיבית) הבאה.

טענה 4.18 שרשרת מרקובית בעלת מספר סופי של מצבים אינה חולפת, כלומר יש לה לפחות מצב נשנה אחד.

הוכחת הטענה: נניח שקבענו מצב התחלתי, ונשמיט אותו מהסימונים. נשים לב כי מספר הביקורים במצב כלשהוא עד זמן n (לא כולל זמן 0) הוא בדיוק n, שכן בכל רגע התהליך נמצא במצב כלשהוא. לכן

$$\sum_i N_i=\infty$$

מתכונות התוחלת נובע כי

$$\infty = \mathbb{E}\left[\sum_i N_i\right] = \sum_i \mathbb{E}\left[N_i\right]$$

כיוון שמספר המצבים הוא סופי, לא יתכן שיתקיימו בו זמנית

$$\mathbb{E}\left[\sum_i N_i\right] = \infty, \qquad \mathbb{E}\left[N_i\right] < \infty \; i \text{ לכל}$$

ולכן לפחות מצב אחד אינו חולף.

טענה זו נותנת גם שיטה נוספת לבדוק אם מצב הוא חולף, על ידי חישוב תוחלת מספר הביקורים בו.

טענה 4.19 מצב i נשנה אם ורק אם $\sum_{n=1}^{\infty} \mathrm{p}_{ii}^{(n)} = \infty$

הוכחה: נתחיל במצב i ונזכר כי לפי הטענה הקודמת, מצב הוא נישנה אם ורק אם $\mathbb{E}\left[N_i \mid X(0)=i\right] = \infty$. נזכר בסימון של פונקציה מציינת, ונחשב

$$\begin{aligned}
\mathbb{E}\left[N_i \mid X(0)=i\right] &= \mathbb{E}\left[\sum_{n=1}^{\infty} 1_i[X(n)] \mid X(0)=i\right] \\
&= \sum_{n=1}^{\infty} \mathbb{E}\left[1_i[X(n)] \mid X(0)=i\right] \\
&= \sum_{n=1}^{\infty} \mathbb{P}\left\{X(n)=i \mid X(0)=i\right\} \\
&= \sum_{n=1}^{\infty} \mathrm{p}_{ii}^{(n)}
\end{aligned}$$

כדרוש.

דוגמה 4.20 נניח שבכל רגע יכולים להגיע למחשב לכל היותר משימה אחת מסוג א' ומשימה אחת מסוג ב'. נניח שההגעות הן בלתי תלויות סטטיסטית. נניח שבכל רגע המחשב מטפל בדיוק במשימה אחת, כאשר למשימות מסוג א' יש עדיפות. כלומר, בכל רגע תצא משימה מסוג א' מהמערכת, אלא אם כן אין במערכת משימות כאילו: במיקרה זה יופנה המחשב למשימות מסוג ב'. המצב כאן הוא הווקטור

$$\underline{X} = \begin{pmatrix} \text{מספר משימות ממתינות מסוג א'} \\ \text{מספר משימות ממתינות מסוג ב'} \end{pmatrix}$$

ברור לפי תאור המערכת כי אם ברגע כלשהוא לא נותרו משימות מסוג א' במערכת, אזי בעתיד תוכל להיות במערכת לכל היותר משימה אחד מסוג זה. אולם יתכן שהתחלנו את פעולת המערכת עם מספר משימות גדול. לכן יתכן שבפרק זמן התחלתי תהיינה במערכת מספר משימות מסוג א', אולם כל מצב מהצורה

$$\underline{X} = \begin{pmatrix} k \\ m \end{pmatrix}$$

כאשר $k > 1$ הוא מצב חולף.

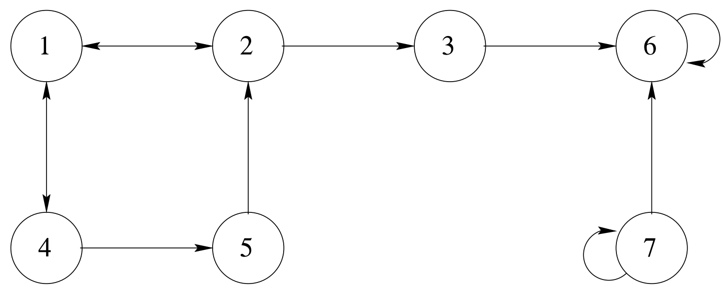
4.5 פרוק מרחב המצב

נניח שמצב i הוא מצב נשנה, ונניח כי $\mathrm{p}_{ij}^{(n)} = 1$ וכן $\mathrm{p}_{ji}^{(m)} = 1$ עבור n, m כלשהם. אזי ברור כי מצב j גם הוא נשנה. אולם יש תוצאות מדויקות יותר בנושא.

הגדרה 4.21 אם $\mathrm{p}_{ij}^{(n)} > 0$ עבור n כלשהוא או ש-$i = j$ אזי נאמר ש-i מוביל ל-j, ונסמן $i \to j$. אם $i \to j$ וגם $j \to i$ נאמר ש-i ו-j מקושרים (communicating), ונסמן $i \leftrightarrow j$.

בסימון הקודם, קיים n כך ש-$\mathrm{p}_{ij}^{(n)} > 0$ אם ורק אם $\rho_{ij} > 0$. מבחינת דיאגרמת המצבים, $i \to j$ אם (ורק אם) קיימת מסילה בדיאגרמה המובילה מ-i ל-j. כתוצאה מכך, $i \leftrightarrow j$ אם ורק אם קיימת בדיאגרמה מסילה מעגלית מ-i לעצמו, העוברת דרך j.

דוגמה 4.22 בדוגמה זו, מצבים $1, 2$ מקושרים, וגם מצבים $1, 4$ מקושרים (אילו זוגות נוספים מקושרים?). מצב 2 מוביל למצב 3, ומצבים $3, 7$ מובילים למצב 6 אולם מצבים $3, 6, 7$ אינם מקושרים לשום מצב. ברור כי מצב 7 הוא מצב חולף (אך האם מצב 1 נשנה? בהמשך נראה כיצד להחליט על כך).



ציור 4.4: מצבים קשורים

טענה 4.23 <u>תכונות יחס הקשירות</u>

1. $i \leftrightarrow i$
2. $i \leftrightarrow j$ אם ורק אם $j \leftrightarrow i$.
3. אם $i \to j$ וגם $j \to k$ אזי $i \to k$.
4. אם $i \leftrightarrow j$ אזי שניהם נשנים או שניהם חולפים.

הוכחה: שלושת הטענות הראשונות נובעות מיידית מההגדרות (ראה דוגמה). נוכיח את 4. כיוון ש-$i \leftrightarrow j$ קיימים m, k כך ש-

$$\mathrm{p}_{ij}^{(m)} > 0 \qquad \text{וגם} \qquad \mathrm{p}_{ji}^{(k)} > 0$$

נניח ש-i נשנה ונבדוק את הקריטריון עבור j

$$\begin{aligned}\sum_{n=1}^{\infty} \mathrm{p}_{jj}^{(n)} \geq \sum_{n=1}^{\infty} \mathrm{p}_{jj}^{(k+n+m)} &= \sum_{i_1, i_2} \mathrm{p}_{ji_1}^{(k)}\, \mathrm{p}_{i_1 i_2}^{(n)}\, \mathrm{p}_{i_2 j}^{(m)} \\ \geq \sum_{n=1}^{\infty} \mathrm{p}_{ji}^{(k)}\, \mathrm{p}_{ii}^{(n)}\, \mathrm{p}_{ij}^{(m)} &= \mathrm{p}_{ji}^{(k)} \left(\sum_{n=1}^{\infty} \mathrm{p}_{ii}^{(n)} \right) \mathrm{p}_{ij}^{(m)}\end{aligned}$$

אם i נשנה אזי ראינו כי

$$\sum_{n=1}^{\infty} \mathrm{p}_{ii}^{(n)} = \infty$$

ולכן גם

$$\sum_{n=1}^{\infty} \mathrm{p}_{jj}^{(n)} = \infty$$

וקיבלנו כי j נשנה. מצד שני, אם i חולף אזי

$$\sum_{n=1}^{\infty} \mathrm{p}_{ii}^{(n)} < \infty$$

ואזי גם

$$\sum_{n=1}^{\infty} \mathrm{p}_{jj}^{(n)} < \infty$$

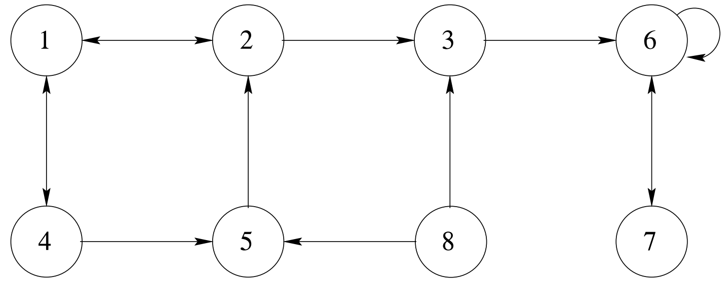
וגם j חולף. בכך הראינו את תכונה 4.

מסקנה מתכונות 1-3: יחס הקשירות הוא רפלקסיבי סימטרי וטרנזיטיבי ולכן הוא יחס שקילות. לכן הוא מחלק את מרחב המצב S ל"קבוצות שקילות". לפי תכונה 4, בכל קבוצה שקילות כל המצבים נשנים או שכולם חולפים.

הגדרה **4.24** קבוצת הקשירות של מצב i היא קבוצת השקילות שלו ביחס ליחס הקשירות. כלומר זוהי קבוצת כל המצבים המקושרים ל-i. קבוצת קשירות A נקראת קבוצה סגורה אם כל מצב ב-A מוביל רק למצבים ב-A.

מהגדרה זו נובע כי בדיאגרמת המעברים, אין קשתות היוצאות מקבוצה סגורה.

דוגמה **4.25** בדוגמה הבאה 8 מצבים. לכל מצב שנבחר נוכל להגדיר קבוצת קשירות. קבוצת הקשירות של מצב 7 מכילה מלבדו גם את מצב 6, ובגלל הסימטריה, קבוצת הקשירות של מצב 6 מכילה גם את מצב 7. לסיכום, איתרנו קבוצת קשירות: $\{6,7\}$. קבוצה זו היא קבוצה סגורה, שכן המצבים $\{6,7\}$ אינם מובילים לאף מצב מחוץ לקבוצה זו. לעומת זאת מצב 3 שייך לקבוצת הקשירות $\{3\}$, המכילה אותו בלבד: זאת כיוון שהוא אינו מקושר לאף מצב אחר. ניתוח זה תופס גם עבור מצב 8. לעומת זאת, קבוצת המצבים $\{1,2,4,5\}$ היא אכן קבוצת קשירות, אך איננה קבוצה סגורה שכן מצב 2, השייך לקבוצה, מוביל למצב 3, שאינו שייך אליה.



ציור 4.5: סיווג מצבים

טענה **4.26** (ללא הוכחה). כל שרשרת סופית נתנת לפירוק למספר סופי של קבוצות סגורות, וקבוצה נוספת (אולי ריקה) של מצבים חולפים. לפחות קבוצת קשירות אחת היא קבוצה סגורה. כל המצבים בכל קבוצה סגורה נשנים (וכאמר, כל המצבים שאינם בקבוצה סגורה חולפים). בשרשרת עם מספר בן מניה של מצבים, מספר הקבוצות הסגורות לא בהכרח סופי, יתכן שאין אף קבוצת קשירות בעלת יותר ממצב אחד, יתכן שאין אף קבוצה סגורה ויתכן שאין אף מצב נישנה, אפילו אין יש קבוצה סגורה.

הגדרה **4.27** שרשרת מרקובית נקראת <u>שרשרת פריקה</u> אם יש לה יותר מאשר קבוצה סגורה אחת. אחרת היא נקראת לא פריקה indecomposeable.

<u>מסקנה:</u> שרשרת מרקובית היא פריקה אם ורק אם יש בה לפחות שני מצבים נשנים שאינם מקושרים.

אם המצב ההתחלתי שייך לקבוצה סגורה, לעולם לא נצא מקבוצה זו. זה מספיק מדוע כל מצביה נישנים במקרה שהשרשרת סופית. אולם יתכן שקיימת קבוצה סגורה נוספת, ואליה לא נגיע; זאת מכיוון שהמצב ההתחלתי הוא כזה. יתכן כמובן להתחיל במצב התחלתי מחוץ לקבוצה סגורה, ולהגיע אליה לאחר מספר צעדים אקראי. מצב התחלתי כזה הוא בהכרח חולף.

טענה **4.28** (ללא הוכחה). כל מצב מחוץ לקבוצה סגורה המוביל למצב בקבוצה, הוא מצב חולף.

אנו רואים, אם כך, שכדי לסווג מצבים לחולפים ונישנים, עלינו לנתח את דיאגרמת המעברים, אולם אין כל חשיבות לערכים של הסתברויות המעבר; תכונות אילו (מצב חולף או נישנה) נקבעות על ידי הקשתות בלבד (כאשר כמובן לא

נצייר קשת כאשר הסתברות המעבר היא 0. כך למשל, בציור 4.5 רק מצבים $6,7$ הם נישנים, ושאר המצבים הם חולפים.

4.6 סטציונריות ושרשרות מרקוביות

ראינו בהמחשות כי הפילוג של שרשרות מרקוב נוטה בדרך כלל להתייצב לאחר מספר צעדים מספיק גדול. כיצד לחבר עובדה זו עם המושג של סטציונריות (הגדרה 3.14)? נראה מייד כי אם מתחילים שרשרת מרקובית עם פילוג מתאים, אזי היא תהיה תהליך סטציונרי.

הגדרה 4.29 פילוג $\underline{\nu}$ ניקרא פילוג סטציונרי או פילוג אינווריאנטי Stationary, invariant עבור שרשרת מרקובית אם $\underline{\nu}(0)=\underline{\nu}$ גורר $\underline{\nu}(n)=\underline{\nu}$ לכל n. כלומר, אם נתחיל את השרשרת עם פילוג זה, הפילוג (החד-מימדי!) ישאר קבוע.

במבט ראשון, סוג הסטציונריות שקיבלנו הוא מוגבל-הוא נוגע רק לפילוג החד מימדי. אולם במקרה המרקובי ההומוגני, זה מספיק!

טענה 4.30 עבור שרשרת מרקוב הומוגנית,

1. הפילוג $\underline{\nu}$ הוא פילוג סטציונרי אם ורק אם הוא מקיים $\underline{\nu}\cdot \mathrm{P}=\underline{\nu}$, כלומר הוא ווקטור עצמי שמאלי של מטריצת המעברים, עם ערך עצמי 1.
2. אם $\underline{\nu}(0)$ הוא פילוג סטציונרי אזי השרשרת המרקובית היא תהליך סטציונרי.

הוכחה: אם $\underline{\nu}(0)=\underline{\nu}$ ו- $\underline{\nu}\cdot \mathrm{P}=\underline{\nu}$ אזי

$$\begin{aligned}\underline{\nu}(n)&=\underline{\nu}(0)\,\mathrm{P}^{n}\\&=\left(\underline{\nu}(0)\,\mathrm{P}\right)\mathrm{P}^{(n-1)}\\&=\underline{\nu}(0)\,\mathrm{P}^{(n-1)}\\&=\left(\underline{\nu}(0)\,\mathrm{P}\right)\mathrm{P}^{(n-2)}\\&=\cdots=\underline{\nu}(0)\,\mathrm{P}=\underline{\nu}(0)\end{aligned}$$

ולכן לפי ההגדרה הפילוג הוא סטציונרי. מצד שני, אם הפילוג הוא סטציונרי אזי לפי ההגדרה

$$\begin{aligned}\underline{\nu}\,\mathrm{P}&=\underline{\nu}(0)\,\mathrm{P}\\&=\underline{\nu}(1)=\underline{\nu}\end{aligned}$$

ובכך הוכחנו את טענה 1. כעת נזכר בהגדרת הסטציונריות 3.14 נקבע $t_1 < t_2$ ו-$\tau > 0$ ונבדוק לפי ההגדרה. נשתמש בהגדרת הסתברות מותנית, במרקוביות ובהומוגניות:

$$\begin{aligned}
&\mathbb{P}\left\{X(t_2+\tau)=j_k, X(t_2-1+\tau)=j_{k-1},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad=\mathbb{P}\left\{X(t_2+\tau)=j_k \mid X(t_2-1+\tau)=j_{k-1},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad\quad\times\mathbb{P}\left\{X(t_2-1+\tau)=j_{k-1},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad=\mathbb{P}\left\{X(t_2+\tau)=j_k \mid X(t_2-1+\tau)=j_{k-1}\right\} \\
&\quad\quad\times\mathbb{P}\left\{X(t_2-1+\tau)=j_{k-1} \mid X(t_2-2+\tau)=j_{k-2},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad\quad\times\mathbb{P}\left\{X(t_2-2+\tau)=j_{k-2},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad=\mathbb{P}\left\{X(t_2)=j_k \mid X(t_2-1)=j_{k-1}\right\}\times\mathbb{P}\left\{X(t_2-2+\tau)=j_{k-2} \mid X(t_2-2)=j_{k-2}\right\} \\
&\quad\quad\times\cdots\times\mathbb{P}\left\{X(t_1+\tau)=j_1\right\}=\prod_{n=2}^{k}\mathrm{p}_{j_{n-1}j_n}\,\nu(t_1+\tau)
\end{aligned}\tag{4.5}$$

אבל ראינו שהפילוג החד-מימדי אינו תלוי בזמן, כלומר

$$\mathbb{P}\left\{X(t_1+\tau)=j_1\right\}=\nu_{j_1}(t_1+\tau)=\nu_{j_1}(t_1)=\mathbb{P}\left\{X(t_1)=j_1\right\}$$

ולכן ההסתברות (4.5) אינה תלויה ב-τ ובפרט ערכה שווה ב-τ וב-0. ביתר פירוט אם נחזור על צעדינו בכיוון ההפוך, נקבל

$$\begin{aligned}
&\mathbb{P}\left\{X(t_2+\tau)=j_k, X(t_2-1+\tau)=j_{k-1},\dots,X(t_1+\tau)=j_1\right\} \\
&\quad=\mathbb{P}\left\{X(t_2)=j_k \mid X(t_2-1)=j_{k-1}\right\} \\
&\quad\quad\times\mathbb{P}\left\{X(t_2-2+\tau)=j_{k-2} \mid X(t_2-2)=j_{k-2}\right\}\times\cdots\times\mathbb{P}\left\{X(t_1)=j_1\right\} \\
&\quad=\mathbb{P}\left\{X(t_2)=j_k, X(t_2-1)=j_{k-1},\dots,X(t_1)=j_1\right\}
\end{aligned}$$

ולכן התהליך הוא סטציונרי.

מתי קיים פילוג סטציונרי? ומתי הוא יחיד? ברור כי לשרשרת חולפת אין פילוג סטציונרי. להילוך שיכור סימטרי (כל צעד הוא ± 1 בהסתברות $1/2$ כל אחד) אין פילוג סטציונרי, כיוון שהפתרון היחיד (עד כדי קבוע) של המשוואה (האין-סופית!)

$$\underline{\nu}\cdot\mathrm{P}=\underline{\nu}$$

הוא הווקטור $\underline{\nu}=(\dots,1,1,1,\dots)$ אשר אינו פילוג.

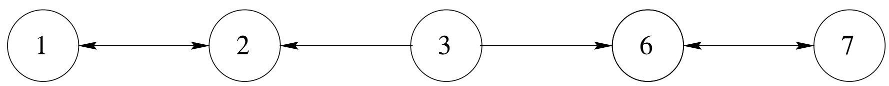
טענה 4.31 (ללא הוכחה). לכל שרשרת מרקוב סופית יש פילוג סטציונרי. השרשרת אינה פריקה אם ורק אם הפילוג הוא יחיד. הפילוג הסטציונרי מקבל ערכים שונים מ-0 רק על מצבים נשנים.

הערה: קיום פילוג סטציונרי נובע מתוצאות באלגברה לינארית: למשוואה

$$\mathrm{P}\,\underline{x} = \underline{x}$$

יש פיתרון (הווקטור $\underline{x} = (\dots, 1, 1, 1, \dots)^T$), כלומר למטריצה P יש וקטור עצמי (ימני) עם ערך עצמי 1. מכך נובע שלמטריצה זו יש גם ווקטור עצמי שמאלי עם ערך עצמי 1. משפט פרון-פרובניוס מתורת המטריצות מבטיח כי, כיוון שאברי המטריצה P הם חיוביים, אזי לווקטור העצמי השמאלי (השייך לערך העצמי הגדול ביותר של P) רכיבים חיוביים. כיוון שמדובר בווקטור סופי, ניתן לנרמל אותו כך שיהיה פילוג, ומטענה 4.30 זהו פילוג סטציונרי.

דוגמה 4.32 בדוגמה שלפנינו יש שתי קבוצות סגורות, ולכן הדוגמה אינה מקיימת את תכונת האי-פריקות. נניח שכל המעברים קורים בהסתברות $1/2$.



ציור 4.6: פילוג סטציונרי

אזי קל לבדוק כי לכל $0 \le \alpha \le 1$, הפילוג

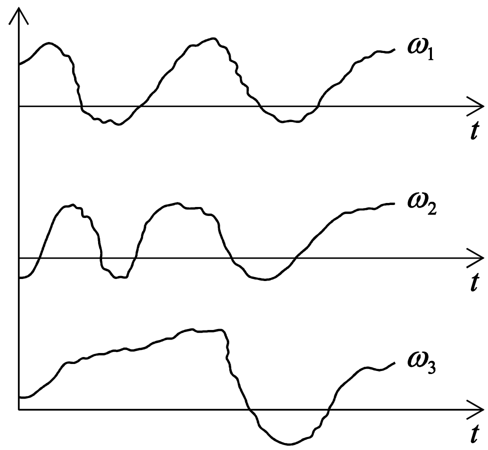
$$\underline{\nu} = (\alpha/2, \alpha/2, 0, (1-\alpha)/2, (1-\alpha)/2) \tag{4.6}$$

הוא פילוג סטציונרי. כלומר, הפילוג הסטציונרי אינו יחיד.

5 תהליכים אקראיים בזמן רציף

5.1 מבוא, הגדרות ודוגמאות

נתאר לעצמנו "∞" מקלטים שהופעלו בזמן t_0, כולם מאותו סוג ומכוונים לאותו תדר. נניח שאין סיגנל כניסה למקלט והרעש העיקרי נוצר במקלט עצמו. נסמן ב-$X(t,\omega)$ את היציאה מהמקלט עבור t נתון. כפונקציה של פרמטר המזל, $X(t,\omega)$ הוא מ"א, לדוגמא, $X(2,\omega)$ הוא מ"א. עבור ω קבוע, בקטע $a \le t \le b$,$X(t,\omega)$ הוא פונקציה של t, פונקצית מדגם (ראה ציור 5.1).



ציור 5.1

במושג תהליך אקראי(ת"א) $\{X_t, a \le t \le b\}$ או $\{X(t,\omega), a \le t \le b\}$ אנו מבינים: אוסף של משתנים אקראים X_t כך שעבור כל n וכל מסרק זמנים $t_1,\ldots,t_n$ ב-$[a,b]$, ידוע חוק ההסתברות של הוקטור האקראי $X_{t_1},\ldots,X_{t_n}$ אנו נשתמש בהמשך באחד מן הסימונים הבאים לציון תהליך אקראי: $X(t,\omega), X(t), X_t$.

<u>דוגמא 1:</u> $X(t,\omega) = A(\omega)t + B(\omega)$ כאשר A,B מ"א. צורות גל טיפוסיות (פונקציות מדגם) במקרה זה הן קוים ישרים.

<u>דוגמא 2:</u> $X(t,\omega) = A(\omega)\sin(2\pi f t + \phi(\omega))$ כאשר A,ϕ מ"א. צורות הגל הטפוסיות במקרה זה הן תנודות הרמוניות בעלות אמפליטודה A ופאזה ϕ.

<u>דוגמא 3:</u> N קבוע (או אקראי) $X(t,\omega) = \sum_{n=1}^{N} X_n(\omega)\sin nt$ כאשר X_n הם מ"א. צורות גל טפוסיות במקרה זה הן סכום משוקלל של תנודות הרמוניות.

<u>חוק ההסתברות של תהליך אקראי:</u>

יהיה $\{X_t, T_1 \le t \le T_2\}$ תהליך אקראי. נעיין באוסף חוקי ההסתברות כדלקמן: עבור כל n וכל $(T_1 \le t_i \le T_2) t_1, t_2, \ldots, t_n$

$$F_{X_{t_1},\ldots,X_{t_n}}(a_1,\ldots,a_n) = \text{Prob}\{X_{t_1} \le a_1,\ldots,X_{t_n} \le a_n\}$$

אוסף זה של פונקציות פילוג נקרא חוק ההסתברות של התהליך האקראי X_t והוא חייב לקיים את חוק הקונסיסט-

נטיות:

$$F_{X_{t_1},\dots,X_{t_n}}(a_1,\dots,a_n) = F_{X_{t_1},\dots,X_{t_{n+1}}}(a_1,\dots,a_n,\infty)$$

תהליך פואסון

נביא עתה דוגמא חשובה של ת"א הנקרא תהליך פואסון. תהליך זה משמש מודל למספר רב של תופעות כגון קריאת מונה גייגר, מעבר מכוניות בנקודה בכביש, כניסת שיחות למרכזית טלפונים, פליטת אלקטרונים וכו'. על מנת להגדיר את תהליך פואסון נתאר לעצמנו ארוע המתרחש באקראי מדי פעם ונסמן ב N_t את מספר המאורעות המתרחשים בפרק הזמן $[0,t)$. לכן התהליך N_t הוא תהליך מניה, דהיינו, הוא מונוטוני לא יורד ויכול לקבל ערכים שלמים $0,1,2,\dots$ בלבד. לגבי חוק ההסתברות של התהליך נניח את ההנחות הבאות:

(א) $\text{Prob}\{(N_{t+\Delta t} - N_t) = 1\} = \lambda\cdot\Delta t + 0(\Delta t)$ כלומר ההסתברות להתרחשות מאורע אחד בלבד בקטע הזמן $(t, t+\Delta t]$ הוא $\lambda\cdot\Delta t + 0(\Delta t)$. כאן $0(\Delta t)$ הוא גודל זניח לגבי Δt כאשר $\Delta t \to 0$,(באופן כללי $f(\varepsilon)$ הוא $0(\varepsilon)$ אם $f(\varepsilon)/\varepsilon \to 0$ כאשר $\varepsilon \downarrow 0$).

(ב) $\text{Prob}\{(N_{t+\Delta t} - N_t) = 0\} = 1-\lambda\cdot\Delta t + 0(\Delta t)$ כלומר ההסתברות שלא התרחש אף מאורע בקטע הזמן $(t, t+\Delta t]$ הוא $1-\lambda\cdot\Delta t + 0(\Delta t)$.

(ג) ארועים בקטעי זמן לא חופפים הם בלתי תלויים. (תהליך כזה נקרא תהליך בעל תוספות בלתי תלויות או הפרשים בלתי תלויים).

מ (א) ו (ב) נובע ש: $\text{Prob}\{(N_{T+\Delta t} - N_t) \geq 2\} = \text{Prob}\,\{\text{שני מאורעות או יותר}\} = 0(\Delta t)$

<u>תזכורת:</u> אם בניסוי בודד יש סיכוי הצלחה p וסיכוי כשלון $(p+q=1)$ אזי ההסתברות ל-k הצלחות ב-n ניסויים בלתי תלויים נתונה ע"י הביטוי

$$\binom{n}{k} p^k q^{n-k} = \frac{n!}{(n-k)!k!} p^k q^{n-k}$$

נחזור לתהליך פואסון: את הקטע $(0,T]$ נחלק ל-$n = T/\Delta$ קטעי זמן. עבור קטעי הזמן הקצרים שנקבל, נפעיל את ההנחות (א) - (ג). נשים לב שעל מנת לקבל $N_T = k$ ארועים בקטע $(0,T]$ צריכים $N-k$ קטעי זמן להיות ללא ארוע ("כשלון"), ואילו ב k קטעי זמן צריך להיות ארוע ("הצלחה"); לכן, כאשר Δ קטן (n גדול) נקבל

$$\text{Prob}\{N_T = k\} \cong \frac{n!}{(n-k)!k!}(\lambda\Delta)^k (1-\lambda\Delta)^{n-k} = \frac{n!}{(n-k)!k!}\left(\frac{\lambda T}{n}\right)^2 \left(1-\lambda\frac{T}{n}\right)^{-k} \to \frac{(\lambda T)^k}{k!} e^{\lambda T}$$

תזכורת

$$\frac{n!}{(n-k)!k!} \to 1\ ; \quad \left(1-\lambda\frac{T}{n}\right)^k \to 1\ ; \quad \left(1-\lambda\frac{T}{n}\right)^n \to e^{\lambda T}$$

ולכן

$$\text{Prob}\{N_T = k\} = \frac{(\lambda T)^k}{k!} e^{-\lambda T}$$

ולכן גם

$$\text{Prob}\{N_{T_2} - N_{T_1} = k\} = \frac{[\lambda(T_2 - T_1)]^k}{k!} e^{-\lambda(T_2 - T_1)}$$

אפשר להוכיח ע"י חשבון ישיר ש-

$$E[N_T] = \lambda T$$

$$E[N_T^2] = (\lambda T)^2 + \lambda T$$

$$\text{Var}(N_T) = (\lambda T)^2 + \lambda T - (\lambda T)^2 = \lambda T$$

במקום להוכיח תוצאות אלה ישירות ע"י שימוש בתוצאות ידועות עבור טורים מסוימים, נחשב כדלקמן: את הקטע $(0, T]$ נחלק לקטעים קטנים ונסמן את נקודות החלוקה ב-$t_0, t_1, \ldots, t_n$ ($t_0 = 0, t_n = T$):

$$E[N_T] = E\left[\sum (N_{t_{i+1}} - N_{t_i})\right] \cong \sum \lambda(t_{i+1} - t_i) = T \cdot \lambda$$

$$E[N_T^2] = E\left[\left(\sum (N_{t_{i+1}} - N_{t_i})\right)^2\right] = E\left[\sum (N_{t_{i+1}} - N_{t_i})^2\right] = E\left[\sum_{i \neq j}\sum (N_{t_{i+1}} - N_{t_i})(N_{t_{j+1}} - N_{t_j})\right]$$

$$= \lambda T + \sum_{\substack{\text{i} \\ i \neq j}} \sum_{\text{j}} \Big(\lambda(t_{i+1} - t_i) \cdot \lambda(t_{j+1} - t_j)\Big) = \lambda T + \lambda^2 T^2$$

אגב, עבור $t_2 > t_1$

$$E[N_{t_2} N_{t_1}] = E[N_{t_1}^2] + E\Big[N_{t_1}(N_{t_2} - N_{t_1})\Big] = (\lambda t_1)^2 + \lambda t_1 + \lambda^2 t_1 (t_2 - t_1) = \lambda t_1 + \lambda^2 t_1 t_2$$

בגלל הנחה (ג), תהליך פואסון הוא תהליך של "תוספות בלתי תלויות", דהיינו: אם $t_1 < t_2 \leq t_3 < t_4$, אזי

$$\text{Prob}\{N_{t_2} - N_{t_1} = k_1, N_{t_4} - N_{t_3} = k_2\} = \text{Prob}\{N_{t_2} - N_{t_1} = k_1\}\, \text{Prob}\{N_{t_4} - N_{t_3} = k_2\}$$

ולכן, עבור מסרק זמנים $t_1 < t_2 < t_3$ נקבל מתוך (ד):

$$\text{Prob}\{N_{t_1} = k_1, N_{t_2} = k_2, N_{t_3} = k_3\}$$

$$= \text{Prob}\{N_{t_1} = k_1\}\, \text{Prob}\{N_{t_2} - N_{t_1} = k_2 - k_1\}\, \text{Prob}\{N_{t_3} - N_{t_2} = k_3 - k_2\}$$

בצורה כזו אנו יכולים לקבל את <u>חוק ההסתברות</u> של $N_{t_1}, \ldots, N_{t_n}$ לכל n ולכל $t_1, \ldots, t_n$ ולכן חוק ההסתברות של תהליך פואסון ידוע. בגלל ההפרשים הבלתי תלויים חוק ההסתברות הוא פשוט. בדרך כלל אין הדבר כך. בהמשך נראה מקרה אחר שבו נוכל לתאר באופן פשוט את חוק ההסתברות של התהליך (תהליך גאוסי).

תהליכים הקשורים לתהליך פואסון

עבור תהליך פואסון N_t נגדיר

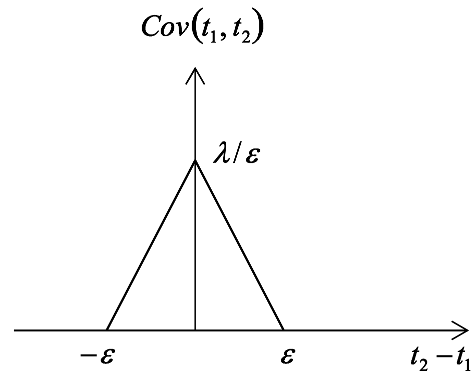
$$Y_\varepsilon(t) \triangleq \frac{N_{t+\varepsilon} - N_t}{\varepsilon}$$

שים לב ש- $Y_0(t)Y\varepsilon(t) \geq 0$, $E[Y_\varepsilon(t)] = \frac{\lambda(t+\varepsilon)-\lambda t}{\varepsilon} = \lambda$; עבור $\varepsilon \to 0$ נקבל ש $Y_0(t)$ הוא סופרפוזיציה של פונקציות דירק. נחזור ל-$\varepsilon > 0$ ונגדיר

$$Z_\varepsilon(t) \triangleq Y_\varepsilon(t) - E[Y_\varepsilon(t)] = \frac{N_{t+\varepsilon} - N_t - \varepsilon\lambda}{\varepsilon}$$

$Y_\varepsilon(t)$ ו-$Z_\varepsilon(t)$ אינם תהליכים של הפרשים בלתי תלויים. נעיין ב- $\mathrm{Cov}(Y_\varepsilon(t_1), Y_\varepsilon(t_2)) = E[Z_\varepsilon(t_1)Z_\varepsilon(t_2)]$ ונבחין בין שני מקרים:

(א) $|t_2 - t_1| \geq \varepsilon$ (ב) $|t_2 - t_1| < \varepsilon$ במקרה (א) ברור ש- $E[Z_\varepsilon(t_1)Z_\varepsilon(t_2)] = 0$. במקרה (ב) וללא הוכחה (החשבון פשוט) מתקבלת התוצאה המופיעה בציור 5.2



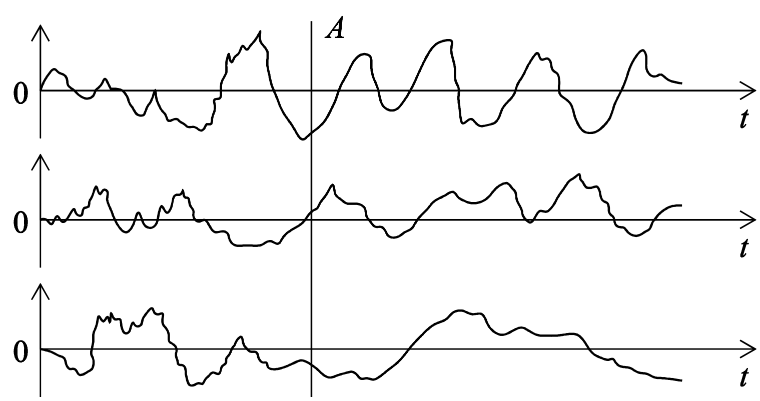
ציור 5.2

<u>שאלות:</u> האם $\mathrm{Prob}\{Z_\varepsilon(t) \leq \alpha\}$ $(t > 0)$ תלוי ב t?

האם $\mathrm{Prob}\{Z_\varepsilon(t) \leq \alpha, Z_\varepsilon(t+5) \leq \beta\}$ $(t > 0)$ תלוי ב t?

5.2 סטציונריות:

נעיין ביציאה מאוסף ("אינסופי") של מגברים החל מרגע הפעלת המגבר. נניח שאין סיגנל וכל היציאה היא רעש שבא מהמגבר עצמו. באופן טיפוסי נקבל את הרישום הבא:



ציור 5.3

החלק מ-$t=0$ ועד לסביבת הזמן $t=A$ מתאר את תהליך התחממות המגבר והוא מתאר לכן תופעת מעבר אקראית. לאחר מכן מגיע המגבר למצב הניתן לתאור כ"מצב מתמיד". נתאר ב-$\{X(t,\omega),\ t\geq 0\}$ את רעש היציאה של המקלט. נעיין ב- $\text{Prob}\{X(t,\omega)\leq a\}$ עבור a קבוע. חוק ההסתברות $F_{X_t}(a)$ בדרך כלל תלוי ב-t, ואכן בדוגמא זו עבור $t<A$ יהיה החוק תלוי ב-t. לעומת זאת עבור $A\ll t$, כלומר, בתחום הזמן של "מצב מתמיד" סביר להניח ש-$F_{X_t}(a)$ יהיה למעשה בלתי תלוי ב-t. שים לב: אין פרוש הדבר ש-$X(t,\omega)$ בלתי תלוי ב-t עבור $A\ll t$ (ראה הציור); רק חוק ההסתברות אינו תלוי ב-t. על מנת שנוכל לטפל בנוחיות בתחום הזמן המתאר את המצב המתמיד נגדיר את המושג של ת"א <u>סטציונרי</u> .

<u>הגדרה:</u> התהליך $\{X(t,\omega),\infty<t<\infty\}$ נקרא ת"א סטציונרי אם עבור כל n , כל $t_1,\ldots,t_n$ וכל τ , חוקי ההסתברות של הוקטור האקראי:

$$\Big(X_{t_1},X_{t_2},\ldots,X_{t_n}\Big)^T$$

ושל הוקטור האקראי:

$$\Big(X_{t_1+\tau},X_{t_2+\tau},\ldots,X_{t_n+\tau}\Big)^T$$

זהים זה לזה, דהיינו:

$$F_{X_{t_1},X_{t_2}\,,\ldots,X_{t_n}}\,(a_1,a_2,\ldots,a_n)=F_{X_{t_1+\tau},X_{t_2+\tau}\,,\ldots,X_{t_n+\tau}}\,(a_1,a_2,\ldots,a_n)$$

בצורה מקוצרת נאמר שת"א הוא סטציונרי אם חוק ההסתברות שלו אינורינטי להזזת ציר הזמן (חוק ההסתברות תלוי בהפרשי זמנים בלבד). תהליך פואסון הוא דוגמא לת"א לא סטציונרי; אפשר לטעון שכל תהליך פיסיקלי איננו סטציונרי מאחר שהוא התחיל בזמן כלשהוא ואכן, המושג של ת"א סטציונרי הוא אידיאליזציה.

<u>דוגמאות:</u>

1. נעיין בשני התהליכים הבאים שהם ת"א מנוונים:

$$\{X(t,\omega)=5\ ,\qquad -\infty<t<\infty\}$$
$$\{X(t,\omega)=5\sin 2\pi t\ ,\qquad -\infty<t<\infty\}$$

התהליך הראשון סטציונרי והשני לא. מדוע? אם $A(\omega)$ מ"א כלשהוא אזי התהליך $X(t,\omega)=A(\omega)$ הוא סטציונרי, מדוע?

2. (ללא הוכחה): לתהליך פואסון נדביק עוד תהליך פואסון בעל ערך שלילי ובזמנים שלילים כאשר החלק בזמן חיובי והחלק בזמן שלילי בלתי תלויים. את התהליך ב-$(-\infty,\infty)$ נסמן ב-N_t. נגדיר $Y(t)\triangleq N_{t+\varepsilon}-N_t$ או $Y_\varepsilon(t)\triangleq(N_{t+\varepsilon}-N_t)/\varepsilon$ אזי $Y(t)$ ו-$Y_\varepsilon(t)$ (ε קבוע) הם ת"א סטציונריים.

3. נעיין בתהליך האקראי

$$X(t,\omega)=A\cos\Big(2\pi f_0t+2\pi\phi(\omega)\Big),\quad\{-\infty<t<\infty\}$$

כאשר f_0,A קבועים (אינם אקראיים) ו-$\phi(\omega)$ מ"א אקראי המפולג באופן אחיד בתחום $[0,1)$

<u>טענה:</u> התהליך $X(t,\omega)$ הוא סטציונרי.

<u>הוכחה:</u> נקבע $n \quad a_1,\ldots,a_n,t_1,\ldots,t_n,$ ונעיין במאורע B המוגדר כדלקמן:

$$B = \Big\{\omega:\ X(t_1,\omega)\le a_1,\ldots,X(t_n,\omega)\le a_n\Big\}$$

נגדיר גם את Φ_B כדלקמן:

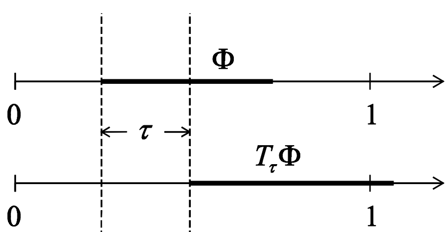
$$\Phi_B = \Big\{\theta:\ \theta\in[0,1)\Big\}\bigcap\Big\{A\cos(2\pi f_0 t_1+2\pi\theta)\le a_1,\ldots,A\cos(2\pi f_0 t_n+2\pi\theta)\le a_n\Big\}$$

דהיינו: Φ_B הוא אוסף כל המספרים θ כך שאם $\phi(\omega)\in\Phi_B$ (כלומר, אם הגרלנו $\phi(\omega)$ כזה השייך ל-Φ) אזי ארע המאורע שקראנו לו B. באופן כללי $\mathrm{Prob}\{B\}=\mathrm{Prob}\{\phi\in\Phi_B\}$ ובמיוחד, אם Φ_B הוא קטע ב-$[0,1)$ אזי $\mathrm{Prob}\{B\}$ הוא ארך הקטע, ואם Φ_B הוא אחוד של קטעים לא חופפים אזי

$$\mathrm{Prob}\{B\}=\{\text{סכום ארכי הקטעים המהווים את }\Phi_B\}$$

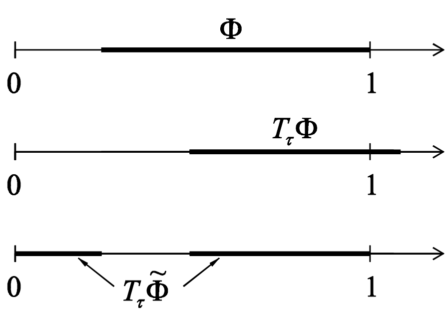
תהייה Φ קבוצת נקודות כלשהיא ב-$[0,1)$ נגדיר את ההזזה של Φ בכמות τ ע"י (ראה ציור 5.4):

$$T_\tau\Phi=\{\phi:\phi-\tau\in\Phi\}$$



ציור 5.4

שים לב שאם גם Φ וגם $T_\tau\Phi$ ב-$[0,1)$ אזי $\mathrm{Prob}\{\phi(\omega)\in\Phi\}=\mathrm{Prob}\{\phi(\omega)\in T_\tau\Phi\}$ וזאת מאחר ו-$\phi(\omega)$ מפולג אחיד בתחום $[0,1)$. עבור מספר ממשי X נסמן ב-$[X]$ את החלק השלם של X (כגון אם $X=2.3$, $[X]=2$) וב-$\tilde{X}$ נסמן $X-[X]$ (0.3 בדוגמא). לכן אם $X=-1.6$ אזי $[X]=-2$ ו-$\tilde{X}=4$. שים לב ש-$\tilde{X}$ תמיד בתחום $[0,1)$. עבור קבוצת נקודות Φ ב-$(-\infty,\infty)$ נסמן ב-$\tilde{\Phi}$ את הקבוצה $\tilde{\Phi}=\{\tilde{\phi}:\phi\in\Phi\}$ (ראה ציור 5.5)



ציור 5.5

כעת אם Φ קבוצת נקודות ב-$[0,1)$ אזי

$$\mathrm{Prob}\Big\{\phi(\omega)\in\Phi\Big\}=\mathrm{Prob}\Big\{\phi(\omega)\in(T_\tau\tilde{\Phi}))\Big\}$$

עבור כל τ ושוב, הדבר נכון כי ϕ מפולג אחיד. נחזור עכשיו ל-

$$\Phi_B=\Big\{\phi:\ A\cos(2\pi f_0 t_i+2\pi\phi)\le a_i,\quad i=1,\ldots,n\Big\}$$

נעיין כעת ב-

$$\Phi_C=\Big\{\phi:\ A\cos(2\pi f_0(t_i+\tau)+2\pi\phi)\le a_i,\quad i=1,\ldots,n\Big\}$$

אזי $\Phi_C=T_{-2\pi f_0\tau}\Phi_B$ ולכן

$$\begin{aligned}\mathrm{Prob}\{X_{t_i+\tau}\le a_i,\quad i=1,\ldots,n\}&=\mathrm{Prob}\{\phi\in\tilde{\Phi}_C\}=\mathrm{Prob}\{\phi\in T_{-2\pi f_0\tau}\tilde{\Phi}_B\}\\&=\mathrm{Prob}\{\phi\in\Phi_B\}=\mathrm{Prob}\{X_{t_i}\le a_i,\quad i=1,\ldots,n\}\end{aligned}$$

ובזאת השלמנו את ההוכחה.

מומנטים

נעזוב לרגע את מושג הסטציונריות ונתרכז במומנטים מסדר ראשון ומסדר שני הקשורים לת"א. המקרים בהם ידוע מפורשות חוק ההסתברות של ת"א נדירים, אולם להרבה בעיות מספיק לדעת את המומנטים מסדר ראשון ושני (וכן כפי שנראה קיים המושג של תהליך אקראי גאוסי שעבורו ידיעת המומנטים מסדר ראשון ושני מגדירה את חוק ההסתברות).

הגדרות:

פונקצית התוחלת:

$$\mu_X(t)\triangleq E[X(t)]$$

פונקצית האוטוקורילציה:

$$\mathbf{R}_X(t_1,t_2)\triangleq E[X_{t_1}X_{t_2}]$$

פונקצית הקווריאנס:

$$\mathbf{K}_X(t_1,t_2)\triangleq E\Big[(X_{t_1}-\mu_X(t_1))(X_{t_2}-\mu_X(t_2))\Big]=\mathbf{R}(t_1,t_2)-\mu_X(t_1)\mu_X(t_2)$$

כפי שנראה בהמשך, לפונקציות הנ"ל יש חשיבות רבה לגבי תהליכים אקראיים. להלן נסכם מספר תכונות פשוטות וחשובות של פונקצית האוטוקורלציה.

1. $\mathbf{R}_X(t_1,t_2)=\mathbf{R}_X(t_2,t_1)$
2. $\mathbf{R}_X(t_1,t_1)\ge 0$

3. עבור כל n וכל $t_1, \ldots, t_n$ המטריצה:

$$\begin{pmatrix} \mathbf{R}_X(t_1,t_1) & \mathbf{R}_X(t_1,t_2) & \ldots & \mathbf{R}_X(t_1,t_n) \\ \mathbf{R}_X(t_2,t_1) & & & \\ \vdots & & & \vdots \\ \mathbf{R}_X(t_n,t_1) & \ldots & \ldots & \mathbf{R}_X(t_n,t_n) \end{pmatrix}$$

מטריצה לא שלילית (מדוע? הוכח).

<u>סטציונריות במובן הרחב</u>

נחזור לסטציונריות; עבור ת"א סטציונרי ברור שמתקיים $F_{X_t}(a) = F_{X_{t'}}(a) \quad \forall a$ לכל t ו-t' לכן

$$E[X_t^m] = E[X_{t'}^m]$$

ולכן המומנטים מכל סדר שהוא הינם בלתי תלויים בזמן. בפרט, המומנט הראשון $\mu_X(t)$ אינו תלוי בזמן, כלומר

$$E[X(t)] = \mu_X(t) = \mu_X(0) = \mu_X$$

שים לב: אם המומנטים מכל סדר שהוא אינם תלויים בזמן אין הדבר גורר סטציונריות של התהליך.

בנוסף $F_{X_{t_1},X_{t_2}}(a_1,a_2) = F_{X_{t_1+\tau},X_{t_2+\tau}}(a_1,a_2)$ ולכן לגבי המומנט השני חייב להתקיים:

$$E[X_{t_1}X_{t_2}] = E[X_{t_1+\tau}X_{t_2+\tau}]$$

לכן $\mathbf{R}_X(t_1,t_2) = \mathbf{R}_X(t_1+\tau, t_2+\tau)$ ואם נבחר $\tau = -t_1$ נקבל:

$$\mathbf{R}_X(t_1,t_2) = \mathbf{R}_X(0, t_2-t_1) = \mathbf{R}_X(|t_2-t_1|)$$

או

$$\mathbf{R}_X(t_1, t_1+\tau) = R_X(\tau) = R_X(-\tau)$$

ולכן גם:

$$\mathbf{K}_X(t_1,t_2) = K_X(|t_1-t_2|);\ \mathbf{K}_X(t_1,t_1+\tau) = K_X(|\tau|)$$

דהיינו: עבור ת"א סטציונרי פונקציות האוטוקורילציה והקווריאנס הן פונקציות של משתנה אחד (והוא הפרש הזמנים) בלבד.

<u>הגדרה:</u> ת"א אקראי נקרא "סטציונרי במובן הרחב" אם $\mu_X(t)$ בלתי תלוי ב-t ו-$\mathbf{R}_X(t_1,t_2)$ תלוי ב-$t_1 - t_2$ בלבד דהיינו:

א. עבור כל t מתקיים $\mu_X(t) = \mu_X(0) = \mu_X$

ב. עבור כל t_2, t_1 מתקיים $\mathbf{R}_X(t_1,t_1+\tau) = \mathbf{R}_X(0,\tau) = R_x(\tau)$

במילים אחרות: ת"א נקרא "סטציונרי במובן הרחב" אם מבחינת המומנטים מסדר ראשון ושני הוא "נראה כאילו היה סטציונרי". ברור שאם ת"א סטציונרי (ושני המומנטים הראשונים קיימים) אזי הוא גם סטציונרי במובן הרחב. ההיפך אינו נכון בדרך כלל; אפשר לראות שישנם ת"א סטציונרים במובן הרחב שאינם סטציונריים.

בדרך כלל קל יותר להוכיח סטציונריות במובן הרחב מאשר סטציונריות. כדי להדגים זאת נעיין ב-$X_t = A(\omega)\cos(2\pi f_0 t+\phi(\omega))$ כאשר A, ϕ מ"א בלתי תלויים ו-ϕ מפולג אחיד בתחום $[0, 2\pi)$. עבור $A(\omega)$ דטרמיניסטי הוכחנו שהתהליך סטציונרי (וההוכחה ניתנת להרחבה ל-A אקראי בלתי תלוי ב-ϕ). נוכיח שהתהליך סטציונרי במובן הרחב. היות ו-ϕ מפולג אחיד בתחום $[0, 2\pi)$:

$$
\begin{aligned}
E[X_t] &= E\Big[A(\omega)\cos(2\pi f_0 t + \phi(\omega))\Big] = E\Big[A(\omega)\Big]E\Big[\cos(2\pi f_0 t + \phi(\omega))\Big] \\
&= E[E(\omega)] \cdot \frac{1}{2\pi}\int_0^{2\pi}\cos(2\pi f_0 t + \phi)d\phi = 0
\end{aligned}
$$

$$
\begin{aligned}
E[X_t, X_{t+\tau}] &= E\Big[A^2(\omega)\cos(2\pi f_0(t+\tau)+\phi)\cdot\cos(2\pi f_0 t+\phi)\Big] \\
&= \frac{1}{2}E\Big[A^2(\omega)\cos 2\pi f_0\tau\Big] + \frac{1}{2}E\Big[A^2\cos(2\pi f_0(2t+\tau)+2\phi)\Big] \\
&= \frac{1}{2}E\Big[A^2(\omega)\Big]\cdot\cos 2\pi f_0\tau + 0
\end{aligned}
$$

ולכן התהליך סטציונרי במובן הרחב. כדאי לזכור שעבור תהליך פשוט זה

$$R_X(\tau) = \frac{E[A^2]}{2}\cos 2\pi f_0\tau$$

<u>הערה:</u> חזור על דוגמא זאת כאשר כמקודם ϕ ו A בלתי תלוים , אולם הפעם $\phi = 0, \pi/2, \pi, 3\pi/2$ בהסתברות 1/4. האם תהליך זה סטציונרי במובן הרחב? האם תהליך זה סטציונרי?

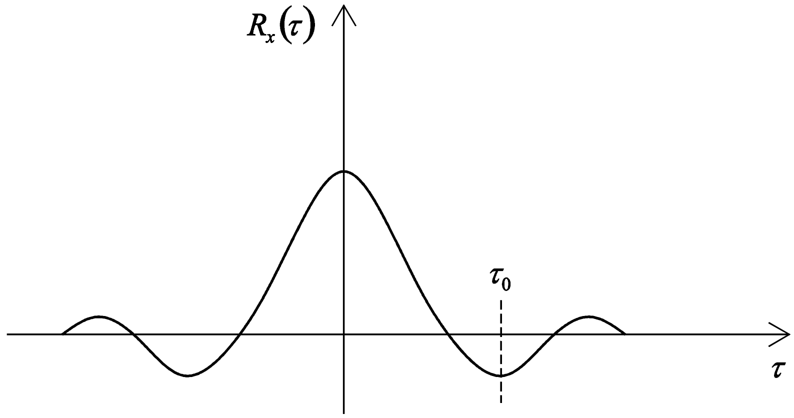
<u>תכונות:</u>

1. אם X_t סטציונרי במובן הרחב אזי $E[(X_{t+\tau} - X_t)^2] = 2[R_X(0) - R_X(\tau)]$.
2. $R_X(0) \geq 0$
3. $R_X(\tau) = R_X(-\tau)$
4. $R_X(0) \geq |R_X(\tau)|$ (הוכח בעזרת $E[X_{t+\tau} \pm X_t)^2] \geq 0$).

מדוע אנחנו מעונינים בפונקצית האוטוקורילציה? היא מסכמת את המומנטים מסדר שני של התהליך וכפי שנראה אח"כ היא מאפשרת לענות על שאלות מענינות. הסבר (חלקי מאוד) על משמעות פונקצית האוטוקורילציה ניתן לראות מהשיקול הבא: נניח סטציונריות, $\mu_X(t) = 0$, ונניח שעבור τ "גדול מאוד", X_t ו- $X_{t+\tau}$ "כמעט" בלתי תלויים סטטיסטית. במקרה זה

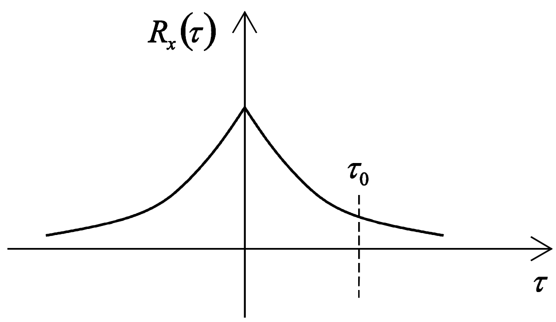
$$R_X(\tau) \underset{|\tau|\to\infty}{\longrightarrow} 0$$

כמו כן ידוע לנו ש-$|R_X(\tau)| \geq R_X(0)$ ולכן מהלך $R_X(\tau)$ יהיה משהו כגון:



ציור 5.6

או



ציור 5.7

נוכל איפוא לומר שבקרוב גס פונקצית האוטוקורילציה "כמעט" מתאפסת עבור $\tau > \tau_0$ ולכן τ_0 מתאר במקרים אלה את "הזכרון האפקטיבי של התהליך". כדי לראות שפונקצית האוטוקורלציה נותנת מעין תאור גס של "הזכרון הטבעי" של התהליך נניח שהקלטנו את $X(t)$ על סרט מגנטי ואנו מריצים את ההקלטה במהירות שהיא α פעמים המהירות המקורית נקבל:

$$Y(t) = X(\alpha t)$$

$$R_Y(\tau) = R_X(\alpha\tau)$$

ואם α גדול אזי $R_Y(\tau)$ מתכווץ והזכרון מתקצר.

קורילציה מצטלבת

לצורך ההמשך נזדקק למושגים הבאים (הקשורים במומנטים מסדר שני של זוגות של תהליכים אקראיים). עבור זוג תהליכים אקראיים X_t ו-Y_t נגדיר את הקורלציה המצטלבת (קרוסקורלציה) כדלקמן:

$$\mathbf{R}_{X,Y}(t_1, t_2) \triangleq E[X_{t_1} Y_{t_2}]$$

ולכן,

$$\mathbf{R}_{X,Y}(t_1,t_2)=\mathbf{R}_{Y,X}(t_2,t_1)$$

כאשר $Z_t=X_t+Y_t$ מתקבל

$$\begin{aligned}\mathbf{R}_Z(t_1,t_2)&=E\Big[(X_{t_1}+Y_{t_1})(X_{t_2}+Y_{t_2})\Big]\\&=\mathbf{R}_X(t_1,t_2)+\mathbf{R}_Y(t_1,t_2)+\mathbf{R}_{X,Y}(t_1,t_2)+\mathbf{R}_{Y,X}(t_1,t_2)\end{aligned}$$

כך ש-$\mathbf{R}_{X,Y}$ מופיע באופן טבעי כאשר אנו רוצים לחשב את $\mathbf{R}_Z$.

<u>הגדרה:</u> זוג תהליכים אקראיים $\{X_t,Y_y,-\infty<t<\infty\}$ נקרא סטציונרי (=סטציונרי במשותף) אם עבור כל n , כל $t_1,\dots,t_n$ וכל τ חוק ההסתברות של הוקטור האקראי ה-$2n$ מימדי

$$\Big(X_{t_1+\tau},X_{t_2+\tau},\dots,X_{t_n+\tau},Y_{t_1+\tau},Y_{t_2+\tau},\dots,Y_{t_n+\tau},\Big)^T$$

בלתי תלוי ב-τ. במקרה הסטציונרי

$$\mathbf{R}_{X,Y}(t_1,t_2)=\mathbf{R}_{X,Y}(t_1+\tau,t_2+\tau)=\mathbf{R}_{X,Y}(0,t_2-t_1)$$

ונגדיר במקרה זה

$$\mathbf{R}_{X,Y}(t_1,t_2)=\mathbf{R}_{X,Y}(t_2-t_1)$$

או

$$\mathbf{R}_{X,Y}(t_1,t_1+\tau)=R_{X,Y}(\tau)=R_{Y,X}(-\tau)$$

<u>הגדרה:</u> זוג התהליכים האקראיים $\{X_t,Y_t,-\infty<t<\infty\}$ נקראים סטציונרים במובן הרחב אם מתקיימים התנאים הבאים עבור כל t וכל τ:

א. $E[Y_t]=E[Y_0];\quad E[X_t]=E[X_0]$

ב. $E[X_{t+\tau}X_t],E[Y_{t+\tau}Y_t]$, ו- $E[Y_{t+\tau}X_t]$ הם פונקציות של τ <u>בלבד</u> .

5.3 תהליך אקראי גאוסי

<u>הגדרה:</u> ת"א $\{X_t,t\in[a,b]\}$ נקרא גאוסי אם עבור כל n וכל $t_1,\dots,t_n\in[a,b]$ הוקטור האקראי $(X_{t_1},\dots,X_{t_n})^T$ הוא ו"א גאוסי.

<u>טענה:</u> אם X_t ת"א גאוסי ב-$[a,b]$ אזי חוק ההסתברות שלו נקבע חד משמעית ע"י פונקציות התוחלת והאוטוקורילציה שלו $\mathbf{R}_X(t_1,t_2),\mu_X(t)\quad t,t_1,t_2\in[a,b],$

<u>הוכחה:</u> עבור כל n וכל $t_1,\dots,t_n$ המקיימים $t_i\in[a,b]$ הוקטור האקראי

$$\underline{Z}=\Big(X_{t_1},X_{t_2},\dots,X_{t_n}\Big)^T$$

הוא וקטור אקראי <u>גאוסי</u> (היות והנחנו ש-$\{X_t, t \in [a,b]\}$ הוא ת"א גאוסי). עבור המומנטים מסדר ראשון ושני של $\underline{Z}$ מתקיים:

$$E[\underline{Z}] = \Big(\mu_X(t_1), \mu_X(t_2), \ldots, \mu_X(t_n)\Big)^T$$

ו-

$$E[\underline{Z}\underline{Z}^T] = \{E[X_{t_i} X_{t_j}]\} = \{\mathbf{R}_X(t_i, t_j)\}$$

לכן עבור כל $(\nu_1, \ldots, \nu_n)$, הפונקציה האופינית של הוקטור $\underline{Z}$ נתונה ע"י:

$$\phi_Z(\nu_1, \ldots, \nu_n) = \exp\left\{ i \sum \nu_i \mu_X(t_i) - \frac{1}{2} \sum_i \sum_j \nu_i \nu_j \mathbf{K}_X(t_i, t_j) \right\}$$

כאשר כזכור $\mathbf{K}_X(t_1,t_2) = \mathbf{R}_X(t_1,t_2) - \mu_X(t_1)\mu_X(t_2)$. הפונקציה האופינית של וקטור אקראי מגדירה חד משמעית את חוק ההסתברות של הוקטור האקראי, לכן $\mu_X(t)$ ו-$\mathbf{R}_X(t_1,t_2)$ מגדירים את חוק ההסתברות של התהליך.

<u>מסקנה:</u> תהליך אקראי גאוסי סטציונרי במובן הרחב הוא סטציונרי.

<u>טענה:</u> אם $\{X_t, -\infty < t < \infty\}$ ת"א גאוסי, גם התהליכים

$$a(t)X_t$$

$$a(t)X_t + b(t)X_{t+c}$$

תהליכים גאוסיים עבור כל $c,\ b(\cdot),\ a(\cdot)$ דטרמיניסטיים (מדוע? הוכח). לכן נצפה גם שאם הגבול

$$\frac{dX_t}{dt} = \lim_{\varepsilon \to 0} \frac{X_{t+c} - X_t}{\varepsilon}$$

קיים, אזי (בהנחות מתאימות) הגבול גם הוא ת"א גאוסי.

יהיה X_t ת"א גאוסי; אזי

$$\int_a^b X_s ds \approx \sum_i X_{s_i}(s_{i+1} - s_i)$$

הוא מ"א גאוסי והתהליכים

$$\int_{-\infty}^{\infty} h(t-\theta)X(\theta)d\theta, \quad \int_{-\infty}^{\infty} g(t,\theta)X(\theta)d\theta$$

הם ת"א גאוסיים כאשר $g(\cdot,\cdot), h(\cdot)$ דטרמיניסטיים.

<u>מסקנה:</u> פעולה לינארית (לא אקראית) על תהליך אקראי גאוסי נותנת תהליך אקראי גאוסי.

<u>הערה:</u> תחת תנאים טכניים ניתן להחליף סדר תוחלת ואנטגרל מהשיקול הבא:

$$E\int_0^t X_s\,ds \approx E\sum X_{t_i}(t_{i+1} - t_i) = \sum EX_{t_i}(t_{i+1} - t_i) \approx \int_0^t EX_s\,ds$$

5.4 מעבר תהליכים אקראיים דרך מערכות לינאריות

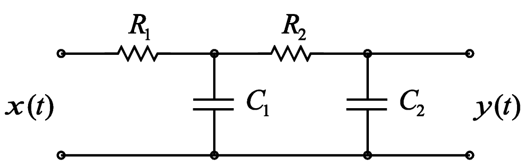
נתונה מערכת כבציור 5.8. תהליך הכניסה הוא $X(t)$ ותהליך היציאה הוא $Y(t)$.

ציור 5.8

שאלה: ידוע חוק ההסתברות של התהליך $\{X(t), -\infty < t < \infty\}$, נתונה מערכת (לא אקראית, לינארית או לא לינארית) שאפיונה ידוע. מהו חוק ההסתברות של תהליך היציאה $\{Y(t), -\infty < t < \infty\}$? בדרך כלל התשובה אינה ידועה אפילו אם מדובר במערכת לינארית קבועה בזמן. באופן מעשי נתעניין הרבה פעמים לא בחוק ההסתברות של $Y(t)$ אלא ב-$E[(Y(t) - E[Y(t)])^2]$ או $E[Y^2(t)]$. למען הקיצור נקרא לגודל זה "הספק היציאה" (כאילו היה מדובר במתח על נגד של אוהם אחד, ואז $E[Y^2(t)]$ הוא ההספק הממוצע הנמסר לנגד). גם לבעיה זו אין פתרון כללי ואנו נסתפק בפתרונה (חישוב $E[Y^2(t)]$) כאשר המערכת לינארית.

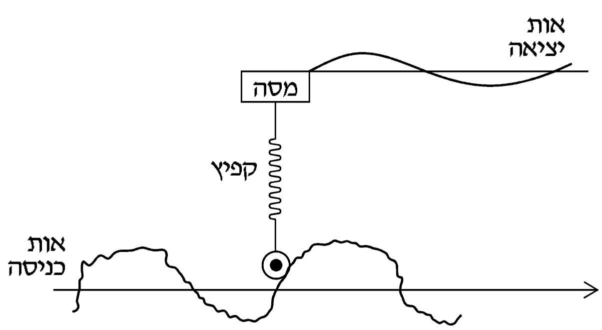
דוגמאות:

(א) מסנן R-C:



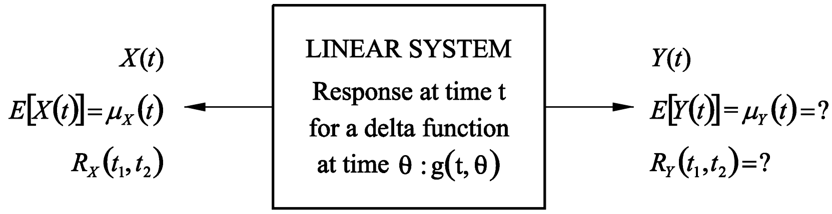
ציור 5.9

(ב) בולם זעזועים:



ציור 5.10

ננסח את השאלה הבאה: נתונה מערכת לינארית לא אקראית שהאיפיון שלה ידוע; מה צריך לדעת על $\{X(t), -\infty < t < \infty\}$, על מנת שנוכל לחשב את $E[Y^2(t)] = ?$. בהמשך נקבל תשובה מלאה לשאלה זו. ברור שידיעת $E[X^2(t)]$ אינה מספיקה (מדוע?). את הבעיה נתאר בצורה ציורית כדלקמן:



ציור 5.11

הקשר בין כניסת המערכת ויציאתה נתון ע"י:

$$Y(t) = \int_{-\infty}^{\infty} X(\theta) g(t, \theta) d\theta \cong \sum_i X(\theta_i) g(t, \theta_i)(\theta_{i+1} - \theta_i) \tag{5.1}$$

לא נתיחס כאן לבעיות של דיוק מתמטי ונרשום:

$$E[Y(t)] \cong \sum_i E\Big[X(\theta_i) g(t, \theta_i)(\theta_{i+1} - \theta_i)\Big] \to \int_{-\infty}^{\infty} \mu_X(\theta) g(t, \theta) d\theta$$

במילים אחרות, בצענו תוחלת על (5.1) והחלפנו את סדר האינטגרציה והתוחלת. כך קבלנו את המומנט מסדר ראשון של $Y(t)$ נעיין כעת במומנטים המצטלבים:

$$\begin{aligned}\mathbf{R}_{X,Y}(t_1, t_2) &= E\left[X(t_1) \int_{-\infty}^{\infty} X(\theta) g(t_2 \theta) d\theta\right] \\ &= E\left[\int_{-\infty}^{\infty} X(t_1) X(\theta) g(t_2, \theta) d\theta\right]\end{aligned}$$

ושוב נחליף סדר האינטגרציה והתוחלת ונקבל

$$\mathbf{R}_{X,Y}(t_1, t_2) = \int_{-\infty}^{\infty} \mathbf{R}_X(t_1, \theta) g(t_2, \theta) d\theta$$

לבסוף, נעיין ב-$\mathbf{R}_Y(t_1, t_2)$. שים לב שידיעת $\mathbf{R}_Y(t_1, t_2)$ כוללת ידיעת "הספק היציאה הממוצע"

$$E[Y^2(t)] = \mathbf{R}_Y(t, t)$$

$$\begin{aligned}E\Big[Y(t_1) Y(t_2)\Big] &= E\left[\int_{-\infty}^{\infty} X(\theta) g(t_1, \theta) d\theta \int_{-\infty}^{\infty} X(\eta) g(t_2, \eta) d\eta\right] \\ &= \left[\iint_{-\infty}^{\infty} X(\theta) X(\eta) g(t_1 \theta) g(t_2, \eta) n\theta d\eta\right] \\ &\cong E\left[\sum_i \sum_j X(\theta_i) X(\eta_i) g(t_1, \theta_i) g(t_2, \eta_j)(\eta_{j+1} - \eta_j)(\theta_{i+1} - \theta_i)\right] \\ &\cong \iint_{-\infty}^{\infty} \mathbf{R}_X(\theta, \eta) g(t_1, \theta) g(t_2, \eta) d\theta d\eta \qquad (5.2)\end{aligned}$$

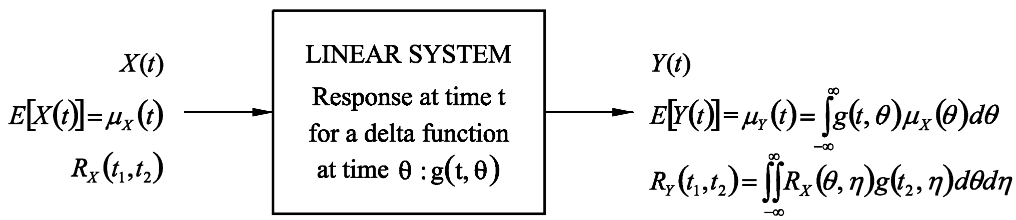
מסקנות:

(א) עבור מערכות ליניאריות לא אקראיות (קבועות בזמן או משתנות בזמן, סיבתיות או לא סיבתיות), ידיעת $\mathbf{R}_X(t_1,t_2),\mu_X(t)$ עבור $-\infty < t,t_1,t_2 < \infty$ מאפשרת את קביעת $\mu_Y(t)$ ואת $\mathbf{R}_Y(t_a,t_b)$ עבור כל t_a,t_b.

(ב) כאשר $X(t)$ ת"א גאוסי גם $Y(t)$ ת"א גאוסי ולכן ידיעת המומנטים מסדר ראשון ושני של $Y(t)$ מגדירה את חוק ההסתברות של תהליך היציאה.

(ג) שים לב שעל מנת לדעת את $\mathbf{R}_Y(t,t)$ צריך לדעת את $\mathbf{R}_X(t_1,t_2)$ עבור כל t_1,t_2.

את התוצאות (5.1) ו (5.2) נסכם בציור הבא:



ציור 5.12

מעבר תהליכים אקראיים סטציונריים במובן הרחב דרך מערכות קבועות בזמן

כאשר המערכת קבועה בזמן היא מאופינת ע"י התגובה להלם $h(t)$ (כלומר $g(t,\theta)=h(t-\theta)$). נוכל לכן לקבל את התוצאות המבוקשות עבור מקרה זה ע"י החלפת $g(t,\theta)$ ב $h(t-\theta)$ בתוצאות הקודמות.

$$Y(t)=\int_{-\infty}^{\infty} X(t-\theta)h(\theta)d\theta$$

ובהחלפת סדר התוחלת והאינטגרציה נקבל

$$E[Y(t)]=\int_{-\infty}^{\infty} E[X(t-\theta)]h(\theta)d\theta=\mu_X\int_{\infty}^{\infty} h(\theta)d\theta$$

$$E\Big[Y(t)Y(t+\tau)\Big]=E\left[\int_{-\infty}^{\infty} X(t-\theta)h(\theta)d\theta\int_{-\infty}^{\infty} X(t+\tau-\eta)h(\eta)d\eta\right]$$

$$=\iint_{-\infty}^{\infty} E\Big[X(t-\theta)X(t+\tau-\eta)\Big]h(\theta)h(\eta)d\theta d\eta$$

$$=\iint_{-\infty}^{\infty} R_X(\tau-\eta+\theta)h(\theta)h(\eta)d\theta d\eta$$

כמו כן

$$E\Big[X(t)Y(t+\tau)\Big] = E\left[\int_{-\infty}^{\infty} X(t)X(t+\tau-\theta)h(\theta)d\theta\right] = \int_{-\infty}^{\infty} R_x(\tau-\theta)h(\theta)d\theta$$

<u>מסקנה</u>: אם $\{X(t), -\infty < t < \infty\}$ סטציונרי במובן הרחב והמערכת לינארית וקבועה בזמן אזי, $\{X(t), Y(t), -\infty < t < \infty\}$ סטציונריים במובן הרחב ואז

$$\mu_Y = \mu_X \int_{-\infty}^{\infty} h(\theta)d\theta \tag{5.3}$$

$$R_Y(\tau) = \iint\limits_{-\infty}^{\infty} R_X(\tau-\eta+\theta)h(\theta)h(\eta)d\theta d\eta \tag{5.4}$$

אם המערכת סיבתית אזי $h(\theta) = 0$ עבור $\theta < 0$ ואז

$$R_Y(\tau) = \int_0^{\infty}\int_0^{\infty} R_X(\tau-\eta+\theta)h(\theta)h(\eta)d\theta d\eta \tag{5.5}$$

השימוש העיקרי של התוצאה עבור $R_Y(\tau)$ הוא "הספק היציאה" ואז

$$R_Y(0) = \iint R_X(\theta-\eta)h(\theta)h(\eta)d\theta d\eta$$

(כאשר גבולות האינטגרציה הם מ-$-\infty$ עד ∞ או מאפס ועד ∞). מעתה והלאה נעסוק רק במערכות קבועות בזמן ובתהליכי כניסה שהם סטציונריים במובן הרחב.

<u>התמרות פוריה</u>

באותות ומערכות למדנו שיש שתי גישות לאפיון מערכות לינאריות <u>קבועות בזמן</u> :

<u>גישה א'</u> : אפיון המערכת ע"י תגובתה לפונקצית דירק וייצוג התוצאה עבור כניסה שרירותית ע"י סופרפוזיציה (ומ-תקבלת קונוולוציה).

<u>גישה ב'</u> : אפיון המערכת ע"י תגובתה לערור הרמוני ($e^{i\omega t}$), יצוג כניסה כללית ע"י סיכום תנודות הרמוניות (התמרת פוריה הפוכה) ושוב סופרפוזיציה. במקרה הדטרמיניסטי הקשר בין הכניסה והיציאה (גישה א') הוא:

$$Y(t) = \int_{-\infty}^{\infty} X(t-\theta)h(\theta)d\theta$$

ואילו במקרה האקראי הקשר בין פונקציות האוטוקורילציה בכניסה וביציאה נתון ע"י

$$R_Y(\tau) = \iint\limits_{-\infty}^{\infty} R_X(\tau+\theta-\eta)h(\theta)h(\eta)d\theta d\eta$$

<u>שאלה</u> : האם לגישה ב' אפשר לתת מובן במקרה של תהליכים אקראיים? בהמשך נראה שהדבר אכן אפשרי וכשם שבמקרה הדטרמיניסטי יש יתרונות רבים לאנליזה במרחב התדר כן גם במקרה האקראי.

<u>חזרה: התמרות פוריה.</u>

נתחיל ב-$X(t)$ לא אקראי ונניח ש-$\int_{-\infty}^{\infty}|X(t)|dt$. התמרת פוריה של $X(t)$ מוגדרת ע"י:

$$\hat{X}(f)=F\{X(t)\}=\int_{-\infty}^{\infty}X(t)e^{-2\pi ift}dt$$

והתמרת פוריה ההפוכה נתונה ע"י

$$X(t)=F^{-1}\{\hat{X}(f)\}=\int_{-\infty}^{\infty}\hat{X}(f)e^{2\pi ift}df$$

(קיימת נקודה עדינה לגבי קיום האינטגרל האחרון אולם נתעלם מבעיה זו). משפט פרסוול בהתמרות פוריה אומר שאם $\int_{-\infty}^{\infty}|X_i(t)|^2dt<\infty$ עבור $i=1,2$ אזי

$$\int_{-\infty}^{\infty}X_1(t)X_2^*(t)dt=\int_{-\infty}^{\infty}\hat{X}_1(f)\hat{X}_2^*(f)df$$

ל-$\int_{-\infty}^{\infty}X^2(t)dt$ נקרא האנרגיה של הפונקציה $X(\cdot)$ (מה האנרגיה של דגם טפוסי של ת"א סטציונרי?)

אם $F\{X(t)\}=\hat{X}(f)$ אזי בתנאים מתאימים קיים

$$F\left\{\frac{dX(t)}{dt}\right\}=2\pi if\hat{X}(f)$$

אם $X(\cdot)$ ממשי אזי

$$F\{X(-t)\}=\hat{X}^*(f)$$

$$F\{X(t+\tau)\}=e^{2\pi if\tau}\hat{X}(f)$$

$$F\left\{\int_{-\infty}^{\infty}X(t-\theta)h(\theta)d\theta\right\}=\hat{X}(f)\cdot\hat{h}(f)$$

<u>צפיפות ספקטרלית</u>

כאשר $(X(t),-\infty<t<\infty)$ ת"א סטציונרי, האנרגיה של פונקצית מדגם טיפוסית היא אינסופית ולכן לא ברור אם בכלל אפשר לבצע התמרת פוריה על פונקציה כזו ואמנם, אנו לא נעשה זאת. אנו נתעניין בהתמרת פוריה של פונקצית האוטוקורילציה. נגדיר, עבור ת"א סטציונריים, $\{Y(t),X(t)\}$ את:

(א) הצפיפות הספקטרלית $S_X(f)$ של $\{X(t),-\infty<t<\infty\}$

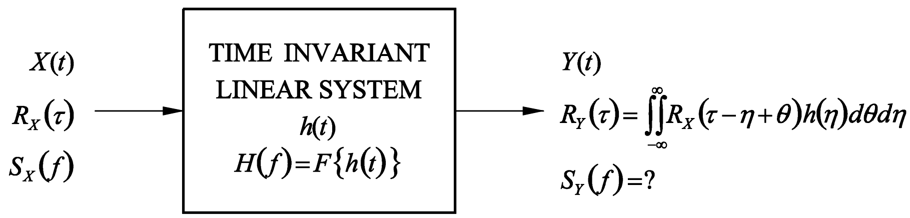
$$S_X(f)\triangleq F\{R_X(\tau)\}$$

(ב) פונקצית הצפיפות הספקטרלית המצטלבת

$$S_{X,Y}(f) \triangleq F\{R_{X,Y}(\tau)\}$$

וזאת בהנחה שלפונקציות $R_{X,Y}(\cdot),\ R_X(\cdot)$ יש התמרת פוריה.

נניח עכשיו $\{X(t), -\infty < t < \infty\}$ סטציונרי במובן הרחב ונניח $E[X(t)] = 0$. נתעניין בצפיפות הספקטרלית ביציאה של מערכת ליניארית כמצויר:



ציור 5.13

אנו נניח כי h ממשית

במקרה הנוכחי:

$$R_{X,Y}(\tau) = E\Big[X(t)Y(t+\tau)\Big] = \int_{-\infty}^{\infty} R_X(\tau-\theta)h(\theta)d\theta$$

ולכן (מדוע?)

$$S_{X,Y}(f) = S_X(f)\cdot H(f)$$

הקונוולוציה הפכה למכפלה. לגבי $S_Y(f)$:

$$R_Y(\tau) = \iint_{-\infty}^{\infty} R_X(\tau-\theta)+\eta)h(\theta)h(\eta)d\theta d\eta$$

על פי ההגדרה $S_Y(f) = \int_{-\infty}^{\infty} R_Y(\tau)e^{-2\pi i f\tau}d\tau$ ולכן בשינוי סדר האינטגרציה

$$\int_{\eta=-\infty}^{\infty}\int_{\theta=-\infty}^{\infty}\int_{\tau=-\infty}^{\infty} e^{-2\pi i f\tau}R_X(\tau+\theta-\eta)h(\theta)h(\eta)d\tau d\theta d\eta = \iint_{-\infty}^{\infty} S_X(f)e^{2\pi i f(\theta-\eta)}h(\theta)h(\eta)d\theta d\eta$$

$$= S_X(f)H(f)\cdot H^*(f)$$

$$= S_X(f)|H(f)|^2$$

כלומר

$$S_Y(f) = S_X(f)|H(f)|^2$$

והספק היציאה הכולל הוא:

$$R_Y(0) = \int_{-\infty}^{\infty} S_X(f)|H(f)|^2 df$$

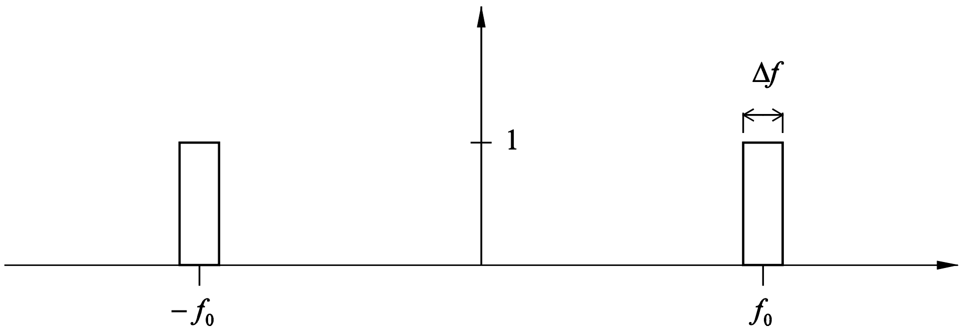
המצב, נכון לעכשיו, הוא כדלקמן: הגדרנו, בצורה שרירותית למדי, את התמרת פוריה של $R_X(\tau)$ וראינו שחוקי המעבר דרך מערכת לינארית קבועה בזמן מקבלים צורה סימפטית משמעות המושג צפיפות ספקטרלית $S_X(f)$. שים לב:

(א) $S_X(-f) = S_X(f)$ וכן $S_X(f)$ ממשית (וזאת מפני ש-$R_X(\tau)$ ממשית וסימטרית)

(ב) $\int_{-\infty}^{\infty} S_X(f)df \geq 0$ (כי הביטוי משמאל הוא $R_X(0)$).

(ג) <u>טענה:</u> $S_X(f) \geq 0$.

<u>הוכחה:</u> נבחר ב-$H(f)$ כדלקמן:



ציור 5.14

הספק היציאה במקרה זה יהיה (בערך) $2\Delta f S_X(f_)) \geq 0$ ולכן עבור כל $f_0 \geq 0, f_0$ $S_X(f_0) \geq 0$.

(ד) הצפיפות הספקטרלית נמדדת ב-$(\text{Volt})^2/\text{Hz}$ או $(\text{Amp})^2/\text{Hz}$

נעיין בדוגמא: $X(t) = A(\omega)\cos(2\pi f_0 t + \phi(\omega))$ כאשר A, ϕ ב"ת ו-ϕ מפולג אחיד בתחום $[0, 2\pi)$. כזכור $R_X(\tau) = 0.5E[A^2]\cos 2\pi f_0 \tau$ ולכן: $S_X(f) =$ (מקדם) $\{\delta(f-f_0) + \delta(f+f_0)\}$.

<u>לסיכום:</u> מושג הצפיפות הספקטרלית נותן תאור של תכולת ההספק בתדרים השונים, דהיינו, גם אם לא ניתן לדבר על התמרת פוריה של פונקצית מדגם של התהליך $\{X(t), -\infty < t < \infty\}$, פונקצית הצפיפות הספקטרלית של התהליך (הסטציונרי במובן הרחב) נותנת את פילוג ההספק לפי ציר התדר. לפיכך, אם $S_X(f)$ הצפיפות הספקטרלית של התהליך, אזי

$$2\int_{f_0}^{f_0+\Delta} S_X(f)df$$

הוא ההספק <u>הממוצע</u> ביציאה ממסננת bandpass אידאלית בעלת רחב סרט Δ בתחום $(f_0, f_0+\Delta)$.

<u>שים לב:</u> כאשר $\phi(t)$ פונקציה דטרמיניסטית בעלת אנרגיה סופית, אנו יכולים לדעת את תכולת <u>האנרגיה</u> בתחום תדרים מסוים ע"י עיון באינטגרל של $|\Phi(f)|^2$ על תחום התדרים המעניין. כאן אנו עוסקים בתכולת <u>ההספק</u>!

כדי להווכח כי $S_x(f)$ אכן מתאר את תכולת ההספק של <u>פונקציות המדגם</u>, נעביר את התהליך X דרך מסנן מעביר סרט $(-f_0, f_0)$, כאשר $S_X(f) = 0$ מחוץ לסרט. אם נגדיר תהליך שגיאה כ-$\varepsilon(t) = X(t) - X * h(t)$ אזי

$$E\varepsilon^2 = \int S_X(f)\,|1 - H(f)|^2\,df = 0\,.$$

הערות:

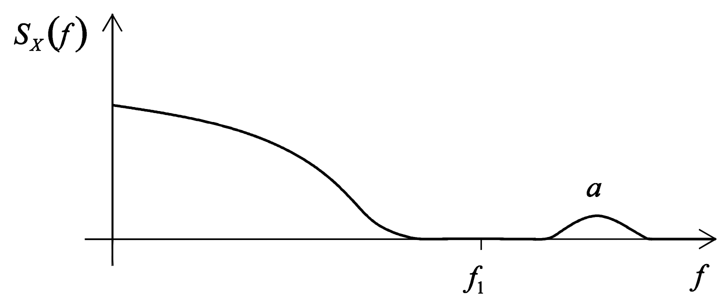
(א) גוזר $Y(t)=dX(t)/dt$ מבליט תדרים גבוהים ומתקיים עבורו $F\{Y(t)\}=2\pi i f F\{X(t)\}$.

(ב) מערכת שהקשר בין הכניסה $X(t)$ והיציאה $Y(t)$ נתון ע"י $Y(t)=\int_{t-\Delta}^{t}X(\theta)d\theta$ מתאימה לתגובה להלם שהיא 1 בין הזמנים אפס ו-Δ ואפס בזמנים אחרים. מתוך עיון בהתמרת פוריה של התגובה להלם נובע מיד שמערכת זו מחליקה (מרסנת תדרים גבוהים).

(ג) בכל מקרה של ציור העוסק בבעיות מהסוג שאנו מטפלים בהם יש לבדוק את משמעות הציר האופקי על מנת לברר אם הוא מבטא זמן או תדר.

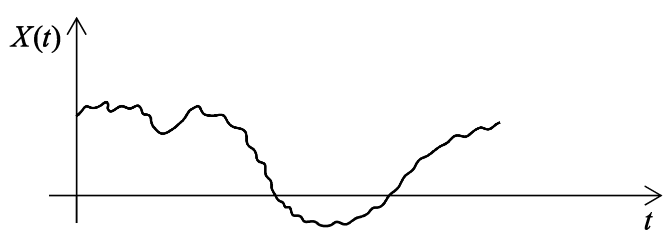
(ד) כזכור הגדרנו את הצפיפות הספקטרלית עבור תהליך אקראי בעל תוחלת אפס. אפשר ליחס צפיפות ספקטרלית גם לתהליכים (סטציונריים במובן הרחב) בעלי תוחלת שונה מאפס. במקרה זה תהיה לצפיפות הספקטרלית פונקצית דירק בתדר אפס. דהיינו, אם לת"א $X(t)$ נוסיף מ"א C בלתי תלוי ב $X(t)$, לקבלת $Y(t)=X(t)+C$, אזי: $S_Y(f)=S_X(f)+E[C^2]\delta(f)$.

<u>דוגמא:</u> נעסוק באות אקראי $X(t)$ סטציונרי במובן הרחב. תהיה $S_X(f)$ הצפיפות הספקטרלית של אות זה (כולל החלק α).



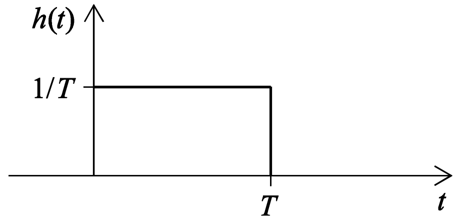
ציור 5.15

אנו מניחים שה"תוספת" המסומנת ב-α נובעת מרעש שאיננו מעוניינים בו. בהנחה ש-$S_X(f)$ (כולל החלק α) מתיחס לתהליך אקראי גאוסי, דגם טיפוסי יראה בצורה:



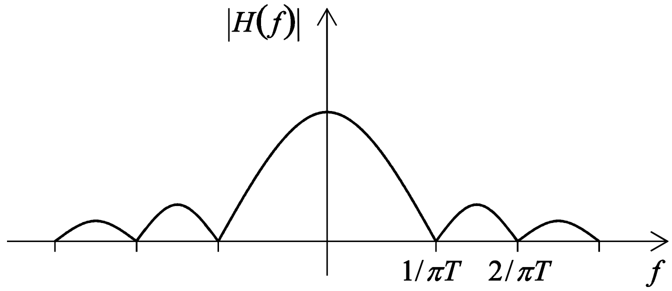
ציור 5.16

כאשר ה"רעידות" של $X(t)$ באות מהרעש, היינו, מהחלק המסומן במישור התדר ב-α. אנו מתבקשים לבצע גזירה. במרחב התדר עלינו להכפיל את $S_X(f)$ ב-$|H(f)|^2=(2\pi f)^2$ והגזירה תגביר את אפקט הרעש בצורה ניכרת, זה ברור הן במרחב הזמן והן במרחב התדר. בכדי להתגבר על אפקט הרעש מוצע לבצע החלקה ע"י העברת האות $X(t)$ דרך מערכת ליניארית קבועה בזמן שתגובתה להלם ניתנת ע"י $h(t)$ כאשר:



ציור 5.17

ואז $H(f) = F\{h(t)\}$ נראה כדלקמן:



ציור 5.18

אם נבחר ב-T גדול מאוד, נפגע בסיגנל עצמו, אם נבחר ב-T קטן מאוד לא נחליק את הרעש. מתוך שיקולים עניניים נראה שכדאי לבחור את T כך ש- $f_1 = 1/\pi T$ כאשר f_1 מופיע בציור לעיל.

שאלה: האם בבעיה זו יש קודם לבצע החלקה ואח"כ גזירה או להפך?

המשפט הבא עוזר לנו להבין את המושג של צפיפות ספקטרלית:

משפט: נניח $\int_{-\infty}^{\infty} |\tau R(\tau)| d\tau < \infty$, נסמן

$$X_T(t) = \begin{cases} X(t) & \text{if} \quad |t| < T \\ 0 & \text{if} \quad |t| \geq T \end{cases}$$

$$\hat{X}_T(f) = F\{X_T(t)\}$$

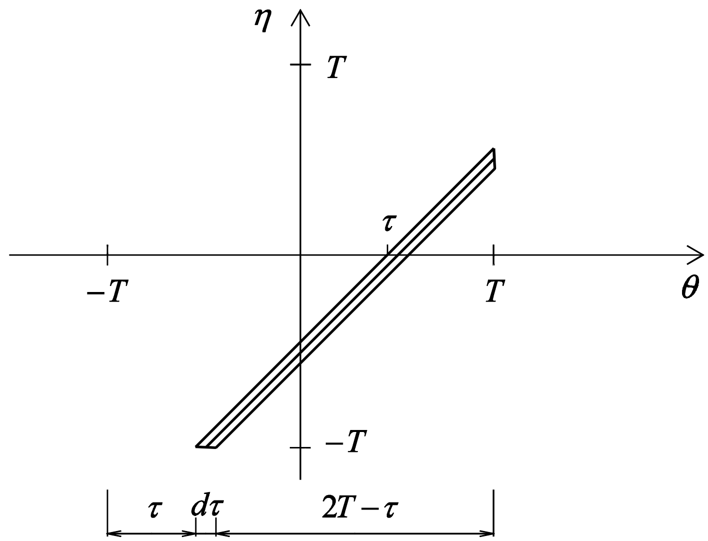
אזי

$$E\left[\frac{1}{2T}\,|\hat{X}_T(f)|^2\right] \underset{T\to\infty}{\longrightarrow} S_X(f)$$

הוכחה:

$$E\left[\frac{1}{2T}|\hat{X}_T(f)|^2\right] = \frac{1}{2T}E\left[\iint_{-T}^{T} X_\theta X_\eta e^{-2\pi i f(\theta-\eta)} d\theta d\eta\right] = \frac{1}{2T}\iint_{-T}^{T} R_X(\theta-\eta) e^{-2\pi i f(\theta-\eta)} d\theta d\eta$$

נסמן $\tau = \theta - \eta$ ונעיין בציור: נבצע קודם אינטגרציה על הפס הצר שעבורו $\tau \leq \theta - \eta \leq \tau + d\tau$



ציור 5.19

לכן

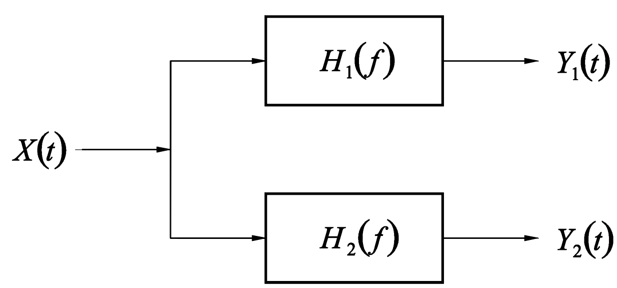
$$E\left[\frac{1}{2T}|\hat{X}_T(f)|^2\right] = \frac{1}{2T}\int_{-2T}^{2T} R_X(\tau)e^{-2\pi i f\tau}(2T-|\tau|)\cdot\sqrt{2}\,\frac{d\tau}{\sqrt{2}}$$
$$= \int_{-2T}^{2T} R_X(\tau)e^{-2\pi i f\tau}\left(1-\frac{|\tau|}{2T}\right)d\tau \underset{T\to\infty}{\longrightarrow} S_X(f) = F\{R_X(\tau)\}$$

מ.ש.ל.

נחזור למושג הצפיפות הספקטרלית המצטלבת

$$S_{X,Y}(f) = F\{R_{X,Y}(\tau)\} = F\{E[X(t)Y(t+\tau)]\}$$

נאמר כי W, Z חסרי קורלציה אם $EW(t_1)Z(t_2)\equiv 0$, או, באופן שקול: $S_{WZ}(f)\equiv 0$. נעיין במיוחד במקרה הבא:



ציור 5.20

$$Y_1(t) = \int_{-\infty}^{\infty} h_1(\theta)X(t-\theta)d\theta$$

$$R_{Y_1,Y_2}(\tau) = E\Big[Y_1(t)Y_2(t+\tau)\Big] = E\left[\iint_{-\infty}^{\infty} X(t-\theta)X(t+\tau-\eta)h_1(\theta)h_2(\eta)d\theta d\eta\right]$$

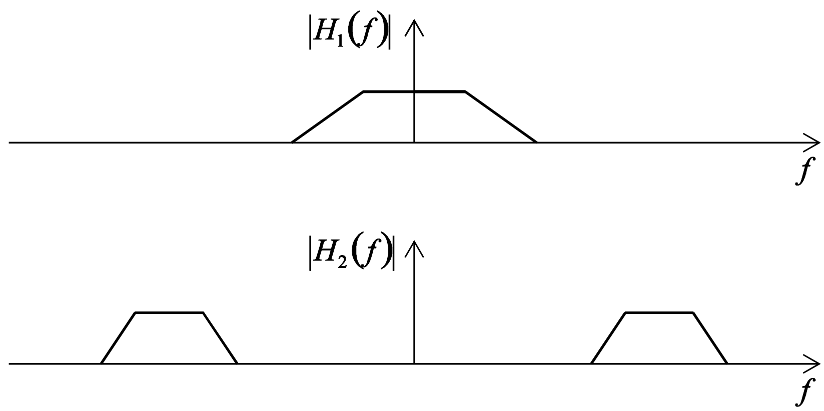
$$= \iint_{-\infty}^{\infty} R_X(\tau+\theta-\eta)h_1(\theta)h_2(\eta)d\theta d\eta$$

לכן:

$$S_{Y_1,Y_2}(f) = \iiint R_X(\tau+\theta-\eta)e^{-2\pi i f\tau}h_1(\theta)h_2(\eta)d\theta d\eta d\tau$$

$$= \iint S_X(f)e^{2\pi i f(\theta-\eta)}h_i(\theta)h_2(\eta)d\theta d\eta = S_X(f)H_1^*(f)H_2(f)$$

במיוחד אם $H_2(f)H_1(f) \equiv 0$, כגון כאשר

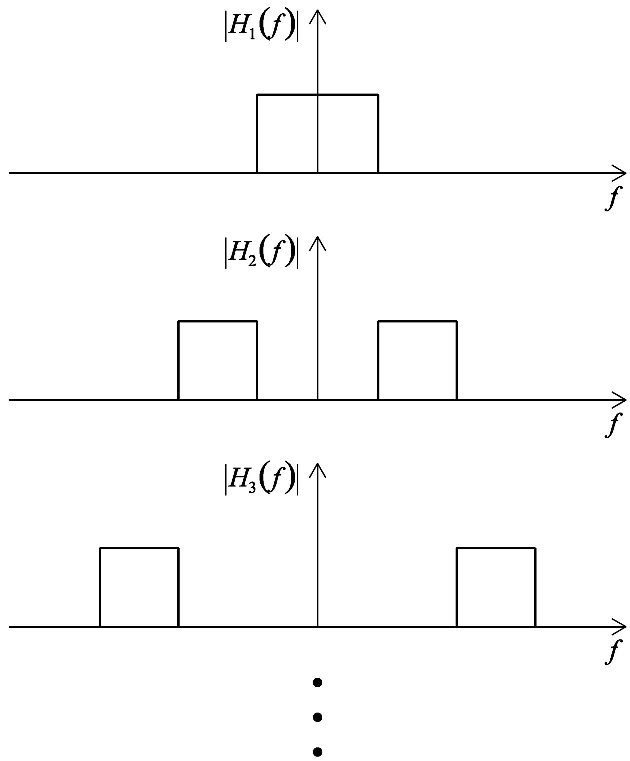


ציור 5.21

אזי התהליכים האקראיים $Y_1(t)$ ו-$Y_2(t)$ חסרי קורילציה (ולכן במקרה של תהליכים גאוסיים במשותף הת"א $\{Y_1(t), -\infty < t < \infty\}$ ו-$\{Y_2(t), -\infty < t < \infty\}$ בלתי תלויים).

הערה: הוכח שהמשתנים האקראיים $X(t)$ ו-$dX(t)/dt$ חסרי קורילציה עבור אותו t, אולם כתהליכים אקראיים הם לא חסרי קורלציה.

נגדיר אוסף של מסננים $H_1(f), H_2(f), \ldots$ ע"י הציור:



ציור 5.22

המסננים מקיימים:

$$\forall i \neq j, \quad H_i(f) \cdot H_j(f) \equiv 0 \; ; \quad \sum_i H_i(f) = 1$$

בעזרת מושג הצפיפות הספקטראלית אנו יכולים לחשב את ההספק ביציאת כל אחד מהמסננים. ההספק ביציאת מסנן ברוחב Δf הוא בערך $2 \cdot \Delta f \cdot S_X(f_0)$, כאשר f_0 הוא התדר המרכזי של המסנן. הנקודה החשובה היא, שאנו יכולים גם ללכת הפוך בשיטה זאת ולשערך את $S_X(f)$ של תהליך $X(t)$ לא ידוע, מתוך מדידת הספקי היציאה של מערך מסננים כנ"ל.

5.5 הזזת ספקטרום

בסעיף זה נעסוק בבעיה הבאה: מה קורה כשמכפילים ת"א סטציונרי במובן הרחב $X(t)$ ב-$\cos(2\pi f_0 t + \phi)$ כאשר ϕ מ"א בלתי תלוי בתהליך $X(t)$ ומפולג אחיד בתחום $[0, 2\pi]$. לפעולה כזו חשיבות רבה במערכות קומוניקציה. נעיין ב-

$$Y(t) = X(t)\cos(2\pi f_0 t + \phi)$$

חשבון פשוט נותן

$$R_Y(\tau) = \frac{1}{2} R_X(\tau) \cos 2\pi f_0 \tau$$

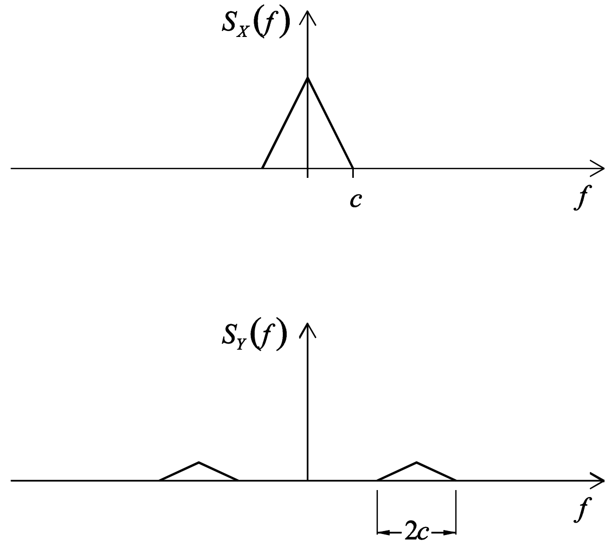
כידוע

$$F\Big\{g(t)\cos 2\pi f_0 t\Big\} = \frac{1}{2}\Big[G(f+f_0) + G(f-f_0)\Big]$$

ולכן

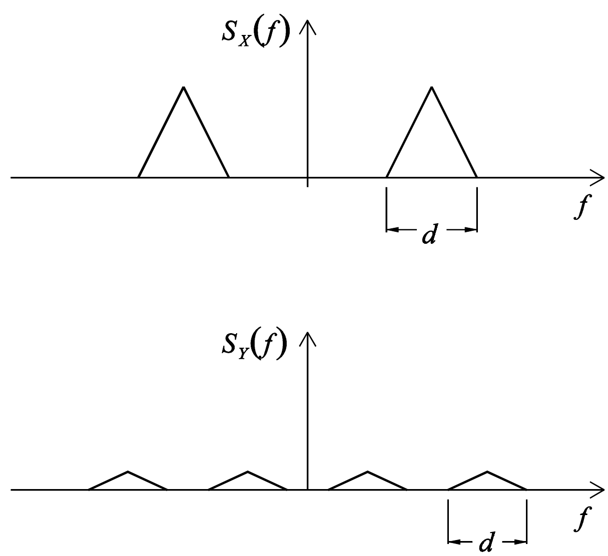
$$S_Y(f) = \frac{1}{4}\Big[S_X(f+f_0) + S_X(f-f_0)\Big]$$

מקרה א':



ציור 5.23א'

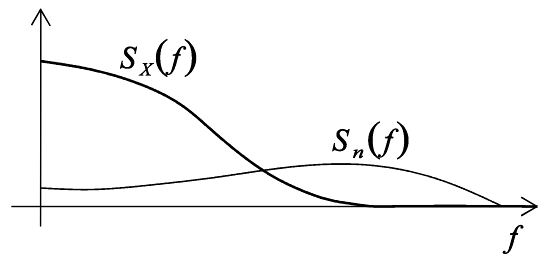
מקרה ב':



ציור 5.23ב'

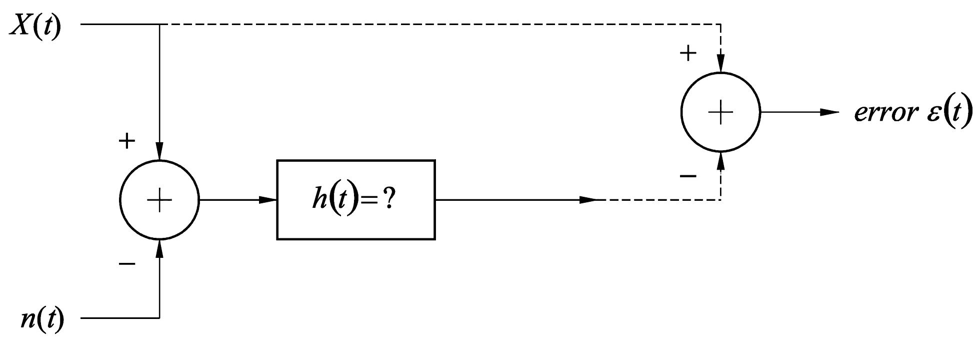
5.6 סינון ליניארי אופטימלי

נעבור כעת לבעיה כללית של סינון ליניארי. נניח שנתונים שני ת"א $\{X(t), -\infty < t < \infty\}$, $\{n(t), -\infty < t < \infty\}$, לכל אחד מהם תוחלת אפס וכל אחד מהם סטציונרי במובן הרחב. כן נניח שהתהליכים חסרי קורילציה ($E[X(t_1)n(t_2)] = 0$ לכל t_1 ו- t_2) ולכן הם סטציונריים במשותף במובן הרחב.



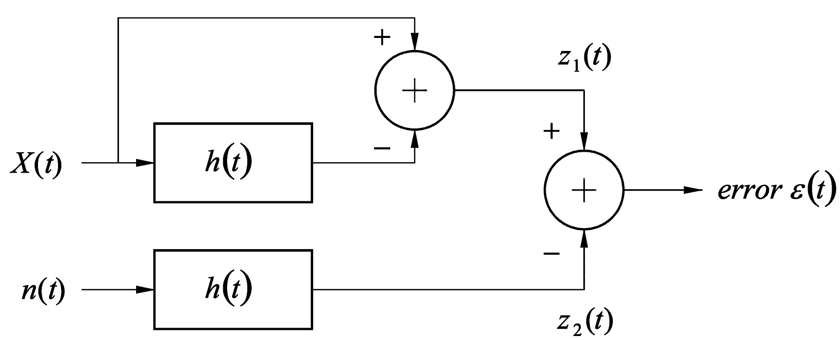
ציור 5.24

נקלט האות $r(t)$ שהוא סכום האות הרצוי $X(t)$ והרעש $n(t)$, כלומר, $r(t) = X(t) + n(t)$. מתוך האות הנקלט $r(t)$ רוצים לסלק את הרעש ולקבל את האות הרצוי $X(t)$ ע"י העברת $r(t)$ דרך מסנן ליניארי שתגובתו להלם היא $h(t)$ וקבלת יציאה נאמנה ככל האפשר ל-$X(t)$. לפעולה זו קוראים סינון (של הרעש) והיא מתוארת בציור הבא:



ציור 5.25

האות המצוי אינו זהה לאות הרצוי היות ויש שגיאה הנובעת הן מהרעש והן מעיוות האות הנגרם ע"י $h(t)$. על מנת להבהיר זאת נצייר את ציור 5.25 בצורה הבאה:



ציור 5.26

ובגלל אי התלות הלינארית בין $X(\cdot)$ לבין $n(\cdot)$ נוכל לרשום

$$E\left[\varepsilon^2(t)\right] = E\left[z_1^2(t)\right] + E\left[z_2^2(t)\right]$$

ולכן

$$E\left[\varepsilon^2(t)\right] = \int_{-\infty}^{\infty} S_n(f)|H(f)|^2 df + \int_{-\infty}^{\infty} S_X(f)|1 - H(f)|^2 df$$

קבלנו ביטוי מפורש עבור השגיאה הריבועית הממוצעת הכוללת. עכשיו נוכל להשוות הצעות שונות ל-$H(f)$. למשל, אם מישהו מציע $H(f) = (1+if/f_0)^{-1}$ (דהיינו, מסננת R-C פשוטה) נוכל לחשב את השגיאה כפונקציה של הפרמטר f_0, לבצע אופטימיזציה ולמצוא את הפרמטר f_0 הטוב ביותר שיגרום לשגיאה ריבועית ממוצעת מינימלית עבור מסנן מהטיפוס הזה.

גישה נועזת יותר היא לשאול מה הוא $h(t)$ האופטימלי שיביא את השגיאה הריבועית הממוצעת למינימום. במקרה זה יש להבחין בין שני מקרים של מציאת $h(t)$ האופטימלי:

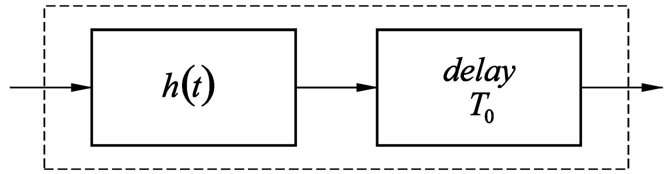
א. אין הגבלות על $h(t)$, דהיינו, אין דרישה ש-$h(t)$ יהיה סיבתי.

ב. קיימת מגבלה ש-$h(t)$ יהיה סיבתי.

אם הפתרון הוא סיבתי אזי ניתן לממש אותו (או קרוב שלו) אם באופן דיגיטלי ואם באופן אנלוגי. בעית הסינון האופטימלי עם מגבלת הסיבתיות נפתרה ע"י Wiener וידועה כמסננת וינר אולם הפתרון הוא קשה ולא נביא אותו בקורס זה. הפתרון ללא מגבלת הסיבתיות הוא קל יחסית ונתרכז בו בהמשך. השאלה הנשאלת מיד היא: האם לפתרון ללא מגבלת הסיבתיות יש מובן פיזיקלי, דהיינו, האם ניתן לממש אותו (או קרוב טוב שלו) באופן אנלוגי או דיגיטלי? אם קבלנו פתרון מהצורה:

ציור 5.27

כאשר עבור $h(t) = 0,\ t < -T_0$ (או זניח) אזי $h(t)$ אינה סיבתית אולם $h(t - T_0)$ סיבתית. כלומר את:



$h(t-T_0)$

ציור 5.28

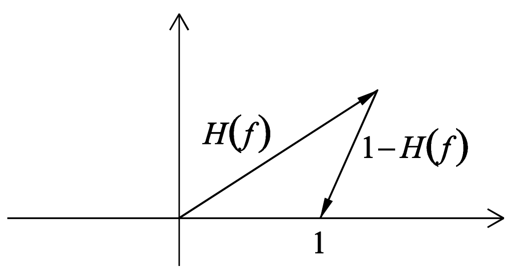
נוכל לממש. לכן, בבעיות בהן אפשר לסבול השהייה (כגון בעיות קומוניקציה בניגוד לבעיות בקרה) יש לפתרון ללא סיבתיות משמעות פיזיקלית. לפתרון זה קוראים "מסננת וינר עם השהייה אינסופית".

נחזור לבעית מציאת המסננת האופטימלית המביאה את השגיאה הריבועית הממוצעת למינימום. השגיאה הריבועית הממוצעת ניתנת כאמור ע"י:

$$E\Big[\varepsilon^2(t)\Big] = \int_{-\infty}^{\infty} S_n(f)|H(f)|^2 df + \int_{-\infty}^{\infty} S_x(f)|1-H(f)|^2 df$$

מצא מסננת $H(f)$ כך ש-$E[\varepsilon^2(t)]$ יהיה מינימום על פני כל המסננות $H(f)$.

עבור f מסוים נעיין ב-$|1-H(f)|$:



ציור 5.29

שים לב שאם עבור אותו f נחליף את $H(f)$ ב-$|H(f)|$ נגרום להקטנת $|1-H(f)|$

$$|1-H(f)| \geq 1 - |H(f)|$$

$$|1-H(f)| \geq |H(f)| - 1\,.$$

ע"י החלפה זו לא נשנה את הביטוי $\int S_n(f)|H(f)|^2 df$ אבל נקטין את האיבר השני $\int S_X(f)|1-H(f)|^2 df$ לכן מותר להניח ש-$H(f)$ האופטימלי ממשי ולא שלילי עבור כל f. הבעיה היא, לכן, למצוא $H(f)$ שעבורה הביטוי

$$E\Big[\varepsilon^2(t)\Big] = \int_{-\infty}^{\infty} S_n(f)\Big[H(f)\Big]^2 df + \int_{-\infty}^{\infty} S_X(f)\Big[1-H(f)\Big]^2 df$$

הוא מינימלי. לצורך זה נחפש עבור כל f את אותו $H(f)$ שעבורו האינטגרנד מינימלי. נסמן (עבור f נתון) $a = H(f)$ את אותו $H(f)$ עבורו האינטגרנד מינימלי. אזי

$$\frac{\partial}{\partial a}\Big\{S_n(f)a^2 + S_X(f)(1-a)^2\Big\} = 0$$

ונקבל

$$a = \frac{S_X(f)}{S_X(f) + S_n(f)}$$

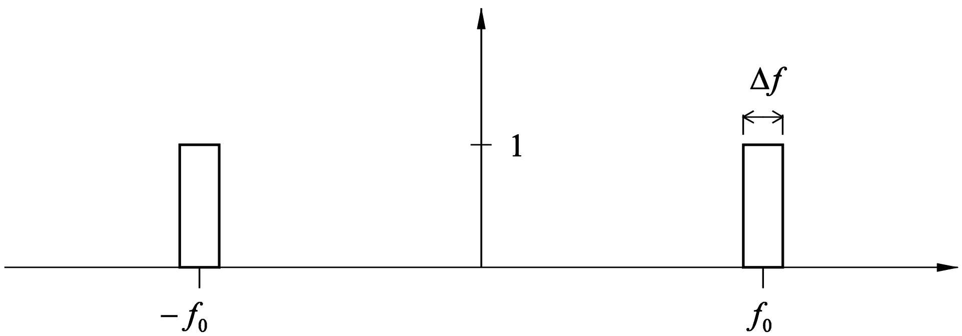
ולכן

$$H_{\text{opt}}(f) = \frac{S_X(f)}{S_X(f) + S_n(f)} \tag{5.6}$$

והשגיאה המינימלית מתקבלת ע"י הצגת $H_{\text{opt}}(f)$ לתוך נוסחת השגיאה הרבועית הממוצעת:

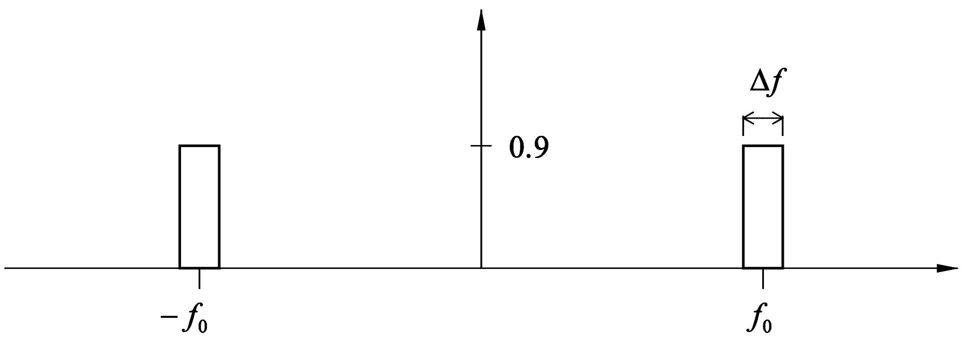
$$\begin{aligned} E\left[\varepsilon_{\min}^2(t)\right] &= \int_{-\infty}^{\infty} \left(\frac{S_n(f)S_X^2(f)}{(S_X(f) + S_n(f))^2} + \frac{S_X(f)S_n^2(f)}{(S_X(f) + S_n(f))^2} \right) df \\ &= \int_{-\infty}^{\infty} \frac{S_X(f)S_n(f)}{S_X(f) + S_n(f)} df \end{aligned}$$

<u>דוגמא:</u> יהיה $X(t)$ תהליך אקראי בעל צפיפות הספק כדלקמן:



ציור 5.30

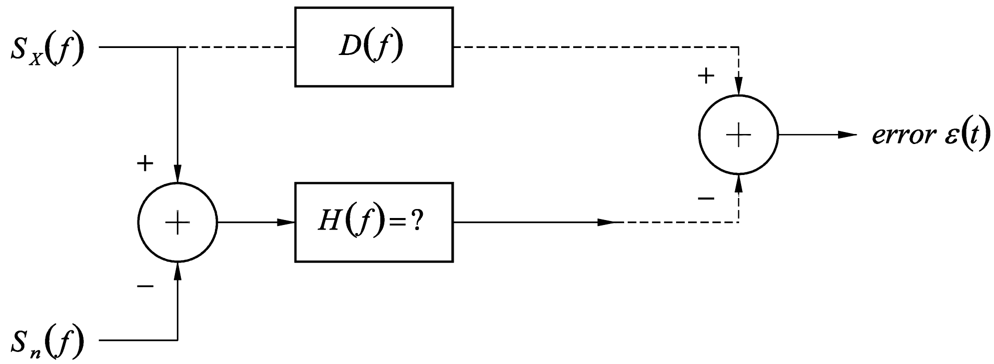
אליו מתווסף רעש $n(t)$ בעל צפיפות הספק $S_n(f) = 1/9$, בלתי תלוי ב $X(t)$. המסננת האופטימלית $H_{\text{opt}}(f)$ לשחזור $X(t)$ מתוך $Y(t) = X(t) + n(t)$, נתונה ע"י (5.6) , ונראית כדלקמן:



ציור 5.31

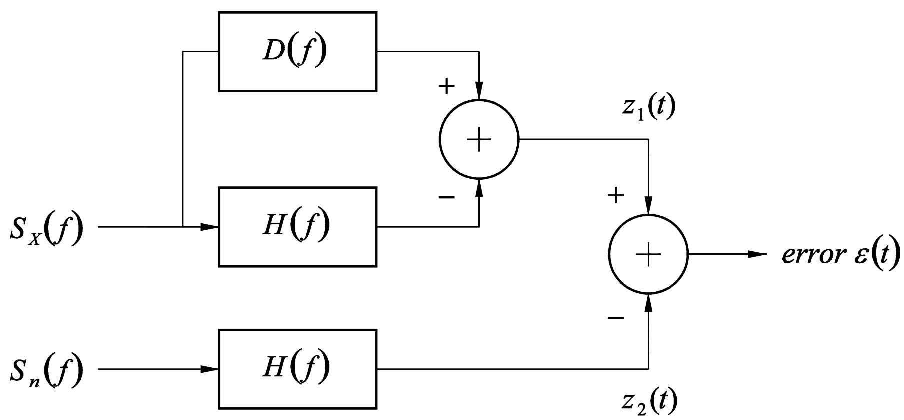
שים לב ש $H_{\text{opt}}(f)$ אינה מעבירה אותות באותם תדרים בהם תכולת התדר של האות $X(t)$ אינה קיימת. מדוע יש הנחתה של 0.9 בתדרים האחרים? (ראה דוגמה ג' בסעיף 2.2).

ניתן להרחיב במקצת את בעית מציאת המסננת האופטימלית. נמשיך להניח $E[n(t)] = E[X(t)] = 0$ וכן ש-$X(t_1), n(t_2)$ חסרי קורילציה. ההרחבה תהיה בכך שבמקום שהסיגנל הרצוי יהיה $X(t)$ נחפש כסיגנל רצוי את $Y(t)$ כאשר $Y(t)$ הוא היציאה של מערכת לינארית עם תגובה להלם $d(t)$ כאשר הכניסה היא $X(t)$. לדוגמא, $Y(t) = dX(t)/dt$ כאשר $d(t)$ הוא גוזר. בציור (במרחב התדר)



ציור 5.32

כאשר $D(f) = F\{d(t)\} = 1$ נקבל את המקרה הקודם. ע"י סופרפוזיציה נוכל להחליף את ציור 5.32 בציור



ציור 5.33

ולכן

$$E\left[\varepsilon^2(t)\right] = \int_{-\infty}^{\infty} S_n(f)|H(f)|^2 df + \int_{-\infty}^{\infty} S_X(f)|D(f) - H(f)|^2 df$$

את המשואה לעיל נרשום בצורה הבאה:

$$E\left[\varepsilon^2(t)\right] = \int_{-\infty}^{\infty} \left|\frac{H(f)}{D(f)}\right|^2 S_n(f)|D(f)|^2 df + \int_{-\infty}^{\infty} S_X(f)|D(f)|^2 \left|1 - \frac{H(f)}{D(f)}\right|^2 df$$

אם במקום לחפש $H(f)$ אופטימלי נחפש $(H(f)/D(f))$ אופטימלי (ואח"כ נכפיל הפתרון ב-$D(f)$) אזי יש לנו בדיוק אותה בעיה כמו במקרה $D(f) = 1$. ולכן:

$$\left(\frac{H(f)}{D(f)}\right)_{\text{opt}} = \frac{S_X(f)|D(f)|^2}{|D(f)|^2[S_X(f) + S_n(f)]} = \frac{S_X(f)}{S_X(f) + S_n(f)}$$

ולכן

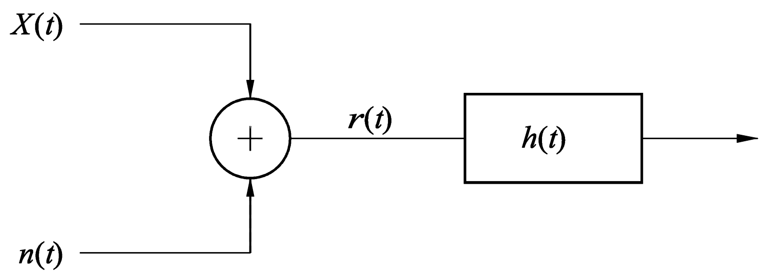
$$H_{\mathrm{opt}}(f)=\frac{D(f)S_X(f)}{S_X(f)+S_n(f)}$$

והשגיאה הריבועית הממוצעת המינימלית המתקבלת היא:

$$E\left[\varepsilon_{\min}^2(t)\right]=\int_{-\infty}^{\infty}\frac{D^2(f)S_X(f)S_n(f)}{S_X(f)+S_n(f)}df$$

עקרון ההשלכה

לסיכום סעיף זה על סינון לינארי אופטימלי, נחשב בחשבון ישיר וקצר את המסנן האופטימלי (כולל המקרה של קורילציה בין הסיגנל לרעש) בעזרת עקרון ההשלכה.



ציור 5.34

כזכור, $r(t)=n(t)+X(t)$. נסמן ב-$d(t)$ את התמרת פוריה ההפוכה של $D(f)$ אזי

$$E\left[\varepsilon^2(t)\right]=E\left[\int_{-\infty}^{\infty}X(t-\theta)d(\theta)d\theta-\int_{\infty}^{\infty}r(t-\theta)h(\theta)d\theta\right]^2$$

ע"י עקרון ההשלכה נקבל שהמסננת האופטימלית חייבת לקיים:

$$E\left[\left(\int_{-\infty}^{\infty}X(t-\theta)d(\theta)d\theta-\int_{-\infty}^{\infty}r(t-\theta)h_{\mathrm{opt}}(\theta)d\theta\right)r(\eta)\right]=0$$

עבור כל η, $-\infty<\eta<\infty$ ולכן:

$$\int_{-\infty}^{\infty}R_{r,X}(t-\theta-\eta)d(\theta)d\theta=\int_{-\infty}^{\infty}R_r(t-\theta-\eta)h_{\mathrm{opt}}(\theta)d\theta$$

נציג $\tau=t-\eta$ ונבצע התמרת פוריה על שני אגפי המשוואה, נקבל:

$$S_{r,X}(f)D(f)=S_r(f)H_{\mathrm{opt}}(f)$$

ולכן:

$$H_{\mathrm{opt}}(f)=\frac{D(f)S_{r,X}(f)}{S_r(f)}$$

הערה 1: שים לב כי לא השתמשנו בקשר $r = X + n$, ולמעשה פתרנו את המקרה הכללי בו r, X ת"א סמ"ר במשותף.

הערה 2: במקרה הפרטי בו $n(t)$ ו- $X(t)$ חסרי קורילציה מתקיים $S_{r,X}(f) = S_X(f)$ וכן $S_r(f) = S_X(f) + S_n(f)$ ומקבלים את התוצאה מהעמוד הקודם.

5.7 כמה מילים על ארגודיות

תהיה $\{X_1(\omega), X_2(\omega), \dots\}$ סדרה של מ"א בלתי תלויים ובעלי פילוג זהה. במקרה זה, חוק המספרים הגדולים אומר שעבור כל פונקציה $f(x), -\infty < x < \infty$ שעבורה $E|f(X_1)| < \infty$ מתקיים

$$\frac{1}{N}\sum_{i=1}^{N} f\Big(X_i(\omega)\Big) \underset{N\to\infty}{\longrightarrow} E\Big[f(X_1)\Big] \tag{5.7}$$

בכדי לתת לתוצאה זו מובן יש צורך להגדיר מה אנחנו מבינים כאשר אומרים שסדרה של משתנים אקראיים מתכנסת. נבטא זאת בצורה הבאה.

נניח ש- $E[|X_1(\omega)|^2] < \infty$ ונרשום $X_i(\omega) = E[X_1(\omega)] + Y_i(\omega)$ אזי ה- $Y_i(\omega)$ בעלי תוחלת אפס, ב"ת ובעלי פילוג זהה.

טענה:

$$E\left(\frac{1}{N}\sum_{1}^{N} X_i(\omega) - E[X_1]\right)^2 \underset{N\to\infty}{\longrightarrow} 0$$

הוכחה: צ"ל

$$E\left(\frac{1}{N}\sum_{1}^{N} X_i(\omega) Y_i\right)^2 \underset{N\to\infty}{\longrightarrow} 0$$

אולם

$$E\left(\frac{1}{N}\sum_{1}^{N} X_i(\omega) Y_i\right)^2 = \frac{1}{N^2}\sum_{1}^{N} E[Y_i^2] = \frac{1}{N}E[Y_1^2] \to 0$$

מכאן גם נובע (ראה אי שוויון צ'ביצ'ב בע"מ 8):

$$\text{Prob}\left\{\frac{1}{N}\sum_{i=1}^{N} |X_i - E[X_1]| > \delta\right\} \le \frac{E[Y_1^2]}{N\delta^2}$$

ולכן הביטוי הולך ל-0 כאשר $N \to \infty$.

שים לב שאם במקום לטפל ב-X_i נטפל ב-$f(X_i)$, ונדרוש $E[|f(X_1)|^2] < \infty$, נקבל:

$$\frac{1}{N}\sum_{i=1}^{N}\Big(E[f(X_i)] - E[f(X_1)]\Big) \to 0$$

תוצאה זאת מהווה "פרוש" אפשרי ל- (5.7).

מתבקשת השאלה אם (5.7) נשאר נכון כאשר יש תלות בין האיברים השונים בסדרה האקראית.

<u>הגדרה:</u> יהיה $\{X(t), -\infty < t < \infty\}$ ת"א סטציונרי. התהליך נקרא ארגודי אם עבור כל K, כל פונקציה (חסומה) f של K משתנים וכל סדרה $t_1, t_2, \ldots, t_k$ מתקיים (בהסתברות 1):

$$\frac{1}{2\pi}\int_{-T}^{T} f\Big(X(t_1+s,\omega),\ldots,X(t_k+s,\omega)\Big)\,ds \underset{T\to\infty}{\longrightarrow} Ef\Big(X(t_1),\ldots,X(t_k)\Big)$$

שוב אנו מתעלמים מהשאלה באיזה מובן מתכנסים המ"א שבאגף שמאל למשתנה האקראי המנוון בצד ימין (משפט ידוע וקשה - המשפט הארגודי - אומר שלאגף שמאל יש אכן גבול אבל המשפט לא אומר מתי הגבול אכן שוה לתוחלת). את התכונה הארגודית נבטא ע"י הסיסמא: עבור תהליכים ארגודיים, ממוצע הזמן שוה לממוצע האנסמבל (=התוחלת).

<u>דוגמאות לתהליכים לא ארגודיים:</u>

1. משתנה אקראי $X_t(\omega) = C(\omega)$

2. $X_t(\omega) = A(\omega)\cos(2\pi f_0 t + \phi(\omega))$ כאשר ϕ בלתי תלוי ב-A ומפולג אחיד בתחום $[0, 2\pi]$. אם $A(\omega)$ מ"א לא מנוון אזי

$$\frac{1}{2T}\int_{-T}^{T} X_t^2 dt \to \frac{A^2(\omega)}{2} \neq E\left[\frac{A^2(\omega)}{2}\right]$$

אפשר להראות שעבור $A(\omega) \equiv$ התהליך אכן ארגודי.

3. יהיו

$$X_t(\omega) = \cos(2\pi f_0 t + \phi)$$
$$Y_t(\omega) = \cos(2\pi f_0 t + \psi)$$

כאשר ϕ, ψ ב"ת ומפולגים אחיד בתחום $[0, 2\pi]$.

נגדיר

$$Z_t(\omega) \triangleq X_t(\omega)Y_t(\omega)$$

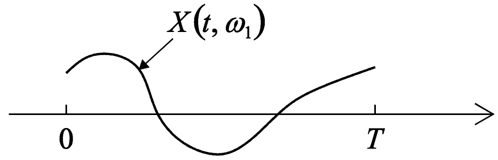
נקל לראות ש-$Z_t(\omega)$ איננו ת"א ארגודי.

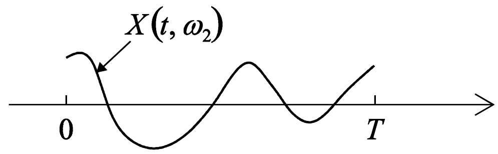
4. תערובת של שני תהליכים שונים שכל אחד ארגודי לא חייב להיות ארגודי.

בהרגשה ברור (פרט אולי למקרים מנוונים) שעל מנת שת"א יהיה ארגודי צריך לדרוש שבמהלך ההתפתחות בזמן של כל דגם, הוא יקבל את כל צורות הגל האפשריות. דהיינו: אם

אזי אם נסתכל על הדגם עם ω_1 עבור זמן מספיק ארוך, נמצא (t_0, t_0+T) שעליו התהליך דומה מאוד לתהליך $X(t, \omega_2)$ בקטע $(0, T)$.

כאמור, הבעיה לקבוע אם ת"א נתון הוא ארגודי היא בעיה קשה. נסתפק לכן בפתרון חלקי (מאד) של בעיית הארגו-דיות:





ציור 5.35

<u>הגדרה:</u> ת"א $\{X(t), -\infty < t < \infty\}$ סטציונרי (או אפילו סטציונרי במובן הרחב) ייקרא ארגודי לגבי הממוצע אם

$$E\left(\frac{1}{2T}\int_{-T}^{T} X(t)dt-\mu\right)^2 \underset{T\to\infty}{\longrightarrow} 0$$

כאשר $E[X(t)]=\mu$, דהיינו, אם

$$E\left(\frac{1}{2T}\int_{-T}^{T}\Big(X(t)-\mu\Big)dt\right)^2 \underset{T\to\infty}{\longrightarrow} 0$$

<u>משפט:</u> אם הקוואריאנס

$$\mathbf{K}_X(t_1,t_2)=E\Big[(X(t_1)-\mu)(X(t_2)-\mu)\Big]$$

מקיים:

$$\frac{1}{4T^2}\int_{-T_2}^{T_2}\int_{-T_1}^{T_1}\mathbf{K}_X(t_1,t_2)dt_1dt_2 \underset{T_1,T_2\to\infty}{\longrightarrow} 0$$

אזי התהליך ארגודי לגבי הממוצע.

<u>הוכחה:</u> מידית.

מתוך משפט זה נוכל לקבל את התוצאה הבאה (לא ניתן כאן את ההוכחה, ראה הערה בהמשך): אם $X(t)$ סטציונרי במובן הרחב ו-$\mathbf{K}_X(t_1,t_1+\tau)=\psi(\tau)$ ואם $\psi(\tau)$ מקיימת

$$\frac{1}{2T}\int_{-2T}^{2T}\psi(\tau)\left(1-\frac{|\tau|}{2T}\right)d\tau \underset{T\to\infty}{\longrightarrow} 0 \qquad (5.8)$$

אזי התהליך ארגודי לגבי הממוצע. ההוכחה מתבססת על הקשר

$$\int_{-T}^{T}\int_{-T}^{T}Q(t_1-t_2)dt_1dt_2=\int_{-2T}^{2T}Q(\theta)(2T-|\theta|)d\theta$$

<u>הערה:</u> ההוכחה לתוצאה זו מצויה בעמ' 63.

ברוח ההגדרה של ארגודיות לגבי הממוצע נגדיר גם:

<u>הגדרה:</u> יהיה $X(t)$ ת"א סטציונרי. נסמן

$$z_\lambda(t) = X(t+\lambda)X(t)$$

אם עבור כל λ ,$z_\lambda(t)$ ארגודי לגבי הממוצע אזי נגיד ש-$X(t)$ ארגודי לגבי הקורילציה. לכן, על מנת ש-$X(t)$ יהיה ארגודי לגבי הקורילציה צריך להתקיים

$$E\left(\frac{1}{2T}\int_{-T}^{T} X(t+\tau)X(t)dt - R_X(\tau)\right)^2 \underset{T\to\infty}{\longrightarrow} 0$$

נסמן ב-$R_z(\tau)$ את פונקצית האוטוקורילציה של $z_\lambda(t)$:

$$R_z(\tau) = E\Big[X(t+\tau+\lambda)X(t+\tau)X(t+\lambda)X(t)\Big]$$

אזי

$$\psi(\tau) = R_z(\tau) - R_X^2(\lambda)$$

צריך לקיים את התנאי (5.8) לעיל.

<u>הערה:</u> על מנת לדעת את $\psi(\tau)$ עלינו לדעת לא רק את המומנטים מסדר שני של $X(t)$ אלא גם את המומנטים מסדר רביעי. במקרה הגאוסי יש קשר בין השניים: מתקיים (כאשר $X(t)$ גאוסי עם תוחלת אפס)

$$\psi(\tau) = R_X(\lambda+\tau)R_X(\lambda-\tau) + R_X^2(\tau)$$

ולכן במקרה זה אם $R_X(\tau)$ יורד לאפס מספיק מהר אזי התהליך ארגודי לגבי הקורילציה.

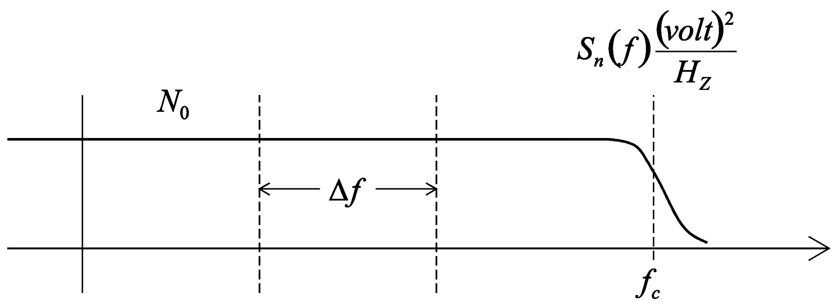
<u>משמעות הארגודיות לגבי מדידות</u>

נתונה מערכת כגון מקלט המיוצר בכמויות. על מנת לבדוק את תגובת המקלטים לאותות המשודרים אליהם בתנאי רעש חיצוני, אנו בונים מערכת מדידה המכילה מקור המייצר את האותות המשודרים. המקור מורכב מגנרטור אות וגנרטור רעש. למערכת זו נחבר את אחד המקלטים המיוצרים ונמדוד את התנהגות המערכת ונוכל מתוך המדידות לקבל תוצאות עבור טיבו של אותו מקלט. אם נחזור על המדידות ונמצע (בצורה נאותה) על פני מספר גדול של מקלטים נקבל הערכה על טיב אותו סוג מקלטים. בכל אותן מדידות, בין אם על מקלט בודד ובין אם על מספר גדול של מקלטים, נשאר (במציאות) אותו גנרטור רעש. בעצם אולי היינו צריכים לבדוק כל מקלט על מספר גנרטורי רעש ולמצע אולם לא עושים זאת. מדוע? התשובה היא שאנו מניחים שהרעש הנוצר מגנרטור הרעש הוא ארגודי ולכן אפשר להסתפק בגנרטור רעש אחד היות וכפי שהוסבר כבר, בהרגשה, על מנת שתהליך יהיה ארגודי כל דגם בודד חייב (בהסתברות 1) לקבל את כל צורות הגל האפשריות שהתהליך יכול לקבל. אותו שיקול תופס גם לגבי האות המשודר. אם המקור המשודר הוא אות אקראי, היינו צריכים לחבר כל מקלט להרבה גנרטורי אות ולמצע. שוב, בהנחה שאות המקור האקראי הוא ארגודי, נוכל להסתפק בגנרטור אות יחיד. שאות המקור האקראי הוא ארגודי, נוכל להסתפק בגנרטור אות יחיד.

6 רעשים

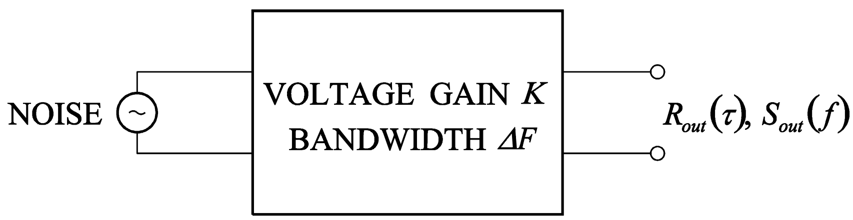
6.1 רעש לבן

בבעיות מעשיות עוסקים, הרבה פעמים, במקרה הבא: קיים מקור רעש בעל תוחלת אפס, שהצפיפות הספקטרלית שלו מתחילה בתדרים נמוכים מאוד ונמשכת עד לתדרים גבוהים מאוד ("מקור רחב סרט") כמצויר:



ציור 6.1

רעש זה מועבר דרך מגבר בעל רוחב סרט ΔF והגברת מתח K. המגבר אולי רחב סרט אולם רוחב הסרט שלו קטן בהרבה מזה של הרעש. במילים אחרות, בתחום התדרים שבו המגבר פועל כמגבר, הצפיפות הספקטרלית למעשה קבועה. הבעיה היא מה עוצמת הרעש האפקטיבית ביציאה מהמגבר.



ציור 6.2

עוצמת הרעש האפקטיבית ביציאה היא

$$\sqrt{R_{\text{opt}}(0)} \cong \sqrt{K^2 N_0 \Delta F 2}$$

שים לב ש-f_c לא מופיעה ביציאה. מה שחשוב זה N_0. אם במקום $S_n(f)$ הנתון היינו מסגננים את הבעיה ומניחים $S_n(f) \equiv N_0$ בכל התדרים היינו מקבלים אותה תוצאה עבור $R_{\text{opt}}(0)$. מקור רעש שהצפיפות הספקטרלית שלו היא N_0 בכל התדרים נקרא "רעש לבן".

שים לב שאם

$$S_n(f) = N_0$$

אזי

$$R_n(\tau) = N_0 \delta(\tau)$$

ואז $R_n(0) \simeq \infty$. בעיה זו מופיעה כבר ב-$S_n(f) = N_0$ כי ההספק נתון ע"י $\int_{-\infty}^{\infty} N_0 df = \infty$.

במלים אחרות: רעש לבן הוא אידיאליזציה של תהליך אקראי פיזיקלי, הוא אינו קיים כתהליך עם הספק ממוצע סופי. יש לראותו כסגנון של תהליך פיזיקלי כפי שתארנו בדוגמא. הרבה פעמים במציאות לא ידוע מהלך $S_n(f)$ בתדרים גבוהים ואפילו f_c אינו ידוע (רק ידוע שהוא מעל לתדר מסויים). כל שידוע הוא שבתדרים המעניינים אותנו הצפיפות הספקטרלית קבועה ברמה N_0 (אגב, במקרה הדטרמיניסטי האם אפשר ליחס אנרגיה סופית לפונקציה $\delta(t)$?). לכן במקרים אלה הסגנון נוח מאוד. עבור מקור רעש לבן

$$\text{הספק היציאה} = \int_{-\infty}^{\infty} N_0 |H(f)|^2 df = \int_0^{\infty} 2N_0 |H(f)|^2 df$$

יש המגדירים צפיפות ספקטרלית חד-צדדית:

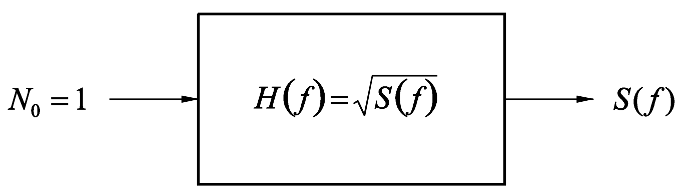
$$S(f)_{\text{חד-צדדי}} = \begin{cases} 0, & f < 0 \\ \\ 2S(f), & f > 0 \end{cases}$$

ואז

$$\text{הספק היציאה} = N_{0\,\text{חד-צדדי}} \cdot \Delta f \cdot K^2$$

בחשבונות תיאורטיים נוח להגדיר את $S(f)$ כפי שהגדרנו (צפיפות ספקטרלית דו-צדדית) ואילו בעבודה הנדסית נוח להשתמש בצפיפות ספקטרלית חד-צדדית. בכל מקרה של עיון בספר או מאמר צריך לודא באיזו הגדרה הם משתמשים. אנחנו נמשיך להשתמש בהגדרה הדו-צדדית.

אפשר לראות כל רעש לא לבן "כאילו נוצר ע"י רעש לבן" כדלקמן:



ציור 6.3

למשל, אם $S(f) = \frac{1}{1+(2\pi f)^2}$ אזי אפשר לבחור באחת האפשרויות

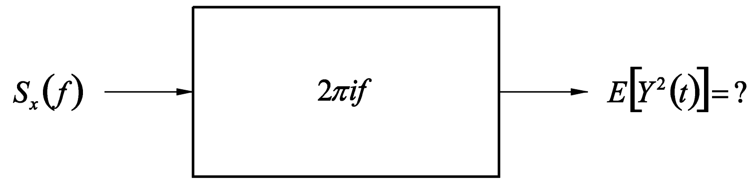
$$H(f) = \begin{cases} \dfrac{1}{1+i2\pi f} \\ \\ \dfrac{1}{1-i2\pi f} \\ \\ \dfrac{1}{\sqrt{1+(2\pi f)^2}} \end{cases}$$

וכל אחת מהן תיתן את $S(f) = \frac{1}{1+(2\pi f)^2}$. שים לב שרק אחת מהשלוש היא סיבתית (איזו?). אפשר להראות (הקריטריון של Paley-Wiener) שאם

$$\int_{-\infty}^{\infty} \frac{|\log S(f)|}{1+f^2} df < \infty$$

אזי קיימת $H(f)$ סיבתית הנותנת ביציאה צפיפות ספקטרלית $S(f)$ כאשר בכניסה רעש לבן.

נעיין בגוזר



ציור 6.4

נעיין במקרים הבאים עבור $S(f)$

$$S(f)=\begin{cases} N_0 \\ \\ \dfrac{1}{1+f^2} \\ \\ \dfrac{1}{1+f^4} \end{cases}$$

אילו ממקורות אלה אפשר לגזור?

מוסר ההשכל הוא שאם מדובר בגוזר הרי שהסגנון של רעש לבן איננו תופס.

<u>אינטגרלים של רעש לבן</u>

יהיה $\{n(t), -\infty < t < \infty\}$ רעש עם צפיפות ספקטרלית N_0 ותהינה $h_1(t), h_2(t), \ldots$ פונקציות דטרמיניסטיות בעלות אנרגיה סופית $\int_{-\infty}^{\infty} h_i^2(t)dt < \infty, i=1,2,\ldots$.

נסמן

$$X_i \int_{-\infty}^{\infty} n(t)h_i(t)dt$$

אזי, ברוח ההגדרה של רעש לבן ($n(t)$ "תהליך בעל רחב סרט גדול מאד")

$$\begin{aligned} E[X_i] &= 0 \\ E[X_i, X_j] &= E\left[\int_{-\infty}^{\infty} n(t)h_i(t)dt \cdot \int_{-\infty}^{\infty} n(s)h_j(s)ds\right] \\ &= \iint_{-\infty}^{\infty} h_i(t)(t)h_j(s)E\Big[n(t)n(s)\Big]dt\,ds \\ &= N_0 \iint_{-\infty} h_i(t)h_j(s)\delta(t-s)dt\,ds \\ &= N_0 \int_{-\infty}^{\infty} h_i(t)h_j(t)dt \end{aligned}$$

ובמיוחד, אם $n(\cdot)$ רעש לבן גאוסי (X_1, X_2) וקטור אקראי גאוסי עם תוחלת אפס ומטריצת קורילציה

$$E[X_1^2] = N_0 \int_{-\infty}^{\infty} h_1^2(s)ds$$

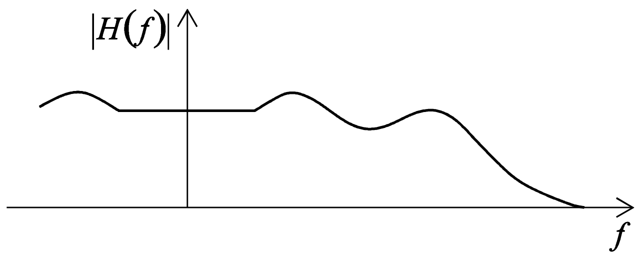
$$E[X_2^2] = N_0 \int_{-\infty}^{\infty} h_2^2(s)ds$$

$$E[X_1 X_2] = N_0 \int_{-\infty}^{\infty} h_1(s)h_2(s)ds$$

מה צריך להיות הקשר בין $h_1(\cdot)$ ל-$h_2(\cdot)$ במקרה זה על מנת ש-X_1, X_2 יהיו מ"א בלתי תלויים?

6.2 רוחב סרט אפקטיבי לרעש

נתונה פונקצית תמסורת $H(f)$ מסוג Low-Pass דהיינו $H(f)$ כמעט אפס בתדרים גבוהים ושונה מאפס בתדרים נמוכים כגון:



ציור 6.5

רוצים להגדיר "רוחב סרט אפקטיבי" לפונקצית תמסורת כזו. הגדרה מקובלת היא ההגדרה הבאה:

$$\Delta F \triangleq \frac{\int_{-\infty}^{\infty} |H(f)|^2 df}{2|H(0)|^2}$$

כלומר רוחב הסרט האפקטיבי לרעש, ΔF, הוא היחס שבין הספק היציאה כאשר בכניסה רעש לבן עם $N_0 = 1$ לבין $2|H(0)|^2$. הגדרה זו נוחה למדי להרבה בעיות, אך כדאי לזכור שיש גם הגדרות אחרות.

__דוגמא:__

$$H(f) = \frac{1}{1 + i\frac{f}{f_0}}$$

אזי f_0 הוא רוחב הסרט של $3\,\text{db}$. $H(0) = 1$ ולכן

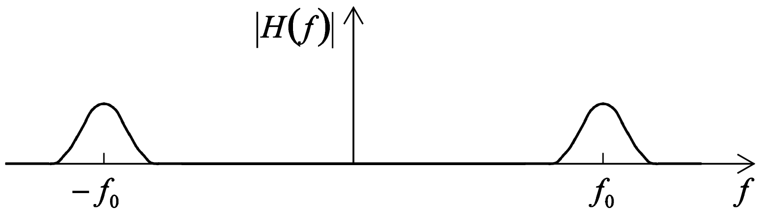
$$\Delta F = \frac{1}{2}\int_{-\infty}^{\infty} |H(f)|^2 df = \frac{f_0}{2}\int_{-\infty}^{\infty} \frac{df/f_0}{1 + (f/f_0)^2} = \frac{\pi}{2} f_0$$

ולכן רוחב הסרט האפקטיבי לרעש הוא $\frac{\pi}{2} f_0$.

במקום להגדיר את רוחב הסרט האפקטיבי ל-$H(0)$ אפשר גם להתיחס לתדר אחר:

$$\Delta F=\frac{\int_{\infty}^{\infty}|H(f)|^2 df}{2|H(f_0)|^2}$$

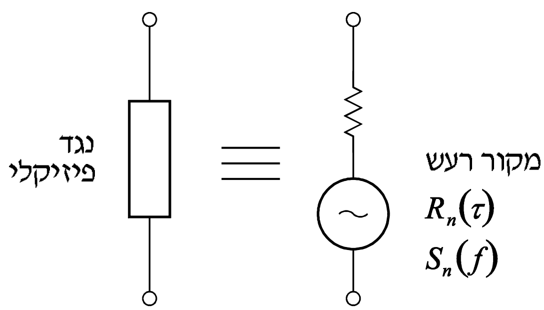
הגדרה זו נוחה במיוחד לפונקציות תמסורת (כולל כמובן מגברים) מטפוס Band Pass, דהיינו לפונקציות תמסורת מהטיפוס:



ציור 6.6

6.3 רעש טרמי (רעש הנגד, רעש Nyquist)

בכל נגד פיזיקלי ישנם אלקטרונים הנמצאים בתנועה אקראית. אלקטרונים אלה חייבים להיות בנגד על מנת שהנגד יוכל אכן להעביר זרם. תנועת האלקטרונים האקראית בנגד יוצרת מתחים על פני הנגד שהממוצע שלהם אפס אולם המתח הרגעי הוא מתח רעש שאיננו אפס. לכן עלינו לראות כל נגד פיזיקלי כמורכב מנגד אידיאלי בצרוף מקור רעש:

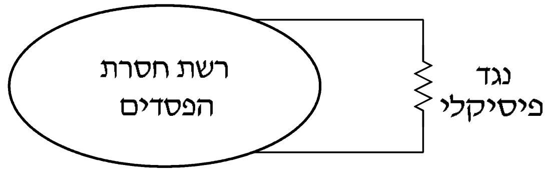


ציור 6.7

בתנאי "שיווי משקל" עם הסביבה נוכל להניח שרעש הנגד הוא תהליך אקראי סטציונרי והבעיה היא למצוא את $S_n(f)$. באופן כללי $S_n(f)$ יהיה אולי פונקציה של התדר, של ההתנגדות הכוללת R, של הטמפרטורה T, של החומרים מהם מורכב הנגד ושל הגיאומטריה שלו. נדגיש שאנו עוסקים בנגד (או רשת נגדים, סלילים וקבלים) ללא מקורות הספק חיצוניים. לפני שנמשיך נשאל את השאלה הבאה: האם גם אלמנטים חסרי הפסדים - קבלים, סלילים, טרנספורמטורים אידיאליים יוצרים רעש על ההדקים שלהם כמו נגד? התשובה היא:

<u>משפט א'</u>: רשת חסרת הפסדים איננה יוצרת רעש.

<u>הוכחה</u>: נעיין ברשת חסרת הפסדים המחוברת לנגד:



ציור 6.8

ונניח שהמערכת נמצאת בשיווי משקל תרמודינמי. אם הרשת חסרת ההפסדים היתה יוצרת מתחי רעש, מתחים אלו היו מחממים את הנגד. באותו זמן מתחי הרעש של הנגד אינם יכולים לחמם את הרשת חסרת ההפסדים מאחר ורשת חסרת הפסדים איננה יכולה לקבל הספק. התוצאה: הנגד היה מתחמם ואילו הרשת היתה מתקררת וזאת בניגוד לחוק השני של התרמודינימיקה. לכן מצב כזה בלתי אפשרי ולכן רשת חסרת הפסדים איננה יכולה לייצר רעש.

<u>משפט ב'</u>: עבור מקור הרעש הטורי עם הנגד:

$$S_n(f) = 2kTR$$

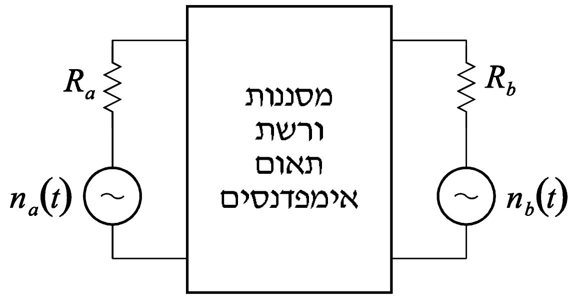
עד לתדירויות "גבוהות מאוד" כאשר k היא הקונסטנטה של בולצמן $1.38\cdot 10^{23}\ \mathrm{Joul}/^oK$. (הערה: "הנוסחה המעשית" היא $S_n(f) = 4kTR$ חד-צדדי ; מכל מקום, תמיד נכון ש-$4kTR\Delta f = \nabla^2 =$ {ממוצע המתח ברבוע בתחום התנדרים Δf).

שים לב שהתוצאה היא אוניברסלית ואיננה תלויה בחומרים או בגיאומטריה של הנגד. לפי תוצאה זו אין "נגד רועש" ו-"נגד שקט" - כל הנגדים טובים או רעים באותה מידה מנקודת המבט של הרעש שהם יוצרים. יש להדגיש שתוצאה זו בהחלט נכונה כאשר מדובר במערכת מבודדת, אולם היא איננה נכונה כשמדובר בנגד שמזרימים דרכו זרמים חיצוניים. נקודה זו תתבהר בהוכחה (החלקית) למשפט זה שנביא.

<u>הוכחה חלקית</u>: נעיין בשני נגדים R_a ו-R_b נחבר אותם זה אל זה באמצעות מעגלי תאום אימפדנסים ומסננות (כאשר מעגלי התאום והמסננות חסרי הפסדים) ונניח שהמערכת נמצאת בשיווי משקל תרמודינמי בינה לבין עצמה ועם הסביבה

ציור 6.9א'

נסכם את המערכת כדלקמן:



ציור 6.9ב'

כאשר $n_b(t), n_a(t)$ הם הרעשים הרגעיים של הנגד אחרי מעבר דרך מסננת צרת סרט, דהיינו:

$$E\left[n_a^2(t)\right] = 2\, S_n^a(f_0) \cdot \Delta f$$

$$E\left[n_b^2(t)\right] = 2\, S_n^b(f_0) \cdot \Delta f$$

בהנחה של תאום אימפדנסים, הזרם שיוצא מהמקור השמאלי הוא $n_a(t)/2R_a$ ומהמקור הימני הוא $n_b(t)/2R_b$. ההספק שהמקור השמאלי מוסר לעומס (הנגד) הימני הוא (נזכור שקיים תאום אימפדנסים):

$$\frac{1}{2} n_a(t) \frac{n_a(t)}{2R_a}$$

בצורה דומה, ההספק שהמקור הימני מוסר לעומס השמאלי הוא:

$$\frac{n_b^2(t)}{4r_b}$$

ולכן, מתוך שיקול של שיווי משקל תרמודינמי, חייב להתקיים איזון ממוצע של הספקים

$$E\left[\frac{n_a^2(t)}{4R_a}\right] = E\left[\frac{n_b^2(t)}{4R_b}\right]$$

או

$$\frac{S_n^a(f_0)}{R_a} = \frac{S_n^n(f_0)}{R_b}$$

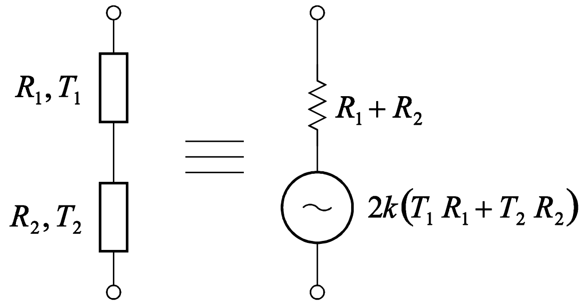
<u>מסקנה:</u> היחס $S_n(f)/R$ הוא יחס אוניברסלי שיכול להיות תלוי בתדר f או בטמפרטורה T אבל איננו יכול להיות תלוי בסוג הנגד (ז.א. לא תלוי בחומר ובגיאומטריה). נקבל ללא הוכחה שע"י אנליזה של נגד עם גיאומטריה פשוטה אפשר להראות

$$S_n(f) = 2kTR$$

ובגלל השיקול שהבאנו, התוצאה מתקיימת עבור כל הנגדים. עד איזו תדירות התוצאה הזו נכונה? בתדרים מספיק גבוהים כאשר מימדי הנגד הם בסדר גודל של אורך הגל ממילא ה"נגד מפסיק להיות נגד". באופן בסיסי התוצאה נכונה עד קרוב לתדירויות אופטיות.

הערות:

א. שים לב שמתחי הרעש מסתכמים "על בסיס של הספקים" דהיינו:

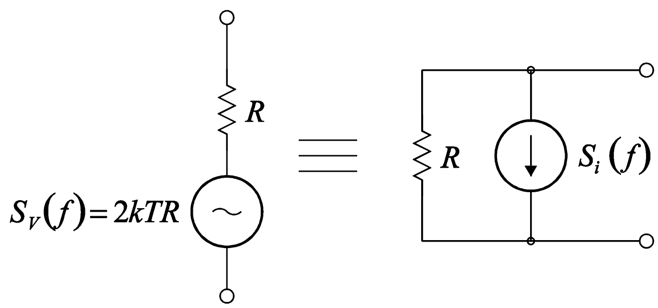


ציור 6.10

ולכן, הטמפרטורה האפקטיבית של שני הנגדים תהיה:

$$T_{\text{eff}} = \frac{T_1 R_1 + T_2 R_2}{R_1 + R_2}$$

ב. מעגל תמורה מקבילי במקום מעגל תמורה טורי



ציור 6.11

מעגל תמורה עם הרעש כמקור מתח טורי, אפשר כידוע, להחליף במעגל תמורה עם הרעש כמקור זרם מקבילי. על מנת למצוא את הקשר בין הצפיפות הספקטרלית של מקור המתח $S_v(f)$ עם זה של הזרם $S_i(f)$, נשווה את המתח ריקם על פני ההדקים: $i(t)R = v(t)$

ולכן

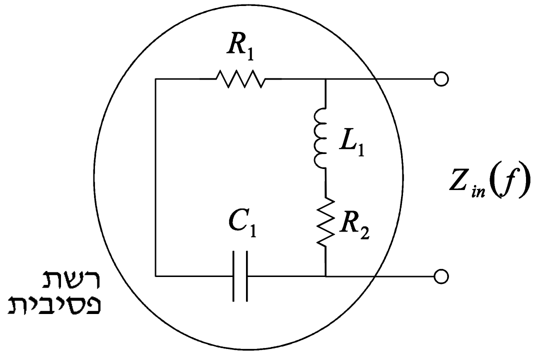
$$R^2 E\Big[i(t+\tau)i(t)\Big] = E\Big[v(t+\tau)v(t)\Big]$$

ומכאן

$$S_i(f) = \frac{S_v(f)}{R^2} = 2kt\,TG$$

(ועבור $S_v(f) = 4kTR$ נקבל $S_i(f) = 4kTG$ כאשר $G = R^{-1}$).

נעבור למשפט ג': נעיין ברשת פסיבית



ציור 6.12

הרשת נמצאת בשיווי משקל תרמודינמי עם סביבתה. מה הרעש על פני הדקי הכניסה של הרשת?

נוכל כמובן להוסיף לכל נגד את מקור הרעש שלו ולחפש את סה"כ הרעש על פני ההדקים, אולם:

<u>משפט ג':</u> אם אימפדנס הכניסה של רשת (פסיבית הנמצאת בשיווי משקל תרמודינמי) הוא

$$Z_{\text{in}}(f) = R_{\text{in}}(f) + iX_{\text{in}}(f)$$

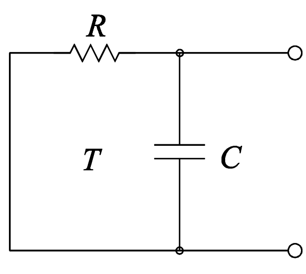
אזי הצפיפות הספקטרלית של מתח הרעש על פני הדקי המעגל יהיה:

$$S(f) = 2kT\,R_{\text{in}}(f)$$

וסה"כ הספק הרעש יהיה $\overline{v}^2 = \int_{\infty}^{\infty} 2kTR_{\text{in}}(f)df$

<u>הוכחה:</u> נחבר את המעגל הנתון עם נגד R באמצעות מסננת צרת סרט מסביב ל-f_0 ועם תאום אימפדנסים. השיקול של שיווי משקל תרמודינמי (ההספק הנמסר שווה להספק המתקבל) נותן לנו מיד את התוצאה המבוקשת.

<u>דוגמא:</u> נעיין במעגל

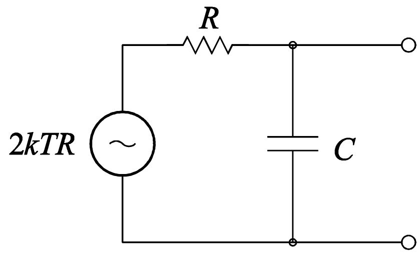


ציור 6.13א'

נוכל לחשב את הרעש על פני ההדקים ע"י מעגל התמורה:

וכאשר $H(f)$ פונקצית התמסורת ממקור מתח הרעש להדקי היציאה נקבל:

$$S(f) = |H(f)|^2 \cdot 2kTR = \left(\frac{\frac{1}{2}\pi ifC}{R + \frac{1}{2\pi ifC}}\right)^2 2kTR = \frac{2kTR}{(2\pi)^2 f^2 R^2 C^2 + 1}$$



ציור 6.13ב'

לפי המשפט שקבלנו

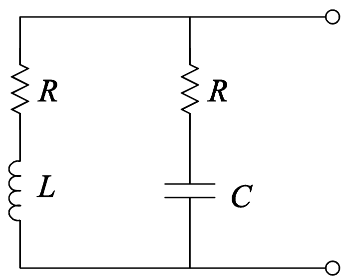
$$Z_{\rm in}(f)=\frac{R\cdot\frac{1}{i2\pi fC}}{R+\frac{1}{i2\pi fC}}=\frac{R(1-i2\pi fRC)}{1+(2\pi fRC)^2}$$

ולכן

$$R_{\rm in}(f)=\frac{R}{1+(2\pi fRC)^2}$$

ומתקבלת, כמובן אותה תוצאה.

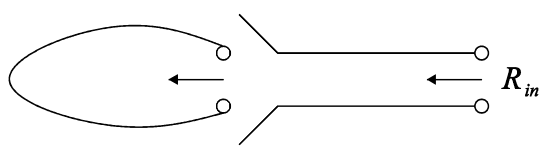
<u>תרגיל:</u>



ציור 6.14

עבור $R=\sqrt{\frac{L}{C}}$ המעגל מתנהג כמו נגד טהור בעל התנגדות R. מצא את הרעש על פני ההדקים בשתי שיטות:
(א) ע"י שמוש במשפט ג'. (ב) ע"י הוספת מקור רעש לכל אחד מהנגדים וחישוב השפעת מקורות אלה על הרעש שעל פני הדקי המעגל.

<u>הערת סיום לסעיף זה:</u> נעיין באנטנה כיוונית שכאנטנת שידור היא קורנת את כל ההספק הנמסר לה לכוון אחד (ראה הציור) ושהיא חסרת הפסדים, (דהיינו, כל ההספק הנמסר לה מוקרן החוצה)



ציור 6.15א'

נניח שאימפדנס הכניסה לאנטנה הוא אוהמי טהור, $R_{\rm in}$. נניח שהאנטנה צופה אל אחד הכוונים הבאים:

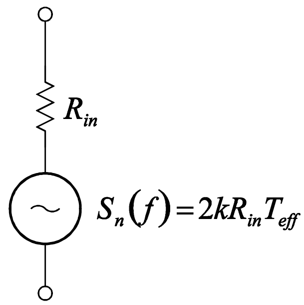
(א) החלל (בניצב לשביל החלב) - (בסדר גודל של 1000^o K ב-100 MHz, יורד לכ-5^o K ב-1000 MHz).

(ב) **שביל החלב** - (**בסדר גודל של** 4000^o K **ב**-100 MHz, **יורד לכ**-15^o K **ב**-1000 MHz).

(ג) **כדור הארץ** - (כ-300^o K).

(ד) **השמש** (נניח אלומת קרינה צרה מספיק כך שתקרין לשמש בלבד) - (**בסדר גודל של אלפי מעלות קלווין**).

מנקודת המבט של רעש תהיה סכימת התמורה של האנטנה



ציור 6.15ב'

כאשר בכל אחד מהמקרים יש להציג את הטמפטורות המתאימות.

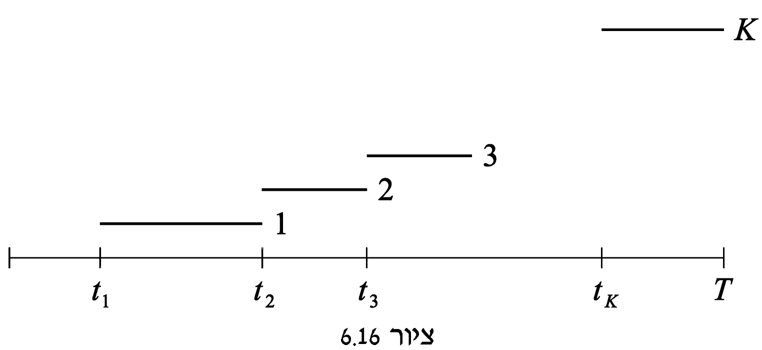
6.4 רעש הדיודה (Shot noise)

לפני שנתיחס לרעש הדיודה, נביא תוספת לתהליכי פואסון. כזכור, תהליך פואסון $N-1$ הוא תהליך של הפרשים בלתי תלויים

$$\text{Prob}\{N_t = k\} = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

כאשר - λ הוא קצב התהליך.

נבצע סימולציה של תהליך אקראי (על מחשב) כדלקמן: נקבע T מסוים, נגריל מספר אקראי K לפי פילוג פואסון עם תוחלת λT. אחרי שהגרלנו את K נבחר באקראי K זמנים ב-$[0,T]$, בלתי תלויים ומפולגים בשווה בתחום הזה. נסמן את הזמנים האלה $(\theta_1, \ldots, \theta_K)$, נסדר את הזמנים בסדר עולה ונסמנם מחדש $(t_1, \ldots, t_K)$ (כעת $t_1 \le t_2 \le \cdots \le t_k$) ונבנה מהם פונקצית מדרגות עולה כדלקמן:

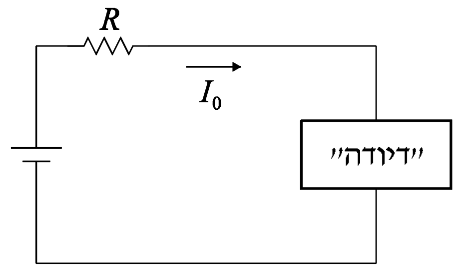


ציור 6.16

<u>טענה</u> (ללא הוכחה): חוק ההסתברות של התהליך האקראי שהמצאנו הוא חוק ההסתברות של תהליך פואסון (טענה זו הגיונית מאוד: אם נחשוב על תהליך פואסון כעל התהליך האקראי שמכונית מגיעה לצומת ואם אנו יודעים ש-K

מכוניות הגיעו לצומת בקטע הזמן $[0,T]$, כל אחת מהמכוניות שיגיעו לצומת לא יודעת על האחרת וסביר להניח שפילוג הגעתה לצומת יהיה פילוג אחיד ב-$(0,T)$).

נעבור כעת לזרם בדיודה: נסמן ב-q את מטען האלקטרון ונניח שאנו מזרימים זרם I_0 (קבוע) דרך הדיודה, משמעות הדבר היא שבכל שניה I_0/q אלקטרונים עוברים מהקטודה לאנודה.



ציור 6.17

אם מטען האלקטרון היה "אפס" (ו"אינסוף" אלקטרונים בשניה) היה עובר דרך הדיודה והנגד זרם D.C. טהור I_0 והצפיפות הספקטרלית של הזרם היתה

$$S(f) = I_0^2\delta(f)$$

מאחר והזרם נגרם ע"י אלקטרונים בעלי מטען סופי נצפה לצפיפות ספקטרלית מהצורה

$$S(f) = I_0^2\delta(f) + S_n(f)$$

מטרתנו היא לחשב את $S_n(f)$.

במודל שלנו, אלקטרון הנפלט ברגע $t=0$ משרה במעגל זרם $i_e(t)$ ולכן:

$$\text{מטען האלקטרון} = q = \int_{-\infty}^{\infty} i_e(t)dt$$

בקטע הזמן T יפלטו בממוצע I_0T/q אלקטרונים. מאחר והאלקטרונים נפלטים מאיזורים שונים בקטודה סביר להניח שיש אי תלות בין זמני הפליטה. נוכל, לכן, להניח שתהליך פליטת האלקטרונים הוא תהליך פואסון. לכן (בהזנחת אפקט הקצוות, ז.א. T גדול)

$$I(t) = \sum_{k=1}^{K} i_e(t-t_k)$$

כאשר K מ"א המפולג לפי פילוג פואסון; $\lambda = I_0T/q$ ו t_k מפולגים אחיד בתחום $[0,T]$. לכן:

$$E\Big[I(t)\Big] = E\left[E\left[\sum_{k=1}^{K} i_e(t-t_k)|K\right]\right] = E\left[K\frac{1}{T}\int_0^T i_e(t-\theta)d\theta\right] = \frac{q}{T}E[K] = \frac{q}{T}\frac{I_0}{q}T = I_0$$

השוויון הראשון הוא מתכונת ההחלקה, בשוויון השני משתמשים בעובדה: $Eg(X) = \int f_X(\theta)g(\theta)d\theta$ (עבור מ"א X ופונקציה g), בשוויון השלישי באה לידי ביטוי הזנחת הקצוות ובשוויון הרביעי מופיעה התוחלת של תהליך פואסון.

נפנה כעת למומנט השני:

$$R(\tau) = E\Big[I(t+\tau)I(t)\Big] = E\left[\sum_{k=1}^{K} i_e(t+\tau-t_k)\sum_{j=1}^{K} i_e(t-t_j)\right]$$

$$= E\left[\sum_{k=1}^{K} i_e(t+\tau-t_k)i_e(t-t_k)\right] + E\left[\sum_{k\neq j}\sum i_e(t+\tau-t_k)i_e(t-t_j)\right]$$

היות וה-t_k מפולגים אחיד בתחום $[-T/2, t/2]$, נמצע תחילה על פני ה-t_k ואח"כ על פני K. לכן, בהזנחת אפקט הקצוות, (E יסמן מצוע על K)

$$R(\tau) = E\left[\sum_{1}^{K}\frac{1}{T}\int_{-T/2}^{T/2} i_e(t+\tau-\theta)i_e(t-\theta)d\theta\right] + E\left[\sum_{k\neq j}\sum\frac{1}{T^2}\int_{-T/2}^{T/2} i_e(t+\tau-\theta)d\theta\int_{-T/2}^{T/2} i_e(t-\eta)d\eta\right]$$

$$\cong \frac{E[K]}{T}\int_{-\infty}^{\infty} i_e(\tau+\theta)i_e(\theta)d(\theta) + \frac{E[K^2-K]}{T^2}q^2$$

עבור פילוג פואסון

$$E[K^2-K] = (\lambda T)^2 \quad ; \quad E[K] = \lambda T$$

לכן

$$R(\tau) = \frac{I_0}{q}\int_{-\infty}^{\infty} i_e(t+\tau)i_e(t)dt + I_0^2$$

נסמן ב-$G_e(f) = F\{i_e(t)\}$, ונקבל ע"י התמרת פוריה

$$S(f) = I_0^2\delta(f) + \frac{I_0}{q}G_e(f)G_e^*(f)$$

יהיה T זמן המעבר של אלקטרון מהקטודה לאנודה, עבור תדרים f המקיימים $\tau f \ll 1$ יתקיים $G_e(f) \cong G_e(0)$ ובמקרה שלנו $G_e(0) = q$ ולכן

$$S(f) \cong I_0^2\delta(f) + I_0\cdot q$$

או

$$S_n(f) = I_0 q$$

התוצאה שקבלנו היא נוסחת רעש הדיודה.

הערות:

(א) שים לב שבעזרת התוצאה שקבלנו וע"י מדידת רעש הדיודה אפשר למצוא באופן נסיוני את מטען האלקטרון q.

(ב) בספרי עזר מוצאים הנוסחא $2I_0q$ מדוע?

(ג) התוצאה שקבלנו היא עבור זרמי דיודה לא גדולים מדי. עבור זרמי דיודה גדולים מופיע אפקט נוסף הנקרא אפקט המטען המרחבי והוא כדלקמן: אם ברגע מסוים עוזבים הרבה אלקטרונים את הקטודה הם יוצרים מטען מרחבי (בין הקטודה לאנודה) המעכב את הפליטה של אלקטרונים נוספים. אפקט זה גורם לרעש דיודה נמוך מהנתון ע"י הנוסחה שקבלנו אולם בזרמי דיודה לא גבוהים מדי הוא זניח.

6.5 רעש טרמי - פתוח מתוך מודל רעש הדיודה

נזכיר כי בדיודה,

$$I = I_0(e^{q\ V/KT} - 1)$$

ולמעשה, ניתן למדל את הזרמים בדיודה כך:

זרם סחיפה	$\xrightarrow{I_0\ e^{q\ V/kT}}$
זרם דיפוזיה	$\xleftarrow{-I_0}$

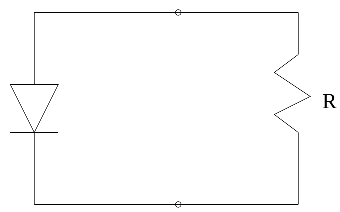
זרם הסחיפה וזרם הדיפזיה הם זרמים פואסניים, בת"ס. לכן,

$$S_n(I) = qI_0(1 + e^{q\ V/kT})$$

ועבור $V = 0$,

$$\boxed{S_n(I) = 2qI_0} \tag{6.1}$$

נחבר כעת את המעגל ראשית, נשים לב כי עבור הדיודה,



$$\frac{\partial I}{\partial V} = \frac{qI_0}{kT} \cdot e^{q\ V/kT}$$

ולכן, סביב מתח 0, מתנהגת הדיודה כמו נגד:

$$\boxed{\left.\frac{\partial I}{\partial V}\right|_{V=0} = \frac{qI_0}{kT} = \frac{1}{R}} \tag{6.2}$$

לכן, עם בחירת R כזה, מתקיים תאום אמפדנסים. לכן,

$$2qI_0 = S_n^{R,I}$$

ומכאן

$$S_n^{R,I} = \frac{2\ kT}{R}$$

ומכאן, במעבר לרעש מתח,

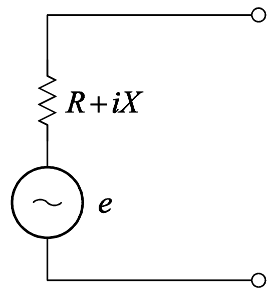
$$\boxed{S_n^{R,V} = 2kTR} \tag{6.3}$$

6.6 אפיון רעש מגבר

נניח שנתונים שני מגברים, האחד עם הגברה גבוהה ורעש גבוה ביציאתו והשני עם הגברה נמוכה ורעש נמוך ביציאתו. איזה מהם אחד לאחר השני איזה משניהם עדיף לשים לפני השני? בבעיות מסוג זה נעסוק בסעיף הנוכחי.

<u>הספק מצוי, הגבר הספק</u>

נעיין בתדר מסוים ונגדיר: הספק מצוי של המקור = הספק מכסימלי שאפשר לקבל מהמקור. הספק זה יתקבל בתנאי התאמת אימפדנסים ולכן עבור הציור יתקבל ההספק המצוי:



ציור 6.18

עבור רעש טרמי בתחום תדרים Δf, נקבל שהספק הרעש המצוי נתון ע"י:

$$E\left[\frac{e^2(t)}{4R(f)}\right] = \frac{2kTR(f)\Delta f \cdot 2}{4R(f)}$$
$$= kT\,\Delta f$$

וזה הספק הרעש המצוי במקור פיזיקלי פסיבי.

<u>הגדרות:</u>

א. הטמפרטורה האפקטיבית של מקור כלשהוא (פסיבי או אקטיבי)

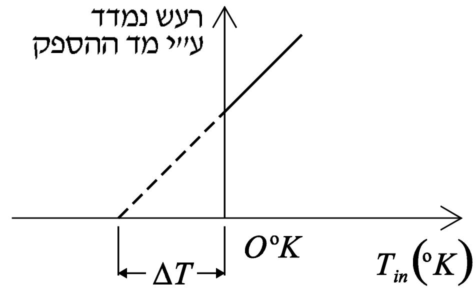
$$T_{\text{eff}} = \frac{1}{k \cdot \Delta f} \cdot [\Delta f \text{ הספק הרעש המצוי ברחב סרט }]$$

ב. הגבר הספק: $G_a = \dfrac{\text{הספק מצוי ביציאה}}{\text{הספק מצוי בכניסה}}$

להלן נביא סקירה מקוצרת על המושגים "טמפרטורת הרעש האפקטיבית של מגבר" ו"ספרת רעש". נעיין במערכת המצוירת; נשנה את הטמפרטורה של המקור (T_{in}) ונצייר את הספק הרעש ביציאה כפונקציה של טמפרטורת המקור. (שים לב, המקור R_0 הוא לא חלק מהמגבר). טמפרטורת הרעש האפקטיבית ΔT מוגדרת כמצויר:

במילים אחרות נחליף המגבר הרועש במגבר זהה אבל ללא רעש ונעלה את טמפרטורת המקור כך שסה"כ לא יהיה שינוי. כמה צריך להוסיף? תשובה ΔT.

ציור 6.19א'



ציור 6.19ב'

לאור ההגדרה, הספק הרעש המצוי ביציאת המגבר ברוחב סרט Δf והגבר G יהיה:

$$G \cdot k \cdot (T_{\text{in}} + \Delta T) \cdot \Delta f$$

לכן: אם המגבר מחובר דרך קו תמסורת ואנטנה חסרת הפסדים והאנטנה צופה ל- $T_{\text{(מקור)}}$, אזי,

אם $\Delta T \gg T_{\text{(מקור)}}$ אזי שיפור ב-ΔT ישפר את הקליטה.

אם $T_{\text{(מקור)}} \gg \Delta T$ אזי שיפור ΔT לא ישפר למעשה.

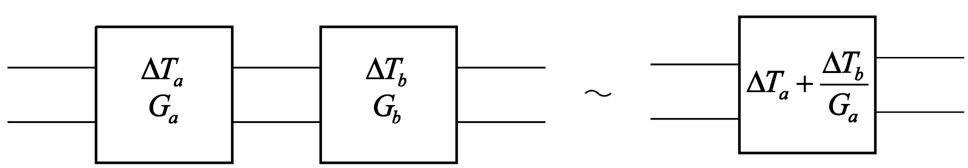
מספרי ΔT טיפוסיים:

מקלט מיקרוגלים לא מעולה $\Delta T \cong 10 \times \mathrm{T}_0 \quad T_0 = 300^{\mathrm{o}}\,\mathrm{K}$

מקלט מיקרוגלים מעולע $-4 \times T_0$

מגבר פרמטרי $\Delta T \sim 80^{o}\mathrm{K}$

<u>טענה:</u> הטמפרטורה האפקטיבית הכוללת של שני מגברים בזה אחר זה (קסקדה):



ציור 6.20

<u>הוכחה:</u> ראשית כפי שכבר הערנו, ההספק המצוי של הרעש ביציאה של מגבר (בודד), הוא $Gk\Delta f(T_{\text{מקור}} + \Delta T)$, (כאשר G - ההגבר המצוי ו- $T_{\text{מקור}}$ טמפרטורת הרעש של המקור). עבור קסקדה, הספק הרעש המצוי ביציאת מגבר

a נתון ע"י $G_a k\Delta f(T_{\text{מקור}} + \Delta T_a)$, ולכן ביציאה של מגבר b יהיה הספק רעש מצוי:

$$G_b G_a k\Delta f(T_{\text{מקור}} + \Delta T_a) + G_b k\Delta f \Delta f \Delta T_b = G_a G_b k\Delta f \left(T_{\text{מקור}} + \Delta T_a + \frac{\Delta T_b}{G_a}\right)$$

ולכן:

$$\Delta T_{ab} = \Delta T_a + \frac{\Delta T_b}{G_a}$$

שים לב שמהנוסחא נוכל להחליט אם עדיף לשים את a לפני b או את b לפני a.

מאפין נוסף לרעש המגבר הוא ספרת הרעש F המוגדרת כדלקמן: נחבר לכניסת המגבר נגד בטמפרטורת הסביבה (300^oK). ההספק המצוי של המקור יהיה כמובן $300kG\Delta f$. נגדיר כעת:

$$F = \frac{\text{הספק הרעש המצוי ביציאת המגבר (כולל רעש המגבר)}}{300k\Delta f G}$$

ולכן F קשור ל ΔT ע"י

$$F = 1 + \frac{\Delta T}{300}$$

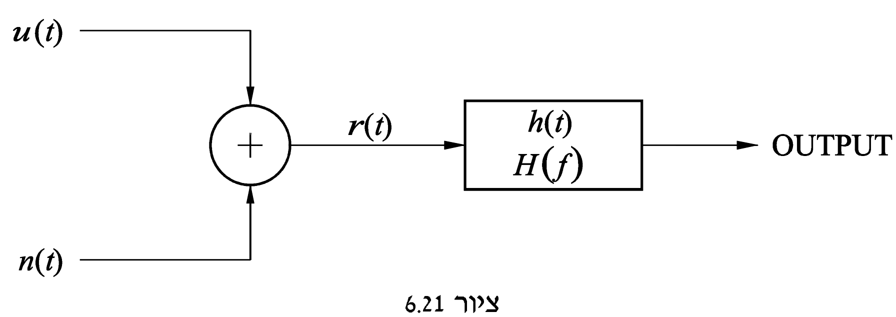
עבור מגבר אידיאלי שאינו יוצר רעש $\Delta T = 0$ ו-$F = 1$. שים לב ש-ΔT מאפיין את רעש המגבר ע"י <u>תוספת</u> הרעש שגורם המגבר, ואילו F מאפיין את רעש המגבר ע"י <u>מקדם כפל</u> לרעש המקור, כאשר המקור בטמפרטורה של 300^oK. מכיון ש-F מוגדר ביחס לטמפרטורה שרירותית ואילו הגדרת ΔT אינה דורשת נקודת יחוס שרירותית, ΔT הוא מושג בסיסי יותר).

<u>תרגיל:</u> הוכח שעבור חיבור שני מגברים בקסקדה

$$F_{ab} = F_a + \frac{F_b - 1}{G_a}$$

6.7 מסננת מתואמת Matched Filter

<u>אות הכניסה:</u> אות דטרמיניסטי בעל אנרגיה סופית, $u(t)$. למען הפשטות נניח ש-$u(t) = 0$ עבור $t > 0$. נסמן $U(f) = F\{u(t)\}$. עיין בציור.



ציור 6.21

<u>רעש הכניסה:</u> רעש לבן, תוחלת אפס צפיפות ספקטרלית N_0.

<u>היציאה:</u> כאשר בכניסה מצוי האות $u(t)$ בלבד (ללא רעש):

$$Y(t) = \int_{-\infty}^{\infty} u(t-\theta)h(\theta)d\theta$$

וברגע $t=0$ תהיה היציאה

$$Y(0) = \int_{0}^{\infty} u(-\theta)h(\theta)d\theta$$

לפי משפט פרסוול נוכל לרשום את $Y(0)$ במרחב התדר

$$Y(0) = \int_{-\infty}^{\infty} H(f)U(f)df$$

<u>הספק הרעש ביציאה</u> (הספק ממוצע):

$$\int_{-\infty}^{\infty} |H(f)|^2 \cdot N_0 df$$

ובמרחב הזמן, ממשפט פרסוול נובע

$$\int_{-\infty}^{\infty} |H(f)|^2 \cdot N_0 df = N_0 \int_{-\infty}^{\infty} h^2(\theta)d\theta$$

נגדיר יחס אות לרעש ביציאה ברגע $t=0$ כדלקמן:

$$\left(\frac{S}{N}\right)_{\substack{\text{out}\\ t=0}} = \frac{(Y(0))^2}{\text{הספק רעש ממוצע ביציאה}}$$

ולכן

$$\left(\frac{S}{N}\right)_{\substack{\text{out}\\ t=0}} = \frac{\left[\int_0^{\infty} u(-\theta)h(\theta)d\theta\right]^2}{N_0 \int_0^{\infty} h^2(\theta)d\theta} = \frac{\left[\int_0^{\infty} U(f)H(f)df\right]^2}{N_0 \int_0^{\infty} |H(f)|^2 df}$$

<u>הבעיה:</u> מצא מערכת לינארית (מאופיינת ע"י $h(\theta)$ או $H(f)$) כך שהיחס $\left(\frac{S}{N}\right)_{\substack{\text{out}\\ t=0}}$ יהיה מכסימלי.

<u>תזכורת:</u> אי השיוויון של שוורץ (לפונקציות קומפלקסיות)

$$\left(\int_{T_1}^{T_2} f_1(t)f_2^*(t)dt\right)^2 \leq \int |f_1(t)|^2 dt \cdot \int |f_2(t)|^2 dt$$

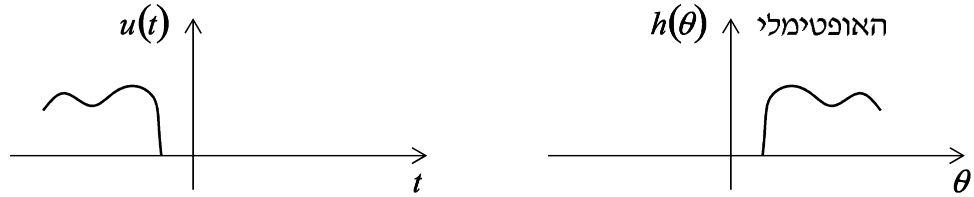
ואי השיוויון הופך לשיוויון כאשר $f_2(t) = \alpha f_1(t)$ (α מקדם פרופורציה שרירותי). לכן, נקבל מאי השיוויון של שוורץ ומהביטויים לעיל

$$\left(\frac{S}{N}\right)_{\substack{\text{out}\\ t=0}} \leq \frac{\int_0^{\infty} (u(-\theta))^2 d\theta \int_0^{\infty} h^2(\theta)d\theta}{N_0 \int_0^{\infty} h^2(\theta)d\theta} = \frac{1}{N_0}\int_0^{\infty} \Big(u(-\theta)\Big)^2 d\theta$$

כמו כן נובע מאי השיוויון של שוורץ שאי השיוויון שקבלנו זה עתה יתקיים כשיוויון כאשר

$$h(\theta) = \alpha u(-\theta)$$

מכאן <u>המסקנה</u>: $h(t)$ האופטימלי הוא תמונת הראי של $u(t)$ לגבי הציר האנכי.



ציור 6.22

$h(t)$ האופטימלי נקרא: המסננת המתואמת.

<u>הערה</u>: בעמוד הקודם מופיע ביטוי במרחב הזמן וביטוי במרחב התדר. במקום להשתמש בביטוי במרחב הזמן, השתמש בבטוי במרחב התדר על מנת לקבל (בעזרת אי השיוויון של שוורץ) ש-$H(f)$ האופטימלי הוא $\alpha U^*(f)$ וע"י התמרת פוריה הפוכה נקבל כמובן שנית את התוצאה ש-$h(t)$ האופטימלי הוא $\alpha u(-t)$.

נספח 1

7 המחשות לתהליכים אקראיים

בפרק זה נתאר מספר כלי המחשה שפיתחנו. ההמחשות משתמשות בתכנת MATLAB (רצוי גרסה 5 ומעלה) בלבד. התאור כאן מתאים למשתמשי חוות ה- PC בפקולטה להנדסת חשמל, טכניון. משתמשים אחרים מוזמנים להעתיק את הקבצים מהאתר הקורס.

הפעל את תכנת MATLAB (רצוי גרסה 5 ומעלה). כדי שתוכל להשתמש בקבצים שיצרנו, תן את הפקודה
path(path,'m:\044202\programs')

<u>הערה:</u> כדי לצייר מספר פונקציות על אותו גרף, תן את הפקודה hold on. כדי לנקות את הגרף: clf. כדי שכל גרף חדש ימחוק את קודמיו: hold off (אופציה זו מבוטלת בתכנות המסתיימות ב-P; ראה בהמשך.

7.1 מבוא: תהליכים פשוטים

דוגמה **7.1** דוגמה זו ממחישה את האותות האקראיים של דוגמה 1.1. כל המ"א בדוגמה זו הם בת"ס ומפולגים יוניפורמית, על תחומים כמצוין. בכל הרצה יבחר ω אחר: לכן כדאי לצייר מספר דגמים על אותו הגרף. כדי להמחיש את התפתחות התהליך בזמן, הוסף את האות "" בסוף הפקודה. משך כל המחשה דינמית כזו-20 שניות.

הפקודה oalin תצייר תהליך אקראי לינארי, עבורו השיפוע הוא מ"א מפולג יוניפורמית על $[0,10]$ וחציית הציר היא בנקודה אקראית בת"ס, שגם היא מ"א מפולג יוניפורמית על $[0,10]$. הפקודה oalinP תמחיש את ציר הזמן.

הפקודה oaiid תצייר רעש לבן, כלומר סידרת משתנים אקראיים בת"ס, מפולגים על $[0,1]$.

הפקודה oarsin(n); תצייר פונקצית סינוס עם משרעת (אמפליטודה) מפולגת על $[0,10]$ פאזה מפולגת על $[0,2\pi]$ ותדירות n. (הסימן ";" <u>אינו</u> מיותר!). להמחשה דינמית: oarsinP(n);

7.2 תהליכים בזמן בדיד

דוגמה **7.2** דוגמה זו ממחישה הילוך אקראי (הילוך שיכור), המתחיל ב-0 וצעדיו מפולגים בצורה אחידה על $[0,1]$. ראה דוגמה 3.8. הפקודה oarw תצייר הילוך שיכור (20 צעדים).

7.3 שרשרות מרקוב

דוגמה **7.3** הפקודה democa ממחישה שרשרת מרקוב הומוגנית עם 3 מצבים. השרשרת מתחילה במצב 3. אפשר לקרוא את הקוד כדי להבין כיצב נוצרת השרשרת. הפקודה democap מציגה את התהליך כפונקציה של הזמן (30 צעדים). הפקודה democb מציגה את פונצית הפילוג, כפונקציה של הזמן. המצב ההתחלתי הוא 3. הפקודה democbb זהה, אך הסתברויות המעבר של השרשרת הן שונות.

7.4 גאוס

דוגמה **7.4** המחשת הפילוג הגאוסי: הפקודה demogaus(n) מגרילה n משתנים גאוסיים סטנדרטים (ממוצע 0 ווריאנס 1), בלתי תלויים סטטיסטית, ומשרטטת היסטוגרמה. מספר האפשרויות (bins) הוא בקירוב שורש n.

נספח 2

8 חזרה נוספת על "מבוא להסתברות"

בפרק זה נאסוף הגדרות ותזכורות הקשורות לתורת ההסתברות, וכן לאלגברה לינארית.

8.1 הסתברות

הגדרה 8.1 <u>מרחב המדגם</u> Ω הוא אוסף של "נקודות" $\{\omega\}$.

אינטואיטיבית נחשוב על מרחב המדגם כעל אוסף כל התוצאות האפשריות של נסוי.

דוגמה 8.2 נניח שרוצים לבנות מודל של זריקת קוביה. כוון שיש שש תוצאות אפשריות, נבחר אוסף (שרירותי!) של שישה עצמים. למשל: אוסף המספרים $\{1,2,3,4,5,6\}$. נוכל גם לבחור את האוסף $\{a,b,c,d,e,f\}$.

ככל שרוצים לבנות מודל של תופעה מסובכת יותר, כך יהיה מרחב המדגם מסובך יותר. לפעמים בונים מרחב מדגם גדול יותר מאשר נחוץ: זאת כדי להשיג מרחב עם מבנה פשוט, או לאפשר הרחבות של המודל בשלב מאוחר יותר.

דוגמה 8.3 כדי לבנות מודל של רולטה, נוכל לבחור כמרחב מדגם את אוסף הנקודות על היקף מעגל היחידה. באוסף זה יותר נקודות מאשר יש תוצאות אפשריות ברולטה. ניתן לתאר כל מאורע (תוצאה ברולטה) כקטע על המעגל.

אם רוצים לבנות מודל עבור רעש יציאה ממגבר בקטע הזמן $[0,1]$, מרחב המדגם יכול להיות אוסף כל הפונקציות הרציפות בקטע הזמן $[0,1]$.

הגדרה 8.4 <u>אוסף המאורעות</u> $\mathbb{F}$ הוא אוסף של תת קבוצות של Ω, המקיימות את התנאים הבאים.

1. $\Omega \in \mathbb{F}$, כלומר המקרה שתמיד קורה הוא מאורע, או-תת הקבוצה Ω של Ω היא מאורע.
2. אם A_1 וכן A_2 הם מאורעות, אזי גם $A_1 \cup A_2$ הוא מאורע, כלומר $A_1 \cup A_2 \in \mathbb{F}$.
3. אם A שייך ל-$\mathbb{F}$ אזי גם המשלים שלו A^c שייך ל-$\mathbb{F}$.
4. איחוד ניתן להמנות של מאורעות אף הוא מאורע. בסימון מתמטי, אם $A_i \in \mathbb{F}$ עבור $i = 1,2,\ldots$ אזי גם $\cup_{i=1}^{\infty} A_i \in \mathbb{F}$.

אוסף תת קבוצות המקיים את כל התנאים לעיל נקרא סיגמה שדה (σ-field) או סיגמה אלגברה (σ-algebra: שני שמות לאותו מושג).

הגדרה 8.5 <u>הקבוצה הריקה</u>, כלומר הקבוצה שאינה מכילה דבר, תסומן ב- $\emptyset$. זוג מאורעות A, B נקראים זרים אם אין להם דבר משותף, כלומר אם $A \cap B = \emptyset$.

הגדרה 8.6 פונקצית הסתברות, או הסתברות $\mathbb{P}\{A\}$ היא פונקציה המוגדרת עבור כל מאורע (ב- $\mathbb{F}$), ומקיימת את התנאים הבאים:

1. $0 \le \mathbb{P}\{A\} \le 1$ לכל $A \in \mathbb{F}$.
2. $\mathbb{P}\{\Omega\} = 1$.
3. לכל אוסף (בן מניה) של מאורעות $\{A_i,\ i = 1, 2, \dots\}$ שהם זרים (כלומר $A_i \cap A_j = \emptyset$ אם $i \neq j$), מתקיים

$$\mathbb{P}\left\{\bigcup_{i=1}^{\infty} A_i\right\} = \sum_{i=1}^{\infty} \mathbb{P}\{A_i\}$$

מ- 2 ו- 3 נובע כי לכל מאורע A מתקיים $\mathbb{P}\{A\} + \mathbb{P}\{A^c\} = 1$.

הגדרה 8.7 לשלשה $\{\Omega, \mathbb{F}, \mathbb{P}\}$ קוראים מרחב הסתברות.

8.2 משתנה אקראי ופילוג

דוגמה 8.8 בדוגמאות 8.2–8.3 תארנו מאורעות פשוטים. עבור רעש היציאה מהמגבר, אפשר לשאול מהי ההסתברות שהאנרגיה תהיה לא גדולה מדי: למשל $\int_0^1 n^2(t)\,dt \le 3$ כאשר $n(t) = n(t,\omega)$ הוא תהליך הרעש (את "משתנה המזל" ω נשמיט בדרך כלל, כדי שהסימון יהיה פשוט יותר: אך הוא קיים ומשפיע!). כאן דרושה יותר זהירות מתמטית כדי לוודא שאכן זהו מאורע (כי אם לא, אזי ההסתברות אינה מוגדרת!).

הגדרה 8.9 משתנה אקראי (בקיצור: מ"א) הוא פונקציה $X(\omega)$ על מרחב המדגם Ω כך שעבור כל מספר ממשי a הקבוצה $\{\omega : X(\omega) \le a\}$ היא מאורע.

שים לב שמקובל להשמיט את המשתנה ω, כלומר מקובל לרשום X כאשר מתכוונים ל-$X(\omega)$. בצורה דומה, לפעמים נרשום בקיצור (למשל) $\{X(\omega) \le a\}$, כאשר אנו מתכוונים לתאר את המאורע $\{\omega : X(\omega) \le a\}$.

דוגמה 8.10 נסמן את המשתנה האקראי ב-X. נניח שהרעש ביציאה ממגבר נתון ע"י $n(t) = n(t,\omega)$, ונניח שהרעש רציף (כפונקציה של הזמן). אזי ניתן להגדיר את המ"א הבאים:

- הערך של רעש המגבר ברגע $t = 2$ למשל (מתח חשמלי, נמדד בוולט V), הוא $X(\omega) = n(2,\omega) = n(t,\omega)|_{t=2}$.
- אפשר להגדיר מ"א מורכבים יותר: למשל האנרגיה של אות הרעש בתחום בזמן $[0,1]$ היא $X(\omega) = \int_0^1 n^2(t,\omega)\,dt$.
- הערך המקסימלי של אות הרעש בתחום $[0,1]$ הוא $Y = \max_{0\le t\le 1} n(t,\omega)$.

הערה מתמטית: ברור כי $\{\omega : X(\omega) \le a\}$ היא קבוצה ב-Ω. אך כדי להראות ש-$X(\omega)$ הוא משתנה אקראי, יש לוודא כי זהו מאורע. לשם כך יש להתייחס להגדרה של $\mathbb{F}$. כל זה ניתן להעשות, אך לא במסגרת הנוכחית.

הגדרה 8.11 פונקצית הפילוג של משתנה אקראי $X(\omega)$ היא פונקציה של משתנה ממשי, המקבלת ערכים בין 0 ו-1, והגדרתה

$$F_X(a) = \mathbb{P}\{X(\omega) \le a\}$$

בהגדרה זו, הסימן X מזהה שאנו עוסקים בפונקציית הפילוג של המשתנה האקראי $X(\omega)$. פונקציית הפילוג $F_X(a)$ היא פונקציה של המשתנה a בלבד.

טענה 8.12 תכונות פונקציית הפילוג $F_X(a)$ (ללא הוכחה):

- פונקצית הפילוג היא מונוטונית (לא יורדת), ורציפה מימין,
- (בסימון לא לגמרי מדויק מתמטית) $F_X(+\infty) = 1$ ו-$F_X(-\infty) = 0$. כלומר, משתנה אקראי מקבל רק ערכים סופיים.
- אם $a_1 < a_2$ אזי

$$\mathbb{P}\{a_1 < X \le a_2\} = F_X(a_2) - F_X(a_1)$$

אם X הוא משתנה בדיד, אזי $F_X(a)$ קבועה על פני קטעים, ועולה בקפיצות (צייר את פונקציית הפילוג של קוביה).

הגדרה 8.13 אם קיימת פונקציה (אינטגרבילית רימן), שנסמן ב-$f_X(\theta)$, כך שלכל a

$$F_X(a) = \int_{-\infty}^{a} f_X(\theta)\, d\theta \tag{8.1}$$

אזי f_X תקרא פונקצית הצפיפות של המ"א X, או הפילוג הסגולי של המ"א X.

8.3 ווקטור אקראי

הגדרה 8.14 ווקטור אקראי $\underline{X}$ במימד N הוא אוסף של N משתנים אקראיים:

$$\underline{X}(\omega) = \{X_1(\omega), X_2(\omega), \ldots, X_N(\omega)\}'$$

כמובן ש- ω הוא משותף לכל הרכיבים (אותו פרמטר מזל לכל רכיבי הווקטור $X_i(\omega)$. הפילוג של ווקטור אקראי הוא הרחבה של מושג הפילוג של משתנה אקראי

הגדרה 8.15 פונקצית הפילוג של ווקטור אקראי $\underline{X}(\omega) = \{X_1(\omega), X_2(\omega), \ldots, X_N(\omega)\}'$ היא פונקציה של ווקטור ממשי $\underline{a} = \{a_1, a_2, \ldots, a_N\}'$, המקבלת ערכים בין 0 ו-1, והגדרתה

$$F_{\underline{X}}(\underline{a}) = \mathbb{P}\left\{X_1(\omega) \le a_1, X_2(\omega) \le a_2, \ldots, X_N(\omega) \le a_N\right\} \tag{8.2}$$

לווקטור אקראי יש פונקציית צפיפות (ראה 8.13) אם קיימת פונקציה $f_{\underline{X}}(\underline{a})$ כך שלכל $\underline{a}$

$$F_{\underline{X}}(\underline{a}) = \int_{-\infty}^{a_1} \cdots \int_{-\infty}^{a_N} f_{\underline{X}}(\underline{a})\, da_1 \cdots da_N$$

ניתן לחשב אותה ע"י

$$f_{\underline{X}}(\underline{a}) = \frac{\partial^N F_{\underline{X}}(\underline{a})}{\partial a_1 \partial a_2 \cdots \partial a_N}$$

8.4 אי תלות סטטיסטית

הגדרה 8.16 אי תלות סטטיסטית.
זוג מאורעות A, B נקראים בלתי תלויים סטטיסטית (בת"ס) אם $\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\} \cdot \mathbb{P}\{B\}$.
זוג משתנים אקראיים X, Y נקראים בלתי תלויים סטטיסטית (בת"ס) אם לכל זוג מספרים a, b המאורעות $\{\omega : X \le a\}$ ו-$\{\omega : Y \le b\}$ הם בת"ס. במילים אחרות, המשתנים האקראיים X, Y הם בת"ס אם לכל a, b מתקיים

$$\mathbb{P}\{X \le a, Y \le b\} = \mathbb{P}\{X \le a\} \cdot \mathbb{P}\{Y \le b\}$$

או, בסימון שונה,

$$F_{X,Y}(a, b) = F_X(a) \cdot F_Y(b)$$

בצורה זהה, שני וקטורים $\underline{X}, \underline{Y}$ נקראים בת"ס אם לכל $\underline{a}, \underline{b}$ מתקיים

$$F_{\underline{X},\underline{Y}}(\underline{a}, \underline{b}) = F_{\underline{X}}(\underline{a}) \cdot F_{\underline{Y}}(\underline{b})$$

אם X, Y בת"ס ו-g היא פונקציה (דטרמיניסטית) כלשהיא (למדקדקים-פונקציית בורל) אזי $X, g(Y)$ גם הם בת"ס.

8.5 תוחלת

הגדרה 8.17 את <u>התוחלת</u> של משתנה אקראי X נסמן באחת מהצורות $\mathbb{E}\,X = \overline{X} = m_X$. התוחלת מוגדרת בצורה הבאה.

- יהי X משתנה אקראי המקבל ערכים בדידים, למשל $\{\alpha_i,\ i = 1, 2, \dots\}$. התוחלת מוגדרת בתנאי ש-

$$\sum_{i=1}^{\infty} |\alpha_i|\, \mathbb{P}\{X_i = \alpha_i\} < \infty$$

ואם תנאי זה מתקיים אזי התוחלת היא

$$\mathbb{E}\,X = \sum_{i=1}^{\infty} \alpha_i\, \mathbb{P}\{X_i = \alpha_i\}$$

- נניח שלמשתנה X יש צפיפות סגולית. אזי התוחלת מוגדרת בתנאי ש-

$$\int_{-\infty}^{\infty} |\alpha| f_X(\alpha)\, d\alpha < \infty$$

ואם תנאי זה מתקיים אזי התוחלת היא

$$\mathbb{E}\,X = \int_{-\infty}^{\infty} \alpha f_X(\alpha)\, d\alpha$$

- באופן כללי, מגדירים תוחלת על ידי קרובים. מקרבים את המשתנה על ידי משתנה בדיד, משתמשים בנוסחה של משתנה בדיד, ומשפרים את הקירוב. ביתר פירוט, לכל מספר n נבחר חלוקה (סופית) של הקטע $(-n, n)$ ל-$k(n)$-קטעים. נדרוש מהחלוקה שתקיים

$$-n = \alpha_1^{(n)} < \alpha_2^{(n)} < \cdots < \alpha_i^{(n)} < \alpha_{i+1}^{(n)} < \cdots < \alpha_{k(n)-1}^{(n)} < \alpha_{k(n)}^{(n)} = n$$

$$\max_i \alpha_{i+1}^{(n)} - \alpha_i^{(n)} \to 0$$

כאשר $n \to \infty$ (לשם כך צריך ש-$k(n)$ יגדל מהר יותר מ-n). התוחלת של משתנה אקראי X מוגדרת בתנאי שקיים קבוע B כך ש-

$$\sum_{i=1}^{k(n)} |\alpha_i^{(n)}| \left[F_X(\alpha_{i+1}^{(n)}) - F_X(\alpha_i^{(n)})\right] < B$$

לכל n גדול מספיק. אם תנאי זה מתקיים, אזי התוחלת היא

$$\mathbb{E}[X] = \lim_{n\to\infty} \sum_{i=1}^{k(n)} \alpha_i^{(n)} \left[F_X(\alpha_{i+1}^{(n)}) - F_X(\alpha_i^{(n)})\right] \tag{8.3}$$

בגלל צורת הסכום, מסמנים גבול זה כאינטגרל, באחת מהצורות הבאות:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} \alpha\, dF_X(\alpha) = \int_{-\infty}^{\infty} \alpha\, F_X(d\alpha)$$

אינטגרל כזה נקרא אינטגרל **סטילצ'ס-לבג** Stiltjes-Lebesgue.

תרגיל 8.18 בדוק שההגדרה הכללית של תוחלת נותנת תוצאה זהה להגדרות הקודמות כאשר המשתנה האקראי מקבל ערכים בדידים, וכאשר למשתנה האקראי יש צפיפות.

טענה 8.19 בהנתן משתנה אקראי X ופונקציה g נגדיר משתנה אקראי $Y = g(X)$. אזי

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \int_{-\infty}^{\infty} \alpha\, dF_Y(\alpha) = \int_{-\infty}^{\infty} g(\alpha)\, dF_X(\alpha)$$

כלומר, כדי לחשב את התוחלת של המשתנה Y אין צורך לחשב את הפילוג שלו: אפשר לבצע את החישוב בעזרת פונקצית הפילוג של X.

לא לכל מ"א יש תוחלת סופית, ולא תמיד ניתן להגדיר תוחלת. לדוגמה, עבור מ"א עם פונקציות פילוג סגולי

$$f_X(\alpha) = \begin{cases} 0 & \text{אם } \alpha \geq 0 \\ \dfrac{2/\pi}{1+\alpha^2} & \text{אם } \alpha < 0 \end{cases}$$

$$f_Y(\alpha) = \frac{1/\pi}{1+\alpha^2}$$

מתקיים כי

$$\int_{-\infty}^{\infty} \alpha f_X(\alpha)\, d\alpha = \int_{-\infty}^{0} \alpha f_Y(\alpha)\, d\alpha = -\infty$$

כך שאפשר להגדיר תוחלת עבור המשתנה X, אולם התוחלת היא אין-סופית. לעומת זאת, כיוון ש-

$$\int_{0}^{\infty} \alpha f_Y(\alpha)\, d\alpha = \infty$$

ולכן עבור Y לא ניתן כלל להגדיר תוחלת.

בהינתן מאורע A נסמן ב-A^c את המאורע המשלים (כלומר $\omega \in A^c$ אם ורק אם $\omega \notin A$). נסמן ב-I_A את הפונקציה המציינת של המאורע A. זהו משתנה אקראי המוגדר כך:

$$I_A(\omega) = \begin{cases} 1 & \text{אם } \omega \in A \\ 0 & \text{אם } \omega \in A^c. \end{cases}$$

טענה 8.20 תכונות התוחלת.

1. לכל מאורע A,

$$\mathbb{E}[I_A] = 0 \cdot \mathbb{P}\{A^c\} + 1 \cdot \mathbb{P}\{A\} = \mathbb{P}\{A\}$$

2. אם $X = C$ קבוע שאינו אקראי, אזי $\mathbb{E}\, X = C$.
3. לינאריות: נניח שלמשתנים X, Y יש תוחלת, ויהיו a, b זוג קבועים. אזי

$$\mathbb{E}\left[aX + bY\right] = a \cdot \mathbb{E}[X] + b \cdot \mathbb{E}[Y]$$

בין אם המשתנים תלויים סטטיסטית ובין אם לאו.

4. אם המשתנים X, Y בלתי תלויים סטטיסטית (ויש להם תוחלת) אזי

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

5. אם $X \geq Y$ (או $\mathbb{P}\{X \geq Y\} = 1$) אזי $\mathbb{E}[X] \geq \mathbb{E}[Y]$, כאשר שוויון יתכן רק אם $\mathbb{P}\{X = Y\} = 1$.

טענה 8.21 אם המשתנה X מקבל ערכים שלמים וחיוביים, אזי

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \mathbb{P}\{X \geq k\}$$

הוכחה: לפי הגדרת התוחלת למקרה זה,

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{k=0}^{\infty} k\,\mathbb{P}\{X=k\} \\
&= \sum_{k=1}^{\infty} \mathbb{P}\{X=k\} + \sum_{k=1}^{\infty}(k-1)\,\mathbb{P}\{X=k\} \\
&= \mathbb{P}\{X\geq 1\} + \sum_{k=2}^{\infty}(k-1)\,\mathbb{P}\{X=k\} \\
&= \mathbb{P}\{X\geq 1\} + \sum_{k=2}^{\infty} \mathbb{P}\{X=k\} + \sum_{k=2}^{\infty}(k-2)\,\mathbb{P}\{X=k\} \\
&= \mathbb{P}\{X\geq 1\} + \mathbb{P}\{X\geq 2\} + \sum_{k=3}^{\infty}(k-2)\,\mathbb{P}\{X=k\} \\
&= \mathbb{P}\{X\geq 1\} + \mathbb{P}\{X\geq 2\} + \dots \\
&= \sum_{k=1}^{\infty} \mathbb{P}\{X\geq k\}
\end{aligned}
$$

8.6 מומנטים

לתוחלות של חזקות של משתנה אקראי קוראים מומנטים.

הגדרה 8.22 יהי X משתנה אקראי. נגדיר

- **מומנט שני:** $\mathbb{E}[X^2]$
- **מומנט מסדר n:** $\mathbb{E}[X^n]$.
- **מומנט מוחלט מסדר n:** $\mathbb{E}[|X|^n]$.
- **מומנט מרכזי מסדר n:** $\mathbb{E}\left[\left(X-\overline{X}\right)^n\right]$
- **שונות** (ווריאנס): $\mathrm{Var}(X)=\mathbb{E}\left[\left(X-\overline{X}\right)^2\right]$. **סטיית התקן** $\sigma_X=\sqrt{\mathrm{Var}(X)}$
- מ"א X נקרא **מנורמל** אם $\mathbb{E}[X]=0$ ו-$\mathbb{E}\left[X^2\right]=1$.

לפי ההגדרה, אם X הוא מ"א כל שהוא בעל מומנט שני סופי, אזי המ"א

$$Z=\frac{X-\mathbb{E}[X]}{\sigma_X} \tag{8.4}$$

הוא מ"א מנורמל.

מתוך חוק ההסתברות של מ"א אפשר לחשב את כל המומנטים שלו. מצד שני, התוחלת (מומנט ראשון) היא קירוב דטרמיניסטי סביר למ"א, והמומנט המרכזי השני מתאר מהו הפיזור סביב הקירוב הדטרמיניסטי. מומנטים גבוהים יותר מספקים מידע נוסף על הפילוג.

הערה: באופן כללי, אם מתקיים התנאי הטכני ש-$(\mathbb{E}[|X|^n])^{1/n}$ אינו עולה מהר מדי, אזי המומנטים מגדירים את חוק הפילוג באופן חד משמעי.

יהיו X, Y זוג משתנים אקראיים.

הגדרה 8.23 הקוואריאנס של זוג מ"א מוגדר ע"י

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\overline{X})(Y-\overline{Y})]$$

אם $\mathrm{Cov}(X,Y)=0$ נאמר שהמ"א הם חסרי קורלציה, או בלתי מתואמים, או בלתי תלויים לינארית (להבדיל מבלתי תלויים סטטיסטית!). מקדם הקורלציה, או מקדם המיתאם ρ בין המ"א מוגדר ע"י

$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}$$

אם X, Y בלתי תלויים לינארית אזי $\rho = 0$. הבה נראה כי $|\rho| \le 1$. כיוון שמקדם המתאם אינו תלוי בממוצע, נניח שלשני המשתנים ממוצע 0. לצורך החישוב נתבונן בביטוי החיובי $\mathbb{E}[(X-\lambda Y)^2]$ ונחפש את הערך של המספר λ עבורו הביטוי יהיה מינימלי. כדי למצוא מינימום, נגזור לפי λ ונשווה ל-0. מהגזירה נקבל כי המינימום מושג עבור

$$\lambda^* = \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(Y)}$$

אם נציב ערך זה, נקבל

$$\begin{aligned}
0 &\le \mathbb{E}[(X-\lambda^* Y)^2] \\
&= \mathbb{E}[X^2] - 2\lambda^*\,\mathrm{Cov}(X,Y) + (\lambda^*)^2\,\mathbb{E}[Y^2] \\
&= \mathbb{E}[X^2] - \frac{(\mathrm{Cov}(X,Y))^2}{\mathrm{Var}(Y)}
\end{aligned}$$

ומכאן ש-

$$\mathrm{Var}(X)\,\mathrm{Var}(Y) \ge (\mathrm{Cov}(X,Y))^2$$

או $|\rho| \le 1$ כנדרש.

כסיכום ביניים, נתאר קשר מפורסם בין מומנטים לבין הסתברויות. קשר זה מראה כיצד מומנטים יכולים לסייע לשם קירוב הסתברויות מסויימות.

משפט 8.24 יהי X משתנה אקראי כלשהוא ותהי g פונקציה חיובית ועולה. נדרוש בנוסף כי $\mathbb{E}[g(X)]$ מוגדר היטב וסופי. אזי לכל מספר α,

$$\mathbb{P}\{X \geq \alpha\} \leq \frac{\mathbb{E}[g(X)]}{g(\alpha)} \tag{8.5}$$

<u>הוכחה:</u> נשתמש בהוכחה זו בסימון הכללי (ראה 8.3) לתוחלת כאינטגרל. לטובת האינטואיציה, אפשר לחשוב על הביטוי $dF_X(x)$ כקיצור ל-$f_X(x)\,dx$. נקבע את α ונחשב

$$\begin{aligned}
\mathbb{E}[g(X)] &= \int_{-\infty}^{\infty} g(x)\,dF_X(x) \\
&\geq \int_{-\infty}^{\infty} I_{\{x\geq\alpha\}}(x) g(x)\,dF_X(x) \\
&\geq \int_{-\infty}^{\infty} I_{\{x\geq\alpha\}}(x) g(\alpha)\,dF_X(x)
\end{aligned}$$

כאשר אי השוויון הראשון מתקיים כי g חיובית, והשני כי היא פונקציה עולה. כעת נשתמש בתכונות התוחלת ונקבל

$$\begin{aligned}
\mathbb{E}[g(X)] &\geq g(\alpha)\int_{-\infty}^{\infty} I_{\{x\geq\alpha\}}(x)\,dF_X(x) \\
&= g(\alpha)\,\mathbb{P}\{X \geq \alpha\}
\end{aligned}$$

ולכן, כפי שטעננו,

$$\mathbb{P}\{X \geq \alpha\} \leq \frac{\mathbb{E}[g(X)]}{g(\alpha)}$$

ממשפט זה אפשר להקיש מספר מסקנות מפורסמות. נביא תחילה את משפט מרקוב (Markov).

טענה 8.25 לכל משתנה אקראי חיובי X ולכל $\alpha > 0$ מתקיים

$$\mathbb{P}\{X \geq \alpha\} \leq \frac{\mathbb{E}[X]}{\alpha}$$

טענה זו נובעת מייד מהבחירה $g(x) = x$, כיוון ש- X חיובי. למדקדקים, שימו לב שכיוון ש-X חיובי, אפשר להגדיר לו תוחלת ללא כל תנאי. כמובן שאים התוחלת אין-סופית, הטענה היא חסרת תועלת...

טענה נוספת הנובעת ממשפט זה נקראת חסם צ'בישב (Chebyshev).

טענה 8.26 לכל משתנה אקראי X ולכל α חיובי מתקיים

$$\mathbb{P}\{|X| \geq \alpha\} \leq \frac{\mathbb{E}[X^2]}{\alpha^2}$$
$$\mathbb{P}\{|X - \mathbb{E}[X]| \geq \alpha\} \leq \frac{\mathrm{Var}[X]}{\alpha^2}$$

השורה הראשונה נובעת מיידית מהמשפט, כיוון שלכל α חיובי,

$$\mathbb{P}\{|X| \geq \alpha\} = \mathbb{P}\left\{|X|^2 \geq \alpha^2\right\}$$

והטענה השנייה נובעת מההגדרה של ווריאנס, ע"י הפעלת המשפט על המשתנה האקראי $X - \mathbb{E}[X]$.

חסם חדש יחסית הנובע מאותו משפט הוא חסם צ'רנוב (Chernoff).

טענה 8.27 לכל משתנה אקראי X, לכל α ולכל $\theta \geq 0$ מתקיים

$$\mathbb{P}\{X \geq \alpha\} \leq \mathbb{E}\left[e^{\theta(X-\alpha)}\right]$$

בתנאי שהתוחלת בצד ימין קיימת.

נשים לב שהפונקציה $e^{\theta \cdot x}$ של המשתנה x היא פונקציה חיובית ועולה. לכן אפשר להפעיל את המשפט ולקבל

$$\begin{aligned}\mathbb{P}\{X \geq \alpha\} &\leq \frac{\mathbb{E}\left[e^{(\theta \cdot X)}\right]}{e^{\theta \cdot \alpha}} \\ &= \mathbb{E}\left[e^{\theta(X-\alpha)}\right]\end{aligned}$$

8.7 פונקציה אפיינית

הפונקציה האפיינית ϕ_X של מ"א X היא פונקציה של המשתנה הממשי ν. היא מוגדרת ע"י

$$\phi_X(\nu) = \mathbb{E}\left[e^{i\nu X}\right] = \int_{-\infty}^{\infty} e^{i\nu\alpha}\, dF_X(x).$$

אם למשתנה יש פילוג סגולי (צפיפות) $f_X(\alpha)$

$$\phi_X(\nu) = \int_{-\infty}^{\infty} e^{i\nu\alpha} f_X(\alpha) d\alpha$$

כלומר, $\phi_X(\nu)$ הוא הצמוד המרוכב של התמרת פוריה של $f_X(\alpha)$. אפשר להראות כי בכל מקרה, הפונקציה האפיינית ϕ_X מגדירה חד משמעית את פונקציית הפילוג $F_X(\cdot)$.

שם לב כי הפונקציה האפיינית מוגדרת היטב (ולכן תמיד קיימת), כי

$$\left|e^{i\alpha}\right| = 1$$

לכל α. אם קיימים מומטים מכל סדר, אזי ניתן לייצג את הפונקציה האפיינית על ידי טור חזקות

$$\phi_X(\nu) = \sum_{k=0}^{\infty} \frac{(i\nu)^k}{k!}\,\mathbb{E}\left[X^k\right]$$

כאשר (תחת תנאים טכניים)

$$\phi_X(0) = 1$$

$$\left.\frac{\partial \phi_X(\nu)}{\partial \nu}\right|_{\nu=0} = i\,\mathbb{E}[X]$$

$$\left.\frac{\partial^k \phi_X(\nu)}{\partial \nu^k}\right|_{\nu=0} = (i)^k\,\mathbb{E}\left[X^k\right]$$

כך שבפרט, ידיעת הפונקציה האפיינית מאפשרת חישוב של המומנטים.

8.8 הסתברות ותוחלת מותנים

הגדרה 8.28 ההסתברות של מאורע A בהנתן מאורע B מוגדרת על ידי

$$\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$$

שים לב כי עבור B קבוע, זוהי הסתברות (כפונקציה של A) המרוכזת בקבוצה B, ולכן יש לה את כל התכונות של הסתברות (הגדרה 8.6). הטענה הבאה שימושית מאד, ונובעת ישירות מההגדרה:

טענה 8.29 עבור מאורעות $\{A_k,\ k = 1, 2, \ldots, K\}$ המקיימים $\mathbb{P}\{A_k\} \neq 0$ מתקיים

$$\mathbb{P}\{A_1 \cap A_2\} = \mathbb{P}\{A_1 \mid A_2\} \cdot \mathbb{P}\{A_2\}$$

$$\mathbb{P}\{A_1 \cap A_2 \mid A_3\} = \mathbb{P}\{A_1 \mid A_2 \cap A_3\} \cdot \mathbb{P}\{A_2 \mid A_3\}$$

$$\begin{aligned}\mathbb{P}\left\{\cap_{k=1}^{K} A_k\right\} &= \mathbb{P}\left\{A_1 \mid \cap_{m=2}^{K} A_m\right\} \cdot \mathbb{P}\left\{A_2 \mid \cap_{m=3}^{K} A_m\right\} \times \cdots \\ &\quad \times \mathbb{P}\{A_{K-2} \mid A_{K-1} \cap A_K\} \cdot \mathbb{P}\{A_{K-1} \mid A_K\} \cdot \mathbb{P}\{A_K\}\end{aligned}$$

$$\begin{aligned}\mathbb{P}\left\{\cap_{k=1}^{K-1} A_k \mid A_K\right\} &= \mathbb{P}\left\{A_1 \mid \cap_{m=2}^{K} A_m\right\} \cdot \mathbb{P}\left\{A_2 \mid \cap_{m=3}^{K} A_m\right\} \times \cdots \\ &\quad \times \mathbb{P}\{A_{K-2} \mid A_{K-1} \cap A_K\} \cdot \mathbb{P}\{A_{K-1} \mid A_K\}\end{aligned}$$

נניח כעת כי X, Y הם משתנים אקראיים, ונגדיר עבור β קבוע את המאורע $B = \{\omega : X(\omega) = \beta\}$. אזי ניתן לרשום

$$\mathbb{P}\{A \mid B\} = \mathbb{P}\{A \mid X(\omega) = \beta\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$$

אולם במקרים רבים (למשל אם למשתנה X יש צפיפות), ההיסתברות של המאורע B היא אפס: $\mathbb{P}\{B\} = 0$. כיצד נגדיר אז את ההסתברות המותנית?

הגדרה 8.30 עבור מאורע A ומשתנה אקראי X נגדיר

$$\mathbb{P}\{A \mid X(\omega) = \beta\} = \lim_{\epsilon \to 0} \mathbb{P}\{A \mid \beta \le X(\omega) \le \beta + \epsilon\}$$

אנו נניח שגבול זה תמיד קיים, ומגדיר הסתברות מותנית כך שכל התכונות מתקיימות.

הגדרה 8.31 <u>הפילוג המותנה</u> של משתנה Y בהנתן המשתנה X מוגדר כ-

$$F_{Y|X}(\alpha \mid \beta) = \mathbb{P}\{Y \le \alpha \mid X = \beta\}$$

צד ימין הוגדר למעלה. אם X משתנה עם צפיפות, נגדיר אותו דרך הגבול.

אם X, Y בלתי תלויים סטטיסטית, אזי ההתניה אינה מוסיפה כל מידע.

טענה 8.32 אים המאורעות A, B הם בלתי תלויים אזי

$$\mathbb{P}\{A \mid B\} = \mathbb{P}\{A\} \tag{8.6}$$

לכן, אם המשתנים X, Y הם בלתי תלויים סטטיסטית, אזי

$$F_{Y|X}(\alpha \mid \beta) = F_Y(\alpha)$$

שתי הטענות נובעות מהגדרת הסתברות מותנית: אם A, B בת"ס אזי

$$\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}} = \frac{\mathbb{P}\{A\}\mathbb{P}\{B\}}{\mathbb{P}\{B\}} = \mathbb{P}\{A\}$$

הטענה השניה נובעת מכך ומהגדרת הפילוג המותנה.

הגדרה 8.33 אם קיימת פונקציה f כך ש-

$$F_{Y|X}(\alpha \mid \beta) = \int_{-\infty}^{\alpha} f_{Y|X}(\theta \mid \beta)\, d\theta$$

אזי f תקרא <u>פילוג סגולי מותנה</u> או <u>צפיפות מותנית</u> של Y בהנתן X.

מההגדרה של צפיפות מותנית נובע כי אם יש ל-X ו-Y צפיפות משותפת $f_{X,Y}(\alpha,\beta)$ אזי אפשר לחשב את הצפיפות המותנית כלהלן:

$$\begin{aligned} f_{Y|X}(\alpha \mid \beta) &= \lim_{\delta\to 0} \frac{F_{Y|X}(\alpha+\delta \mid \beta) - F_{Y|X}(\alpha \mid \beta)}{\delta} \\ &= \lim_{\delta\to 0} \frac{\mathbb{P}\{\alpha \le Y \le \alpha+\delta \mid X=\beta\}}{\delta} \\ &= \lim_{\delta\to 0}\left[\lim_{\epsilon\to 0} \frac{\mathbb{P}\{\alpha \le Y \le \alpha+\delta, \beta \le X \le \beta+\epsilon\}}{\delta\cdot\mathbb{P}\{\beta \le X \le \beta+\epsilon\}}\right] \\ &= \frac{\delta\cdot\epsilon\cdot f_{X,Y}(\alpha,\beta)}{\delta\cdot\epsilon\cdot f_X(\beta)} \\ &= \frac{f_{X,Y}(\alpha,\beta)}{f_X(\beta)} \end{aligned}$$

לסיכום, אם יש צפיפות משותפת אזי

$$\boxed{f_{Y|X}(\alpha \mid \beta) = \frac{f_{X,Y}(\alpha,\beta)}{f_X(\beta)}} \tag{8.7}$$

ההיסתברות המותנית $\mathbb{P}\{Y \le \alpha \mid X=\beta\}$ וכן גם הפילוג המותנה $F_{Y|X}(\alpha \mid \beta)$ והצפיפות המותנית $f_{Y|X}(\alpha \mid \beta)$ הם כולם פונקציות של המשתנה β. ניתן כעת להציב במקום המספר β, את המשתנה האקראי X. יתקבל, כמובן, משתנה אקראי חדש (תלוי ב-ω) המוגדר על ידי

$$\mathbb{P}\{Y \le \alpha \mid X\} = \mathbb{P}\{Y \le \alpha \mid X=\beta\}|_{\beta=X}$$

מהגדרה זו נובע שהשוויון

$$F_{Y|X}(\alpha \mid X) = \mathbb{P}\{Y \le \alpha \mid X=\beta\}$$

יתקיים אם ורק אם ω מקיים $X(\omega)=\beta$. בצורה דומה יש להבין את המשמעות של $f_{Y|X}(\alpha \mid X)$.

כאמור, עבור β קבוע הפילוג המותנה הוא בעל כל התכונות של פילוג "רגיל", אפשר להגדיר תוחלת מותנית בדרך שהגדרנו תוחלת.

הגדרה 8.34 התוחלת המותנית של משתנה אקראי Y בהנתן שהמשתנה X מקבל את הערך β מוגדרת על ידי

$$\mathbb{E}[Y \mid X=\beta] = \int_{-\infty}^{\infty} \alpha\, dF_{Y|X}(\alpha \mid \beta) = \int_{-\infty}^{\infty} \alpha F_{Y|X}(d\alpha \mid \beta)$$

אים יש פילוג סגולי מותנה, אזי

$$\mathbb{E}[Y \mid X=\beta] = \int_{-\infty}^{\infty} \alpha f_{Y|X}(\alpha \mid \beta)\, d\alpha$$

עבור פונקציה כלשהיא g מתקיים

$$\mathbb{E}\left[g(Y) \mid X=\beta\right] = \int_{-\infty}^{\infty} g(\alpha) F_{Y|X}(d\alpha \mid \beta) = \int_{-\infty}^{\infty} g(\alpha) f_{Y|X}(\alpha \mid \beta)\, d\alpha$$

כאשר השוויון האחרון מתקיים במידה וקיים פילוג סגולי מותנה.

כמו במקרה של תוחלת רגילה, התוחלת המותנית שהגדרנו היא פונקציה של משתנה ההתנייה β, ולכן כמו במקרה של הסתברות (פילוג או צפיפות) מותנים, אפשר להציב את המשתנה המתנה X במקום β.

$$\mathbb{E}\left[Y \mid X\right] = \mathbb{E}\left[Y \mid X=\beta\right]\big|_{\beta=X}$$

הערה למדקדקים: ההגדרה המדויקת של תוחלת מותנית היא מופשטת יותר (ומדויקת יותר). התוחלת המותנית (כמישתנה אקראי) מוגדרת רק עד כדי מאורע שהסתברותו אפס: לכן כל שוויון שמופיעה בו תוחלת מותנית יהיה נכון פרט לקבוצת ω שהסתברותה אפס. בהמשך לא נתייחס לנקודה זו.

עבור β קבוע התוחלת המותנית מחושבת כמו תוחלת רגילה. לכן היא יורשת את תכונות התוחלת: אולם יש לתוחלת המותנית גם תכונות יחודיות.

טענה 8.35 <u>תכונות התוחלת המותנית</u>. התוחלת המותנית מוגדרת לכל משתנה שיש לו תוחלת סופית. בתאור התכונות לעיל נניח תמיד שמתקיימים התנאים הדרושים כדי שהתוחלת המותנית תהיה קיימת. בדומה לתוחלת הרגילה,

1. לכל מאורע A ולכל משתנה אקראי X,

$$\mathbb{E}[I_A \mid X] = \mathbb{P}\left\{A \mid X\right\}$$

2. אם $Y=C$ קבוע שאינו אקראי, אזי $\mathbb{E}[Y \mid X]=C$.

3. לינאריות: יהיו a,b זוג קבועים. אזי

$$\mathbb{E}\left[aZ+bY \mid X=\beta\right] = a\cdot\mathbb{E}[Z \mid X=\beta] + b\cdot\mathbb{E}[Y \mid X=\beta]$$

$$\mathbb{E}\left[aZ+bY \mid X\right] = a\cdot\mathbb{E}[Z \mid X] + b\cdot\mathbb{E}[Y \mid X]$$

4. אם $Z\geq Y$ (או $\mathbb{P}\left\{Z\geq Y \mid X=\beta\right\}=1$) אזי $\mathbb{E}[Z \mid X=\beta] \geq \mathbb{E}[Y \mid X=\beta]$.

   התכונות הבאות הן מיוחדות לתוחלת המותנית.

5. $\mathbb{E}[X \mid X=\beta]=\beta$ ולכן גם $\mathbb{E}[X \mid X]=X$.

6. עבור פונקציות g,h כלשהן,

$$\mathbb{E}\left[g(X)\cdot h(Y) \mid X=\beta\right] = g(\beta)\,\mathbb{E}\left[h(Y) \mid X=\beta\right]$$

ובאופן כללי יותר, אם h תלוייה בשני המשתנים X,Y

$$\mathbb{E}\left[g(X)\cdot h(X,Y)\mid X=\beta\right]=g(\beta)\,\mathbb{E}\left[h(\beta,Y)\mid X=\beta\right]$$

לכן, לפי ההגדרות, מתקיים בשני המקרים בהתאמה

$$\mathbb{E}\left[g(X)\cdot h(Y)\mid X\right]=g(X)\,\mathbb{E}\left[h(Y)\mid X\right]$$

$$\mathbb{E}\left[g(X)\cdot h(X,Y)\mid X\right]=g(X)\,\mathbb{E}\left[h(X,Y)\mid X\right]$$

.7

$$\mathbb{E}\left[\mathbb{E}(Y\mid X)\right]=\mathbb{E}[Y]$$

ואם המשתנים X,Y בלתי תלויים סטטיסטית אזי

$$\mathbb{E}[Y\mid X]=\mathbb{E}[Y]$$

.8

$$\mathbb{E}\left[g(X)\cdot h(X,Y)\right]=\mathbb{E}\left[g(X)\cdot\mathbb{E}(h(X,Y)\mid X)\right]$$

הערות: תכונה 5 אומרת כי אים ערכו של X כמתנה נקבע לערך מסויים, אז אמנם אפשר להתייחס ל-X כקבוע (גם בתפקידו כמותנה, לא רק כמתנה). תכונות 7 ו-8 נקראות "החלקה".

תכונה 5 נובעת מההגדרות, שכן הפילוג המשותף של X עם עצמו הוא "דלתה". את תכונה 6 נראה בעזרת הפילוג המותנה. לצורך ההוכחה נגדיר משתנה חדש Z שהגדרתו היא $Z=X$. הגדרה זו תאפשר להפריד בין תפקידים שונים של אותו המשתנה X. בסימון החדש, עלינו להראות כי

$$\mathbb{E}\left[g(Z)\cdot h(Z,Y)\mid X=\beta\right]=g(\beta)\,\mathbb{E}\left[h(\beta,Y)\mid X=\beta\right]$$

לפי ההגדרה,

$$\begin{aligned}\mathbb{E}\left[g(Z)\cdot h(Z,Y)\mid X=\beta\right]&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g(z)h(z,y)F_{Y,Z|X}(dz\,dy\mid\beta)\\&=\int_{-\infty}^{\infty}g(\beta)h(\beta,y)F_{Y,Z|X}(dy\mid\beta)\end{aligned}$$

מכיוון מהגדרת הפילוג המותנה המשותף (ראה בפרט את הנוסחה לצפיפות המותנית), הפילוג מתרכז ב-$X=Z$ כך שהצפיפות המותנית היא פונקצית דלתה. בשלב זה המשתנה Z אינו רלוונטי מבחינת פונקציית הפילוג המותנית: הוא

אינו משפיע על הארגומנטים. בנוסף, $g(\beta)$ הוא קבוע מבחינת האינטגרל, ולכן אפשר להוציאו אל מחוץ לאינטגרל. אם כך

$$\begin{aligned}\mathbb{E}\left[g(Z)\cdot h(Z,Y)\mid X=\beta\right]&=g(\beta)\cdot\int_{-\infty}^{\infty}h(\beta,y)F_{Y|X}(dy\mid\beta)\\&=g(\beta)\,\mathbb{E}\left[h(\beta,Y)\mid X=\beta\right]\end{aligned}$$

כאשר השוויון האחרון נובע מהגדרת התוחלת המותנית.

שאר השוויונים ב-6 נובעים מההגדרה. השוויון הראשון ב-7 נובע מתכונות הפילוג המותנה-ראה בפרט את נוסחה (8.7) עבור הצפיפות המותנית. השוויון השני נובע מטענה 8.32. לבסוף, 8 נובע מהטענות הקודמות.