רגרסיה לינארית בייסיאנית

תזכורת: $D = (\underbrace{1,0,0,1,0,\dots}_{\text{1000 הטלות מטבע}})$

$$\theta = \begin{pmatrix} \theta_0 & \text{סיכוי ל-0} \\ \theta_1 & \text{סיכוי ל-1} \end{pmatrix}$$

יכולים להיות אלגוריתמים שונים $\hat{\theta}(D)$ לשערך את θ בהינתן D, אבל אם אנחנו משתמשים ב-

$$BMSE(\hat{\theta}(D)) = \mathbb{E}_{\theta^{GT}, D} \|\hat{\theta}(D) - \theta^{GT}\|^2$$

אז האלגוריתם האופטימלי (שמשיג את השגיאה הנמוכה ביותר) הוא

$$\hat{\theta}^{MMSE}(D) = \mathbb{E}(\theta | D)$$

דוגמה: $D = \{x_i, y_i\}$ כאשר

x_i כמות הגשם שירדה ב-10 ימים האחרונים

y_i כמות הגשם שירדה ביום ה-i

מטרה: לחשב $y_\theta(x)$ כמות הגשם שירד? מחר.

אנחנו מניחים ש $y_\theta(x)$ הוא מהצורה הבאה

$$y_\theta(x)=\sum_n \theta_n h_n(x)$$

עם $h_n(x)$ פונקציות בסיס.

למשל- $y_\theta(x)=\sum_{n=0}^{4}\theta_n x^n$

מהו אלג' לומד טוב? $\hat{\theta}(D)$

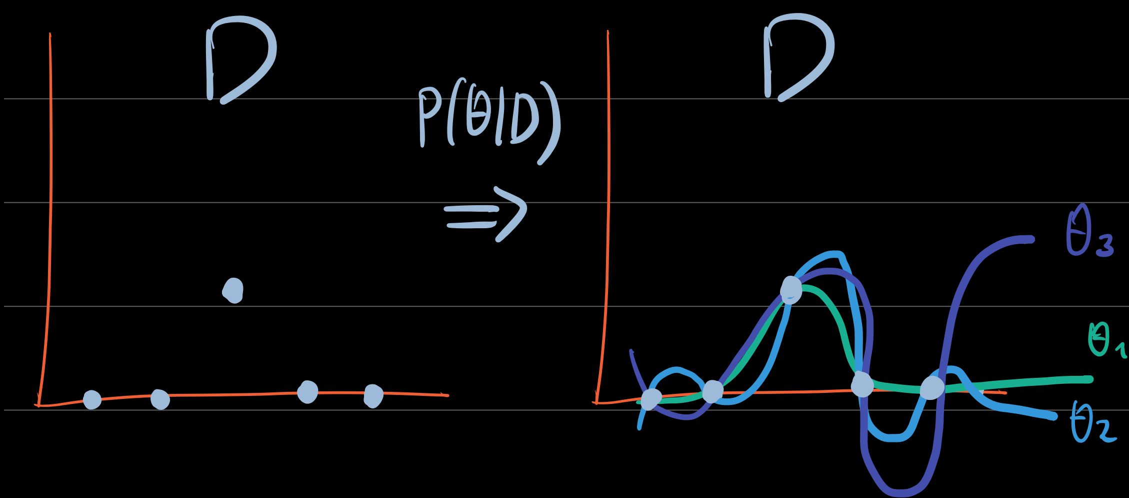
ארגז שמקבל נתונים מ-50 ימים האחרונים ומחזיר מקדמי פולינומים $\hat{\theta}(D)$

<u>שיטת ההערכה-</u>

נחזיר כמה פעמים ונחשב ממוצע.
1. נגריל θ^{GT}
2. נגריל $D|\theta^{GT}$ ונגריל 50 מקדמים מהפולינום
3. נחשב את $\|\hat{\theta}(D)-\theta^{GT}\|^2$

<u>הערה:</u> אם טיב הניבוי מוגדר ע"י $\mathbb{E}([y(x_0)-\hat{y}_D(x_0)]^2)$
האלגוריתם האופטימלי מחשב את $\mathbb{E}(y(x_0)|D)$

כלומר - במקרה זה $\hat{\theta}(D)$ הוא בעצם $y(x_0)$ והחישוב הוא אותו דבר.

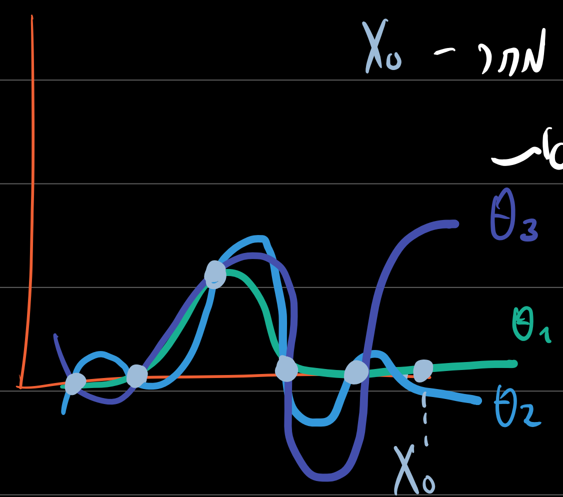


ואם נרצה לדעת מה יקרה מחר - X_0

הדבר הטוב ביותר יהיה לקחת את

התוחלת של $y(x_0)|D$ על

פני כל ה- θ.



הנחה- $\theta \sim N(\mu_\theta, \Sigma_\theta)$

$$D|\theta \sim N(\mu, \Sigma)$$

מוסיפים רעש נורמלי ל-$y(x_i)$

$$P(y_i|x_i,\theta) \sim N\left(\sum_n \theta_n h_n(x_i), \sigma^2\right)$$

מתוך ההנחות הנ"ל -

$$P(D|\theta) = \prod_i P(y_i|x_i,\theta) = $$

$$= \prod_i \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{\left(y_i - \sum_n \theta_n h_n(x_i)\right)^2}{2\sigma^2}\right)$$

עכשיו נרצה על פי ההנחות הנ"ל לחשב התפלגות של θ בהינתן D.

המודל שהצגנו קודם.

<u>המודל הגנרטיבי בכתיב הווקטורי</u> -

$$\theta \sim N(\mu_\theta, \Sigma_\theta) \quad (1$$

$$y|x_i,\theta \sim N(H\cdot\theta, \sigma^2 I) \quad (2$$

כאשר $H_\theta \in P\times N$ עם $H_{i,n} = h_n(x_i)$

כלומר

$$H = \begin{bmatrix} x_1^0 & x_1^1 & x_1^2 & \dots & x_1^N \\ x_2^0 & x_2^1 & x_2^2 & \dots & x_2^N \\ \vdots & & & & \vdots \\ x_P^0 & x_P^1 & x_P^2 & \dots & x_P^N \end{bmatrix}$$

$$\begin{bmatrix} \sigma^2 & & 0 \\ & \ddots & \\ 0 & & \sigma^2 \end{bmatrix}$$

משפט:

אם $\theta \sim N(\mu_\theta, \Sigma_\theta)$ ו- $y|X,\theta \sim N(H\cdot\theta, \sigma^2 I)$

אז-

$$\mu_{\theta|D} = \left(\frac{1}{\sigma^2}H^TH+\Sigma_\theta^{-1}\right)^{-1}\left(\frac{1}{\sigma^2}H^Ty+\Sigma_\theta^{-1}\mu_\theta\right)$$

$$\Sigma_{\theta|D}=\left(\frac{1}{\sigma^2}H^TH+\Sigma_\theta^{-1}\right)^{-1}$$

הוכחה: קודם כל נכתוב את $P(\theta,D)$

$$P(\theta,D)=P(\theta)\cdot P(D|\theta)$$

$$=\frac{1}{Z_1}\exp\left(\frac{-1}{2}(\theta-\mu_\theta)^T\Sigma_\theta^{-1}(\theta-\mu_\theta)\right)\cdot\frac{1}{Z_2}\exp\left(\frac{-1}{2\sigma^2}\|H\theta-y\|^2\right)$$

ולכן $P(\theta|D)=\frac{P(\theta,D)}{P(D)}$

נשים לב ש- $P(\theta,D)$ היא מהצורה $\frac{1}{Z}\exp(-J(\theta))$

כאשר $J(\theta)$ היא תבנית ריבועית של θ.

למה:

אם $P(\theta)=e^{-J(\theta)}$ כאשר $J(\theta)$ פונקציה ריבועית

אז $\theta\sim N(\mu,\Sigma)$ כאשר $\mu=\arg\min_{\theta}\left(J(\theta)\right)$

ו- $\Sigma^{-1}=\frac{\partial^2 J}{\partial^2\theta}$

הוכחה: תהי $J(\theta)=\frac{1}{2}(\theta-\mu)^T\Sigma^{-1}(\theta-\mu)$

אז בוודאי ש- $\arg\min_{\theta} J(\theta)=\mu$ (תבנית ריבועית אי שלילית). קל גם לראות ש- $\frac{\partial^2 J}{\partial^2\theta}=\Sigma^{-1}$

אם כן - ההתפלגות מוגדרת ע"י -

$$P(\theta)=\frac{1}{Z}\exp\left(J(\theta)\right)$$
$$=\frac{1}{Z}\exp\left(-\frac{1}{2}(\theta-\mu)^T\Sigma^{-1}(\theta-\mu)\right)$$

כלומר $\theta\sim N(\mu,\Sigma)$

נחזור להוכחת המשפט. במקרה שלנו -

$$J=\frac{1}{2}(\theta-\mu_\theta)^T\Sigma_\theta^{-1}(\theta-\mu_\theta)+\frac{1}{2\sigma^2}(H\theta-y)^T(H\theta-y)$$

נגזור ונשווה לאפס למציאת $\mu_{\theta|D}$

$$\frac{\partial J}{\partial \theta} = (\theta - \mu_\theta)^T \Sigma_\theta^{-1} + \frac{1}{\sigma^2}(H\theta - y)^T H$$

$$\frac{\partial J}{\partial \theta} := 0 \Rightarrow \theta^T \Sigma_\theta^{-1} - \mu_\theta^T \Sigma_\theta^{-1} = -\frac{1}{\sigma^2}\theta^T H^T H + \frac{1}{\sigma^2} y^T H$$

$$\Rightarrow \theta^T \Sigma_\theta^{-1} + \frac{1}{\sigma^2}\theta^T H^T H = \frac{1}{\sigma^2} y^T H + \mu_\theta^T \Sigma_\theta^{-1}$$

$$\Rightarrow \theta^T \left(\Sigma_\theta^{-1} + \frac{1}{\sigma^2} H^T H\right) = \frac{1}{\sigma^2} y^T H + \mu_\theta^T \Sigma_\theta^{-1}$$

$$\Rightarrow \theta = \left(\frac{1}{\sigma^2} H^T H + \Sigma_\theta^{-1}\right)^{-1} \left(\Sigma_\theta^{-1}\mu_\theta + \frac{1}{\sigma^2} H^T y\right)$$

נגזור פעם שניה למציאת $\Sigma_{\theta|D}$

$$\frac{\partial^2 J}{\partial^2 \theta} = \Sigma_\theta^{-1} + \frac{1}{\sigma^2} H^T H$$

$$\Rightarrow \Sigma_{\theta|D} = \left(\Sigma_\theta^{-1} + \frac{1}{\sigma^2} H^T H\right)^{-1}$$