(67656) עיבוד אותות מתקדם | סיכום

סוכם מהרצאותיו של ד״ר אור אורדנטליך 2020-2021

ע"י ברק רמני - שגיאות והערות ניתן לשלוח ל- barak.ramni@mail.huji.ac.il

תאריך עריכה אחרון - 23 בינואר 2022

תוכן העניינים

1 תהליכים מרקוביים

1.1 שלשה מרקובית

תהליך מרקובי הוא תהליך שבו התלות בעבר היא יחודית, תהליך שבו אם נרצה לנבא את העתיד על סמך העבר הדבר היחיד שמעניין הוא הדגימה האחרונה.

שלשה מרקובית יהיו X, Y, Z שלשה של משתנים אקראיים - נאמר שהם מסמנים שלשה מרקובית ונסמן $X \to Y \to Z$ אם מתקיים

$$F_{Z|Y,X}(Z|Y,X) = F_{Z|Y}(Z|Y)$$

כלומר, אם X, Z בת"ס בהינתן Y

נבחין כי כיוון החצים משמעותי ובעצם Y שובר את התלות בין X-ל-Z

במקרה הבדיד נקבל התנהגות דומה $X \to Y \to Z$ שלשה מרקובית אם

$$Pr(Z = z|Y = y, X = x) = Pr(Z = z|Y = y)$$

דוגמה - יהיו $X, N_1, N_2 \overset{iid}{\sim} \mathcal{N}(0, \sigma^2)$

בנוסף נגדיר שלשה חדשה

$$\begin{aligned}
&X \\
&Y = X + N_1 \\
&Z = X + N_1 + N_2
\end{aligned}$$

נטען שהשלשה הנ"ל הינה מרקובית.

התפלגות Z - Z הוא צ"ל של שלוש מ"א גאוסים בת"ס לכן Z גאוסי בעצמו. התוחלת שלו היא 0 כי תוחלת של סכום היא סכום התוחלות, השונות היא $3\sigma^2$ כי הם iid. כלומר

$$Z \sim \mathcal{N}(o, 3\sigma^2)$$

נראה את התפלגות $Z|X = x$

$$\begin{aligned}
[Z|X = x] &= x + [N_1 + N_2|X = x] \\
&= x + N(0, 2\sigma^2) \\
&\sim \mathcal{N}(x, 2\sigma^2)
\end{aligned}$$

מכיוון שהפילוג $[Z|X = x] \neq [Z]$ נקבל כי Z, X תלויים סטטיסטית

כעת נתבונן בהתפלגות $[Z|Y = y]$, מהגדרת Z, Y בעצם קיבלנו ש-$Z = Y + N_2$

$$\begin{aligned}
[Z|Y = y] &= y + [N_2|Y = y] \\
&\overset{*}{=} y + \mathcal{N}(0, \sigma^2) \\
&\sim \mathcal{N}(y, \sigma^2)
\end{aligned}$$

* כי $Y = X + N_1$ ו-N_2 בת"ס ב-X, N_1 $\Leftarrow$ N_2, Y בת"ס ופרט גם X

כעת נתבונן ב- $[Z|Y=y,X=x]$

$$\begin{aligned}[Z|Y=y,X=x] &= y+[N_2|Y=y,X=x]\\ &\sim \mathcal{N}\left(y,\sigma^2\right)\end{aligned}$$

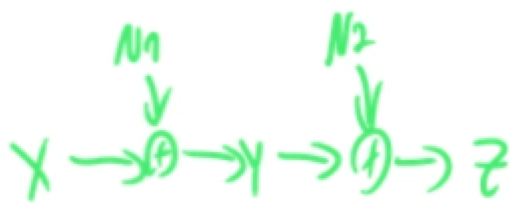
כלומר קיבלנו ש-

$$[Z|Y=y]=[Z|Y=y,X=x]$$

מסקנה - ההתניה ב-X אינה מוסיפה מידע, נשארנו באותו פילוג $\Leftarrow$ בהינתן Y מתקיים ש-X ו-Z בת"ס וזה ההגדרה של שרשרת מרקוב

$$F_{Z|X,Y}\left(Z|X,Y\right)=F_{Z|Y}\left(Z|Y\right)$$
$$\Downarrow$$
$$X\rightarrow Y\rightarrow Z$$

אינטואיציה ציורית



כשמצריים את המבנה מיד ברור כי ברגע שמתנים ב-Y, Z זה בעצם Y ועוד רעש שאינו תלוי ב-Y

טענה אם השלשה $Z\rightarrow Y\rightarrow X$ היא מרקובית בסדר הזה, אזי $X\rightarrow Y\rightarrow Z$ היא שלשה מרקובית בסדר הזה. כלומר,

$$Pr\left(X=x|Y=y,Z=z\right)=Pr\left(X=x|Y=y\right)$$

הוכחה נוכיח עבור מקרים בדידים, ההוכחה עבור המקרה הרציף זהה עד כדי החלפת כל פונ' הסתברות בפונק' צפיפות.

$$\begin{aligned}Pr\left(X=x|Y=y,Z=z\right) &\overset{Bayes}{=} \frac{Pr\left(X=x,Z=z|Y=y\right)}{Pr\left(Z=z|Y=y\right)}\\ &\overset{Bayes}{=} \frac{Pr\left(X=x|Y=y\right)\cdot Pr\left(Z=z|Y=y,X=x\right)}{Pr\left(Z=z|Y=y\right)}\\ &\overset{Markov}{=} \frac{Pr\left(X=x|Y=y\right)\cdot \cancel{Pr\left(Z=z|Y=y\right)}}{\cancel{Pr\left(Z=z|Y=y\right)}}\\ Pr\left(X=x|Y=y,Z=z\right) &= Pr\left(X=x|Y=y\right)\end{aligned}$$

מסקנה אין חשיבות לחצים ומעתה נסמן $X-Y-Z$ אך חשוב מי נמצא באמצע, כלומר המשתנה האמצעי הוא זה שכאשר מתנים בו התלות נשברת.

1.2 תהליך מרקובי

הגדרה תהליך אקראי $\{X(t)\}$ בזמן בדיד או רציף ייקרא מרקובי (מסדר ראשון - בקורס זה נעסוק רק בסדר ראשון) אם לכל $t_1 < t_2$ מתקיים כי

$$X(t_2) - X(t_1) - \{X(t), t < t_1\}$$

הם שלשה מרקובית.

כלומר, אם התהליך הוא מרקובי, כדי לדעת את התנהגות t_2 מספיק רק להבין מה קרה ב-t_1 והעבר לא מעניין.

דוגמאות לתהליכים מסוג זה - תהליך ווינר, הילוך שיכור, תהליך פואסון, תהליכים אוטו-רגרסיבים
לתהליך כללי $\{X_n\}_{n=1}^{\infty}$על $[J]$ לכל $0 \leq n_0 < n_1 < ... < n_l$ מתקיים - פיתוח של כלל השרשת / חוק בייס

$$Pr(X_{n_0} = x_{n_0}, ..., X_{n_L} = x_{n_L}) = Pr(X_{n_0} = x_{n_0}) \cdot \prod_{l=1}^{L} Pr\left(X_{n_l} = x_{n_l} | X_{n_{l-1}} = x_{n_{l-1}}, \ldots, X_{n_0} = x_{n_0}\right)$$

במקרה הכי פשוט

$$Pr(A = a, B = b, C = c) = Pr(A = a) \cdot Pr(B = b | A = a) \cdot Pr(C = c | B = b, A = a)$$

בשרשרת מרקוב ניתן לפשט את הביטוי ל-

$$Pr(X_{n_0} = x_{n_0}, ..., X_{n_L} = x_{n_L}) = Pr(X_{n_0} = x_{n_0}) \cdot \prod_{l=1}^{L} Pr\left(X_{n_l} = x_{n_l} | X_{n_{l-1}} = x_{n_{l-1}}\right)$$

ובפרט

$$\begin{aligned} Pr(X_0 = x_0, X_1 = x_1, ..., X_n = x_n) &= Pr(X_0 = x_0) \cdot Pr(X_1 = x_1 | X_0 = x_0) \cdot \ldots \cdot Pr(X_n = x_n) \\ &= Pr(X_0 = x_0) \cdot \prod_{l=1}^{L} Pr(X_l = x_l | X_{l-1} = x_{l-1}) \end{aligned}$$

נגדיר מטריצה Q המטריצות $Q_n \in [0,1]^{JxJ}$ כך שהאיבר ה-(i,j) שלהן מקיים

$$Q_n(i,j) = Pr(X_n = j | X_{n-1} = i) \quad i,j \in [J]$$

מטריצות אלו מקיימות שהשורה ה-i היא ההסתברות המותנית של המצב בזמן n בהינתן שבזמן $n-1$ היינו במצב i.
ז"א שכל אחת מהשרות במטריצה מהווה וקטור הסתברות - כל איבר בוקטור הוא אי שלילי וסכום השורה שווה לאחד (מטריצה סטוכסטית)

אחת שהגדרנו את המטריצות הללו בעצם קיבלנו ש-

$$Pr(X_0 = x_0, X_1 = x_1, ..., X_n = x_n) = Pr(X_0 = x_0) \cdot \prod_{l=1}^{L} Q_l(x_{l-1}, x_l)$$

כלומר, קיבלנו תיאור סטטיסטי מלא של התהליך והוא דורש רק את הפילוג בזמן 0 - $[Pr(X_0 = i)]_{i=1}^{J}$ ואת מטריצות הסתברויות המעבר $Q_1, Q_2, ...$

1.3 שרשראות מרקוב הומוגניות

מקרה הרבה יותר פשוט ובו נתעסק בקורס, הוא המקרה שבו המטריצות Q אינן תלויות בזמן - כלומר בכל נק' זמן ההחלטה הסטטיסטית של איך לעבור ממצב i למצב הבא מתקבלת באותו אופן - תהליך כזה נקרא שרשרת מרקוב הומוגנית

הגדרה שרשרת מרקוב $\{X_n\}_{n=0}^{\infty}$על $[J]$ תיקרא הומוגנית אם

$$\begin{aligned} Pr\left(X_n = j | X_{n-1} = i\right) &= Pr\left(X_1 = j | X_0 = i\right) \\ &= Q\left(i, j\right) \end{aligned}$$

והתיאור הסטטיסטי המלא של התהליך ניתן ע"י הפילוג ההתחלתי $\{Pr\left(X_0 = 1\right)\}_{i=1}^{J}$ ומטריצת הסתברויות המעבר Q

הגדרה ווקטור π_n יהיה וקטור ההסתברויות של התהליך בזמן n כלומר

$$\pi_n \triangleq \left[Pr\left(X_n = 1\right), Pr\left(X_n = 2\right), \ldots, Pr\left(X_n = J\right)\right]$$

בפרט וקטור ההתפלגויות ההתחלתי יסומן כ-π_0

טענה (בלי הוכחה) עבור שרשרת מרקוב הומוגנית עם פילוג התחלתי π_0 ומטריצת הסתברויות מעבר Q מתקיים

$$\pi_n = \pi_0 \cdot Q^n$$

ואם נרצה לחשב

$$Pr\left(X_n = j | X_{n-k} = i\right) = Q^k\left(i, j\right)$$

מסקנה לשרשרת מרקוב הומוגנית עם π_0 ו-Q לכל זמן - $0 \le n_0 \le n_1 \le \ldots \le n_L$ מתקיים

$$Pr\left(X_{n_0} = x_{n_0}, ..., X_{n_L} = x_{n_L}\right) = \pi_0 \cdot Q^{n_0}\left(x_{n_0}\right) \cdot \prod_{l=1}^{L} \left(Q^{n_l - n_{l-1}}\right)_{\left(x_{n_{l-1}}, x_{n_l}\right)}$$

וזו נוסחה כללית להסתברות התהליך בכל אוסף של זמנים.

1.3.1 דוגמא - השפן, המטבע, העוגה והאופה.

מאפייה מוכרת עוגות גזר.
ביום 0 יש 3 עוגות במלאי.
החל מיום 1, בכל יום נכנס השפן למאפייה ומטיל מטבע הוגן. אם התוצאה היא עץ - הוא קונה עוגה, אם יצא פאלי הוא לא קונה.
אם בסוף היום יש רק עוגה אחת במלאי האופה מכין עוד 2 עוגות אחרי סגירת המאפייה, כך שבתחילת היום הבא יהיו 3 עוגות.
נסמן ב-X_n את מספר העוגות במלאי ביום ה-n

1. הסבר מדוע התהליך $\{X_n\}_{n=0}^{\infty}$ הוא שרשרת מרקבו הומוגנית על המצבים $\{1, 2, 3\}$

צריך להראות שמה שיקרה ב-X_n בהינתן כל מה שקרה בעבר רק תלוי ב-X_{n-1} וזה קורה כי הארנב מטיל מטבע באופן בלתי תלוי בכל ההטלות הקודמות וגם במצב X_{n-1}, החלטת המעבר תלויה רק בתוצאת הטלת המטבע ותוצאת ההטלה לא תלויה בעבר. ולכן תהלך מרקובי

הומוגניות - בכל זמן n להטלת המטבע אותו פילוג. בנוסף הדינמיקה לא תלויה בזמן, כלומר, בכל יום ההחלטה האם לקנות עוגה נקבעת ע"י תוצאת הטלת המטבע. כמו כן פעולת האופה גם אינה משתנה

2. כתבו את π_0 ואת מטרית המעברים Q

π_0 מוגדר - נתון כי ביום 0 יש 3 עוגות במלאי ולכן וקטור הפילוג ההתחלתי הוא -

$$\pi_0 = [0, 0, 1]$$

*הקורדינטה הראשונה - $Pr\left(X_0 = 1\right)$, השנייה - $Pr\left(X_0 = 2\right)$ והשלישית - $Pr\left(X_0 = 3\right)$

לגבי המטריצה Q גודלה $3x3$

$$Q = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{array}{l} \leftarrow Pr\left(X_n = 0 | X_{n-1} = 1\right) \\ \leftarrow Pr\left(X_n = 0 | X_{n-1} = 2\right) \\ \leftarrow Pr\left(X_n = 0 | X_{n-1} = 3\right) \end{array}$$

3. חשבו את

(א) $Pr\left(X_3 = 2, X_5 = 3, X_6 = 1, X_{100} = 2\right)$

(ב) $Pr\left(X_2 = 1, X_4 = 3, X_5 = 2\right)$

נשתמש בנוסחה שפיתחנו -

$$Pr\left(X_{n_0} = x_{n_0}, ..., X_{n_L} = x_{n_L}\right) = \pi_0 \cdot Q^{n_0}\left(x_{n_0}\right) \cdot \prod_{l=1}^{L} \left(Q^{n_l - n_{l-1}}\right)_{\left(x_{n_{l-1}}, x_{n_l}\right)}$$

(א) נבחין כי במקרה א' בהתחלה נצטרך לחשב את Q^3 כי מתחילים מזמן X_3, לאחר מכן יש הפרש של 2 בין X_3 ל-X_5 ולכן נחשב את Q^2 ובמעבר האחרון נצטרך לחשב את Q^{94} - לא סביר ולכן נחפש התחכמות.

נמצא את הקאץ'! יש לנו מעבר ממצב 3 למצב 1 בין היום ה-5 ליום ה-6 וזה לא אפשר כי כל יום השפן יכול לקנות מקסימום עוגה אחת. ולכן $Pr\left(\ldots\right) = 0$

(ב) כאן נצטרך לחשב במפורש.

$$Pr\left(\ldots\right) = \pi_0 \cdot Q^2\left(1\right) \cdot Q^2\left(1, 3\right) \cdot Q\left(3, 2\right)$$

נחשב את Q^2 ואת $\pi_0 \cdot Q^2$

$$Q^2 = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

$$\begin{aligned} \pi_0 \cdot Q^2 &= [0, 0, 1] \cdot \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix} \\ &= \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] \end{aligned}$$

ולסיכום

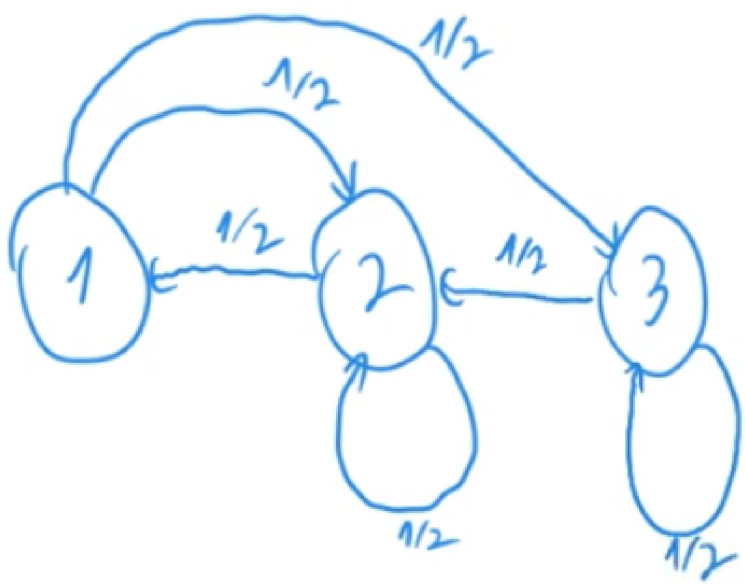
$$\begin{aligned} Pr\left(\ldots\right) &= \frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{2} \\ &= \frac{1}{32} \end{aligned}$$

1.4 ייצוג גרפי של שרשראות מרקוב וסיווג מצבים

נרצה לייצג את מטריצה Q בצורה גרפית ונגדיר תכונות של צמתים בגרף וכן נחלק את הגרף למחלקות.
הגרף תלוי אך ורק במטריצת Q ולא בפילוג ההתחלתי.
לשם הדוגמא נשתמש במטריצת המעברים מהדוגמא של השפן

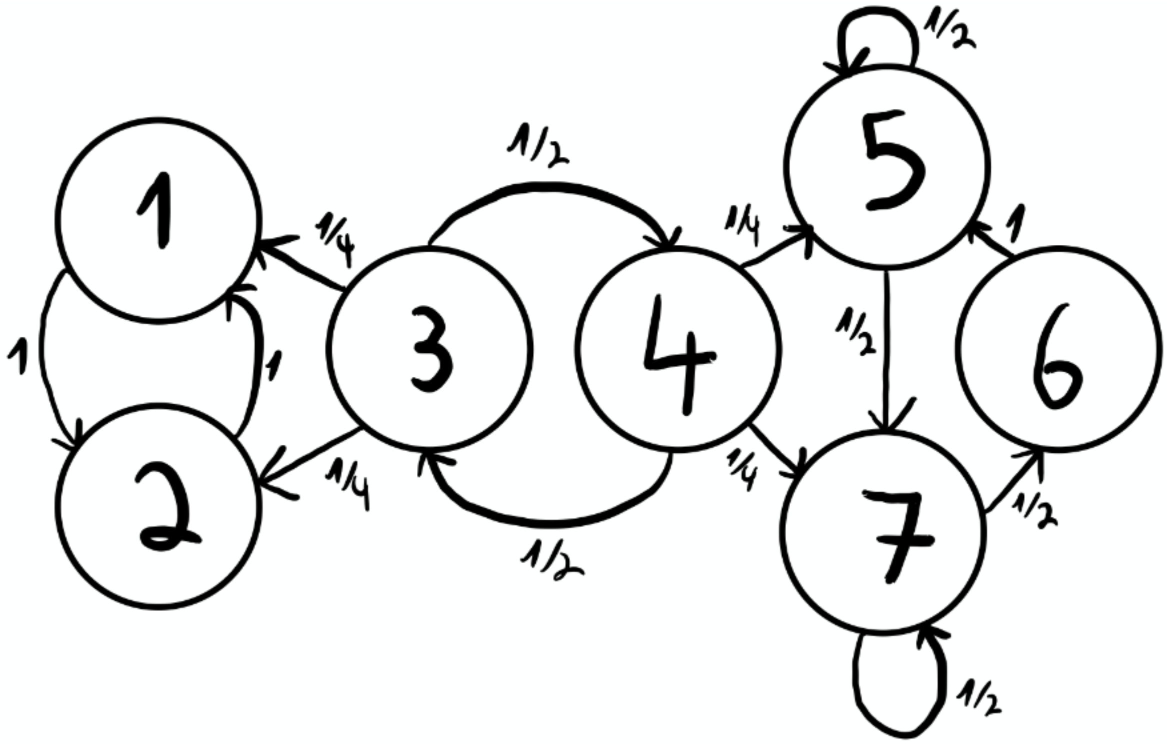
$$Q = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

ראשית נצייר עיגול לכל אחד מהמצבים, במטריצה הזו יש שלושה מצבים ולכן שלושה עיגולים
לאחר מכן נצייר חצים המייצגים את המעבר ממצב למצב - ונקבל



גרף מכוון וממושקל עם שלושה צמתים, כל משקולת מתארת את ההסתברות לעבור ממצב למצב.

1.4.1 תכונות והגדרות לגרפים



נגישות נאמר שמצב j נגיש ממצב i ונסמן $j \to i$ אם קיים מסלול מ-i ל-j, לדוגמא $1 \to 3$ אבל $3 \not\to 1$ כי מ-1 לא ניתן להגיע ל-3.

קשירות נאמר שמצב j ו-i קשירים ונסמן $j \longleftrightarrow i$ אם $j \to i$ וכן $i \to j$ כלומר, הקשירות היא דו-כיוונית. לדוגמא $2 \longleftrightarrow 1$ וכן $1 \not\longleftrightarrow 3$

נבחין כי מתקיים טרנזטיביות, כלומר אם $r \longleftrightarrow i$ וכן $j \longleftrightarrow r$ אזי $j \longleftrightarrow i$

מחלקה אוסף מקסימלי של מצבים קשירים, כלומר, כל המצבים במחלקה קשירים זה לזה ולא קשירים לאף מצב מחוץ למחלקה.

לדוגמא - $\{5,6,7\}$, $\{3,4\}$, $\{1,2\}$ מהווים מחלקה

מצב נישנה מצב יקרא נישנה אם הוא נגיש מכל אחד מהמצבים שנגישים ממנו - מצב i נישנה אם $j \longleftrightarrow i \Rightarrow j \to i$.

ובצורה פשוטה, אם ניתן להגיע ממצב 1 למצב 2 אז ניתן לחזור מ-2 ל-1 ובדיאגרמה $\{1,2,5,6,7\}$ הם מצבים נישנים. $\{3,4\}$ הם לא מצבים נישנים.

מצב חולף מצב יקרא חולף אם הוא לא נישנה - $\{3,4\}$

מחלקה נשנית מחלקה עם מצב נשנה $\{5,6,7\}$, $\{1,2\}$ היא מחלקה נישנת כי כל אחד מהמצבים בה נשנים

מחלקה חולפת היא מחלקה שאינה נשנת - או מחלקה עם מצב חולף. בדוגמא שלנו $\{3,4\}$ היא מחלקה חולפת.

זמני חזרה של מצב i אוסף כל הזמנים t כך ש- $Pr\left(X_t = i | X_0 = i\right) > 0$ בגרף שלנו החצים מסמנים מעברים בהסתברות גדולה מאפס, ולכן זמני החזרה של המצב הם אוסף המסלולים שמתחילים ב-i ומסתיימים ב-i

לדוגמא - במצב 5 כיוון שיש חץ מ-5 לעצמו כל אורך של מסלול אפשרי $\{1,2,3,....\} : 5$.

מצב 6 - המסלול הכי קצר הוא באורך 3, אך אפשר לקבל מסלול בכל אורך שנרצה $6:\{3,4,....\}$

מצב 1 - המסלול הכי קצר באורך 2, אבל בקפיצות של 2 ולכן $1:\{2,4,6,...\}$

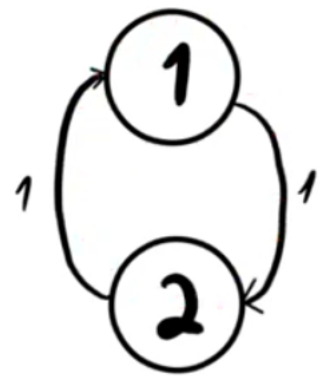
מצב 2 - דומה למצב 1.

מחזור של מצב ה-GCD של זמני החזרה של המצב - המחלק הגדול המשותף של המסלול.

לכל המצבים במחלקה נשנית יש את אותו מחזור ולכן ניתן לקרוא למחזור הנל **המחזור של המחלקה**

מחלקה שזמן המחזור שלה הוא 1 תיקרא **א-פריודית** ומחלקה שזמן המחזור שלה גדול מ-1 תקרא **פריודית**

1.5 התנהגות אסיפטוטית של שרשראות מרקוב



דוגמא 1 זו שרשרת עם 2 מצבים. המצבים הללו מהווים **מחלקה נשנית** אחת והמחלקה היא **מחזורית** - $1,2:\{2,4,6...\}$ עם מחזור של **2**.

נבחין כי לכל זמן זוגי מתקיים ש-$X_n = X_0$ כי בכל זמן זוגי נחזור לנק ההתחלה, ובזמנים האי-זוגיים נקבל $X_n \neq X_0$

אם $X_0 \sim \pi_0 = (p, 1-p)$ נקבל ש-

$$\pi_{2n} = (p, 1-p) = \pi_0$$
$$\pi_{2n+1} = (1-p, p)$$

המקרה היחיד בו הפילוג בין הזמנים הזוגים והאי-זוגיים הוא זהה יקרה כאשר $p = \frac{1}{2}$, כלומר, $\pi_0 = \left(\frac{1}{2}, \frac{1}{2}\right)$ ואז נקבל ש- $\pi_n = \left(\frac{1}{2}, \frac{1}{2}\right) \ \forall n$

דוגמא 2 גם כאן יש 2 מצבים אך נוסף לנו חץ מאחד לעצמו ולכן יש לנו מחלקה 1 **נשנית א-פריודית.**

איך מתנהג π_n?

$$\begin{aligned}\pi_n\left(1\right) &= \frac{1}{2}\cdot\pi_{n-1}\left(1\right)+1\cdot\pi_{n-1}\left(2\right)\\ &= \frac{1}{2}\cdot\pi_{n-1}\left(1\right)+\left(1-\pi_{n-1}\left(1\right)\right)\\ &= 1-\frac{1}{2}\pi_{n-1}\left(1\right)\end{aligned}$$

קיבלנו משוואת הפרשים נסמן $a_n=\pi_n\left(1\right)$

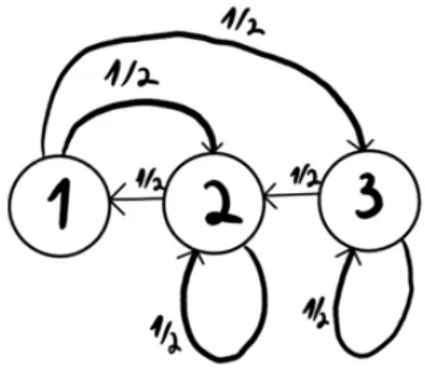
$$a_n=1-\frac{1}{2}a_{n-1}$$

והפתרון שלה הוא (אין לי מושג איך הוא פתר את זה)

$$\begin{aligned}a_n &= \pi_n\left(1\right)\\ &= \frac{2}{3}+\left(-1\right)^n\cdot\frac{\pi_0\left(1\right)-\frac{2}{3}}{2^n}\end{aligned}$$

למעשה קיבלנו ביטוי שמקשר בין π_0 - הפילוג ההתחלתי, לפילוג שלה בכל זמן. נוסחה לפילוג השולי של השרשרת בזמן n ורואים שהתלות בתנאי ההתחלה היא אך ורק דרך האיבר הימני שכופל אותו $\frac{1}{2^n}$ ולכן ב-∞ ההסתברות שואפת ל-$\frac{2}{3}$ ו-$\pi_n\to\left[\frac{2}{3},\frac{1}{3}\right]$

נבחין שאם היינו מאתחילם את השרשת אם הפילוג $\pi_0=\left[\frac{2}{3},\frac{1}{3}\right]$ אז נקבל $\forall n$ $\pi_n=\left[\frac{2}{3},\frac{1}{3}\right]$



דוגמא 3 השרשרת שמתארת את הבעיה עם השפן והמאפיה, בשרשת זו יש מחלקה 1 **נשנית, א-פריודית.**

ראינו ש-

$$Q^2 = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

למעשה ניתן להראות באינדוקציה $\forall n > 2$ מתקיים כי

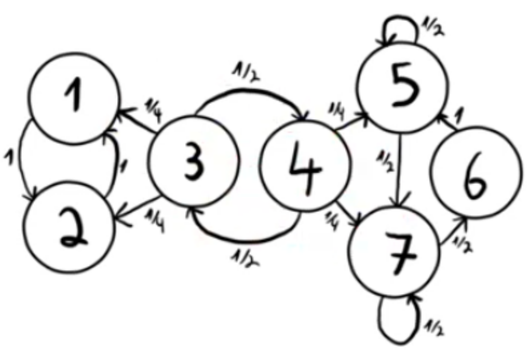
$$Q^n = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{bmatrix}$$

ובפרט, לכל π_0 מתקיים ש-

$$\begin{aligned} \pi_n &= \pi_0 \cdot Q^n \ \forall n \geq 2 \\ &= \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] \end{aligned}$$

ואם נאתחל אץ השרשת עם $\pi_0 = \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right]$ נקבל

$$\pi_n = \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] \ \forall n \geq 0$$



דוגמא 4 בשרשת זו יש 2 מחלקות **נשנות** מחלקה - $\{1,2\}$ שהיא מחלקה **פריודית** ומחלקה $\{5,6,7\}$ - שהיא מחלקה **א-פריודית**. ויש לנו מחלקה **חולפת** שהיא $\{3,4\}$

אם X_0 נמצא באחת המחלקות הנשנות התהליך יישאר שם לנצח.

למשל אם $\pi_0 = [*,*,0,0,0,0,0]$ כלומר התחלנו בהסתברות כלשהי במצב 1 או 2 ושאר המצבים בהסתברות 0, בהכרח $\pi_n = [*,*,0,0,0,0,0]$.

אם $\pi_0 = [0,0,0,0,*,*,*]$ כלומר התחלנו במחלקה של $\{5,6,7,\}$ נקבל ש-$\pi_n = [0,0,0,0,*,*,*]$

אם מתחילים במחלקה חולפת לא נישאר שם, כלומר, אם $\pi_0 = [0,0,*,*,0,0,0]$ בטוח אחרי מספיק זמן נגיע ל-$\pi_n = [*,*,0,0,\dot{*},\dot{*},\dot{*}]$

סכום ההסתברויות של * היא ההסתברות שהתחלנו ב-$\{3,4\}$ וסיימנו ב-$\{1,2\}$ וסכום ההסתברויות של $\dot{*}$ היא ההסתברות שהתחלנו ב-$\{3,4\}$ וסיימנו ב-$\{5,6,7\}$כנל לגבי $\dot{*}$.

שחכת עבר נאמר ששרשרת מרקוב שוכחת את העבר אם

$$\lim_{n\to\infty} \pi_n = \lim_{n\to\infty} \pi_0 \cdot Q^n$$

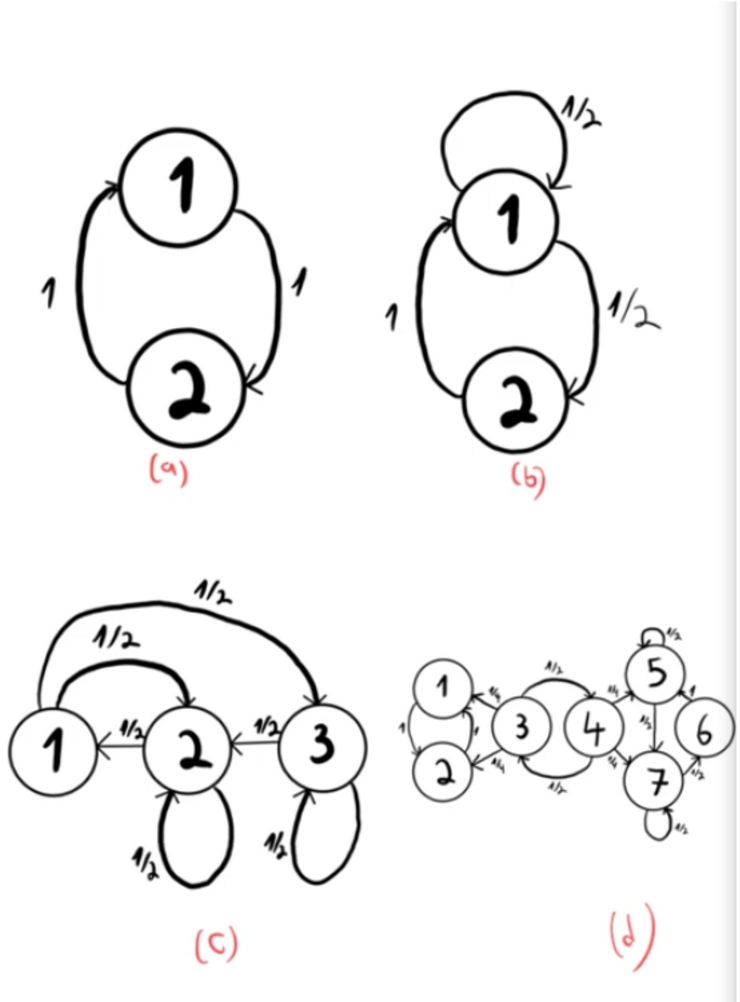
הגבול קיים ובלתי תלוי ב-π_0 ובמקרה הזה נסמן את הגבול ב-π_∞

$$\pi_\infty = \lim_{n\to\infty} \pi_n$$

נבחין כי שכחת העבר שקולה לתכונה ש-

$$\lim_{n\to\infty} Q^n = \begin{bmatrix} \pi_\infty \\ \pi_\infty \\ \vdots \\ \pi_\infty \end{bmatrix}$$

נתבונן שוב בדוגמאות מההרצאות הקודמות -



שרשרת a - לא שוכחת את העבר. π_n מחזורי ויש לו ערכים שונים בזמנים זוגיים/אי-זוגיים כתלות ב-π_0

שרשרת b - עבור n מספיק גדול הסיכוי להיות במצב 1 מתקרב ל-$\frac{2}{3}$ ולהיות במצב 2 מתקרב ל-$\frac{1}{3}$ בלי תלות בתנאי ההתחלה ולכן השרשרת כן שוכחת את העב

שרשרת c- לכל $n \geq 2$ המטריצה Q^n קבועה ופרט יש לה גבול והוא מתבטא בכך שכל השורות זהות ולכן השרשת שוכחת את העבר

שרשרת d - חשוב להבחין איפה מתחילים, אם מתחילים במחלקה-$\{1,2\}$ או $\{5,6,7\}$ נשאר שם לנצח כלומר, לא ניתן לדבר על שחכת העבר ושרשרת זו לא שוכחת את העבר.

פילוג סטציונרי וקטור הסתברות - וקטור שכל האיברים אי-שליליים ומסתכמים ל-1 π^{stat} ייקרא פילוג סטציונרי של שרשרת מרקוב עם מטריצת מעברים Q אמ"מ

$$\pi^{\text{stat}} = \pi^{\text{stat}} \cdot Q$$

כלומר אם נכפול את וקטור השורה π^{stat} במטריצה Q נקבל את אותו וקטור.

כמה הערות לפילוג סטציונרי:

1. לפי הגדרה -

$$\begin{aligned}\pi^{\text{stat}}\cdot Q^{n} &= \pi^{\text{stat}}\cdot\left(Q\cdot Q^{n-1}\right)\\ &= \left(\pi^{\text{stat}}\cdot Q\right)\cdot Q^{n-1}\\ &= \pi^{\text{stat}}\cdot Q^{n-1}\\ &\vdots\\ &= \pi^{\text{stat}}\end{aligned}$$

והמסקנה היא שאם $\pi_0=\pi^{\text{stat}}$ אז $\pi_n=\pi^{\text{stat}}\ \forall n\geq 0$

2. סטציונריות של התהליך - ראינו נוסחה כללית של ההסתברות שתהליך מרקובי הומוגני יקבל ערכים מסויימים בכל אוסף של זמנים - פילוג סטטיסטי מלא של התהליך

$$Pr\left(X_{n_0}=x_{n_0},...,X_{n_L}=x_{n_L}\right)=\pi_0\cdot Q^{n_0}\left(x_{n_0}\right)\cdot\prod_{l=1}^{L}\left(Q^{n_l-n_{l-1}}\right)_{\left(x_{n_{l-1}},x_{n_l}\right)}$$

נבחין כי אם $\pi_0=\pi^{\text{stat}}$ אז

$$Pr\left(X_{n_0}=x_{n_0},...,X_{n_L}=x_{n_L}\right)=\underbrace{\pi_0\cdot Q^{n_0}}_{\pi^{\text{stat}}\left(x_{n_0}\right)}\left(x_{n_0}\right)\cdot\prod_{l=1}^{L}\left(Q^{n_l-n_{l-1}}\right)_{\left(x_{n_{l-1}},x_{n_l}\right)}$$

ולמעשה אין תלות ב-x_{n_0} וכל שאר האיברים לא תלויים בזמנים, אלא רק בהפרשים. ולכן אם מתחילים את השרשת עם הפילוג הסטציונרי נקבל שההסתברות שהשרשרת "תבקר" במצבים מסויימים באוסף של זמנים לא תלויה בזמנים, אלא, אך ורק בהפרשי הזמנים - **תהליך סטציונרי במובן הצר** מהקורס שיטות סטטיסטיות.
ולכן אם $\pi_0=\pi^{\text{stat}}$ אזי התהליך $\{X_n\}_{n=0}^{\infty}$ הוא stationary strict-sense

3. וקטור ההסתברות π הוא סטצוינארי אמ"מ הוא מקיים

$$Q^T\pi^T=\lambda\cdot\pi^T$$

עם $\lambda=1$

1.6 משפט פרון-פרובינוס - Perron-Frobenius

לא נוכיח את המשפט אך ניתן את הטענות שלו

1. **תמיד קיים** פילוג סטציונרי, כלומר, תמיד קיים וקטור הסתברות π^{stat} שמקיים

$$\pi^{\text{stat}}=\pi^{\text{stat}}\cdot Q$$

2. אם השרשרת שמאופיינת ע" מטריצת המעברים Q יש מחלקה נשנית **אחת ויחידה** אזי π^{stat} הוא יחיד.

3. אם השרשרת שמאופיינת ע" מטריצת המעברים Q יש r מחלקות נשנות אזי קיימים r וקטורי הסתברות סטציונארים ונסמן אותם $\pi_1^{stat},\ldots,\pi_r^{stat}$ הוקטורים הללו בלתי תלויים לינארית ולמעשה אפילו אורתוגונלים! כל אחד מהווקטורים האלו π_i^{stat} מתאים למחלקה נשנית אחרת "ונתמך" על איבריה.

הכוונה בנתמך היא שכל האיברים של הוקטור הזה שלא מתאים למצבים במלחקה שווים ל-0, לדוגמא מחלקה שכוללת את המצבים $\{1,2\}$ תראה כך - $[*,*,0,...,0]$

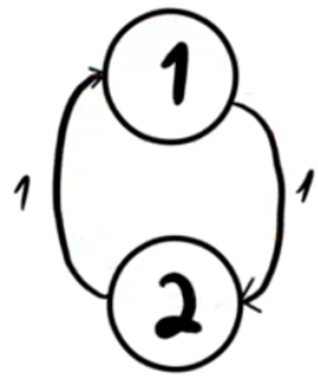
בנוסף לכל המספרים האי-שליליים $\alpha_1,\ldots,\alpha_r$ שמסתכמים ל-1, כלומר, $\sum_{i=1}^{r}\alpha_i=1$ גם $\pi=\sum_{i=1}^{r}\alpha_i\cdot\pi_i^{stat}$ סטציונארי.

4. אמ"מ ל-Q יש **מחלקה נשנית אחת ויחידה** ומחלקה זו **א-פריודית** אזי השרשרת **שוכחת את העבר** ומתקיים $\pi^{\text{stat}} = \pi_\infty$ במקרה הזה $\{X_n\}_{n=0}^{\infty}$ הוא אסימפטוטית סטציונרי לכל π_0

כל תהליך מרקוב שמאותחל עם פילוג סטציונארי אז הוא סטציונרי במובן הצר, כלומר, הפילוג בכל אוסף הזמנים תלוי רק בהפרשי הזמנים.

טענה 4 של משפט פרון-פרובינוס קובעת שאם השרשרת עומדת בתנאים הנ"ל אם נאתחל את השרשת בפילוג π_0 שהוא לא פילוג סטיצונרי אז השרשרת לא תהיה סטצינראית במובן הצר **אך אם נסתכל אחרי מספיק זמן** הפילוג של השרשרת בזמן n יהיה קרוב מאוד ל-π_∞ והפילוג של אוסף הזמנים יהיה תלוי כמעט לחלוטין בהפרשי הזמנים. (הספקנו לשכוח כמעט לחלוטין את π_0)

1.6.1 דוגמאות לשימוש במשפט Perron-Frobenius



1. לשרשרת זו מחלקה נשנית 1 עם מחזור 2

- מטענה 2 של המשפט - קיים פילוג סטציונרי יחיד
- טענה 4 לא מתקיימת כי המחלקה לא א-פריודית. ולכן $\pi_0 = (p, 1-p)$ ו-$\pi_n = (p, 1-p)$, $\pi_{2n+1} = (1-p, p)$ ו-$\pi^{\text{stat}} = \left(\frac{1}{2}, \frac{1}{2}\right)$ והשרשת לא שוכחת את העבר לא משנה אם איזה פילוג נכנסו (גם אם נכנסו עם $\pi_0 \left(\frac{1}{2}, \frac{1}{2}\right)$

לשרשרת זו מחלקה נשנית 1 עם מחזור א-פריודית

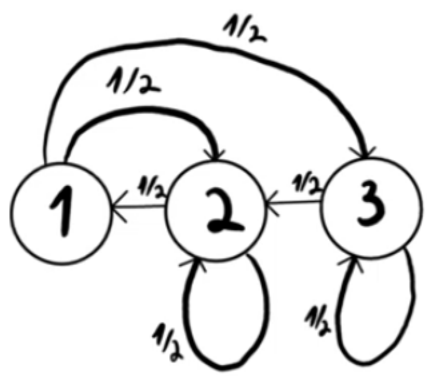
2.
- מטענה 2 של המשפט - קיים פילוג סטציונרי יחיד
- מטענה 4 של המשפט - השרשרת שוכחת את העבר ו- $\pi_\infty = \pi^{\text{stat}}$

ראינו שלכל π_0 מתקיים ש- $\pi_n \xrightarrow{n\to\infty} \left[\frac{2}{3}, \frac{1}{3}\right] = \pi_\infty$

נוודא שאכן $\pi_\infty = \left(\frac{2}{3}, \frac{1}{3}\right)$ הוא סטציונרי

$$
\begin{aligned}
\pi_\infty \cdot Q &= \left[\frac{2}{3}, \frac{1}{3}\right] \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \\
&= \left[\frac{2}{3}, \frac{1}{3}\right] \\
&= \pi_\infty
\end{aligned}
$$

לא בדקנו שזה הפילוג הסטציונרי היחיד - אבל נאמין לפרון שהמשפט שלו עובד

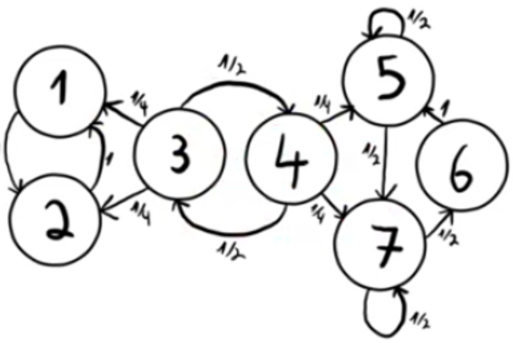


3. לשרשרת זו מחלקה נשנית 1 עם מחזור א-פריודית
- מטענה 2 של המשפט - קיים פילוג סטציונרי יחיד
- מטענה 4 של המשפט - השרשרת שוכחת את העבר ו- $\pi_\infty = \pi^{\text{stat}}$

ראינו שלכל $n \geq 2$ ולכל π_0 מתקיים ש- $\pi_n = \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right]$ ובפרט $\pi_n \xrightarrow{n\to\infty} \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] = \pi_\infty$

נוודא שאכן $\pi_\infty = \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right]$ הוא סטציונרי

$$
\begin{aligned}
\pi_\infty \cdot Q &= \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] \cdot \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\
&= \left[\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right] \\
&= \pi_\infty
\end{aligned}
$$



4. לשרשרת זו 2 מחלקה נשנית אחת עם מחזור א-פריודית ואחת עם מחזור פריודי

- מטענה 3 של המשפט - יש שני וקטורים סטציונרים בת"ל, שכל אחד מהם נתמך על מחלקה נשנית אחרת.
- מטענה 4 של המשפט - השרשרת **לא** שוכחת את העבר

נדגים את טענה 4 - אם נתחיל ממחלקה $\{1,2\}$ נשאר במחלקה זו לנצח, כלומר, האתחול של השרשת משפיע על ההתנהגות עד זמן ∞ ולכן היא לא שוכחת את העבר.

החלק ה-3 מבטיח שיש 2 וקטורים סטציונרים שיראו כך

$$\pi_1^{stat} = [*, *, 0, 0, 0, 0, 0]$$
$$\pi_2^{stat} = [0, 0, 0, 0, *, *, *]$$

נמצא את הוקטורים הנ"ל -נתחיל מ-π_1^{stat} כאשר - $\pi_1^{'}$ הוא וקטור עם 2 מצבים.

$$\left[\pi_1^{'}, 0, 0, 0, 0, 0\right] \cdot Q = \left[\pi_1^{'}, 0, 0, 0, 0, 0\right]$$
$$\pi_1^{'} \cdot Q^{(1)} = \pi_1^{'} \qquad Q^{(1)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$
$$\Downarrow$$
$$\pi_1^{'} = \left[\frac{1}{2}, \frac{1}{2}\right]$$

$Q^{(1)}$ זה מטריצת המעברים עבור המחלקה $\{1,2\}$

ולכן נקבל כי $\pi_1^{stat} = \left[\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0\right]$

כעת נחשב את הוקטור π_2^{stat} נחפש וקטור מהצורה - $\pi_2^{stat} = [0, 0, 0, 0, *, *, *]$

$$\pi_2^{'} \cdot Q^{(2)} = \pi_2^{'} \qquad Q^{(2)} = \begin{array}{cc} \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 1 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix} & \begin{array}{c} \leftarrow 5 \\ \leftarrow 6 \\ \leftarrow 7 \end{array} \end{array}$$
$$\pi_2^{'} \cdot Q^{(2)} = \pi_2^{'} \cdot I$$
$$\Downarrow$$
$$\pi_2^{'} \cdot \left(Q^{(2)} - I\right) = [0, 0, 0]$$

כעת נסמן $\pi_2^{'} = [a,b,c]$

$$[a,b,c] \cdot \begin{bmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ 1 & -1 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{bmatrix} = [0,0,0]$$

שלוש משוואות בשלושה נעלמים

$$a = c$$
$$a = b$$
$$b = c$$

כלומר - $a = b = c$ **אבל!** הוקטור $[a,b,c]$ הוא וקטור הסתברות ולכן הסכום $a+b+c=1$

משמע - $a = b = c = \frac{1}{3}$

ולסיכום –

$$\pi_1^{stat} = \left[\frac{1}{2}, \frac{1}{2}, 0, 0, 0, 0, 0\right]$$
$$\pi_2^{stat} = \left[0, 0, 0, 0, \frac{2}{5}, \frac{1}{5}, \frac{2}{5}\right]$$

ניזכר שהטענה ה-3 של משפט פרון-פרובינוס קובע שאם ניקח

$$\pi = \alpha \pi_1^{stat} + (1-\alpha)\,\pi_2^{stat}$$

נקבל של-π פילוג סטציונרי לכל $0 \le \alpha \le 1$ - מלינאריות.

והוקטור π(שהוא סטציונרי) יראה כך

$$\pi = \left[\frac{\alpha}{2}, \frac{\alpha}{2}, 0, 0, \frac{2(1-\alpha)}{5}, \frac{1-\alpha}{5}, \frac{2(1-\alpha)}{5}\right]$$

1.7 תרגול 1 - שרשראות מרקוב

1.7.1 [שאלת רובוט]

Question 2 *A robot company manufactures two models for house cleaning robots:*

- *The robot* ***HUJI1*** *cleans a room for two hours and then passes to clean another room*
- *The robot* ***HUJI2*** *cleans a room for one hour, and then moves to clean another room for the next hour, with probability p, or stays to clean the same room for another hour, with probability $1-p$. The random decisions this robot makes every hour are statistically independent.*

The robot works in an apartment with two rooms - room 1 and room 2. Its initial state $X_0 \in \{1,2\}$ is drawn uniformly (with probability $1/2$ in the first room and with probability $1/2$ in the second room). Let X_n be the location of the robot at the ***beginning*** *of the $(n+1)$th hour.*

1. *For the model HUJI1, is $\{X_n\}_{n=0}^{\infty}$ a homogeneous Markov chain? Is $\{X_{2n}\}_{n=0}^{\infty}$?*
2. *For the model HUJI2, is $\{X_n\}_{n=0}^{\infty}$ a homogeneous Markov chain? Is $\{X_{2n}\}_{n=0}^{\infty}$?*

A room is considered clean if the robot has worked there for at least one hour.

3. *For each one of the models, find the probability that the house is clean after 4 hours.*

A detective that knows the model of the robot enters the house after 2.5 hours and wants to find out in which of the two rooms the robot has started. Denote by $X_0 \in \{1,2\}$ the room in which the robot has started, and by $Y \in \{1,2\}$ the room in which the robot was found after 2.5 hours. We will denote by $\bar{Y}$ the other room. Assume $p = 1/4$.

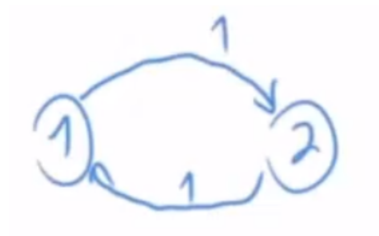
4. *For each of the two models, find the optimal estimator of X_0 from Y in the sense of minimizing error probability.*
5. *Assume the detective comes in not after 2.5 hours, but after $2n + 1/2$ hours for very large n. For each of the two models, can we find an estimator of X_0 from Y that satisfies $\lim_{n\to\infty} \Pr(\hat{X}_0(Y) \neq Y) < 1/2$?*

פתרון

1. האם התהליך X_n הוא שרשרת מרקוב הומוגנית -

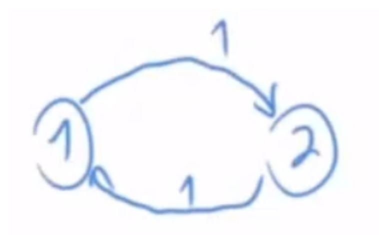
נשים לב שעבור HUJI1 בזמנים אי-זוגיים השרשרת נראת כך -

ובזמנים זוגיים כך -



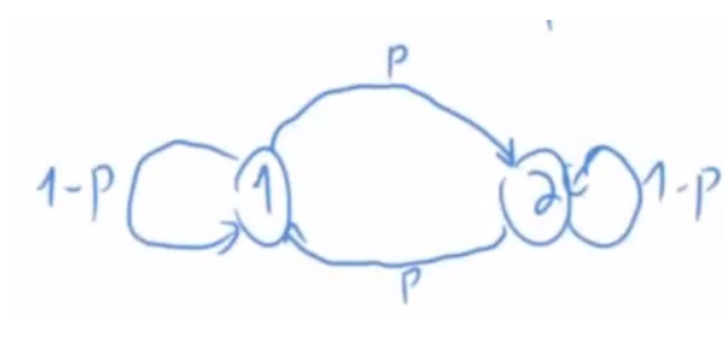
אנחנו רואים שהתנהגות השרשראות תלויה בזמן ולכן היא **לא** שרשרת מרקוב הומוגנית.

עבור התהליך X_{2n} - ידוע כי הרובוט קופץ כל שעתיים בדיוק מחדר לחדר ולכן השרשרת נראת כך -



ולכן השרשרת **כן** מארת תהליך מרקוב הומוגני

2. עבור הרובוט HUJI2 בכל שעה הוא מקבל החלטה לפי הטלת מטבע - כלומר לא תלוי בזמן, בנוסף, הוא מטיל את אותו מטבע כל פעם - הדינמיקה אינה תלויה בזמן, ולכן כאשר מדובר בתהליך X_n התהליך הוא מרקובי הומוגני כאשר השרשרת נראת כך -



עבור התהליך X_{2n} - כלומר דגימה של הרובוט כל שעתיים.

אנו יודעים שבכל שעתיים יש 2 הטלות מטבע, קרי, או שהרובוט נשאר במקום או שהוא ימצא במצב השני. התהליך הוא מרקובי כי המעבר מחדר לחדר תלוי אך ורק ב-2 הטלות המטבע וההטלות הן בת"ל והוא הומוגני כי ההטלות מתבצעות באותה צורה

הסתברות המעברים ניתן לחשב ע"י

$$\begin{aligned} Pr\left(X_2 = j|X_0 = i\right) &= Pr\left(X_{2(n+1)} = j|X_{2n} = i\right) \\ &= Q^2\left(i,j\right) \end{aligned}$$

נחשב את Q ואת Q^2

$$\begin{aligned} Q &= \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} \\ Q^2 &= \begin{bmatrix} 1-2p\left(1-p\right) & 2p\left(1-p\right) \\ 2p\left(1-p\right) & 1-2p\left(1-p\right) \end{bmatrix} \end{aligned}$$

3. חדר נקי מוגדר כנקי אם הרובוט עבד שם לפחות שעה.

לכל אחד מהדגמים נמצא את ההסתברות שהבית נקי אחרי 4 שעות

(א) עבור רובוט HUJI1 הרובוט התחיל באחד משני החדרים ואחרי שעתיים הוא עבר לחדר השני, ולכן עבור הדגם הזה הבית נקי בהסתברות 1. התהליך למעשה לא אקראי.

(ב) עבור הרובוט HUJI2 יהיה יותר קל להסתכל מה הסיכוי שהבית לא נקי- כלומר

$$\{X_0 = 1, X_1 = 1, X_2 = 1, X_3 = 1\}$$

או אם הוא התחיל בחדר 2

$$\{X_0 = 2, X_1 = 2, X_2 = 2, X_3 = 2\}$$

והרובוט לא משנה את מצבו במשך 4 שעות כלומר, הוא לא משנה את מצבו 3 פעמים ברציפות והסיכוי לכך הוא $(1-p)^3$ ולכן המשלים הינו

$$1-(1-p)^3$$

חזרה זריזה על שערוך- נניח מ"א Z כך ש - $Z \in [J] = \{1,2,...,J\}$

וכן מ"א נוסף, תלוי ב-Z, $u \in U$

בנוסף יש לנו משערך $F(u) : U \to [J]$

$$\begin{aligned} P_e &\triangleq Pr\left(F\left(u\right) \neq Z\right) \\ &= \sum_{u\in U} Pr\left(F\left(u\right) \neq Z, U = u\right) \\ &= \sum_{u\in U} Pr\left(U = u\right) \cdot Pr\left(F\left(u\right) \neq Z|U = u\right) \\ &= \sum_{u\in U} Pr\left(U = u\right) \cdot \left[1 - Pr\left(Z = F\left(u\right)|U = u\right)\right] \\ &\geq \sum_{u\in U} Pr\left(U = u\right) \cdot \left[1 - \max_{j\in J} Pr\left(Z = j|U = u\right)\right] \end{aligned}$$

המשערך האופטימלי במובן מינימום הסתברות שגיאה הוא

$$\hat{Z}_{MAP}(u) = \underset{j\in J}{\operatorname{argmax}} Pr(Z=j|U=u)$$

והסתברות השיגאה שלו

$$\begin{aligned}P_{e,MAP} &= Pr\left(\hat{Z}_{MAP}(u) \neq u\right)\\ &= \sum_{u\in U} Pr(U=u)\cdot\left[1-\max_{j\in J} Pr(Z=j|U=u)\right]\end{aligned}$$

4. לבית מגיע בלש ונכנס אחרי שעתיים וחצי, נסמן ב-Y את המיקום שבו הבלש מצא את הרובוט, הבלש ירצה להבין איפה היה הרובוט ב-X_0 בנוסף נסמן את $\bar{Y}$ המשלים של Y

בנוסף נניח $p=\frac{1}{4}$ הבלש ירצה לשערך באמצעות Y את X_0 והוא ירצה שהמשערך יהיה טוב במובן של מינימום הסתברות שגיאה - הוא מנחש איפה הרובוט התחיל ורוצה להיות צודק בהסתברות גבוהה ככל האפשר / לטעות בהסתברות קטנה ככל האפשר.

ראינו משערך הכי טוב בין שני משתנים שהם תלויים זה בזה באופן כללי ובמקרה שלנו X_0 מעניין אותנו (משחק את תפקיד Z)

$$\hat{X}_{0,MAP} = \underset{j\in\{1,2\}}{\operatorname{argmax}} Pr(X_0=j|Y=y)$$

בנוסף נבחין כי Y - מציין איפה הרובוט נמצא אחרי שעתיים וחצי ולכן נבדוק איפה הוא יהיה בתחילת השעה השלישית - $Y=X_2$

ועבור HUJI1 אנחנו יודעים כי הוא מדלג - הוא עובר כל שעה. ולמעשה $X_0=\bar{X}_2=\bar{Y}$ ועבור דגם זה אנו יודעים בצורה דטרמיניסטית שאם מצאנו אותו ב-X_2 אז ב-X_0 הוא היה בחדר המשלים

$$\begin{aligned}\hat{X}_{0,MAP} &= \bar{Y}\\ P_{e,MAP} &= 0\end{aligned}$$

עבור HUJI2 נחשב

$$Pr(X_0=i|X_2=j) = \frac{Pr(X_2=j|X_0=i)\cdot Pr(X_0=i)}{Pr(X_2=j)}$$

מהגדרת השאלה אנו יודעים כי

$$\begin{aligned}Pr(X_0=i) &= \frac{1}{2}\quad \forall i\in\{1,2\}\\ Pr(X_2=j) &= \left(\pi_0\cdot Q^2\right)(j)\\ &= \frac{1}{2}\end{aligned}$$

ונותר לחשב את

$$\begin{aligned}Pr(X_2=j|X_0=i) &= Q^2(i,j)\\ Q^2 &= \begin{bmatrix} 1-2p(1-p) & 2p(1-p)\\ 2p(1-p) & 1-2p(1-p)\end{bmatrix}\end{aligned}$$

שזה למעשה

$$
\begin{aligned}
Pr\left(X_0=i|Y=j\right) &= Pr\left(X_0=i|X_2=j\right)\\
&= \begin{cases} 1-2p\left(1-p\right)|_{p=\frac{1}{4}}=\frac{5}{8} & i=j\\ 2p\left(1-p\right)|_{p=\frac{1}{4}}=\frac{3}{8} & i\neq j\end{cases}
\end{aligned}
$$

ולכן נבחר את $i=j$, כלומר ננחש שהרובוט התחיל במקום בו מצאנו אותו. כי ההסתברות לצדוק הוא $\frac{5}{8}$ והשגיאה שלנו $P_{e,MAP}\left(Y\right)=\frac{3}{8}$

2 ארגודיות

2.1 הגדרת הבעיה

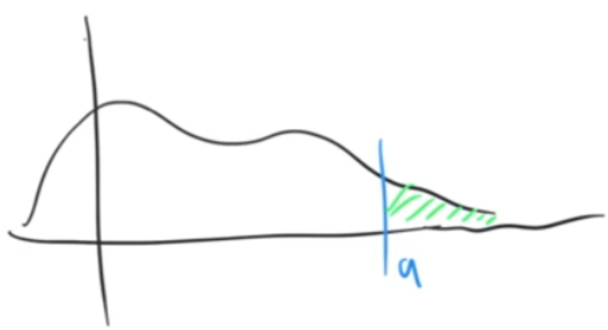
נניח שיש לנו גישה לראליזציה $\{X_n\}_{n=1}^{N}$ של שרשרת מרקוב Q **אבל** Q אינה ידועה לנו.
נרצה לשערך את Q מתוך המדידות, $\{X_n\}_{n=1}^{N}$.
מסתכלים על דגימות מתוך שרשרת מרקוב הומוגנית שיש לה Q אך אנו לא יודעים אותה. נרצה להשתמש במדידות כדי לשערך את Q.

דוגמא לשימוש - נסתכל על מחיר של מנייה ונרצה להחליט האם להשקיע בה או לא. נניח שהמשתנה X_n מייצג את אחוז העליה/ירידה של המניה. ונניח שיש 7 מצבים להתנהגות המניה בכל יום - $\{-3\%,...,0\%,...,+3\%\}$ והגיוני לחשוב שיש תלות בין הימים - זכרון, כלומר, מה שיקרה מחר תלוי במה שקרה היום ואתמול וכו. אבל לצורך בעיה זו נמדל כתהליך מרקובי מסדר ראשון, כלומר יש תלות רק יום אחורה. בדוגמא זו עשינו הרבה הנחות - למשל הומוגניות.

2.2 כלים שימושיים: הערכות זנב, ריכוז מידה, WLLN - Weak law of large numbers, CLT - Central limit theorem

הערכות זנב

אם נסתכל על פונקציית פילוג, הזנב זה כל מה שקורה מעל לנק' מסויימת ב-pdf - הסיכוי שמ"א מקבל ערך יותר גדול מנק' מסויימת

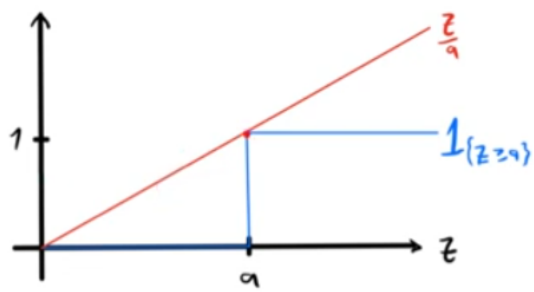


אי שיוויון מרקוב

משפט עבור מ"א אי שלילי Z, לכל $a>0$ מתקיים

$$Pr\left(Z\geq a\right)\leq\frac{\mathbb{E}\left(Z\right)}{a}$$

הוכחה נעזר בתמונה



$$Pr\left(Z\geq a\right)=\mathbb{E}\left(1_{\{Z\geq a\}}\right)$$

האינדיקטור בכחול בתמונה.
ניתן לחסום את האינדיקטור ע"י $\frac{Z}{a}$ ניתן לעשות זאת כי הנחנו שהמ"א הוא אי-שלילי.
ולכן

$$\begin{aligned}Pr\left(Z\geq a\right)&=\mathbb{E}\left(1_{\{Z\geq a\}}\right)\\&\leq\mathbb{E}\left(\frac{Z}{a}\right)=\frac{\mathbb{E}\left(Z\right)}{a}\end{aligned}$$

א"ש צ'בישב לכל $a>0$

$$Pr\left(\left|Z-\mathbb{E}\left(Z\right)\right|\geq a\right)\leq\frac{var\left(Z\right)}{a^{2}}$$

הוכחה נפעיל את א"ש מרקוב על המ"א $\left(\left|Z-\mathbb{E}\left(Z\right)\right|\right)^{2}$ ניתן לעשות זאת כי הוא אי-שלילי ועומד בתנאים.

$$\begin{aligned}Pr\left(\left|Z-\mathbb{E}\left(Z\right)\right|\geq a\right)&=Pr\left(\left|Z-\mathbb{E}\left(Z\right)\right|^{2}\geq a^{2}\right)\\\text{markov}&\leq\frac{\mathbb{E}\left(\left|Z-\mathbb{E}\left(Z\right)\right|^{2}\right)}{a^{2}}\\&=\frac{var\left(Z\right)}{a^{2}}\end{aligned}$$

ממוצע אימפירי

יהיו $Z_1, ..., Z_n$ מ"א iid שמתפלגים כמו המ"א Z
נרצה להתבונן בגודל הבא (נבחין שזה ממוצע אימפירי)

$$\bar{Z}_n \triangleq \frac{1}{n}\sum_{i=1}^{n} Z_i$$

נסתכל על התוחלת והשונות של גודל זה
עבור התוחלת החישוב פשוט, נעזר בלינאריות התוחלת ונקבל

$$\begin{aligned}\mathbb{E}\left(\bar{Z}_n\right) &= \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left(Z_i\right)\\ &= \frac{1}{n}n\mathbb{E}\left(Z\right)\\ &= \mathbb{E}\left(Z\right)\end{aligned}$$

ועבור השונות

$$\begin{aligned}var\left(\bar{Z}_n\right) &= var\left(\frac{1}{n}\sum_{i=1}^{n} Z_i\right)\\ &= \frac{1}{n^2}var\left(\sum_{i=1}^{n} Z_i\right)\\ iid &= \frac{1}{n^2}\sum_{i=1}^{n} var\left(Z_i\right)\\ &= \frac{n\cdot var\left(Z\right)}{n^2}\\ &= \frac{var\left(Z\right)}{n}\end{aligned}$$

מסקנה השונות של ממוצע אימפירי יורדת ככל שנמצע יותר מ"א בת"ל - החוק החלש של המספרים הגדולים.

החוק החלש של המספרים הגדולים.

נזכיר כי

$$\bar{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i$$

משפט לכל $\epsilon > 0$ מתקיים $\lim_{n\to\infty} Pr\left(\left|\bar{Z}_n - \mathbb{E}\left(Z\right)\right| > \epsilon\right) = 0$

כלומר אם נסתכל על $\bar{Z}_n$ עבור n מאוד גדול, הסיכוי שהוא יימצא במרחק יותר גדול מ-ϵ מהתוחלת הולך וקטן, וככל שנגדיל את n ההסתברות תדעך ל-0

הוכחה לפי צ'בישב

$$Pr\left(\left|\bar{Z}_n-\mathbb{E}\left(Z\right)\right|>\epsilon\right)=Pr\left(\left|\bar{Z}_n-\mathbb{E}\left(\bar{Z}_n\right)\right|>\epsilon\right)$$
$$chevishev\leq\frac{var\left(\bar{Z}_n\right)}{\epsilon^2}$$
$$\lim_{n\to\infty}=\frac{var\left(Z\right)}{n\cdot\epsilon^2}=0$$

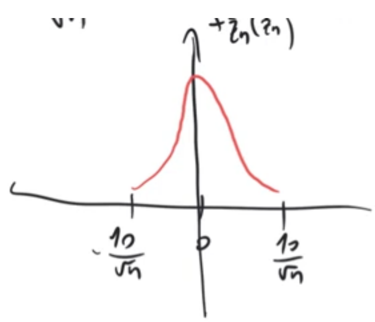
אפיון עדין של $\bar{Z}_n-\mathbb{E}\left(Z\right)$

ראינו ש-$\bar{Z}_n-\mathbb{E}\left(Z\right)$ הולך וקטן ככל ש-n גדל, נתבונן באפיון יותר עדין של הפרש זה

$$var\left(\bar{Z}_n\right)=\frac{var\left(Z\right)}{n}$$

כלומר, סטיית התקן היא סדר גודגל של $\frac{1}{\sqrt{n}}$

נניח ש-$\bar{Z}_n$ הוא בעל תוחלת 0 ושונות 1 ונרצה להסתכל על הפילוג בתחום מסויים לדוגמא $\left[-\frac{10}{\sqrt{n}},\frac{10}{\sqrt{n}}\right]$ כלומר נרצה להעריך רק בתחום קבוע $\left[-\frac{a}{\sqrt{n}},\frac{a}{\sqrt{n}}\right]$, מסתבר שאם מסתכלים על המ"א בקנה מידה הנכון הוא נראה מאוד דומה **לגאוסיין** וזה נקרא **CLT** או **משפט הגבול המרכזי.**



נגדיר

$$\Phi\left(x\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{\frac{-t^2}{2}}dt$$

זהו ה- CDF of $\mathcal{N}\left(0,1\right)$

ה-CLT

$$\sqrt{n}\left(\bar{Z}_n-\mathbb{E}\left(Z\right)\right)\xrightarrow{d}\mathcal{N}\left(0,var\left(Z\right)\right)$$

משפט הגבול המרכזי

נניח $0<var\left(Z\right)<\infty$, לכל $t\in\mathbb{R}$

$$\lim_{n\to\infty}Pr\left(\sqrt{n}\left(\bar{Z}_n-\mathbb{E}\left(Z\right)\right)\leq t\right)=\Phi\left(\frac{t}{\sqrt{var\left(Z\right)}}\right)$$

נסביר - $\bar{Z}_n - \mathbb{E}(Z)$ - מ"א סטייה של הממוצע האמפרי מהתוחלת, מוכפל ב-$\sqrt{n}$ ההכפלה היא כדי לקבל קנה מידה סופי כמו שראינו קודם. והסיכוי שהוא קטן או שווה מ-t זה בעצם ה-cdf שלו.
כלומר אם ניקח n מספיק גדול, ה-cdf של המ"א הזה ישאף לפונקציית שגיאה, כלומר ל-cdf של $\mathcal{N}(0, var(Z))$
זהו ה-clt והוא מתאר את לנו את הסטייה של הממוצע האמפירי מהתוחלת בקנה מידה של $\frac{1}{\sqrt{n}}$

חסם צ'רנוף

טענה חסם צ'רנוף - לכל $t > 0$

$$Pr(Z \geq a) \leq e^{-ta} \cdot \mathbb{E}\left(e^{tz}\right)$$

הוכחה

$$\begin{aligned} Pr(Z \geq a) &\overset{*}{=} Pr\left(e^{tz} \geq e^{ta}\right) \\ markov &\leq \frac{\mathbb{E}\left(e^{tz}\right)}{e^{ta}} \end{aligned}$$

* כי e^{tx} היא מונוטונית עולה

מסקנה לכל $t > 0$ ו-$a \in \mathbb{R}$

$$\begin{aligned} Pr\left(\bar{Z}_n - \mathbb{E}(Z) \geq a\right) &\leq e^{-nta} \cdot \left[\mathbb{E}\left(e^{t(Z-\mathbb{E}(Z))}\right)\right]^n \\ &= \left(e^{-ta}\mathbb{E}\left(e^{tz}\right)\right)^n \end{aligned}$$

אם נרצה לשאול מה הסיכוי שהממוצע האמפירי סוטה סטייה קבועה מהתוחלת, הסיוכי הזה דועך אקספ'

הוכחה

$$\begin{aligned} Pr\left(\bar{Z}_n - \mathbb{E}(Z) \geq a\right) &= Pr\left(\frac{1}{n}\sum(Z_i - \mathbb{E}(Z)) \geq a\right) \\ &= Pr\left(\sum(Z_i - \mathbb{E}(Z)) \geq na\right) \\ chernof &\leq e^{-nta} \cdot \mathbb{E}\left(e^{t\sum Z_i - \mathbb{E}(Z)}\right) \\ &= e^{-nta} \cdot \mathbb{E}\left(\prod_{i=1}^{n} e^{t(Z_i - \mathbb{E}(Z))}\right) \\ &= e^{-nta} \cdot \prod_{i=1}^{n} \mathbb{E}\left(e^{t(Z_i - \mathbb{E}(Z))}\right) \\ iid &= e^{-nta} \cdot \mathbb{E}\left(e^{t(Z-\mathbb{E}(Z))}\right)^n \end{aligned}$$

2.3 ארגודיות - עכשיו באמת.

נרצה לשערך את Q מפונקצית מדגם עם ריאליזציה $\{X_n\}_{n=0}^{\infty}$

חימום נדמיין קופסא - סימולטור שמחזיק את Q, ניתן להכניס לקופסא פילוג התחלתי π_0 והסימולטור פולט פונ' מדגם של התהליך מזמן 0 עד m

חימום: $X_n \in (J)$

ניתן לגשת לסימולטור כמה פעמים שרוצים ובכל גישה נקבל פונק' מדגם בלתי תלויה.
ונרצה לשערך באמצעות סימולטור זה את איברי Q, לדוגמא - $Q\left(i,j\right)$
נזכיר ש-

$$Q\left(i,j\right)=Pr\left(X_n=j|X_{n-1}=i\right)$$

אסטרטגיה -

1. להגריל הרבה מ"א בת"ל עם פילוג $\left[X_n|X_{n-1}=i\right]$ וזה בידיוק השורה ה-i של המטריצה $Q\left(i,\bullet\right)$
נפעיל את הסימולטור עם $X_0=i:\pi_0=\left[0,0,...,0,\underset{i}{1},0,...,0\right]$ ונקבל ש-

$$\begin{aligned}X_1&\sim Pr\left(X_1=\bullet|X_0=i\right)\\&=Q\left(i,\bullet\right)\end{aligned}$$

נסמן את המ"א שייצרנו כך - $X_1\left(1\right),X_1\left(2\right),...,X_1\left(n\right)$ פנינו לסימולטור n פעמים וכל פעם האינדקס 1 אומר שמסתכלים על תהליך X שדגמנו אותו בדגימה הראשונה ומה שבתוך הסוגריים מייצג על איזה פונקציית מדגם אנחנו עובדים. כלומר פנינו לסימולטור שלנו n פעמים וכל פעם אתחלנו עם ת"ה דטרמניסטי $X_0=i$ וכל פעם קיבלנו פונ' מדגם בת"ל ולקחנו רק את הדגימה הראשונה ממנה

2. נחשב את המ"א

$$1_{\{X_1(1)=j\}},...,1_{\{X_1(n)=j\}}$$

נגדיר

$$\hat{Q}\left(i,j\right)=\frac{1}{n}\sum_{t=1}^{n}1_{\{X_1(t)=j\}}$$

כל המ"א בסכום הזה הם בת"ס כי הם נוצרו מהסימולטור ומ-WLLN נקבל

$$\begin{aligned}\hat{Q}\left(i,j\right)&\to\mathbb{E}\left(1_{\{X_1(t)=j\}}\right)\\&=\mathbb{E}\left(1_{\{X_1(t)=j\}}|X_{n-1}=i\right)\\&=Pr\left(X_n=j|X_{n-1}=i\right)\\&=Q\left(i,j\right)\end{aligned}$$

ה-WLLN אומר שאם ניקח n מספיק גדול

$$\lim_{n\to\infty}Pr\left(\left|\hat{Q}\left(i,j\right)-Q\left(i,j\right)\right|>\epsilon\right)=0$$

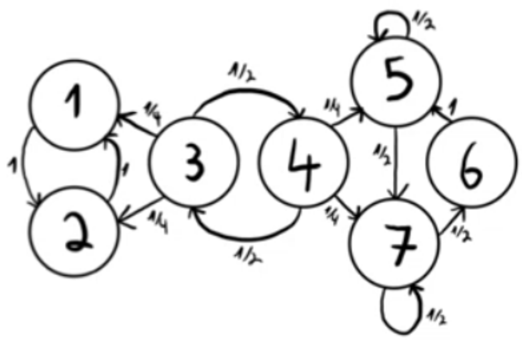
כלומר, עבור מספיק גישות לסימולטור נצליח לשערך את $Q\left(i,j\right)$ ברזולוציה טובה כרצוננו.

בעיה במציאות אין בד"כ גישה לסימולטור שיכול לייצר המון ריאליזציות בת"ל של שרשרת מרקוב עם מטריצה Q.

המקרה הנפוץ יותר - יש גישה לפונ' מדגם בודדת מאוד ארוכה שממנה נרצה לשערך את מטריצת המעברים Q

המטרה להבין מתי שערוך Q מתוך פונק' מדגם בודדת מאוד ארוכה אפשרית או בלתי אפשרית.

בשביל המקרה **הבלתי אפשרי** נתבונן בדוגמא -



נגיד ונרצה לשערך את המטריצה המתאימה לו, המטריצה תהיה עם הרבה 0, כי כל קורדינטה שמייצגת מעבר בין 2 מצבים שאין בינהם קשת תהיה 0 וכשיש קשת הקורדינטה תהיה הערך שעל הקשת.
נרצה לשערך את המטריצה שמתאימה לגרף מתוך פונק' מדגם בודדת.
במקרה הזה **אין לנו סיכוי** להצליח. למה?
נגיד הגרלנו פונקצית מדגם $3, 4, 3, 5, 5, 7, 6, 5, 7, 7, 6, ...$, אנו אף פעם לא רואים את מצבים $1, 2$ כי ברגע שנכנסו למחלקה נשנית $\{5, 6, 7\}$ לא נצא ממנה ולעולם לא נראה את $\{1, 2\}$ ולכן אין לנו סיכוי לשערך את המצבים $1, 2$

מסקנה בשרשרת עם יותר ממחלקה נשנית אחת על כל ריאליזצית מדגם שנגריל תמיד נצפה אך ורק במה שקורה בתוך מחלקה נשנית אחת. את יתר המחלקות הנשנות נפספס ולכן אין לנו סיכוי לשערך את Q

בדוגמא שלנו אם נתבונן במעבר מ-$\{5, 7\} \to 4$ או $\{1, 2\} \to 3$ מתוך 4 החצים הללו כל הילוך אינסופי יעבור רק **באחד** מהם.

2.3.1 תהליך ארגודי

הגדרה תהליך אקראי סטציונרי $\{X_n\}_{n=0}^{\infty}$ יקרא **ארגודי** אם לכל פונק' $\mathcal{X}^k \to \mathbb{R}$ ולכל $\epsilon > 0$

$$\lim_{n\to\infty} \Pr\left(\left|\hat{G}_n - \mathbb{E}\left[g\left(X_1, \ldots, X_k\right)\right]\right| > \epsilon\right) = 0$$

כאשר $\hat{G}_n$ זה משערך של הפונקציה מתוך דגימה אחת ממש ארוכה - ממוצע הפונקציה בזמן

$$\hat{G}_n = \frac{1}{n}\sum_{t=0}^{n-1} g\left(X_t, X_{t+1}, \ldots, X_{t+k-1}\right)$$

מיצוע בזמן $\longleftrightarrow$ מיצוע סטטיסטי

2.3.2 שרשרת מרקוב ארגודית

נאמר ששרשרת מרקוב היא ארגודית אמ"מ היא שוכחת את העבר. כלומר אמ"מ יש לה מחלקה אחת נשנית א-פריודית.
נשים לב שההגדרה לשרשרת מרקוב ארגודית **אינה** מקרה פרטי של ההגדרה הכללית לארגודיות
ההגדרה המרקובית תלויה רק ב-Q אבל אינה תלויה ב-π_0, למעשה לא דרשנו בכלל סטציונריות!

נסתכל על הפונקציה

$$\begin{aligned} g\left(X_{t-1},X_t\right) &= 1_{\{X_{t-1}=i,X_t=j\}} \\ \hat{G}_n &= \hat{\pi}_n\left(i,j\right) \\ &= \frac{1}{n}\sum_{t=1}^{n}1_{\{X_{t-1}=i,X_t=j\}} \end{aligned}$$

נרצה להראות ש-

$$\begin{aligned} \hat{G}_n &\to \mathbb{E}\left(g\left(X_{t-1},X_t\right)\right) \\ &= Pr\left(X_{t-1}=i,X_t=j\right) \\ &= Pr\left(X_{t-1}=i\right)\cdot Pr\left(X_{t-1}=j|X_{t-1}=i\right) \\ &= \pi^{\text{stat}}\left(i\right)\cdot Q\left(i,j\right) \end{aligned}$$

אסטרטגיה: נראה ש

$$\begin{aligned} \mathbb{E}\left(\hat{\pi}_n\left(i,j\right)\right) &= \pi^{\text{stat}}\left(i\right)\cdot Q\left(i,j\right) \\ var\left(\hat{\pi}_n\left(i,j\right)\right) &\stackrel{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

לשם פשטות נניח

$$\pi_0=\pi^{\text{stat}}$$

כלומר שהתהליך שלנו סטציונרי לחלוטין
חישוב התוחלת

$$\begin{aligned} \mathbb{E}\left(\hat{\pi}_n\left(i,j\right)\right) &= \mathbb{E}\left(\frac{1}{n}\sum_{t=1}^{n}1_{\{X_{t-1}=i,X_t=j\}}\right) \\ &= \frac{1}{n}\sum_{t=1}^{n}\mathbb{E}1_{\{X_{t-1}=i,X_t=j\}} \\ &= \frac{1}{n}\sum_{t=1}^{n}Pr\left(X_{t-1}=i,X_t=j\right) \\ &= \pi^{\text{stat}}\left(i\right)\cdot Q\left(i,j\right) \end{aligned}$$

חישוב השונות מזעזע ומי שרוצה להתעמק זה נמצא ב-lecture notes עמוד 19-20

3 תהליכים אקראיים - חזרה

תהליך אקראי הוא אוסף של מ"א $X\left(t\right)$ לכל $t\in I$
לדוגמא $I=\left[0,1\right]$, ובכל זמן בקטע 0-1 מוגדר מ"א $X\left(t\right)$

נזכיר שמ"א הוא מיפוי ממרחב המדגם ל-$\mathbb{R}$

$$X : s \to \mathbb{R}$$

תהליך אקראי הוא מיפוי ממרחב המדגם וציר הזמן I ל-$\mathbb{R}$

$$X\left(t,\omega\right) : I \times S \to \mathbb{R}$$

למעשה נוכל לחשוב על תהליך אקראי כהגרלה של פונקציה לכל $\omega_0 \in S$ נקבל פונקציה אחרת $X\left(t,\omega_0\right)$
דרך נוספת היא להתבונן על אוסף מ"א כך שאם נקפיא את ציר הזמן מתקיים $X\left(t_0,\omega\right)$ הוא מ"א

מקרה פרטי שאנו מכירים טוב מאוד(?) אם ציר הזמן I הוא בדיד וסופי למשל

$$I = \{1,2,3,4,5\}$$

במקרה הזה התהליך האקראי הוא פשוט וקטור אקראי שגודלו $|I|$ וכאשר I אינו סופי ניתן לחשוב עליו כוקטור אקראי אין סופי - לדוגמא כל המספרים השלמים.
יש שני צירים בהם התהליך יכול להתקיים רציף או בדיד - סה"כ 4 קומבינציות.

1. זמן - דוגמא לרציף $I = [0,\infty)$ ובדיד $I = \mathbb{N}$
2. "אמפליטודה" - רציף $X\left(t\right) \in \mathbb{R}$ ואמפ' בדידה $X\left(t\right) \in \mathbb{Z}$

סימונים של תהליך רציף/בדיד הוא $\{X\left(t\right)\}_{t\in I}$ כך שאם I רציף התהליך רציף בזמן ואם I בדיד התהליך בדיד בזמן, כאשר I בדיד, בד"כ נסמן $\{X_n\}_{n\in I}$

3.1 אפיון סטטיסטי של תהליכים אקראיים

נזכיר כי עבור וקטור אקראי הסטטיסטיקה המלאה נתונה ע"י ה-CDF, כלומר $\underline{X} = \left(x_1,...,x_k\right)^T$כדי לאפיין את הפילוג צריך לתת לכל אוסף הערכים את ה-CDF בנקודה שלו -

$$F_{\underline{\mathrm{X}}}\left(x_1,..,x_k\right) \triangleq \Pr\left(X_1 \le x_1,...X_k \le x_k\right)$$

עבור תהליך אקראי האפיון הסטטיסטי המלא שלו כולל את ה-CDF לכל ו"א שמתקבל ע"י דגימה של התהליך בזמנים שונים.
אם התהליך שלנו הוא $\{X\left(t\right)\}_{t\in I}$ אפיון סטטיסטי מלא יכלול את הפונקציות

$$F_{X(t_1),...,X(t_k)}\left(x_{t_1},...,x_{t_k}\right) \quad \forall k\in\mathbb{N},\ \forall t_1<t_2<...<t_k\ S.T.\ t_1,...,t_k\in\mathbb{N}$$

בעצם מה שהדבר הזה אומר זה שהאפיון הסטטיסטי המלא של התהליך לא משנה באיזה אוסף של זמנים נדגום את התהליך, כל k זמנים שניקח שנמצאים באינטרוול I נוכל לדגום את התהליך ונקבל וקטור אקראי עם k דגימות ונוכל לתאר את ה-CDF שלו.
אבל! ה-CDF צרכים להיות עקביים (קונסיסטנטים) למשל אם נתאר את ה-CDF של $F_{X(3),X(4)}\left(x_3,x_4\right)$ ואת ה-CDF של $F_{X(3)}\left(x_3\right)$ והם צרכים להיות עקביים במובן שאם נסתכל על ה-CDF המשותף ביחס ל-$X\left(3\right)$, נדרוש לקבל

$$F_{X(3),X(4)}\left(x_3,\infty\right) = F_{X(3)}\left(x_3\right)$$

3.2 תהליך אקראי גאוסי

הגדרה תהליך אקראי $\{X(t)\}_{t\in I}$ ייקרא גאוסי אמ"מ כל אוסף של דגימות מהתהליך הוא וקטור אקראי גאוסי.

כלומר, לכל $t_1,...t_k \in I$ הוקטור $[X(t_1),...,X(t_k)]^T$ הוא ו"א גאוסי (וא"ג)
נתבונן ב-PDF

$$f_{X(t_1),...,X(t_k)}(x_1,\ldots,x_k) = f_{X(t_1),...,X(t_k)}(\text{x}) = \frac{1}{\sqrt{(2\pi)^k |\mathbf{C_t}|}} \cdot \exp\left[-\frac{1}{2}\cdot(x-\boldsymbol{\eta}_\mathbf{t})^T \mathbf{C_t}^{-1}(x-\boldsymbol{\eta}_\mathbf{t})\right]$$

כאשר

$$\begin{aligned}\boldsymbol{\eta}_\mathbf{t} &= [\mathbb{E}[X(t_1)],...,\mathbb{E}[X(t_k)]]^T\\ \mathbf{C}_{(i,j)} &= cov[X(t_i),X(t_j)]\\ &= \mathbb{E}[(X(t_i)-\mathbb{E}[X(t_i)])\cdot(X(t_j)-\mathbb{E}[X(t_j)])]\end{aligned}$$

היות וה-PDF הוא מהצורה הנ"ל, בכדי להגיע לאיפיון מלא של וקטור דגימות זה התוחלת של וקטור הדגימות וה-cov, כמובן שהם צרכים להיות קונסיסטנטים, כפי שהגדרנו קודם.

3.3 סטציונריות

תהליך סטציונרי הוא תהליך אינווריאנטי להזזות בציר הזמן
תהליך אקראי $\{X(t)\}_{t\in I}$ ייקרא סטציונארי אמ"מ לכל $t_1,...,t_k \in I$ ולכל T מתקיים

$$F_{X(t_1),...,X(t_k)}(x_1,\ldots,x_k) = F_{X(t_1+T),...,X(t_k+T)}(x_1,\ldots,x_k)$$

כלומר הסטטיסטיקה של התהליך לא משתנה בזמן, ונקבעת רק ע"פ הפרשי הזמנים.

תהליך סטציונרי אסימפטוטי אם $\lim_{T\to\infty} F_{X(t_1+T),...,X(t_k+T)}(x_1,\ldots,x_k)$ הגבול קיים התהליך ייקרא סטציונרי והפילוג המתקבל בגבול ייקרא הפילוג הסטצינרי -

דוגמה לכך היא שרשרת מרקוב ששוחכת את העבר, בזמן מספיק גדול היא מתכנסת לפילוג סטציונרי וזה עקבי עם ההגדרה הנ"ל.

3.4 סטטיסטיקה מסדר 2 של תהליך אקראי וסטציונריות במובן הרחב

עבור תהליך אקראי $\{X(t)\}_{t\in I}$

פונ' התוחלת תקרא $\eta_x(t)$ והיא מוגדרת להיות $\forall t\in I \quad \eta_X(t) = \mathbb{E}[X(t)]$

פונק' האוטוקורלציה נסמן ב-$R_X(t_1,t_2)$והיא תוגדר להיות $\forall t_1,t_2\in I \quad R_X(t_1,t_2) = \mathbb{E}[X(t_1)\cdot X(t_2)]$

פונק' האוטוקווארינס נסמן $C_X(t_1,t_2)$ ונגדיר אותה להיות

$$\begin{aligned}C_X(t_1,t_2) &= \mathbb{E}[(X(t_1)-\eta_X(t_1))(X(t_2)-\eta_X(t_2))]\\ &= R_X(t_1,t_2) - \eta_X(t_1)\eta_X(t_2)\end{aligned}$$

מקדם הקורלציה של התהליך $\{X(t)\}_{t\in I}$ נסמן ב-$\rho_X(t_1,t_2)$ ונגדיר

$$\rho_X(t_1,t_2) = \frac{C_X(t_1,t_2)}{\sqrt{C_X(t_1,t_1)\,C_X(t_2,t_2)}}$$

4 ההגדלים הנ"ל הם שמאפיינים את הסטטיסטיקה מסדר שני של התהליך, למעשה צריך לדעת את פונ' התוחלת ופונק' האוטוקורלציה

3.5 סטציונריות במובן הרחב - WSS

הגדרה תהליך אקראי $\{X(t)\}_{t\in I}$ הוא WSS אמ"מ הסטטיסטיקה שלו מסדר שני אינווריינטית להזזה בזמן, כלומר,

$$\begin{aligned}\eta_X(t_1) &= \eta_X(t_1+T)\\ &\Rightarrow \eta_X(t) = \eta_X \quad \forall t\in I\end{aligned}$$

וכן מתקיים

$$\begin{aligned}R_X(t_1,t_2) &= R_X(t_1+T,t_2+T) \quad \forall t_1,t_2\in I\\ &\Rightarrow R_X(t_1,t_2) = R_X(t_2-t_1)\end{aligned}$$

* למשל אם נבחר $T=-t_1$ נקבל

$$\begin{aligned}R_X(t_1,t_2) &= R_X(t_1-t_1,t_2-t_1)\\ &= R_X(t_2-t_1)\\ &\triangleq R_X(\tau)\end{aligned}$$

לסיכום תהליך הוא WSS אמ"מ

$$\begin{aligned}\eta_X(t) &= \eta_X\\ R_X(t,t+\tau) &= R_X(\tau)\end{aligned}$$

נשים לב ש-

$$\begin{aligned}R_{X(0)} &= \mathbb{E}[X(t)\cdot X(t)]\\ &= \mathbb{E}\left[X^2(t)\right]\\ &\geq 0\end{aligned}$$

כלומר $R_{X(0)}$ הוא המומנט השני של X בכל זמן ולכן הוא גם אי שלילי.
וכן מתקיימות 2 תכונות עבור תהליך WSS

1. פונק' האוטוקורלציה סימטרית

$$\begin{aligned}R_X(\tau) &= R_X(-\tau)\ \forall\tau\\ R_X(t_1,t_2) &= R_X(t_2,t_1)\\ R_X(t_1-t_2) &= R_X(t_2-t_1)\end{aligned}$$

2.

$$R_X(0)\geq |R_X(\tau)|\ \forall\tau$$

הוכחה בעזרת א"ש קושי שורץ

WSS **אסימפטוטית** $\{X(t)\}_{t\in I}$ ייקרא WSS אסימפטוטית אמ"מ הגבולות הבאים קיימים

$$\begin{aligned}&\lim_{t\to\infty}\eta_X(t)\\ &\lim_{T\to\infty}R_X(t_1+T,t_2+T)\end{aligned}$$

כלומר, אם נתבונן מספיק רחוק ערך התוחלת נהיה קבוע (או כמעט קבוע)

3.6 תהליך וינר - $Wiener$ / תנועה בראונית

3.6.1 תהליך הילוך שיכור בדיד

תהליך אקראי iid $\{W_n\}_{n=0}^{\infty}$ כאשר כל דגימה של תהליך כזה מפולגת אחיד בתחום $\{-1,1\}$
וכן נגדיר תהליך AR, כך ש$X_0=0$ ו-$n\geq 1$ $X_n=X_{n-1}+d\cdot W_n,$
תהליך AR עם $\alpha=1$ ובמקרה זה תהליך זה אינו סטציונארי.

$$X_n=d\cdot\sum_{k=1}^{n}W_n,\quad n\geq 1$$

נגדיר את ה-**תוספות**

$$\begin{aligned}X\left(n_1,n_2\right)&\triangleq X_{n_2}-X_{n_1}\\&=d\sum_{k=1}^{n_2}W_k-d\sum_{k=1}^{n_1}W_k\\&=d\sum_{k=n_1+1}^{n_2}W_k\end{aligned}$$

נבחין שעבור 4 זמנים, למשל, $i\leq j\leq l\leq m$ שהתוספת $X\left(i,j\right)=d\sum_{k=i+1}^{j}W_k$ בת"ס ב-$X\left(l,m\right)=d\sum_{k=l+1}^{m}W_k$
ובעצם קיבלנו ש-

$$\begin{aligned}X\left(i,j\right)&=d\sum_{k=i+1}^{j}W_k\\&=f\left(\{W_k\}_{l+1}^{m}\right)\end{aligned}\qquad\qquad\begin{aligned}X\left(l,m\right)&=d\sum_{k=l+1}^{m}W_k\\&=f\left(\{W_k\}_{i+1}^{j}\right)\end{aligned}$$

היות ו-$i\leq j\leq l\leq m$ נקבל שהאינטרוולים הבאים זרים

$$(i+1,j)\cap(l+1,m)$$

כלומר הדגימות של W שקובעות את התוספת של $X\left(i,j\right)$ הן בת"ל בדגימות של W שקובעת את התוספת $X\left(l,m\right)$ וכיוון ש-W הוא iid אז התוצאה של הפעלת הפונקציה הם מ"א בלתי תלויים
ובנוסף התוחלת של תוספת היא 0 -

$$\mathbb{E}[X(i,j)]=\mathbb{E}\left[d\sum_{k=i+1}^{j}W_k\right]=0\text{ all for }i\leq j$$

ומתקיים -

$$\eta_n=\mathbb{E}\left[X_n\right]=\mathbb{E}\left[d\sum_{k=1}^{n}W_k\right]=0$$

כעת נתבונן בפונקציית האוטוקורלציה, נניח $m \geq n$

$$\begin{aligned} R(n,m) &\triangleq \mathbb{E}\left[X_n X_m\right] = \mathbb{E}\left[X_n\left(X_n + X_m - X_n\right)\right] = \mathbb{E}\left[X_n\left(X_n + X(n,m)\right)\right] \\ &= \mathbb{E}\left[X_n^2\right] + \mathbb{E}\left[X_n X(n,m)\right] \\ &= \mathbb{E}\left[X_n^2\right] + \mathbb{E}[X(0,n)X(n,m)] \\ &= \mathbb{E}\left[X_n^2\right] + \mathbb{E}[X(0,n)]\mathbb{E}[X(n,m)] \\ &= \mathbb{E}\left[X_n^2\right] \\ &= Var\left[X_n\right] \\ &= Var\left[d\sum_{k=1}^{n} W_k\right] \\ &= d^2 \cdot n \end{aligned}$$

אם היינו מניחים ש-$m \geq n$ היינו מקבלים $d^2 m$ ולכן ככל נקבל

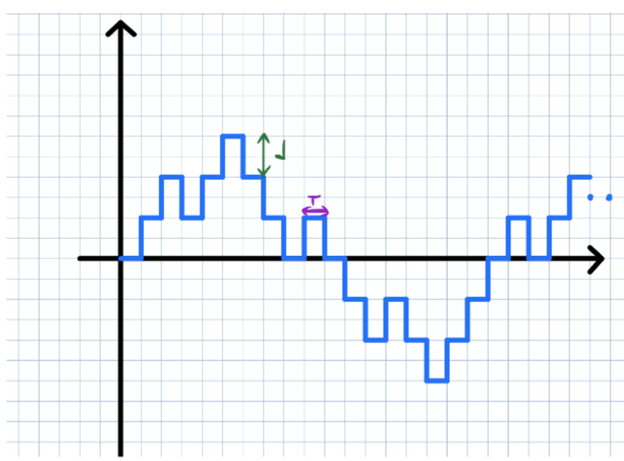
$$R\left(n,m\right) = d^2 \cdot \min\left(n,m\right)$$

3.6.2 תהליך הילוך שיכור בזמן רציף

נתבסס על $\{X_n\}_{n=1}^{\infty}$ ונבנה תהליך בזמן רציף - $\{X_{T,d}\left(t\right)\}_{t\geq 0}$ שמשנה את ערכו כל T שניות בהתאם לערך של $\{X_n\}$ נראה את המשוואה

$$X_{T,d}(t) = \sum_{k=0}^{\infty} X_k \cdot 1_{\{t\in[kT,(k+1)T)\}}$$

נסביר את המשוואה באמצעות הדוגמא הבאה



הציור ממחיש את התהליך בצורה טובה, התהליך קבוע בכל T שניות, והוא משנה את ערכו לפי מה שהילוך השיכור הבדיד עשה, בכל שלב התהליך או קופץ d למעלה או למטה כתלות במה שהילוך השיכור הבדיד החליט.
התהליך רציף כי בכל ערך בציר הזמן יש ערך, אבל מצד שני הוא מאוד בדידי כי הוא מוגדר היטב מהתהליך הבדיד.

כעת נגדיר

$$N(t) = \left\lfloor \frac{t}{T} \right\rfloor$$

מספר הקפיצות שביצע $\{X_{T,d}(t)\}$ עד זמן t וקיבלנו ש-

$$\begin{aligned} X_{T,d}(t) &= X_{N(t)} \\ Continuous \quad & Discrete \end{aligned}$$

מתוך השיוויון הנ"ל אפשר לחשב את הסטטיסטיקה של התהליך הרציף מתוך הסטטיסטיקה של התהליך הבדיד

$$\begin{aligned} \eta_t &= \mathbb{E}\left[X_{T,d}(t)\right] \\ &= \mathbb{E}\left[X_{N(t)}\right] \\ &= 0 \\ R(t_1,t_2) &= \mathbb{E}\left[X_{N(t_1)}X_{N(t_2)}\right] \\ &= d^2 \cdot \min\left\{N(t_1), N(t_2)\right\} \\ &= d^2 \cdot \min\left\{\left\lfloor \frac{t_1}{T} \right\rfloor, \left\lfloor \frac{t_2}{T} \right\rfloor\right\} \end{aligned}$$

ועבור $t_1, t_2 \gg T$ נקבל כי

$$R(t_1,t_2) \approx \frac{d^2}{T} \min\{t_1,t_2\}$$

3.6.3 הרצפה

נרצה לקבל תהליך "חלק" מתוך $X_{T,d}(t)$ ויש לנו 2 בעיות

- חוסר חלקות באמפליטודה - כל פעם התהליך קופץ בקוונטות של d, אם ניקח את $d \to 0$ נקבל בגבול קפיצות חלקות.
- חוסר חלקות בזמן - התהליך משתנה רק כל T שניות, וכדי לפתור בעיה זו נצמצם את אינטרוול T לכמה שיותר קטן ונקבל משהו שנראה רציף $T \to 0$

ראינו ש-$R_x(t_1,t_2)$ כאשר $t_1, t_2 \gg T$ מתקיים

$$R(t_1,t_2) \approx \frac{d^2}{T} \min\{t_1,t_2\}$$

אבל אם ניקח גם את T וגם את d לאפס באותו קצב, כלומר, $T = d \to 0$ נקבל ש-$R(t_1,t_2) = 0$ וזה תהליך לא מעניין, ואם ניקח $d = T^{\frac{1}{4}} \to 0$ נקבל

$$R(t_1,t_2) \approx \frac{T^{\frac{1}{2}}}{T} \min\{t_1,t_2\} \to \infty$$

ותהליך זה מתפוצץ וגם הוא אינו מעניין
כדי לקבל תהליך חלק ולא טריואלי צריך לבחור T ו-d כך שהיחס בינהם נשאר קבוע אפילו כש-T הולך לאפס ולכן נבחר

$$d = \sqrt{\alpha T} \quad \alpha > 0$$

3.6.4 תהליך $X_\alpha(t)$

נגדיר את התהליך $X_\alpha(t)$

$$X_\alpha(t) = \lim_{T\to 0} X_{T,d=\sqrt{\alpha T}}(t)$$

נאפיין את הסטטיסטיקה מסדר שני של התהליך

$$\begin{aligned}\eta_X(t) &= \mathbb{E}\left[X_\alpha(t)\right] = 0 \\ R(t_1,t_2) &= \mathbb{E}\left[X_{\alpha(t_1)} X_{\alpha(t_2)}\right] \\ &= \alpha\cdot\min\{t_1,t_2\}\end{aligned}$$

הפילוג של התהליך $X_\alpha(t)$

$$X_\alpha(t) = \lim_{T\to 0} X_{T,d=\sqrt{\alpha T}}(t)$$

נסתכל על הפילוג של $X_{T,d=\sqrt{\alpha T}}$ גלפני לקיחת הגבול $T\to 0$

$$\begin{aligned}X_{T,d=\sqrt{\alpha T}} &= \sqrt{\alpha T}\sum_{k=1}^{N(t)} W_k \\ &= \sqrt{\alpha t}\cdot\sqrt{\frac{T}{t}}\sum_{k=1}^{N(t)} W_k \\ &= \sqrt{\alpha t}\cdot\frac{1}{\sqrt{t/T}}\frac{\sqrt{N(t)}}{\sqrt{N(t)}}\sum_{k=1}^{N(t)} W_k \\ &= \sqrt{\alpha t}\cdot\sqrt{\frac{\lfloor t/T\rfloor}{t/T}}\frac{1}{\sqrt{N(t)}}\sum_{k=1}^{N(t)} W_k\end{aligned}$$

כעת נתבונן בגבול $T\to 0$

$$\sqrt{\alpha t}\cdot\underbrace{\sqrt{\frac{\lfloor t/T\rfloor}{t/T}}}_{=1}\underbrace{\frac{1}{\sqrt{N(t)}}\sum_{k=1}^{N(t)} W_k}_{=\mathcal{N}(0,1)}$$

נבחין כי כאשר $T\to 0$ אז $N(t)\to\infty$ ולפי משפט הגבול המרכזי כאשר

$$\frac{1}{\sqrt{N(t)}}\sum_{k=1}^{N(t)} W_k \xrightarrow{CLT} \mathcal{N}(0,1)$$

ולכן קיבלנו שמתקיים

$$\begin{aligned}X_\alpha(t) &\sim \sqrt{\alpha t}\cdot\mathcal{N}(0,1) \\ &\sim \mathcal{N}(0,\alpha t)\end{aligned}$$

והמסקנה היא שהפילוג השולי של כל דגימה $X_\alpha(t)$ הוא גאוסי $\mathcal{N}(0,\alpha t)$

ומה לגבי התוספות?

כל תוספת $X(t_1,t_2)$ היא סכום של $N(t_2)-N(t_1)$ משתנים אקראיים בת"ל מהסדרה $\{W_n\}$ עד כדי כפל בסקלר ולכן אין באמת הבדל בפילוג התוספת $X_\alpha(t)\sim X_\alpha(t_1,t_1+t)$ למעשה מה שחשוב זה **כמות** האיברים שמשתתפים בסכום **והמסקנה** היא שהתוספת

$$X_\alpha(t_1,t_2)\sim\mathcal{N}(0,\alpha(t_2-t_1))$$

בנוסף עבור אינטרואלים זרים $(t_1,t_2)\cap(t_3,t_4)=\emptyset$ התוספות של

$$\begin{aligned}&X_{T,d=\sqrt{\alpha T}}(t_1,t_2)\\&X_{T,d=\sqrt{\alpha T}}(t_3,t_4)\end{aligned}$$

הם בת"ס בגבול ש-$T\to 0$ כי הן תלויות ב-$\{W_k\}$ זרים.
ובפעול עבור $t_1\le t_2\le t_3\le t_4$

$$X_\alpha(t_1,t_2)\perp X_\alpha(t_3,t_4)$$

עד כה ראינו שהתוספות של התהליך בת"ס לאינטרוולים זרים ומתפלגות נורמלי $X_\alpha(t_1,t_2)\sim\mathcal{N}(0,\alpha(t_2-t_1))$
וכעת נוכל להסיק שהתהליך $X_\alpha(t)$ הוא תהליך אקראי גאוסי.
איך נסיק? נתבונן ב-k דגימות של התהליך

$$\begin{bmatrix}X_\alpha(t_1)\\X_\alpha(t_2)\\\vdots\\X_\alpha(t_k)\end{bmatrix}=\begin{bmatrix}1&0&0&\cdots&0\\1&1&0&\cdots&0\\&&\vdots&&\\1&1&\cdots&1&0\\1&1&\cdots&1&1\end{bmatrix}\cdot\begin{bmatrix}X_\alpha(0,t_1)\\X_\alpha(t_1,t_2)\\\vdots\\X_\alpha(t_{k-1},t_k)\end{bmatrix}$$

לדוגמא אם נרצה מה ערך התהליך בזמן k נחבר את כל התוספות עד t_k
בעתם יש לנו פה וקטור דגימות של התהליך שנתון ע"י טרנספורמציה לינארית דטרמיניסטית כפול וקטור שמכיל אך ורק תוספות זרות!
כל אחת מהקורדינטות של התוספות היא גאוסית שמתפלגת נורמילת ובגלל שהתוססות זרות כל אחד מהאיברים בלתי תלוי סטטיסטית באחרים וזה למעשה ו"א עם איברים גאוסיים בת"ס ולכן הוא בעצמו ו"א גאוסי!
וכשיש לנו טרנספורמציה לינארית של ו"א גאוסי, נקבל ו"א גאוסי!

3.6.5 תהליך וינר - הגדרה פורמלית

תהליך וינר $X(t)$ עם פרמטר $\alpha>0$ הוא תהליך אקראי עם $X(0)=0$ והתוספות שלו גאוסיות בתס

$$X(t_1,t_2)\sim\mathcal{N}(0,\alpha(t_2-t_1))\quad\forall_{0\le t_1\le t_2}$$

וכן מתקיים

$$\begin{aligned}\eta_X(t)&=0\ \forall_{t>0}\\R_X(t_1,t_2)&=\alpha\cdot\min\{t_1,t_2\}\quad\forall_{t_1,t_2\ge 0}\end{aligned}$$

4 רעש גאוסי לבן

4.1 הנגזרת של תהליך ווינר

יהיה $\{X(t)\}_{t>0}$ תהליך ווינר עם פרמטר $\alpha>0$

נגדיר את הביטוי הבא -

$$\begin{aligned} X_\epsilon(t) &= \frac{X(t+\epsilon) - X(t)}{\epsilon} \\ &= \frac{X(t, t+\epsilon)}{\epsilon} \end{aligned}$$

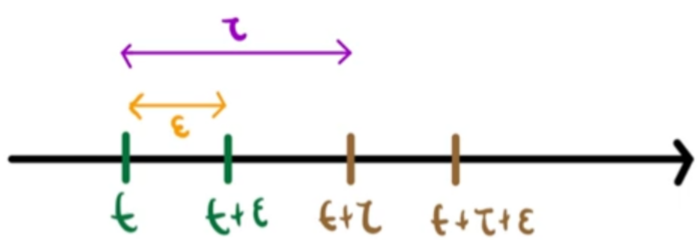
מכיוון ש-$\{X(t)\}_{t>0}$ הוא תהליך גאוסי, גם $\{X_\epsilon(t)\}_{t>0}$ אז כדי לחשב את הסטטיסטיקה המלאה שלו צריך רק לחשב את התוחלת והאוטוקורלציה!

$$\begin{aligned} \mathbb{E}[X_\epsilon(t)] &= \frac{1}{\epsilon}\left[\cancelto{0}{\mathbb{E}[X(t+\epsilon)]} - \cancelto{0}{\mathbb{E}[X(t)]}\right] \\ &= 0 \end{aligned}$$

$$\begin{aligned} R_{X_\epsilon}(t, t+\tau) &= \mathbb{E}[X_\epsilon(t) X_\epsilon(t+\tau)] \\ &= \mathbb{E}\left[\frac{X(t, t+\tau)}{\epsilon}\frac{X(t+\tau, t+\tau+\epsilon)}{\epsilon}\right] \\ &= \frac{1}{\epsilon^2}\mathbb{E}[X(t, t+\epsilon) X(t+\tau, t+\tau+\epsilon)] \end{aligned}$$

כעת נחלק ל-2 מקרים ונניח ש-$\tau > 0$:

1. $\tau > \epsilon$

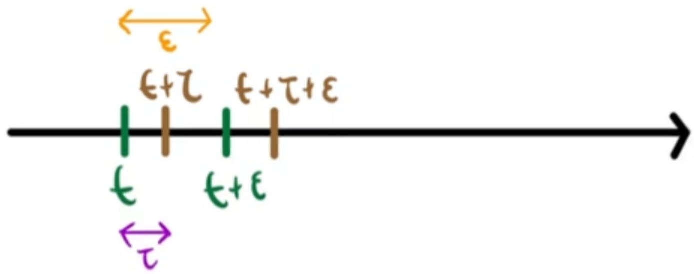


מה בעצם יש לנו כאן?

יש לנו 2 אינטרוולים שזרים ובתהליך ווינר כשהאינטרוולים זרים התוספות בת"ס - $(t, t+\epsilon)$ ו-$(t+\tau, t+\tau+\epsilon)$ ומכיוון שלכל אחת מהתוספות יש תוחלת אפס והאוטוקורלציה זה מכפלת התוספות נקבל ש-

$$\boxed{R_{X_\epsilon}(t, t+\tau) = 0 \quad \forall t}$$

2. $0 < \tau \leq \epsilon$



כאן האינטרוולים נחתכים ולכן נחלק את כל הסיפור הזה ל-3 אינטרוולים זרים - $(t,t+\tau)$, $(t+\tau,t+\epsilon)$ ו-$(t+\epsilon,t+\tau+\epsilon)$ ונבחין כי

$$
\begin{aligned}
X\left(t,t+\epsilon\right) &= X\left(t,\overset{1}{t}+\tau\right)+X\left(t+\overset{2}{\tau},t+\epsilon\right) \\
X\left(t+\tau,t+\tau+\epsilon\right) &= X\left(t+\overset{2}{\tau},t+\epsilon\right)+X\left(t+\epsilon,\overset{3}{t}+\tau,t+\epsilon\right)
\end{aligned}
$$

נזכיר כי שלושת האינטרוולים $(1,2,3)$ הם זרים ומכיוון שהתהליך ווינר עם תוספות בת"ס התוספות 1,2,3 הם בתס כלומר

$$
\begin{aligned}
\mathbb{E}\left[X\left(t,t+\epsilon\right)X\left(t+\tau,t+\tau+\epsilon\right)\right] &= \mathbb{E}\left[\left((1)+(2)\right)\cdot\left((2)+(3)\right)\right] \\
&= \cancel{\mathbb{E}\left[\left((1)\cdot(3)\right)\right]}+\cancel{\mathbb{E}\left[\left((1)\cdot(2)\right)\right]}+\cancel{\mathbb{E}\left[\left((2)\cdot(3)\right)\right]}+\mathbb{E}\left[\left((2)\cdot(2)\right)\right] \\
&= \mathbb{E}\left[\left((2)\cdot(2)\right)\right] \\
&= \mathbb{E}\left[X^2\left(t+\tau,t+\epsilon\right)\right] \\
because \Downarrow &= \alpha\left(\epsilon-\tau\right) \\
X\left(t+\tau,t+\epsilon\right) &\sim \mathcal{N}\left(0,\alpha\left(\epsilon-\tau\right)\right)
\end{aligned}
$$

כלומר, נקבל ש-

$$
\begin{aligned}
R_{X_\epsilon}\left(t,t+\tau\right) &= \frac{1}{\epsilon^2}\left(\alpha\left(\epsilon-\tau\right)\right) \\
&= \frac{\alpha}{\epsilon}\left(1-\frac{\tau}{\epsilon}\right)
\end{aligned}
$$

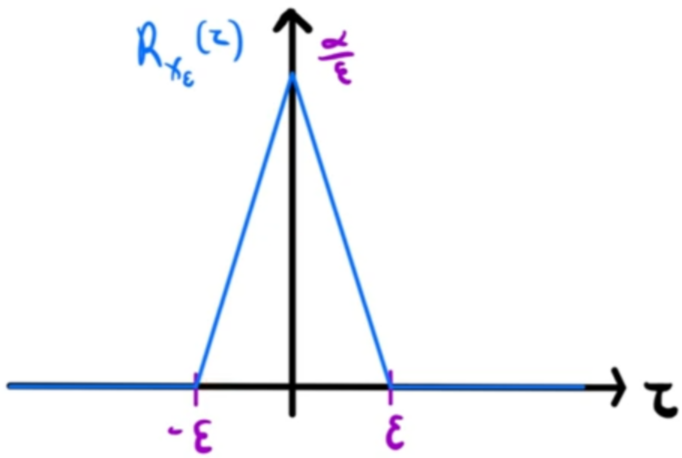
לסיכום עבור $\tau\geq 0$

$$
R_{X_\epsilon}\left(t,t+\tau\right)=R_{X_\epsilon}\left(\tau\right)\begin{cases}\frac{\alpha}{\epsilon}\left(1-\frac{\tau}{\epsilon}\right) & \tau\leq\epsilon \\ 0 & \tau>\epsilon\end{cases}
$$

באותה צורה בדיוק ניתן להראות ש

$$
R_{X_\epsilon}\left(t,t+\tau\right)=R_{X_\epsilon}\left(\tau\right)\begin{cases}\frac{\alpha}{\epsilon}\left(1-\frac{|\tau|}{\epsilon}\right) & |\tau|\leq\epsilon \\ 0 & |\tau|>\epsilon\end{cases}
$$

קיבלנו שפונקציית האוטוקורלציה נראת ככה -



4.2 תהליך הנגזרת

$$X'(t) = \frac{d}{dt}X(t)$$
$$= \lim_{\epsilon=0} X_\epsilon(t)$$

התהליך $X'(t)$ הוא גאוסי מכאן נראה ש-

$$R_{X'}(\tau) = \lim_{\epsilon=0} X_\epsilon(\tau)$$
$$\mathbb{E}[X'(t)] = 0$$

מה יקרה בגבול של ϵ במשולש שלנו? הפיק יתפוס גובה והבסיס ל-0, אבל האינטגרל נשאר קבוע:

$$\int_{-\infty}^{\infty} R_{X_\epsilon}(\tau)\, d\tau = \frac{1}{2}(2\epsilon)\frac{\alpha}{\epsilon}$$
$$= \alpha\ \forall\epsilon > 0$$

והגבול של אינטגרל זה הוא דלתא!

$$\lim_{\epsilon=0} R_{X'}(\tau) = \alpha\delta(\tau)$$

כלומר, $\{X'(t)\}$הוא תהליך אקראי גאוסי עם תוחלת 0 ופונקציית אוטוקורלציה

$$R_{X'}(\tau) = \alpha\delta(\tau)$$

תהליך זה נקרא רעש גאוסי לבן!

תהליך כזה הוא אינו פיזיקלי!
אם $\{N(t)\}$ הוא רעש גאוסי לבן מתקיים

- לכל $t_1 \neq t_2$ $N(t_1)$ בת"ל $N(t_2)$

• בנוסף אם נתבונן על השונות נקבל שהיא אין-סוף

$$\begin{aligned}\mathbb{E}\left[N^2(t)\right] &= R_N(0) \\ &= \alpha\delta(0) \\ &= \infty\end{aligned}$$

מסקנה - רעש גאוסי לבן אינו פיזיקלי!

למרות זאת זה מודל מאוד שימושי!

5 ספקטרום צפיפות הספק - Power Spectrum Density

5.1 אותות דטרמיניסטים

עבור אות נתון $\{X(t)\}$ ממשי, ניתן להגדיר את התמרת הפוריה

$$X(\omega) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt$$

ונוכל להגדיר את

$$A(\omega) = \left|X^2(\omega)\right|$$

תכונות

- $A(\omega)$ הוא ממשי
- $A(\omega) = A(-\omega)$ - פונקציה זוגית - זה מתקיים כי $|X(\omega)| = |X(-\omega)|$
- פרסבל -

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} A(\omega)d\omega = \int_{-\infty}^{\infty}\left|x^2(t)\right| dt$$

כעת ניתן לאפיין את האנרגיה של האות $x(t)$ בקטע תדר קטן סביב ω_0, כלומר בתחום התדר אם ניקח אינטרוול תדר קטן $d\omega$- התמרת פוריה תהיה פחות או יותר קבועה, ונקבל שהתמורה של קטע זה יהיה $\frac{1}{2\pi}A(\omega_0)\, d\omega$

ניתקל בבעיה עבור אות מחזורי או אות שאינו דועך ולכן נגדיר את האות הסופי בזמן $x_T(t)$ בצורה הבאה

$$x_T(t) \begin{cases} x(t) & 0 \le t \le T \\ 0 & o.w. \end{cases}$$

לאות זה ניתן לחשב את האנרגיה כי הוא סופי ואינטגרבילי ולכן התמרת הפוריה קיימת, נסמן אותה

$$\begin{aligned}X_T(\omega) &= \mathcal{F}\{x_T(t)\} \\ &= \int_{-\infty}^{\infty} x_T(t)\, e^{-j\omega t} dt \\ &= \int_{-\infty}^{T} x_T(t)\, e^{-j\omega t} dt\end{aligned}$$

כלומר, לכל $T < \infty$ התמרת פורייה של $X_T(\omega)$של $\{x_t(t)\}$קיימת

כעת נגדיר גם פונקציית צפיפות ההספק של האות להיות

$$\tilde{S}_x(\omega) = \lim_{T\to\infty}\frac{1}{T}\left|X_T(\omega)\right|^2$$

5.2 תהליכים אקראיים

עבור תהליך אקראי $\{X\left(t\right)\}$ נגדיר

$$X_{T}\left(t\right)\begin{cases}X\left(t\right) & 0\leq t\leq T\\0 & o.w.\end{cases}$$

והתמרת הפוריה שלו תהיה

$$\begin{aligned}X_{T}\left(\omega\right) & =\mathcal{F}\left\{X_{T}\left(t\right)\right\}\\ & =\int_{0}^{T}x_{T}\left(t\right)e^{-j\omega t}dt\end{aligned}$$

היות ש-$\{X\left(t\right)\}$ אקראי, גם $\{X_{T}\left(\omega\right)\}$ הוא תהליך אקראי בתחום התדר.

5.2.1 נגדיר power spectrum density

של התהליך האקראי $\{X\left(t\right)\}$להיות

$$S_{X}(\omega)=\lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left|X_{T}(\omega)\right|^{2}$$

לכל ω נקבל גודל דטרמניסטי
באותו אופן עבור תהליך אקראי בזמן בדיד $\{X_{n}\}_{n=0}^{\infty}$ נגדיר את

$$N_{n}^{N}=\begin{cases}X_{n} & 0\leq n\leq N-1\\0 & o.w.\end{cases}$$

ונגדיר

$$\begin{aligned}X^{N}\left(\omega\right) & =\mathcal{F}\left\{X_{n}^{N}\right\}\\ & =\sum_{n=-\infty}^{\infty}X_{n}^{N}e^{-i\omega n}\\ & =\sum_{n=0}^{N-1}X_{n}^{N}e^{-i\omega n}\end{aligned}$$

ונגדיר את ה-PSD של תהליך אקראי בזמן בדיד להיות

$$S_{X}\left(\omega\right)=\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left|X^{N}\left(\omega\right)\right|^{2}$$

(הרצאה 11, $24:00$)

5.2.2 משפט ווינר-חינצ'ין

עבור תהליך אקראי WSS בזמן רציף

$$\begin{aligned}(*)\,S_{X}\left(\omega\right) & =\int_{-\infty}^{\infty}R_{x}\left(\tau\right)e^{-j\omega\tau}d\tau\\ & =\mathcal{F}\left\{R_{x}\left(\tau\right)\right\}\end{aligned}$$

הערה! תחת תנאים טכניים קלים שנראה בהמשך - ככלל כל התהליכים שנראה בקורס יעמדו בתנאים עבור תהליך אקראי WSS בזמן בדיד $\{X_n\}$ נקבל

$$(**)\,S_X\left(\omega\right)=\sum_{-\infty}^{\infty}R_x\left[k\right]e^{-j\omega k}$$
$$=\mathcal{F}\left\{R_x\left[k\right]\right\}$$
$$R_X\left[k\right]=\mathbb{E}\left[X_nX_{n+t}\right]$$

מעתה נסמן את $(*)$ ואת $(**)$ כהגדרה של ה-PSD של תהליך WSS

הוכחה - קלילה למקרה הבדיד צריך להוכיח -

$$\lim_{N\to\infty}\frac{1}{N}\mathbb{E}\left|X^N\left(\omega\right)\right|^2=\sum_{-\infty}^{\infty}R_x\left[k\right]e^{-j\omega k}$$

עבור N קבוע נקבל

$$\frac{1}{N}\mathbb{E}\left|X^N\left(\omega\right)\right|^2=\frac{1}{N}\mathbb{E}\left|\sum_{0}^{N-1}X_ne^{-j\omega k}\right|^2$$
$$|a|^2=a\cdot a^*$$
$$\Downarrow$$
$$=\frac{1}{N}\mathbb{E}\left[\sum_{m=0}^{N-1}X_me^{-j\omega m}\left(\sum_{\ell=0}^{N-1}X_\ell e^{-j\omega\ell}\right)^*\right]$$
$$=\frac{1}{N}\mathbb{E}\left[\sum_{m=0}^{N-1}X_me^{-j\omega m}\sum_{\ell=0}^{N-1}X_\ell e^{j\omega\ell}\right]$$
$$=\frac{1}{N}\mathbb{E}\left[\sum_{m=0}^{N-1}\sum_{\ell=0}^{N-1}X_mX_\ell e^{-j\omega(m-\ell)}\right]$$
$$=\frac{1}{N}\sum_{m=0}^{N-1}\sum_{\ell=0}^{N-1}\mathbb{E}\left(X_mX_\ell\right)e^{-j\omega(m-\ell)}$$
$$=\frac{1}{N}\sum_{m=0}^{N-1}\sum_{\ell=0}^{N-1}R_X[m-\ell]e^{-j\omega(m-\ell)}$$
$$*lec\,11\,\text{40:00}=\sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)R_X[k]e^{-j\omega k}$$

$$=\frac{1}{N}\mathbb{E}\left|\sum_{m=0}^{N-1}X_me^{-j\omega m}\right|^2$$

כעת נציין את התנאים הטכניים

- אם $R_{X[k]}$ דועך, ספציפית אם $\sum_{k=-\infty}^{\infty}\left|R_{X[k]}\right|<\infty$ מתקיים

$$\lim_{N\to\infty}\sum_{k=-N}^{N}\left(1-\frac{|k|}{N}\right)R_{X[k]}e^{-j\omega k}=\mathcal{F}\left\{R_{X[k]}\right\}$$

לסיכום
עבור WSS רציף נקבל

$$(*)\, S_X(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left|X_T(\omega)\right|^2$$

ועבור WSS בדיד נקבל

$$(**)\, S_X(\omega) = \sum_{-\infty}^{\infty} R_x[k] e^{-j\omega k} = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left|X^N(\omega)\right|^2$$

5.2.3 תכונות של PSD

1. $S_X(\omega)$ היא פונקציה ממשית אי-שלילית
2. $S_X(\omega) = S_X(-\omega)$
3. נספק של התהליך האקראי מקיים ש - במקרה הרציף:

$$\mathbb{E}\left[X^2(t)\right] = R_X(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_X(\omega)\, d\omega$$

ובמקרה הבדיד

$$\mathbb{E}\left[X_n^2\right] = R_X(0) = \frac{1}{2\pi}\int_{-k}^{k} S_X(\omega)\, d\omega$$

הערה! סטטסטיקה מסדר שני של תהליך אקראי זה ה-$\mathbb{E}[X(t)]$ וה-$R_X(t, t+\tau)$
עבור תהליך WSS זה שקול ל-$\mathbb{E}[X(t)]$ ו- $S_X(\omega)$ שבמקרה שלנו

$$R_X(\tau) = \mathcal{F}^{-1}\{S_X(\omega)\}$$

למעשה כל מה שצריך זה התוחלת ואת $S_X(\omega)$

5.2.4 דוגמא

רעש גאוסי לבן:

$$\{N(t)\}_{t\geq 0}$$

אם מתקיים $\mathbb{E}\left[N\left(t\right)\right]=0$ וכן $R_{N}\left(\tau\right)=\alpha\cdot\delta\left(\tau\right)$ איך יראה הספקטרום של תהליך כזה?

$$\begin{aligned}S_{N}\left(\omega\right) & =\mathcal{F}\left\{ R_{N}\left(\tau\right)\right\} \\ & =\mathcal{F}\left\{ \alpha\cdot\delta\left(\tau\right)\right\} \\ & =\alpha\cdot\mathcal{F}\left\{ \delta\left(\tau\right)\right\} \\ & =\alpha\ \forall\omega\in\mathbb{R}\end{aligned}$$

ומכאן השם רעש לבן כי בתחום התדר נקבל פונקציה קבועה בגובה α, בדיוק כמו אור לבן
נהוג להחליף את α ב- $\frac{N_{0}}{2}$

5.3 תהליכים $JWSS$ וקרוס ספקטרום

נאמר שזוג תהליכים אקראיים $\left\{ X\left(t\right)\right\}$ ו-$\left\{ Y\left(t\right)\right\}$ הם $JWSS$ אם מתקיימים שתי התכונות הבאות -

- כל אחד מהם הוא WSS בעצמו
- פונקציית הקרוס קורלציה תלוייה רק בהפרש הזמנים, כלומר,

$$\begin{aligned}R_{X,Y}\left(t,t+\tau\right) & =\mathbb{E}\left[X\left(t\right)Y\left(t+\tau\right)\right]\\ & =R_{XY}\left(\tau\right)\end{aligned}$$

הערה:

פונקציית הקרוס-קורלציה אינה פונקציה זוגית עבור $JWSS$ אבל,

$$\begin{aligned}R_{XY}\left(\tau\right) & =\mathbb{E}\left[X\left(t\right)Y\left(t+\tau\right)\right]\\ & =\mathbb{E}\left[Y\left(t+\tau\right)X\left(t\right)\right]\\ & =R_{YX}\left(-\tau\right)\end{aligned}$$

עבור $\left\{ X\left(t\right)\right\}$ ו-$\left\{ Y\left(t\right)\right\}$ $JWSS$ נגדיר את פונק' הקרוס ספקטרום להיות
במקרה הרציף -

$$\begin{aligned}S_{XY}\left(\omega\right) & =\mathcal{F}\left\{ R_{XY}\left(\tau\right)\right\} \\ & =\int_{-\infty}^{\infty}R_{XY}\left(\tau\right)e^{-j\omega\tau}d\tau\end{aligned}$$

ובמקרה הבדיד -

$$\begin{aligned}S_{XY}\left(\omega\right) & =\mathcal{F}\left\{ R_{XY}\left[k\right]\right\} \\ & =\sum_{-\infty}^{\infty}R_{XY}\left[k\right]e^{-j\omega k}\end{aligned}$$

6 מעבר תהליך אקראי במערכת LTI

6.1 חזרה על מערכות LTI - המקרה הדטרמניסטי

מערכת מערכת - נסמנה ב-T היא מיפוי מפונקציה בזמן, למשל $\left\{ x\left(t\right)\right\}$ לפונק' אחרת בזמן $\left\{ y\left(t\right)\right\}$

לדוגמא

$$y(t) = x^2(t) \ \ \forall t$$

נבחין שמערכת זו היא לא לינארית! ואנחנו נתעניין במערכות כן לינאריות בקורס זה

מערכת לינארית מערכת שלכל $\{x_1(t)\}$ ו-$\{x_2(t)\}$ ולכל $\alpha, \beta \in \mathbb{R}$ מתקיים

$$T\{\alpha x_1(t) + \beta x_2(t)\} = \alpha T\{x_1(t)\} + \beta T\{x_2(t)\}$$

בעצם המערכת הלינארית "עובדת" בנפרד על כל אות

מערכת time invariant היא מערת עבורה אם $T\{x(t)\} = \{y(t)\}$ אזי לכל $\tau \in \mathbb{R}$ מתקיים

$$T\{x(t-\tau)\} = \{y(t-\tau)\}$$

כלומר הזזה של זמן כניסה מתרגמת להזזה של אותו זמן במוצא.

מערכת LTI היא גם לינארית וגם TI

התגובה להלם - impulse response התגובה של מערכת LTI להלם מוגדרת כ-

$$\{h(t)\} = T\{\delta(t)\}$$

בנוסף נזכיר תכונה של δ

$$\int_{-\infty}^{\infty} f(x)\,\delta(x-x_0)\,dx = f(x_0)$$

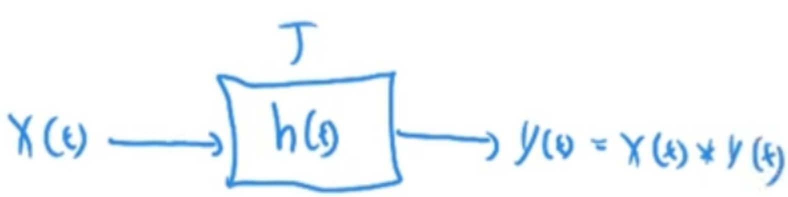
בעצם הדלתא "דוגמת" את הפונקציה

תכונה בסיסית של מערכת LTI לכל אות כניסה $\{x_1(t)\}$ אות המוצא יהיה - קונבולוציה

$$\begin{aligned} y(t) &= T\{x(t)\} \\ &= \int_{-\infty}^{\infty} \underbrace{h(\tau)}_{T\{\delta(\tau)\}}\ x(t-\tau)\,d\tau \\ &= [h*x](t) \\ &= h(t)*x(t) \end{aligned}$$

6.2 תגובה של מערכת LTI לכניסה אקראית

לעיתים נצייר את המערכת



לעיתים קרובות לא נדע מה יהיה הכניסה למערכת ולכן נסמן כניסה זו כתהליך אקראי ונסמן אותו כ-$X(t)$ ובאותו אופן את המוצא $Y(t)$

- מעתה נניח שתהליך הכניסה שלנו הוא $\{X(t)\}$ הוא WSS
- נראה ש-$\{Y(t)\}$ הוא WSS ולמעשה $\{Y(t)\},\{X(t)\}$ הם JWSS
- נקבל נוסחאות פשוטות לחישוב
  - $\eta_Y, R_Y(\tau), R_{XY}(\tau)$ מתוך הסטטיסטיקה מסדר שני של $\{X(t)\}$, כלומר, מ-$\{h(\tau)\}, R_X(\tau), \eta_X$

6.2.1 חישוב התוחלת

[להכניס תמונה]

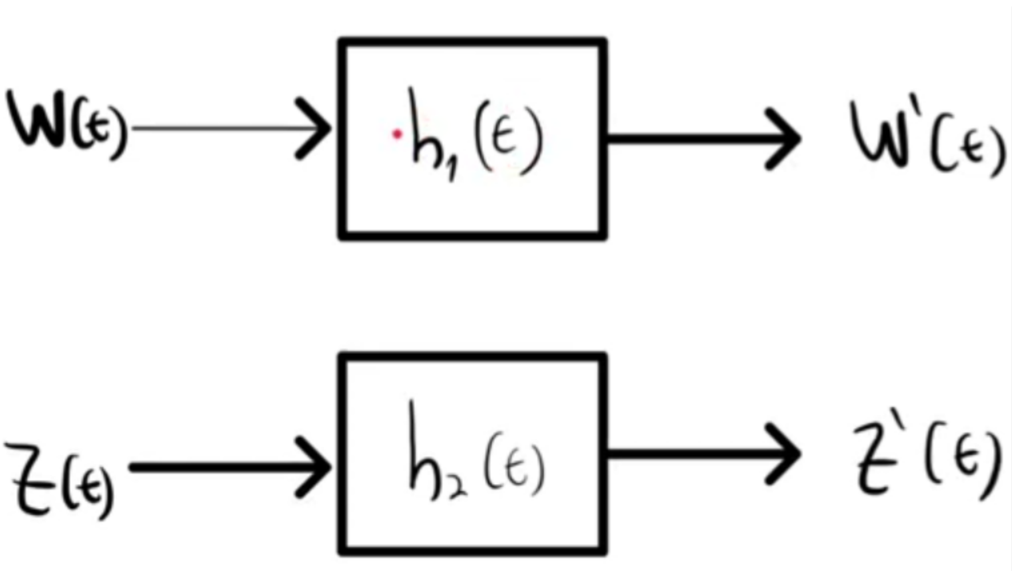
$$\begin{aligned}\mathbb{E}[Y(t)] &= \mathbb{E}[[h*x](t)] \\ &= \mathbb{E}\left[\int_{-\infty}^{\infty} h(\tau)\,x(t-\tau)\,d\tau\right] \\ * &= \int_{-\infty}^{\infty} h(\tau)\,\mathbb{E}[x(t-\tau)]\,d\tau \\ ** &= \int_{-\infty}^{\infty} h(\tau)\,\eta_X\,d\tau \\ &= \eta_X\int_{-\infty}^{\infty} h(\tau)\,d\tau\end{aligned}$$

* מהחת פוביני - ניתן להחליף סדר של אינטגרל תחת תנאים מסויימים, בקורס שלנו נניח שניתן ולא ניכנס לדקויות, אך זה לא תמיד נכון

** היות ו-X הוא תהליך WSS התוחלת שלו לכל τ, t קבועה ולא תלויה בהם
התוחלת לא תלויה ב-t!

6.2.2 חישוב האוטוקורלציה והקרוסקורלציה - $R_{XY}(\tau), R_Y(\tau)$

נעשה חישוב כללי וממנו נוציא מקרה פרטי של התוצאה
נתבונן במערכת הבאה



למערכת הראשונה נכניס תהליך אקראי WSS עם פונקציית אוטוקורלציה של $R_W\left(\tau\right)$ ובאותו אופן למערכת השנייה $R_Z\left(\tau\right)$ שני התהליכים ביחד הם $JWSS$ עם פונקציית קרוסקורלציה $R_{WZ}\left(\tau\right)=\mathbb{E}\left[W\left(t\right)Z\left(t+\tau\right)\right]$

במוצא נקבל את W' ואת Z' ונרצה לחשב את הקרוסקורלציה שלהם $R_{W'Z'}(\tau)$

$$\begin{aligned}
R_{W'Z'}(\tau) &= \mathbb{E}\left[W'(t)\,Z'(t+\tau)\right] \\
&= \mathbb{E}\left[\left[W*h_1\right](t)\cdot\left[Z*h_2\right](t+\tau)\right] \\
&= \mathbb{E}\left[\left(\int_{-\infty}^{\infty} h_1(\alpha)\,W(t-\alpha)\,d\alpha\right)\cdot\left(\int_{-\infty}^{\infty} h_2(\beta)\,Z(t+\tau-\beta)\,d\beta\right)\right] \\
&= \mathbb{E}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(\alpha)\,h_2(\beta)\,W(t-\alpha)\,Z(t+\tau-\beta)\,d\alpha d\beta\right] \\
* &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_1(\alpha)\,h_2(\beta)\cdot R_{WZ}(\tau-\beta+\alpha)\,d\alpha d\beta \\
** &= \int_{-\infty}^{\infty} h_1(\alpha)\,d\alpha\,\underbrace{\left[\int_{-\infty}^{\infty} h_2(\beta)\cdot R_{WZ}(\tau-\beta+\alpha)\,d\beta\right]}_{\text{convolution}} \\
&= \int_{-\infty}^{\infty} h_1(\alpha)\left[h_2*R_{WZ}\right](\tau+\alpha)\,d\alpha \\
*** &= \int_{-\infty}^{\infty} h_1(-\gamma)\left[h_2*R_{WZ}\right](\tau-\gamma)\,d\gamma \\
\text{let}\, g(t) &= h_1(-t) \\
&\Downarrow \\
^{*}\text{convolution} &= \int_{-\infty}^{\infty} g(\gamma)\left[h_2*R_{WZ}\right](\tau-\gamma)\,d\gamma \\
&= \left[g*\left[h_2*R_{WZ}\right]\right](\tau) \\
&= h_1(-\tau)*R_{WZ}(\tau)*h_2(\tau) \\
R_{W'Z'}(\tau) &= h_1(-\tau)*R_{WZ}(\tau)*h_2(\tau)
\end{aligned}$$

*שינוי סדר בלי לשאול שאלות - פוביני

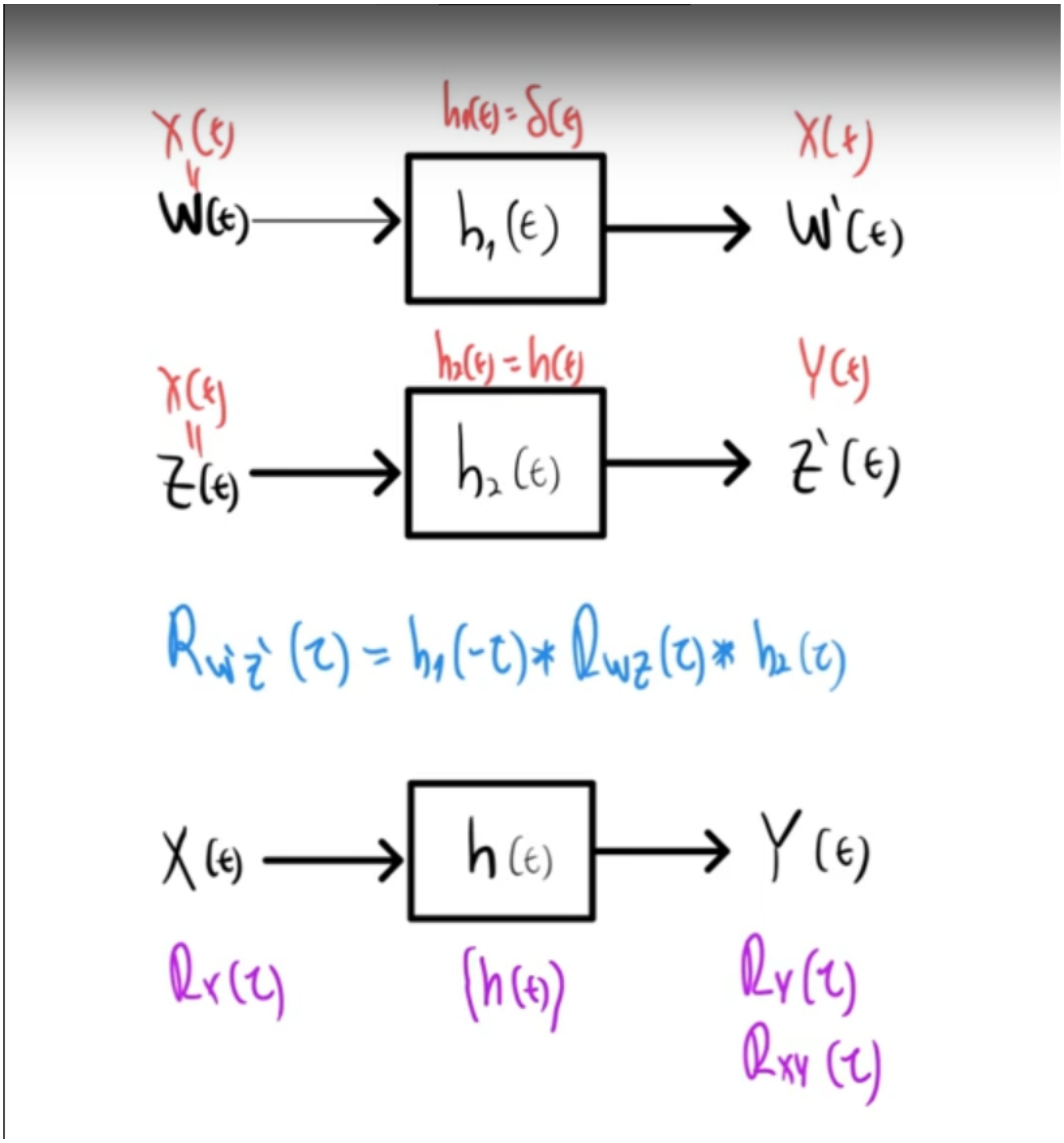
** בעצם 2 קונבולוציות

*** החלפת משתנה, האינטגרל בתחום $(-\infty,\infty)$ ולכן זה לא משנה את סימן האינטגרל

6.2.3 חישוב האוטוקורלציה והקרוסקורלציה - $R_{XY}(\tau), R_Y(\tau)$ - עכשיו באמת!

אם נתון לנו $\{h(t)\}$ ואת $\{R_X(\tau)\}$

נתבונן במערכת הבאה ונשתמש בפיתוח שעשינו -



למעשה נרצה לחשב מהפיתוח שעשינו - $R_{W'Z'}=h_1(-\tau)*R_{WZ}(\tau)*h_2(\tau)$ את $R_{XY}(\tau)=\delta(-\tau)*R_X(\tau)*h(\tau)$

$$
\begin{aligned}
R_{W'Z'} &= h_1(-\tau)*R_{WZ}(\tau)*h_2(\tau)\\
&\Downarrow\\
R_{XY}(\tau) &= \delta(-\tau)*R_X(\tau)*h(\tau)\\
&= R_X(\tau)*h(\tau)
\end{aligned}
$$

אזהרה!

אצלנו בקורס נסמן ונשתמש

$$R_{XY}(\tau)=\mathbb{E}\left[X(t)\,Y(t+\tau)\right]$$

לפעמים בספרות מסמנים דווקא הפוך

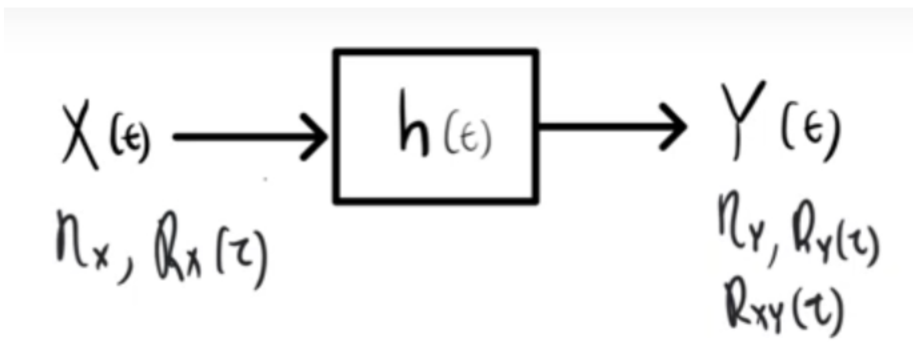
$$R_{XY}(\tau) = \mathbb{E}\left[X(t+\tau)\,Y(t)\right]$$

כעת נחשב את $R_Y(\tau)$

באותו אופן רק שכעת נסמן את $h_1(t) = h(t)$ ונקבל ש-

$$R_Y(\tau) = h(-\tau) * R_X(\tau) * h(\tau)$$

לסיכום



יש לנו תהליך אקראי $X(t)$ שהוא WSS שהתוחלת שלו הוא η_X ואוטוקורלציה $R_X(\tau)$ התהליך עובר במערכת LTI עם תגובה להלם $h(t)$ ומוצא המערכת הוא $Y(t)$ ראינו ש-$Y(t)$ הוא גם WSS ולמעשה $JWSS$ עם $X(t)$ וחישבנו את

$$\eta_Y = \eta_X \int_{-\infty}^{\infty} h(\tau)\,d\tau$$

$$R_Y(\tau) = h(-\tau) * R_X(\tau) * h(\tau)$$

$$R_{XY}(\tau) = R_X(\tau) * h(\tau)$$

במקרה הבדיד הנוסחאות הם בדיוק אותו הדבר רק שהאינטגרל מוחלף בסכום

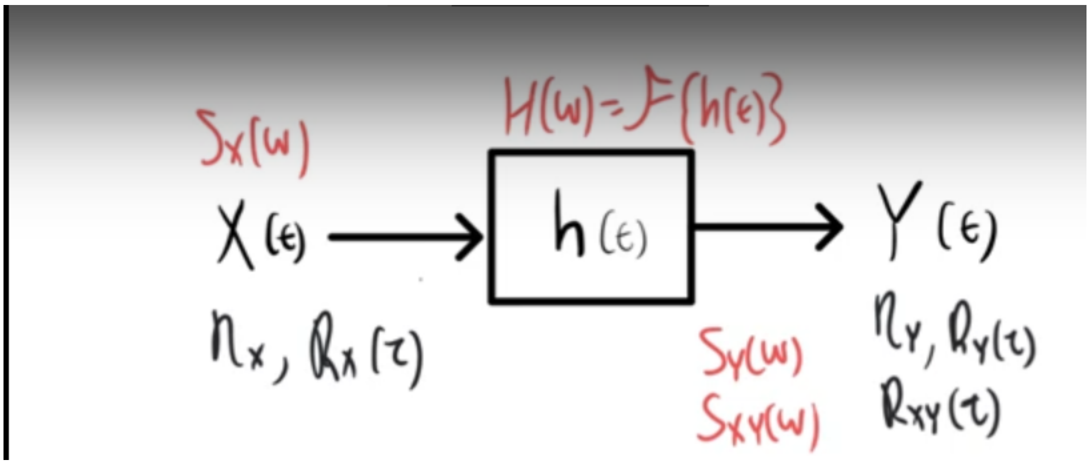
$$\eta_Y = \eta_X \sum_{\ell=-\infty}^{\infty} h_\ell$$

$$R_{XY}[k] = R_X[k] * h[-k]$$

$$R_Y[k] = h[-k] * R_X[k] * h[k]$$

6.2.4 עבור PSD ו-קרוס-ספקטרום

אם נרצה לחשב כפונקציה של הספקטרום $X\left(t\right)$, כלומר את $S_X\left(\omega\right)$ נתבונן בהתמרת פוריה שלו $H\left(\omega\right)=\mathcal{F}\left\{h\left(t\right)\right\}$



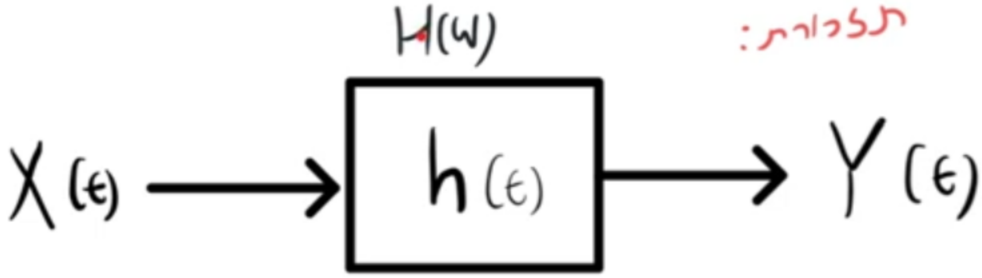
נכתוב בתחום התדר מה קורה -

$$\begin{aligned}\eta_Y&=\eta_X\cdot H\left(0\right)\\ S_{XY}\left(\omega\right)&=\mathcal{F}\left\{R_{XY}\left(\tau\right)\right\}\\ &=H\left(\omega\right)\cdot S_X\left(\omega\right)\\ S_Y\left(\omega\right)&=\mathcal{F}\left\{R_Y\left(\tau\right)\right\}\\ &=\mathcal{F}\left\{h\left(-\tau\right)\right\}\cdot\mathcal{F}\left\{R_X\left(\tau\right)\right\}\cdot\mathcal{F}\left\{h\left(\tau\right)\right\}\\ &=H^*\left(\omega\right)\cdot S_X\left(\omega\right)\cdot H\left(\omega\right)\\ &=\left|H\left(\omega\right)\right|^2\cdot S_X\left(\omega\right)\end{aligned}$$

6.3 ייצור תא"ג עם סטטיסטיקה רצויה וסינון צר סרט

תזכורת -

תזכורת:



$$\eta_Y = H(0)\cdot \eta_X$$

$$S_{XY}(\omega) = H(\omega)\cdot S_X(\omega)$$

$$S_Y(\omega) = H^*(\omega)\cdot H(\omega)\cdot S_X(\omega) = |H(\omega)|^2\cdot S_X(\omega)$$

דוגמא - נתבונן בתהליך $AR(\alpha)$ בזמן בדיד, כאשר $0<\alpha<1$,
נגדיר $\{X_n\}$ המוגדר להיות

$$X_n = \alpha X_{n-1} + Z_n$$

כאשר Z_n הוא רעש לבן בזמן בדיד, כלומר, מתקיים ש-

$$\begin{aligned} \mathbb{E}[Z_n] &= 0 \ \forall n \\ R_Z &= \sigma^2 \cdot \delta[k] \\ S_Z(\omega) &= \sigma^2 \ \forall \omega \end{aligned}$$

כלומר כל זוג דגימת של Z_n הן חסרות קורלציה
נרצה להראות שניתן לייצר את X_n ע"י מערכת עם כניסה Z_n ומערכת עם תגובה להלם $H(\omega)$

$$Z_n \to \boxed{H(\omega)} \to X_n$$

נרצה לפתור את משוואת ההפרשים, נעביר אגפים ונפתור

$$\begin{aligned} X_n - \alpha X_{n-1} &= Z_n \\ X_n * [\delta_n - \alpha\delta_{n-1}] &= Z_n \\ H(\omega) &= \frac{1}{\mathcal{F}\{\delta_n - \alpha\delta_{n-1}\}} \\ &= \frac{1}{1-\alpha e^{-j\omega}} \end{aligned}$$

ונרצה לחשב את ה-PSD כלומר את $S_X(\omega)$ ונקבל (מהתמונה למעלה, שהיא התזכורת מסומן ב-$S_Y(\omega)$

$$\begin{aligned} S_X(\omega) &= H^*(\omega) \cdot H(\omega) \cdot S_Z(\omega) \\ &= \frac{1}{1-\alpha e^{j\omega}} \cdot \frac{1}{1-\alpha e^{-j\omega}} \cdot \sigma^2 \\ &= \frac{\sigma^2}{1+\alpha^2-\alpha\left(e^{j\omega}+e^{-j\omega}\right)} \\ &= \frac{\sigma^2}{1+\alpha^2-2\alpha\cos(\omega)} \end{aligned}$$

נתבונן בביטוי שקיבלנו

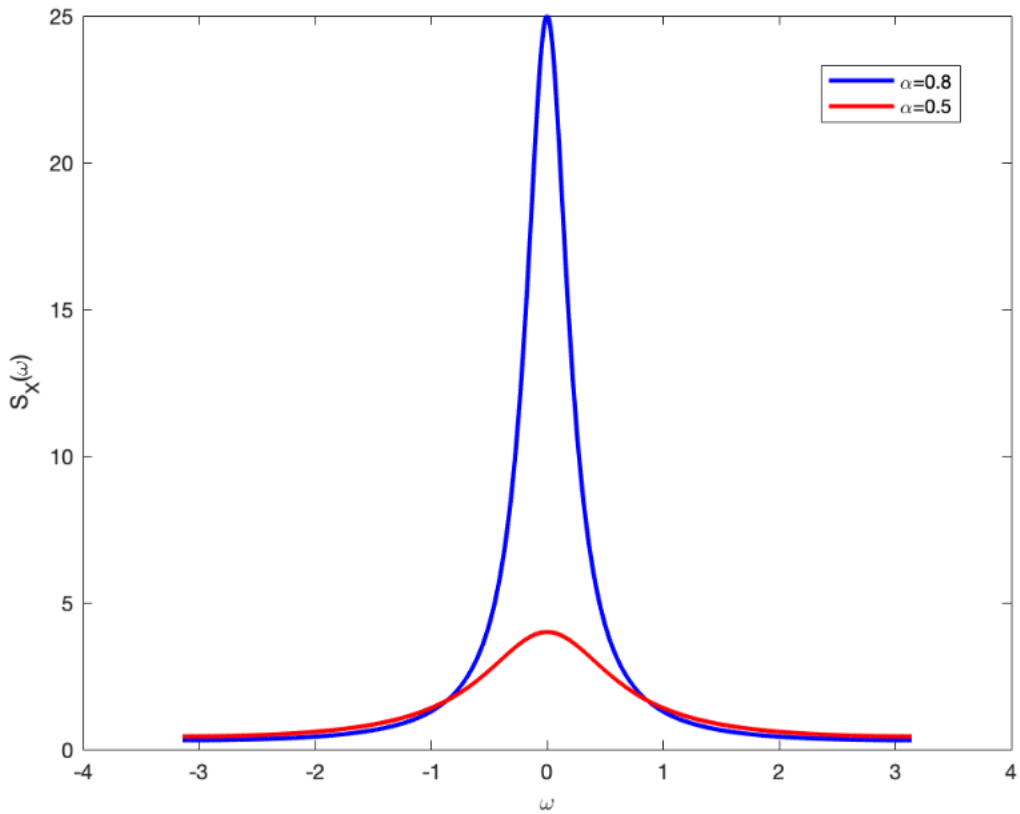


Figure 4: PSD of an AR(α) process.

ככל ש-α גדול יותר התהליך צר סרט - אפשר לחשוב על זה שלתהליך יש הרבה זכרון וזה מתרגם ב-PSD לרוחב סרט קטן וכש-α קטן התהליך יהיה שטוח ולא יהיה זכרון בכלל.

6.4 ייצור של תא"ג עם סטטיסטיקה נתונה

תזכורת - איך מייצרים

$$\underline{X} \sim \mathcal{N}\left(\underline{\eta}, \sum\right)$$

מתוך וקטור אקראי גאוסי iid

$$\underline{X} \sim \mathcal{N}\left(\underline{0}, \underline{I}\right)$$

מתכון:

- מוצאים מטריצה $\underline{A}$ שמקיימת

$$\sum = \underline{A} \cdot \underline{A}^T$$

ומחשבים את

$$X = \underbrace{\underline{A} \cdot \underline{Z}}_{\mathcal{N} \sim \left(\underline{\eta}, \sum\right)} + \underline{\eta}$$

היום - נרצה לענות על השאלה, איך מייצרים תהליך אקראי גאוסי $\{X(t)\}$ עם $R_X(\tau)$ ותוחלת 0 מתוך רעש גאוסי לבן $\{Z(t)\}$ עם תוחלת 0 ו-$R_Z(\tau) = \frac{N_0}{2}\delta(\tau)$

נשתמש במערכת LTI

$$Z(t) \longrightarrow \overset{H(\omega)}{\boxed{h(t)}} \longrightarrow X(t)$$

איך לבחור את $\{h(t)\}$?!
זה מסתבר די פשוט -

$$\begin{aligned} R_X(\tau) &= h(-\tau) * R_Z(\delta) * h(\tau) \\ &= \frac{N_0}{2} \underbrace{h(-\tau) * h(\tau)}_{\text{we want this to be } \frac{2}{N_0} \cdot R_X(\tau)} \end{aligned}$$

אז נמצא מסנן $\{h(t)\}$ שמקיים

$$R_X(\tau) = \frac{2}{N_0} \cdot h(-\tau) * h(\tau)$$

בנוסף נקבל ש-$\{X(t)\}$ הוא תא"ג כמעבר של תא"ג במערכת $LTI \Leftarrow$ ל-$\{X(t)\}$ יש את הסטטיסטיקה הרצויה.

בתחום התדר נקבל

$$\begin{aligned} S_X(\omega) &= \mathcal{F}\{R_X(\tau)\} \\ &= |H(\omega)|^2 \cdot S_Z(\omega) \\ &= |H(\omega)|^2 \cdot \frac{N_0}{2} \end{aligned}$$

כלומר כל $H\left(\omega\right)$ מהצורה

$$H\left(\omega\right)=\sqrt{\frac{2}{N_0}}\cdot\sqrt{S_X\left(\omega\right)}\cdot e^{-jA(\omega)}$$

ייתן

$$\left|H\left(\omega\right)\right|^2=\frac{2}{N_0}S_X\left(\omega\right)$$

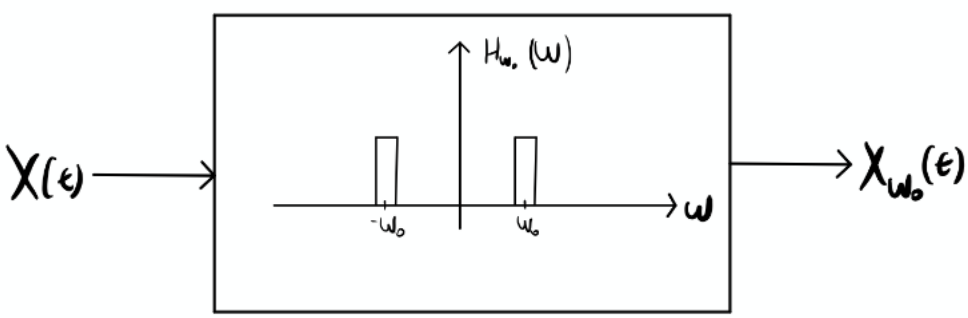
וזה נכון לכל בחירה של תגובת פאזה $A\left(\omega\right)$ $\Leftarrow$ יש הרבה מססנים שונים $H\left(\omega\right)$ שייתנו את ה-PSD הרצוי.

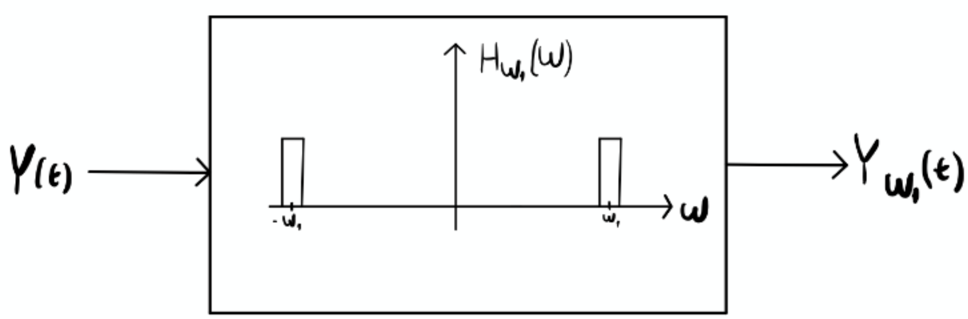
האם יש עדיפות של מסנן אחד על פני האחר? (האם יש עדיפות לפאזה אחת על פני אחרת?)

אז פרקטית - במערכות אמיתיות נרצה מסנן $\{h\left(t\right)\}$ יהיה סיבתי ומה הכוונה, שהתגובה להלם תגיב רק החל מזמן 0 תחת תנאים טכניים מסויימים ניתן למצוא $A\left(\omega\right)$ שיבטיח סיבתיות - משפט הפירוק הספקטרלי - פאלי וינר.

6.5 סינון צר סרט של תהליכים סטציונריים

נתונים שני תהליכים $\{X\left(t\right)\}$,$\{Y\left(t\right)\}$ שהם $JWSS$ ונניח שכל הסטטיסטיקה מסדר שני שלהם ידועה - $\{S_X\left(\omega\right)\},\{S_Y\left(\omega\right)\},\{S_{XY}\left(\omega\right)\}$
וכעת נניח שמעבירים את התהליך $X\left(t\right)$ במסנן צר סרט, כלומר מעביר את התדרים בחלום הקטן סביב ω_0, רוחב החלון הוא $2\pi\Delta$ והגובה הוא 1
ובאותו אופן את התהליך $Y\left(t\right)$ נעביר סביב התדר ω_1 רוחב החלון הוא $2\pi\Delta$ והגובה הוא 1

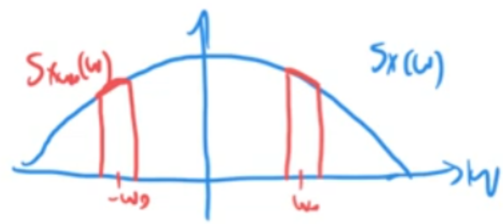




נרצה להבין מה הסטטיסטיקה המשותפת של התהליכים דרל המסננים
ניזכר בנוסחאות הכלליות שפיתחנו -

$$
\begin{aligned}
S_{X_{\omega_0}}(\omega) &= |H_{\omega_0}(\omega)|^2 \cdot S_X(\omega) \\
* &= H_{\omega_0}(\omega) \cdot S_X(\omega)
\end{aligned}
$$

*תגובת התדר למסנן היא תמיד חיובית וממשית והעלה בריבוע של 1 לא משנה כלום.
שרטוט לאינטואיציה, בחול לפני המסנן באדום - אחרי המסנן



עבור התדר התהליך השני נקבל

$$
\begin{aligned}
S_{Y_{\omega_1}}(\omega) &= |H_{\omega_1}(\omega)|^2 \cdot S_Y(\omega) \\
&= H_{\omega_1}(\omega) \cdot S_Y(\omega)
\end{aligned}
$$

ניזכר בנוסחר הכללית לקרוס-ספקטרום

$$
\begin{aligned}
&R_{X_{\omega_0}Y_{\omega_1}(\tau)} = h_{\omega_0}(-\tau) * R_{XY}(\tau) * h_{\omega_1}(\tau) \\
&\text{fourier after} \Downarrow \\
&S_{X_{\omega_0}Y_{\omega_1}}(\omega) = H^*_{\omega_0}(\omega) \cdot H_{\omega_1}(\omega) \cdot S_{XY}(\omega) \\
&\qquad * = H_{\omega_0}(\omega) \cdot H_{\omega_1}(\omega) \cdot S_{XY}(\omega)
\end{aligned}
$$

* כי H ממשי ולכן אין משמעות לצמוד.

עתה נסיק -

- $X_{\omega_0}(t)$ הוא התהליך אחרי שהעברתי אותו במסנן, נרצה להבין מה ההספק -

$$\begin{aligned}
\text{why?} \to \mathbb{E}\left[X_{\omega_0}^2(t)\right] &= R_{X_{\omega_0}}(0) \\
&= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty} S_{X_{\omega_0}}(\omega)\, d\omega \\
* &= 2\cdot\frac{1}{2\pi}\int\limits_{\omega_0-\Delta\pi}^{\omega_0+\Delta\pi} S_{X_{\omega_0}}(\omega)\, d\omega \\
&= 2\cdot\frac{1}{2\pi}\int\limits_{\omega_0-\Delta\pi}^{\omega_0+\Delta\pi}\left[\overbrace{H_{\omega_0}(\omega)}^{=1\text{range our in}} \cdot S_X(\omega)\right] d\omega \\
&= 2\cdot\frac{1}{2\pi}\int\limits_{\omega_0-\Delta\pi}^{\omega_0+\Delta\pi} S_X(\omega)\, d\omega \\
\Delta \to 0 \Rightarrow \quad & 2\cdot\frac{1}{2\pi}S_X(\omega_0)\int\limits_{\omega_0-\Delta\pi}^{\omega_0+\Delta\pi} 1 d\omega \\
&= 2\Delta S_X(\omega_0)
\end{aligned}$$

* כי PSD סימטרי אז ניתן להכפיל ב-2 ולעשות את האינטגרל רק בצד אחד.

- ראינו ש -

$$S_{X_{\omega_0}Y_{\omega_1}}(\omega) = H_{\omega_0}(\omega)\cdot H_{\omega_1}(\omega)\cdot S_{XY}(\omega)$$

מה קורה כאשר האינטרוולים בהם "חיים" המסננים זרים - למשל אצלנו ω_0, ω_1 מספיק רחוקים כך שאין חפיפה בינהם.

מה ניתן להגיד על הקורלציה בין התהליכים $X_{\omega_0}(t)$ ו-$Y_{\omega_1}(t)$, הקרוס-ספקטרום של תהליכים אלו נתון ע"י המכפלה של המסננים כפול הקרוס-ספקטרום המקורי. אך כשאין חפיפה המכפלה בין המסננים תתאפס, ונקבל שני ת"א שהקרוס-ספקטרום שלהם הוא אפס ולכן גם הקרוס קורלציה מתאפסת (כי הקרוס-קורלציה זה התמרה של הקרוס-ספקטרום)

$$\begin{gathered}
S_{X_{\omega_0}Y_{\omega_1}}(\omega) = 0\ \forall\omega \\
\Downarrow \\
R_{X_{\omega_0}Y_{\omega_1}(\tau)} = 0\ \forall\tau
\end{gathered}$$

במקרים אלו תהליכים $\{X_{\omega_0}(t)\}$ ו-$\{Y_{\omega_1}(t)\}$ חסרי קורלציה.

בגבול $\Delta \to 0$ לכל $\omega_0 \neq \omega_1$ $\{X_{\omega_0}(t)\}$ ו-$\{Y_{\omega_1}(t)\}$ חסרי קורלציה.

אפילו אם ניקח $\{X(t)\} = \{Y(t)\}$ נקבל ש-$\{X_{\omega_0}(t)\}, \{X_{\omega_1}(t)\}$ חסרי קורלציה עבור $\omega_0 \neq \omega_1$

- מסקנה שלישית -

אם $\{X(t)\}, \{Y(t)\}$ הם גאוסיים במשותף, כלומר לא רק $JWSS$, אלא, גם גאוסיים. ח"ק = אי תלות. ונקבל

$$\begin{aligned}
\forall\omega_0 \neq \omega_1 &: \{X_{\omega_0}(t)\} \perp \{Y_{\omega_1}(t)\} \\
&: \{X_{\omega_0}(t)\} \perp \{X_{\omega_1}(t)\}
\end{aligned}$$

והמסקנה מהדבר הזה - שנרחיב עליה בהמשך - אם נרצה לשערך את $\{X(t)\}$ מתוך $\{Y(t)\}$ - בהנחה שהם גאוסיים במשותף ו-$JWSS$ זה **שקול** לשערך בתדר, כלומר, $\{X(\omega)\}$ מתוך $\{Y(\omega)\}$,

אבל לכל $\omega_0 \neq \omega_1$ מתקיים שהזוגות בלתי תלויים

$$\left(X(\omega_0), Y(\omega_0)\right) \perp \left(X(\omega_1), Y(\omega_1)\right)$$

כלומר, אם נרצה לשערך את $X(\omega_0)$ אין שום סיבה להשתמש באינפורמציה בתדרים אחרים, נרצה להשתמש ב-$Y(\omega_0)$

ניתן לעשות את השערוך בכל תדר בנפרד - וזה מסנן ווינר!

7 מסנן וינר

7.1 שערוך אופטימלי במובן $LMMSE$

נתונים שני תהליכים אקראיים $JWSS$ $\{X(t)\}$ ו-$\{Y(t)\}$, כך ש-$\{X(t)\}$ זה מה שמעניין אותנו ו-$\{Y(t)\}$ המדידות - "מה שרואים"

נרצה לבנות מכונה שמקבלת בכניסה את $\{Y(t)\}$ עוברת במערכת ובמוצא יוצא משערך $\left\{\hat{X}(t)\right\}$

$$\{Y(t)\} \rightarrow \boxed{T} \rightarrow \left\{\hat{X}(t)\right\}$$

דוגמא לתהליך כזה יכול להיות - $\{X(t)\}$ אות EKG ו-$\{Y(t)\}$ מדידה רועשת שלו, כלומר, $Y(t) = X(t) + Z(t)$.

נגיד נרצה לקחת מדידה רועשת ולסנן ממנה כמה שיותר רעשים - כדי שהרופא יוכל להסיק מנתונים לא רועשים.

נתמקד בשערוך של $X(t)$ - כלומר דגימה ספציפית בזמן קבוע t מתוך $\{Y(t)\}$

ואם נרצה לבנות משערך ל - $\{X(t)\}$ פשוט נשערך את $X(t)$ בכל הזמנים ונאסוף את כל הדגימות האלו למשערך $\left\{\hat{X}(t)\right\}$ לתהליך כולו

לעיתים הגישה של לשערך את כל התהליך ע"י שערוך של כל זמן בנפרד היא אופטימלית. וזה תלוי באיך נמדוד את השגיאה.

קרטריון מאוד מקובל (ובו נעסוק בקורס שלנו) היא נורמה $L2$, כלומר,

$$\begin{aligned} D &= \frac{1}{T}\mathbb{E}\int_0^T \left[X(t) - \hat{X}(t)\right]^2 dt \\ &= \frac{1}{T}\int_0^T MSE\left[\hat{X}(t)\right] dt \end{aligned}$$

מתי לא אופטימלי לשערך כל זמן בנפרד? למשל אם נקבע קרטריון כזה - (דיון לא כך חשוב, מה שחשוב שאנו רוצים לשערך את $\{X(t)\}$ בכל זמן בנפרד)

$$\begin{aligned} D &= \mathbb{E}\max_{t\in[0,T]}\left|X(t) - \hat{X}(t)\right| \\ &\neq \max_{t\in[0,T]}\mathbb{E}\left|X(t) - \hat{X}(t)\right| \end{aligned}$$

7.1.1 נגדיר את הבעיה

נתונה סטטיסטיקה משותפת של זוג תהליכים $JWSS$ $\{X(t)\}$ ו-$\{Y(t)\}$

רוצים למצוא את המשערך $\hat{X}(t)$ ל-$X(t)$ מתוך המדידות $\{Y(t)\}$ שמביא למינימום את ה-MSE $\mathbb{E}\left[\left(X(t) - \hat{X}(t)\right)^2\right]$

מאוד דומה למשערך של סקלר X מתוך וקטור $\underline{Y}$ - כמו שעשינו בקורס של עמי - $\hat{X}_{opt} = \mathbb{E}\left[X|\underline{Y}\right]$

אנחנו נרצה לשערך את $X\left(t\right)$מתוך אוסף סופי של דגימות של $\left\{ Y\left(t\right)\right\}$ למשל - $Y\left(t_{1}\right),...,Y\left(t_{k}\right)$ אפשר לקרוא לדגימות האלו וקטור של $\underline{Y}$ וזה בדיוק מה שראינו עם עמי

$$\hat{X}_{opt}\left(t\right)=\mathbb{E}\left[X\left(t\right)\mid\left\{ Y\left(t_{1}\right),...,Y\left(t_{k}\right)\right\} \right]$$

ולכן אין שום סיבה שכאשר נעבור לתהליך אקראי במקום אוסף סופי של דגימות משהו ישתנה!

אז מה המשערך האופטימלי במובן $MMSE$ של $X\left(t\right)$ מתוך $\left\{ Y\left(t\right)\right\}$? בדיוק כמו מהווקטור!

$$\hat{X}_{opt}\left(t\right)=\mathbb{E}\left[X\left(t\right)\mid\left\{ Y\left(t\right)\right\} \right]$$

לא נוכיח את זה אבל ההוכחה דומה לההוכחה של שערוך סקלר מווקטור.

אבל מה, יש בעיה! כמעט אף פעם לא ניתן לחשב את התוחלת המותנית הזו.

וכאשר לא ניתן לחשב את המשערך האופטימלי, נלך לפתרון שהוא תת-אופטימלי. ובתחום ההנדסה אוהבים לחפש את הפתרון הלינארי, וזה מה שעשה ווינר וזה מה שנתמקד בו בקורס שלנו.

נמצא את המשערך הלינארי הטוב ביותר במובן ה-$MSSE$, כלומר $LMSSE$ של $X\left(t\right)$ מתוך $\left\{ Y\left(t\right)\right\}$

חזרה לבעיה שלנו -

נתונה סטטיסטיקה משותפת של זוג תהליכים $JWSS$ $\left\{ X\left(t\right)\right\}$ ו-$\left\{ Y\left(t\right)\right\}$ עם תוחלות 0

נרצה לשערך דגימה של $\left\{ X\left(t\right)\right\}$ מתוך כל התהליך $\left\{ Y\left(t\right)\right\}$ ונרצה לעשות זאת עם משערך לינארי -

$$\hat{X}\left(t\right)=\int_{-\infty}^{\infty}g\left(s\right)Y\left(s\right)ds$$

נרצה למצוא את המשערך הטוב ביותר, כלומר, שמביא למינימום את ה-MSE כלומר המשערך -$MMSE$

$$\mathbb{E}\left[\left(X\left(t\right)-\hat{X}\left(t\right)\right)^{2}\right]$$

מה התהליך שאנחנו צרכים לעשות - למצוא פונקציה $\left\{ g\left(s\right)\right\}$ שהיא "הכי טובה" כדי שיהיה לנו נח נסמן $\left\{ g\left(s\right)=h\left(t-s\right)\right\}$ (אינטגרל הקונבולוציה?)

כלומר נסתכל על משערך

$$\begin{aligned}\hat{X}\left(t\right) & =\int_{-\infty}^{\infty}h\left(t-s\right)Y\left(s\right)ds\\
 & =\left[h\star Y\right]\left(t\right)
\end{aligned}$$

וזה בדיוק אינטגרל הקונבולוציה בזמן t!

וכעת נחפש את - $\left\{ h_{opt}\left(s\right)\right\}$ הטובה ביותר, כלומר שמביאה למינימום את

$$\begin{aligned}\hat{X}_{LMMSE}\left(t\right) & =\int_{-\infty}^{\infty}h_{opt}\left(t-s\right)Y\left(s\right)ds\\
 & =\left[h_{opt}\star Y\right]\left(t\right)
\end{aligned}$$

עתה נמצא את h_{opt}, איך נמצא אותו?

נשתמש בעקרון האורתוגונליות!

עבור שערון לינארי של סקלר X מתוך וקטור $\underline{Y}$ המשערך האופטימלי במובן ה-$LMMSE$ מקיים

$$\mathbb{E}\left[\underbrace{X\left(t\right)-\hat{X}_{LMMSE}\left(t\right)}_{\text{error}}\right]\cdot\left[\underline{\text{Y}}\text{ of func linear}\right]=0$$

המשערך יכול להיות אופטימלי אמ"מ השגיאה הזו היא ח"ק עם כל פונקציה לינארית של המדידות - ניתן אינטואיציה למה: אם נמצא פונקציה לינארית של המדידות שיש לה קורלציה עם השגיאה, נוכל לשפר את המשערך - איך? נשתמש בפונקציה

הזו שיש לה קורלציה עם השגיאה כדי לשערך מחדש את השגיאה ובכך להקטין את השגיאה ואם הצלחתי להקטין השגיאה המשערך לא היה אופטימלי!
במקרה שלנו - עבור שערוך לינארי של $X(t)$ מתוך $\{Y(t)\}$ המשערך האופטימלי מקיים את עקרון האורתוגונליות הבא:

$$\mathbb{E}\left[\underbrace{X(t)-\hat{X}_{LMMSE}(t)}_{e(t)}\right]\cdot\left[\underbrace{\int_{-\infty}^{\infty}\varphi(s)\,Y(s)\,ds}_{\{Y(t)\}\text{ of func linear}}\right]=0$$

לא נוכיח זאת, אבל נאמין שזה עובד.
נבחין שהדרישה שכתבנו שקולה ל-

$$\mathbb{E}\left[\left(X(t)-\hat{X}_{LMMSE}(t)\right)\cdot Y(v)\right]=0\quad\forall v$$

השיוויון הראשון נכון לכל $\varphi(s)$ ובפרט ל-$\varphi(s)=\delta(v)$ וכעת נציב את הביטוי למשערך האופטימלי - נזכיר שהוא היה $\hat{X}(t)=\int_{-\infty}^{\infty}h_{opt}(t-s)\,Y(s)\,ds$ ונקבל ונדרוש שהשיוויון יתקיים לכל v

$$\begin{aligned}
\mathbb{E}\left[\left(X(t)-\hat{X}_{LMMSE}(t)\right)\cdot Y(v)\right]&=0\\
&\Downarrow\\
\mathbb{E}\left[\int_{-\infty}^{\infty}h_{opt}(t-s)\,Y(s)\,ds\cdot Y(v)\right]&=\mathbb{E}\left[X(t)\cdot Y(v)\right]\\
&{*}\text{magic!}\Downarrow\\
\int_{-\infty}^{\infty}h_{opt}(t-s)\,\mathbb{E}\left[Y(s)\,Y(v)\right]ds&=\mathbb{E}\left[Y(v)\cdot X(t)\right]\\
\int_{-\infty}^{\infty}h_{opt}(t-s)\,R_Y(s-v)\,ds&=R_{YX}(t-v)\\
\int_{-\infty}^{\infty}h_{opt}(t-s)\,R_Y(t-v-(t-s))\,ds&=R_{YX}(t-v)\\
\int_{-\infty}^{\infty}h_{opt}(\tau)\,R_Y(t-v-\tau)\,d\tau&=R_{YX}(t-v)\\
\left[h_{opt}\star R_Y\right](t-v)&=R_{YX}(t-v)\\
&{**}\Downarrow\\
\left[h_{opt}\star R_Y\right](\tau)&=R_{YX}(\tau)\quad\forall\tau
\end{aligned}$$

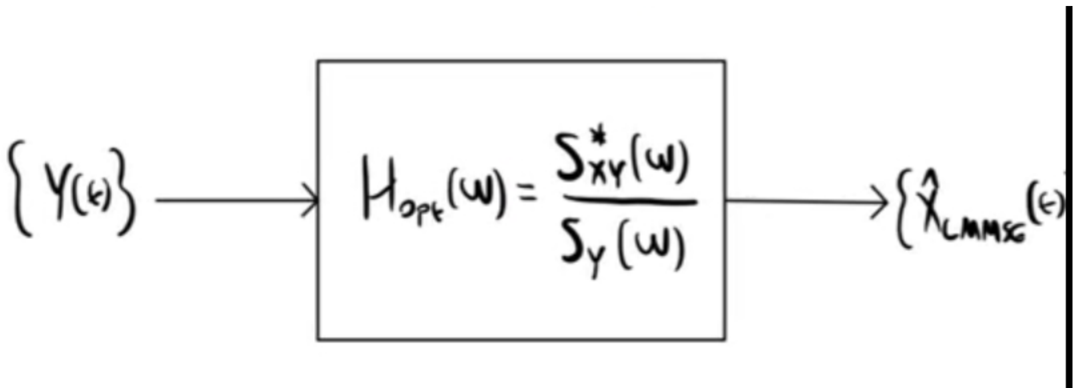
* בקורס הנהחו שמותר תחת תנאים שלא הגדרנו לשנות סדר של תוחלת ואינטגרל
** כי הגדרנו שצריך להתקיים לכל v אז ניתן להחליף את $t-v$ ב-τ
אוקיי, אז מצאנו תנאי מספיק והכרחי שאכן h_{opt} היא אכן הפונקציה הכי טובה במובן ה-$MMSE$ אבל מסובך כל עיניין הקונבולוציה הזה.
אז נעבור לתחום התדר! בתחום התדר הכל יותר פשוט.

$$\begin{aligned}
H_{opt}(\omega)\cdot S_Y(\omega)&=S_{YX}(\omega)\\
H_{opt}(\omega)&=\frac{S_{YX}(\omega)}{S_Y(\omega)}\\
&=\frac{S_{XY}^{*}(\omega)}{S_Y(\omega)}\quad\forall\omega
\end{aligned}$$

נשים לב למשהו מאוד מאוד מעניין - התחלנו עם h_{opt} הכי טוב ע"מ לשערך את $X\left(t\right)$ וכתבנו אותו כקונבולוציה עם דגימה בזמן ספציפי והיינו מצפים שהוא יהיה שונה בכל זמן, כלומר שונה לכל t. ולכן היינו מצפים שיהיה תלוי בזמן.

אבל! הזמן t לא מופיע!! בפועל לא משנה באיזה זמן נרצה לשערך את $X\left(t\right)$ מה שנעשה זה לקחת את התהליך Y ולעשות לו קונבולוציה עם המסנן h_{opt} ולדגום את **התוצאה** בזמן שאנחנו מחפשים.

למעשה חיפשנו משערך ל-t ספציפי וקיבלנו שאם נרצה לשערך את כל התהליך $X\left(t\right)$ מכל התהליך $Y\left(t\right)$ נפעל כך -



ניקח את המדידות - $\left\{Y\left(t\right)\right\}$ נעביר אותם דרך המסנן h_{opt} - ובמוצא נקבל את המשערך הלינארי הטוב ביותר.

המשערך האופטימלי ל-$\left\{X\left(t\right)\right\}$ מתוך $\left\{Y\left(t\right)\right\}$ ניתן ע"י מערכת LTI. למסנן זה נקרא מסנן ווינר - (זה אותו ווינר של תהליך ווינר)

בתחום הזמן זה יראה כך

$$\left\{Y\left(t\right)\right\}\left\{\hat{X}_{LMMSE}\left(t\right)\right\}\rightarrow\boxed{h_{opt}}\rightarrow$$

7.1.2 הקשר לשיערוך סקלרי

אם יש לנו X ו-Y מ"א עם תוחלת אפס אז

$$\hat{X}_{LMMSE}=\frac{\mathbb{E}\left[X\cdot Y\right]}{\mathbb{E}\left[Y^{2}\right]}\cdot Y$$

איך נראה משערך לתהליך אקראי בתחום התדר? (אז אי אפשר להגדיר התמרת פוריה לתהליכים WSS כי האנרגיה היא אינסופית בהסתברות - 1, או משהו כזה שאני לא זוכר) בכל מקרה נתעלם מסוגיה זו רק לשם קבלת אינטואיציה והבנה

אז נתבונן במערכת שלנו ונכניס לה את - $\left\{Y\left(\omega\right)\right\}$

$$\left\{Y\left(\omega\right)\right\}\rightarrow\boxed{H_{opt}\left(\omega\right)=\frac{S_{XY}^{*}\left(\omega\right)}{S_{Y}\left(\omega\right)}}\rightarrow\hat{X}_{LMMSE}\left(\omega\right)=\frac{S_{XY}^{*}\left(\omega\right)}{S_{Y}\left(\omega\right)}\cdot Y\left(\omega\right)$$

וזה מאוד דומה לצורה של המשערך הסקלרי.

בעצם מסנן ווינר עושה שערוך סקלרי של $X\left(\omega\right)$ מתוך $Y\left(\omega\right)$ עבור כל תדר בנפרד

7.2 שגיאת השערוך

נגדיר

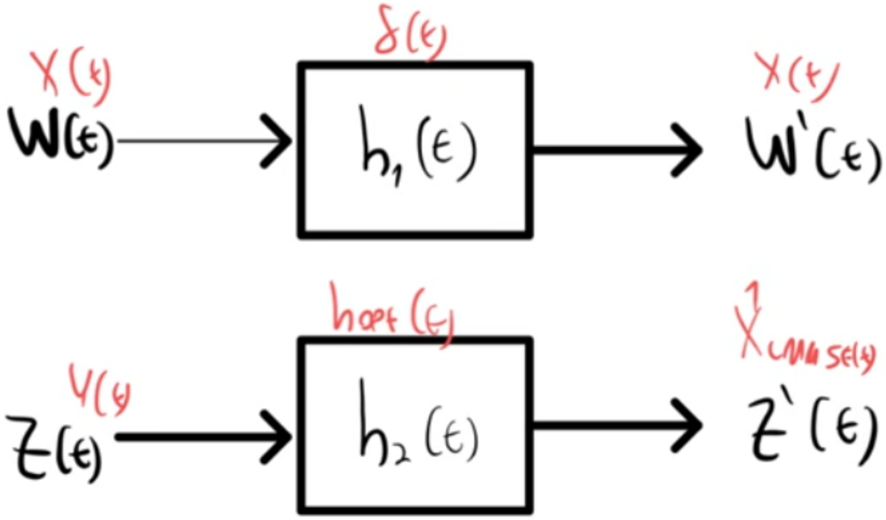
$$e\left(t\right)=X\left(t\right)-\hat{X}_{LMMSE}\left(t\right)\quad\forall t$$

והתהליך $\{e\left(t\right)\}$ הוא תהליך שגיאת השערוך ונרצה להבין איך תהליך זה מתנהג

נטען שהתהליך של שגיאת השערוך הוא WSS ואפילו $JWSS$ עם $X\left(t\right)$, JWSS areand $\{X\left(t\right)\}$ $\{e\left(t\right)\}$ איך נראה זאת?

ניזכר במערכת היחוס מלפני 2 הרצאות ונציב $X\left(t\right)=W\left(t\right)$ ו-$h_1\left(t\right)=\delta\left(t\right)$ כל ש-$W'\left(t\right)=X\left(t\right)$ וניקח $Z\left(t\right)=Y\left(t\right)$, $h_2\left(t\right)=h_{opt}\left(t\right)$ ו-$Z'\left(t\right)=\hat{X}_{LMMSE}\left(t\right)$

כלומר,



ראינו שאם הכניסות X,Y הם $JWSS$ אז גם המוצא $\left\{\hat{X}_{LMMSE}\left(t\right)\right\}$, $\{X\left(t\right)\}$ הוא $JWSS$ וקל לראות שזה גורר ש-$\{X\left(t\right)\}$ ו-$\left\{e\left(t\right)=X\left(t\right)-\hat{X}_{LMMSE}\left(t\right)\right\}$ גם הוא $JWSS$ כדי להראות זאת צריך לחשב את הקרוס-קורלציה בין $e\left(t\right)$ ל-$X\left(t\right)$ ונציב את $e\left(t\right)$ ומשם החישוב די טריוויאלי...

$$**********proff*********$$

כעת נחשב את התוחלת

$$\begin{aligned}
\eta_e &= \mathbb{E}\left[e\left(t\right)\right] \\
&= \mathbb{E}\left[X\left(t\right)\right]-\mathbb{E}\left[\hat{X}_{LMMSE}\left(t\right)\right] \\
* &= \eta_X-\eta_Y\cdot H_{opt}\left(0\right) \\
&= 0
\end{aligned}$$

*התוחלת של מוצא מערכת LTI שהכניסה שלו היא תהליך WSS $Y\left(t\right)$

נחשב את $R_e(\tau)$ פונק' האוטוקורלציה של $\{e(t)\}$ וכדי לחשב זאת נשתמש בטריק!

$$\begin{aligned}
X(t) &= e(t) + \hat{X}_{LMMSE}(t) \\
&\Downarrow \\
R_X(\tau) &= \mathbb{E}\left[X(t)\,X(t+\tau)\right] \\
&= \mathbb{E}\left[\left(e(t) + \hat{X}_{LMMSE}(t)\right)\left(e(t+\tau) + \hat{X}_{LMMSE}(t+\tau)\right)\right] \\
&= R_e(\tau) + \overset{=0*}{\cancel{\mathbb{E}\left[e(t)\cdot \hat{X}_{LMMSE}(t+\tau)\right]}} + \cancel{\mathbb{E}\left[\hat{X}_{LMMSE}(t)\cdot e(t+\tau)\right]} + R_{\hat{X}_{LMMSE}}(\tau) \\
&= R_e(\tau) + R_{\hat{X}_{LMMSE}}(\tau)
\end{aligned}$$

* מתאפס כי $\hat{X}_{LMMSE}(t+\tau)$ היא פונקציה של המדידות - פונקציה לינארית והשגיאה אורתוגונלית לכל פונקציה של מדידות!! באופן דומה לאיבר המשולב השני.

וכאשר נעביר אגפים נקבל

$$R_e(\tau) = R_X(\tau) - R_{\hat{X}_{LMMSE}}(\tau)$$

וכעת נתבונן בתדר, ב-PSD

$$S_e(\omega) = S_X(\omega) - S_{\hat{X}_{LMMSE}}(\omega)$$

נחשב את $S_{\hat{X}_{LMMSE}}(\omega)$ את $\hat{X}_{LMMSE}$ קיבלנו עי העברת $Y(t)$ במסנן ווינר וראינו ש-PSD של מוצא מערכת LTI הוא

$$\begin{aligned}
S_{\hat{X}_{LMMSE}}(\omega) &= \left|H_{opt}(\omega)\right|^2 \cdot S_Y(\omega) \\
&= \left|\frac{S^*_{XY}(\omega)}{S_Y(\omega)}\right|^2 \cdot S_Y(\omega) \\
&= \left|\frac{S_{XY}(\omega)}{S_Y(\omega)}\right|^2 \cdot S_Y(\omega) \\
&= \frac{\left|S_{XY}(\omega)\right|^2}{S_Y(\omega)}
\end{aligned}$$

כלומר, קיבלנו ש-

$$\begin{aligned}
S_e(\omega) &= S_X(\omega) - S_{\hat{X}_{LMMSE}}(\omega) \\
&= S_X(\omega) - \frac{\left|S_{XY}(\omega)\right|^2}{S_Y(\omega)}
\end{aligned}$$

כעת נוכל לחשב את ההספק של השגיאה!

$$\begin{aligned}
\mathbb{E}\left[e^2(t)\right] &= R_e(0) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} S_e(\omega)\,d\omega
\end{aligned}$$

וכאן סיימנו את הפיתוח לזמן רציף

בזמן בדיד הנוסחאות למשערך יהיו

אם $\{X_n\}$ ו-$\{Y_n\}$ שני תהליכים $JWSS$ עם תוחלת 0 המשערך האופטימלי הלינארי במובן $MMSE$ ל-X בזמן n מתוך תהליך Y יהיה

$$\hat{X}_{n,LMMSE} = h_{n,\,opt} \star Y_n \ \forall n$$

$$h_{n,\,opt} = \mathcal{F}^{-1}\left\{H_{opt}(\omega) = \frac{S^*_{XY}(\omega)}{S_Y(\omega)}\right\}$$

$$e_n = X_n - \hat{X}_{n,\,LMMSE}$$

התהליך $\{e_n\}$ הוא $JWSS$ עם התהליך $\{X_n\}$ ובדיוק כמו המקרה הרציף

$$\mathbb{E}[e_n] = 0$$
$$S_e(\omega) = S_X(\omega) - \frac{|S_{XY}(\omega)|^2}{S_Y(\omega)}$$
$$\mathbb{E}\left[e_n^2\right] = R_e(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_e(\omega)\,d\omega$$

7.3 הערות

1. מה קורה כשיש תוחלות שונות מ-0?

כלומר, $\{X(t)\}$ ו-$\{Y(t)\}$ הם $JWSS$ עם תוחלות שונות מאפס,

תמיד אפשר לאפס את התוחלות ע"י החסרת התוחלת, כלומר,

$$X'(t) = X(t) - \eta_X$$
$$Y'(t) = Y(t) - \eta_Y$$

כמובן ש-$\{X'(t)\}$ ו-$\{Y'(t)\}$ הם $JWSS$

כעת נשערך את $\{X'(t)\}$ מתוך $\{Y'(t)\}$ באמצעות מסנן ווינר ונוסיף לתוצאה את η_X גרפית זה נראה כך-

$$Y(t) \longrightarrow \overset{\overset{\eta_Y}{\downarrow}}{\ominus} \xrightarrow{Y'(t)} \boxed{H_{opt}(\omega) = \frac{S^*_{X'Y'}(\omega)}{S_{Y'}(\omega)}} \xrightarrow{X'_{LMMSE}(t)} \overset{\overset{\eta_X}{\downarrow}}{\oplus} \longrightarrow \hat{X}_{LMMSE}(t)$$

2. עבור תהליכים גאוסיים במשותף

$\{X(t)\}$ ו-$\{Y(t)\}$ הם $JWSS$ וגאוסיים במשותף, נטען שבמקרה כזה שיערוך $MMSE$ ו-$LMMSE$ הם שקולים, כלומר, המשערך האופטימלי הוא לינארי.

למה?

$$\mathbb{E}[X(t) \mid \{Y(t)\}] = \int_{-\infty}^{\infty} \varphi(s)\,Y(s)\,ds$$

עבור פונקציה כלשהי $\{\varphi(s)\}$

במקרה הוקטורי/סקלרי - ראינו עם עמי שזה פונקציה לינארית של המדידות

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

ואם זה המצב, במקרה הגאוסייני $\{Y(t)\}$ הוא פשוט פונקציה לינארית של $\{X(t)\}$

$\Leftarrow$ במקרה הגאוסי מסנן ווינר הוא לא רק המשערך הלינארי הטוב ביותר, אלא, המשערך הטוב ביותר בלי אילוצים - כלומר, אפילו אם לא נגביל למשערך לינארי במקרה ש-$\{X(t)\}$ ו-$\{Y(t)\}$ הם גאוסיים במשותף הדבר הכי טוב שניתן לעשות הוא להשתמש במסנן ווינר!

7.4 דוגמא!

נניח 2 תהליכים $\{Z(t)\}, \{X(t)\}$ שהם $JWSS$ והם גם חסרי קורלציה, כלומר,

$$R_{XZ}(\tau) = 0 \quad \forall\tau$$

וכן מתקיים

$$\begin{aligned} \eta_X &= 0 \\ \eta_Z &= 0 \end{aligned}$$

כעת נגדיר תהליך אקראי חדש

$$Y(t) = X(t) + Z(t) \quad \forall t$$

נרצה למצוא את המשערך הלינארי הטוב ביותר ל-$X(t)$ מתוך התהליך $\{Y(t)\}$.
איך ניגשים לזה? אם נצליח להראות ש-$\{Y(t)\}, \{X(t)\}$ הם $JWSS$ אז המשערך הטוב ביותר הוא מבנן ווינר.
אז נתחיל - נראה שהם $JWSS$

$$\begin{aligned} \mathbb{E}[Y(t)] &= \mathbb{E}[X(t) + Z(t)] \\ &= 0 \\ R_Y(t, t+\tau) &= \mathbb{E}[(X(t) + Z(t)) \cdot (X(t+\tau) + Z(t+\tau))] \\ &= R_X(\tau) + \cancel{R_{ZX}(\tau) + R_{XZ}(\tau)} + R_Z(\tau) \\ &= R_X(\tau) + R_Z(\tau) \end{aligned}$$

כלומר - $Y(t)$ הוא WSS כי התוחלת שלו0, לא תלויה בזמן ופונק' האוטוקורלציה שלו לא תלויה בזמן, אלא בהפרש זמנים.
כעת נשאר להראות ש-$\{Y(t)\}$ ו-$\{X(t)\}$ הם $JWSS$ נחשב -

$$\begin{aligned} R_{XY}(t, t+\tau) &= \mathbb{E}[X(t)(X(t+\tau) + Z(t+\tau))] \\ * &= R_X(\tau) + \cancel{R_{XZ}(\tau)} \\ &= R_X(\tau) \end{aligned}$$

* כי X, Z הם חסרי קורלציה!
על הדרך נחשב את ה-PSD

$$\begin{aligned} S_Y(\omega) &= \mathcal{F}\{R_X(\tau) + R_Y(\tau)\} \\ &= S_X(\omega) + S_Z(\omega) \\ S_{XY}(\omega) &= \mathcal{F}\{R_{XY}(\tau)\} \\ &= S_X(\omega) \end{aligned}$$

ואיך יראה המסנן האופטימלי?

$$\underbrace{\{Y(t)\}}_{\underbrace{X(t)}_{\text{signal}}+\underbrace{Z(t)}_{\text{noise}}} \longrightarrow \boxed{H_{opt}=\frac{S_{XY}^{*}(\omega)}{S_{Y}(\omega)}} \longrightarrow \hat{X}_{LMMSE}(t)$$

כלומר נותר לחשב את

$$\begin{aligned} H_{opt}(\omega) &= \frac{S_{X}^{*}(\omega)}{S_{Y}(\omega)} \\ * &= \frac{S_{X}(\omega)}{S_{Y}(\omega)} \\ &= \frac{S_{X}(\omega)}{S_{X}(\omega)+S_{Z}(\omega)} \end{aligned}$$

* היות ו-$S_X(\omega)$ היא PSD ו-PSD היא ממשית ולכן אין משמעות לצמוד.

וסיימנו - מצאנו את המסנן האופטימלי.

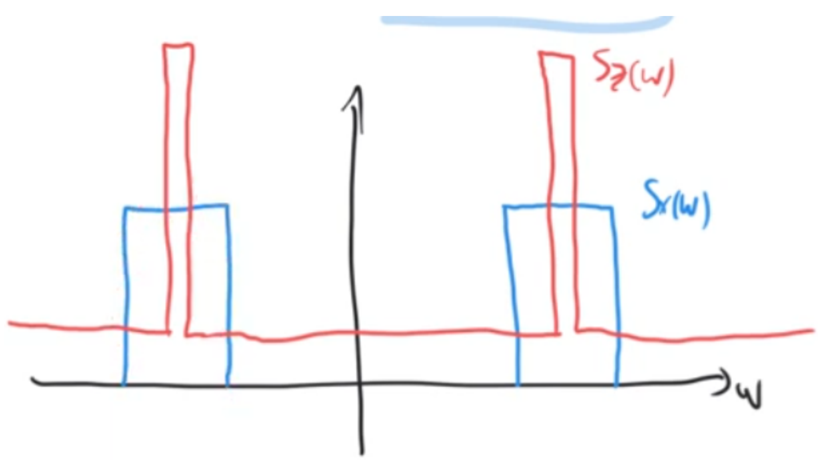
נתבונן בנוסחה של המסנן כדי לקבל מעט אינטואיציה על מה קורה פה, אז נכתוב את המסנן בצורה מעט שונה

$$\begin{aligned} H_{opt}(\omega) &= \frac{\frac{S_X(\omega)}{S_X(\omega)}}{\frac{S_X(\omega)+S_Z(\omega)}{S_X(\omega)}} \\ &= \frac{1}{1+\frac{S_Z(\omega)}{S_X(\omega)}} \end{aligned}$$

כלומר, בתדרים שבהם ל-Z (ההספק לרעש) הרבה יותר קטן מההספק של X- (הסיגנל שלנו) $H_{opt}(\omega)=1$ כלומר המסנן יעביר את הסיגנל והרעש כמו שהם

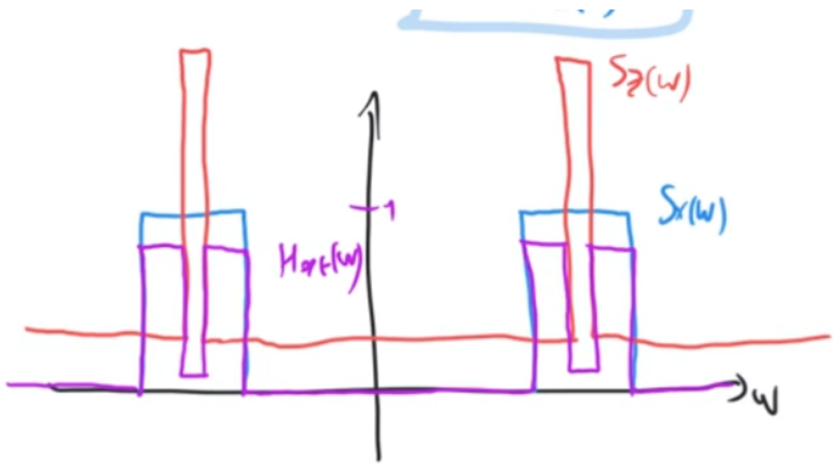
אבל כאשר Z יותר גדול מ-X המסנן יקבל ערך מאוד קטן, $H_{opt}(\omega)\ll 1$, ובפועל התדרים שבהם הרעש חזק המסנן ינחית ובתדרים בהם בסיגנל חזק המסנן יעביר!

ציור:



ה-PSD של הסיגנל מצוייר בכחול סוג של filter pass band ויש לנו את ה-PSD של הרעש בצבע אדום הוא כמעט רעש לבן (כי הוא כמעט שטוח) למעא תדר אחד שבוא הוא מאוד חזק

ובמקרה הרעש מאוד חזק איפה שהסיגנל שלנו חיי.
מה המסנן יעשה?
איפה שיש רק רעש במסנן יהיה לנו ∞, כלומר הוא לא יעביר כלום, (למה להעביר רק רעש?)
איפה שיש מעט רעש וסיגנל המסנן יעביר כמעט את כל הסיגנל
איפה שיש הרבה רעש ומעט סיגנל המסנן ינחית וישאר מעט סיגנל (לפי היחס שחישבנו במסנן)
תמונה - ההתנהגות בסגול -



כעת נותר לנו לכתוב את השגיאה של המסנן

$$
\begin{aligned}
S_e(\omega) &= S_X(\omega) - \frac{|S_{XY}(\omega)|^2}{S_Y(\omega)} \\
&= S_X(\omega) - \frac{S_X^2(\omega)}{S_Y(\omega)} \\
&= S_X(\omega)\left(1 - \frac{S_X(\omega)}{S_Y(\omega)}\right) \\
&= S_X(\omega)\left(\frac{S_Y(\omega) - S_X(\omega)}{S_Y(\omega)}\right) \\
&= \left(\frac{S_X(\omega)\cdot S_Z(\omega)}{S_X(\omega) + S_Z(\omega)}\right)
\end{aligned}
$$

8 הגדרות ומשפטים.