הרצאה 2

שערוך אצווה(BATCH) בשיטת מינימום ריבועים בשיטת (minimum squares)
מודל לינארי:
מניח מדידה:

$$\underbrace{Z(k)}_{scalar\ measurement} = \underbrace{h(\eta)\theta}_{paramenetr\ vector} + \underbrace{\mu(k)}_{mesurement\ noise}$$

where $E[\mu(k)] = 0$

אנחנו מניחים כי ממוצע הרעש הוא אפס (no bias)
בה"כ Z, K, μ וקטורים.

ידוע לנו כי H דטרמיניסטי, Z הם המדידות הקיימות לנו. כיצד נוכל לשערך את וקטור הפרמטרים טתא? וכיצד מוגדר טתא אופטימלית?
טתא אופטימלית יכולה להיות מוגדרת בדרכים שונות לפי דרישה.
נביט במקרה בו :

$$assumption \quad \underbrace{\theta(k)}_{parameter\ vector\ till\ time\ k} \rightarrow Z(k) = H(k)\theta + V(k) \underbrace{\approx}_{we\ use\ mean\ of\ noise\ as\ closest\ approx} Z(k) + H(k)\theta(k)$$

בשיטת מינימום ריבועים ניתן לתת משקל לכל אחד מהפרמטים. פונקציית המחיר J:

$$J[\theta(k)] = \vec{Z}^T(k)W(w)\vec{Z}(k)$$

$$def: \quad \underbrace{\vec{Z} = Z(k) - \hat{Z}(k)}_{Equation\ error\ (residual)}$$

הערה: *שגיאת השערוך*היא Z ובמקרה הלינארי הזה היא גם *שגיאת המשוואה*
הערה: J הוא סקלר.
עבור W = I (מטריצת המשקל) מטריצת היחידה, נקבל את שיטת מינימום ריבועים (LS) , אחרת שיטת מינימום ריבועים משקלית (WLS).
משום שנרצה למצוא את המשקל J כך שנקבל תטא מינימלית נגדיר

$$def: \hat{\theta}(k)_{LS} = arg(min(J[\theta(k)]))$$

הערה: W צריכה להיות PD(positive definite) אחרת מדידות מסוימות לא יילקחו בחשבון.

החוק שלוקח מדידות ומייצר מהם שערוך נקרא משערך.
(where theta is a random variable (or random process in time $\hat{\theta}(k)$

$$J[\theta(k) = [Z(w)] - H(k)\hat{\theta}(k)]^T W(k)[Z(w) - H(w)\hat{\theta}(w)] =$$

$$Z^T(k)W(k)H(k) - 2Z^T(k)W(k)H(k)\hat{\theta}(k) + H^T(k)W(k)H(k)\hat{\theta}(k)$$

REMINDER: $\frac{d(\vec{a}^T\vec{x})}{dx} = \vec{a}$ and $\frac{d\vec{X}^TA\vec{X}}{dx}$

$$Z^T(k)W(k)Z(k) - 2Z^T(k)W(k)H(k)\hat{\theta}(k) + \hat{\theta}^T H^T(k)W(w)H(k)\hat{\theta}(k)$$

נגזור לקבלת J מינימלית שמתאימה לטתא מינימלית

$$\frac{dJ[\hat{\theta}(k)]}{d\theta(k)} = -2[Z^T(k)W(k)H(k)\hat{\theta}(k)]^T + 2[Z^T(k)W(k)H(k)\hat{\theta}(k)] = 0$$

$$Z^T(k)W(k)H(k)\hat{\theta}(k) = Z^T(k)W(k)H(k)$$

Where H is full rank , we arrive at

$$\hat{\theta}(k)_{WLS}(k) = [H(k)^T W(k) H(k)]^{-1} H(k)^T W(k) Z(k)$$

זוהי מדידה לא טובה כי H - עולה בריבוע משמע אין

$$\hat{\theta}(k)_{LS}(k) = [H(k)^T H(k)]^{-1} H(k)^T Z(k)$$

כאשר H איננה ריבועית לא קיימת מטריצה הופכית של H ועל כן נשתמש pseudo inverse. (קירוב למט' הופכית לא)
נבדוק שהתשובה שלנו היא אכן *מינימום* לפי נגזרת שניה (הסיאן):

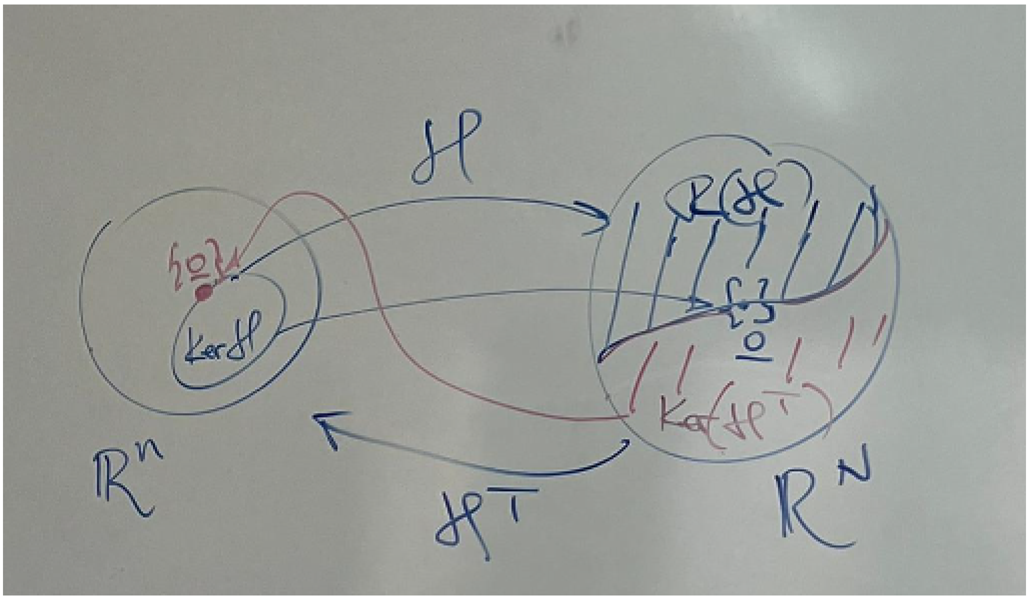
$$\frac{d^2(J(\hat{\theta}(k)))}{d^2\hat{\theta}} = 2Z^T(k)W(k)H^T(k) > 0 \rightarrow minimum.$$

כדי לפשט את החישובים נניח כי W = I :

$$H^T(k)H(k)\hat{\theta}(k)_{WLS} = H^T(k)Z(k)$$

נוציא את H^T מחוץ לסוגריים ונקבל:

$$[H^T(k)][\underset{\hat{Z}(k)}{Z(k) - H(k)\hat{\theta}(k)}] = 0 \longrightarrow H^T(k)\vec{Z}(k) = 0 \longrightarrow \vec{Z} \in \underset{\perp R[H^T(k)]}{\ker H^T(k)}$$



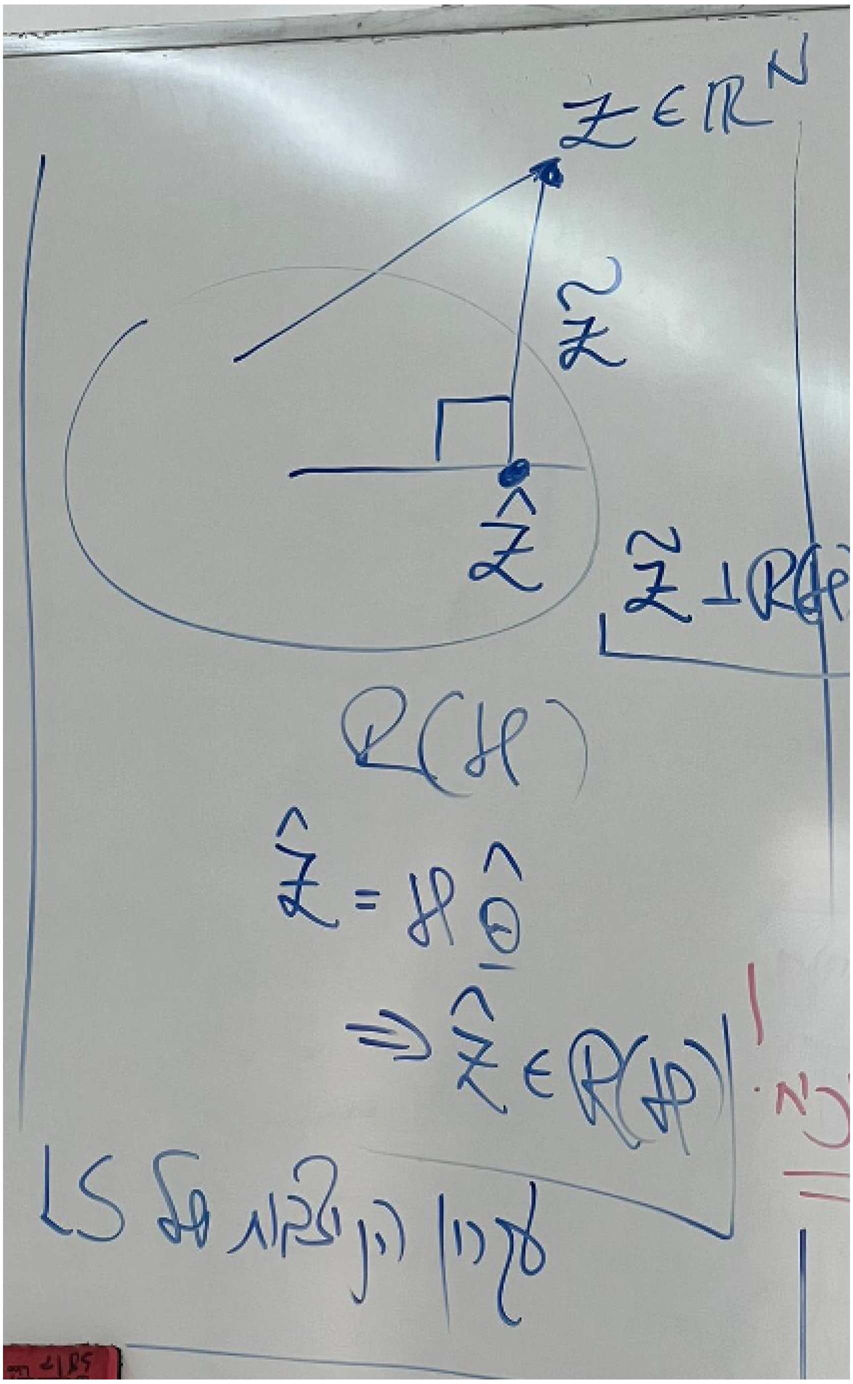
טענה (הוכחה לפי הכלה משני הכיוונים):

$$\ker H^T(k) = R(H(k)^{\perp})$$

טענה שקולה היא ש:

$$\vec{Z}(k) \in R(H)^{\perp}$$

עבור Z המדידה שלנו נרצה למצוא את המרחק הקטן ביותר (מינימום מרחק), זוהי בעיית אורטוגונליות, נוריד אנך ונקבל שגיאה מינימלית.



אושמן המליץ על ספר Lawson & Hanson solving least square problems לפתרון נומרי של בעיות מינימום ריבועים.

דוגמה לשערוך טתא ממדידות חסרות קורלציה:

$$Z(k) = \theta + V(k),\ K = 1, \dots, N$$

where H is a matrix of ones

$$\hat{\theta}_{LS} = (H^T H)^{-1} H^T Z = \frac{1}{N} \sum_{i=1}^{N} Z(i)$$

בעיית מינימום ריבועים רקורסיבית(RLS)
שיטה: נוסיף לוקטור המקורי את המשוואה בזמן k+1

$$Z(k+1) = h^T(k+1)\theta + V(k+1))$$

$$+$$

$$Z(k) = H^T(k)\theta + V(k))$$

$$\hat{\theta}_{WLS}(k+1) = [H(k+1)^T W(k+1) H(k+1)]^{-1} H(k+1)^T W(k+1) Z(k+1)$$

נניח כי W אלכסונית - בשיטה זו ההנחה הכרחית משום שבלעדיה לא ניתן להניח את שלמות שיטת מינימום הריבועים הרקורסיבית.

$$W(k+1) = \begin{matrix} W(k+1) & 0 \\ 0 & W(k) \end{matrix}$$

$$P \underset{def}{=} H^T(k+1) W(k+1) H(k+1) = [h(k) \quad H^T(k)] \begin{bmatrix} W(k+1) & 0 \\ 0 & W(k) \end{bmatrix} [h(k) \quad H^T(k)]]^T$$

$$= h(k+1) W(k+1) h^T(k+1) + H^T(k) W(k) H(k) \underset{def}{=} P^{-1}(k)$$

לבסוף הקשר הנומרי בין המדידה החדשה למידע הקודם:

$$P^{-1}(k+1) = P^{-1} + h(k+1) w(k+1) h^T(k+1)$$