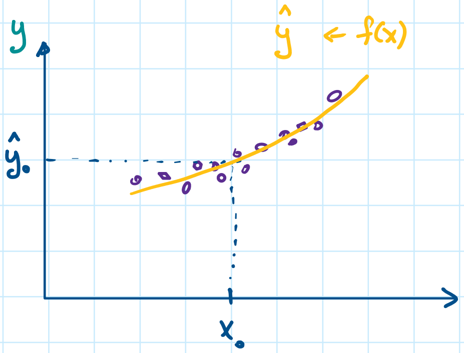
Lecture4 - Prediction

Thursday, 28 November 2024 16:12

חיזוי - רעיון כללי

מטרה: לקבל חיזוי $\hat{y}_0$ עבור x_0



הנחת יסוד:

מודל $\hat{y} = \underbrace{f(x)}_{\text{לא תמיד ידוע}} + \underbrace{\text{רעש}}_{\text{אקראי}}$

$f_1(x) = b$

קו מגמה $f_2(x) = ax + b$

<u>מטרה</u>: ערכים של a,b

חיזוי מבוסס מודל ← ע"י ערכים של a,b מקבלים חיזוי

<u>מדד הצלחה</u> שגיאה ריבועית

$$\underbrace{SE}_{\text{square error}} = E\big[(\underset{\text{ערך "אמיתי"}}{Y} - \underset{\text{חיזוי}}{\hat{Y}})^2\big]$$

חיזוי ע"י קבוע $\hat{Y} = b$

<u>מטרה</u>: למצוא ערך של b, עבורו SE היא מינימלי MSE (minimum)

$SE(b) = E[(Y-b)^2]$

↓

ממוצע את כל התוצאות האפשריות של ניסוי אקראי.

פתיחת סוגריים

$$se(b) = E[(Y-b)^2] = E[Y^2 - 2bY + b^2] = E[Y^2] - 2bE[Y] + b^2$$

חיפוש מינימום

$$\frac{d\,se(b)}{d\,b} = -2E[Y] + 2b = 0$$

<u>סיכום</u>:

<u>חיזוי</u> $b_{opt} = E[Y]$

<u>שגיאת חיזוי</u> $mse_{min} = E[(Y - E[Y])^2] = \text{Var}[Y]$

חיזוי לינארי $\hat{Y} = ax + b$

שמות נוספים:
- רגרסיה לינארית
- קו מגמה לינארי

שגיאה ריבועית

$$se(a,b) = E\left[(\hat{Y} - Y)^2\right]$$

מינימום ע"י נגזרת

כלל שרשרת

$$[f(g(x))]' = f'(g(x))\,g'(x)$$

נגזרת הפנימית

$$\frac{\partial mse(a,b)}{\partial a} = E\left[2(Y - aX - b)(-X)\right] = -2E[XY] + 2aE[X^2] + 2bE[X] = 0$$

$$\frac{\partial mse(a,b)}{\partial b} = E\left[2(Y - aX - b)(-1)\right] = -2E[Y] + 2aE[X] + 2b = 0$$

זוג משוואות ב-2 נעלמים

סיכום ע"י מטריצה

$$\begin{bmatrix} E[X^2] & E[X] \\ E[X] & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} E[XY] \\ E[Y] \end{bmatrix}$$

$$a_{opt} = \frac{E[XY] - E[X]\,E[Y]}{E[X^2] - E^2[X]} = \frac{\text{Cov}[X,Y]}{\text{Var}[X]}$$

$$b_{opt} = E[Y] - \frac{\text{Cov}[X,Y]}{\text{Var}[X]}E[X]$$

שונות משותפת

מדד כמותי להצלחה של חיזוי לינארי.

תשובה סופית עבור a,b

דרך רישום נוספת

$$\hat{Y} = a_{opt}x + b_{opt} = \frac{\text{Cov}[X,Y]}{\text{Var}[X]}x + E[Y] - \frac{\text{Cov}[X,Y]}{\text{Var}[X]}E[X] = E[Y] + \frac{\text{Cov}[X,Y]}{\text{Var}[X]}\big(x - E[X]\big)$$

השוואה לחיזוי ע"י קבוע

$$mse_{min} = E[(Y - E[Y])^2] = \text{Var}[Y]$$

סיכום:

$$mse_{min} = E\left[\left(Y - (a_{opt}X + b_{opt})\right)^2\right] = \text{Var}[Y]\left(1 - \rho_{XY}^2\right)$$

השיפור הוא פי $(1 - \rho_{XY}^2)$

$$\rho_{XY} = \frac{\text{Cov}[X,Y]}{\ldots}$$

$\rho_{XY} = 0$ אין קשר לינארי ⇐ משמעות למקדם a

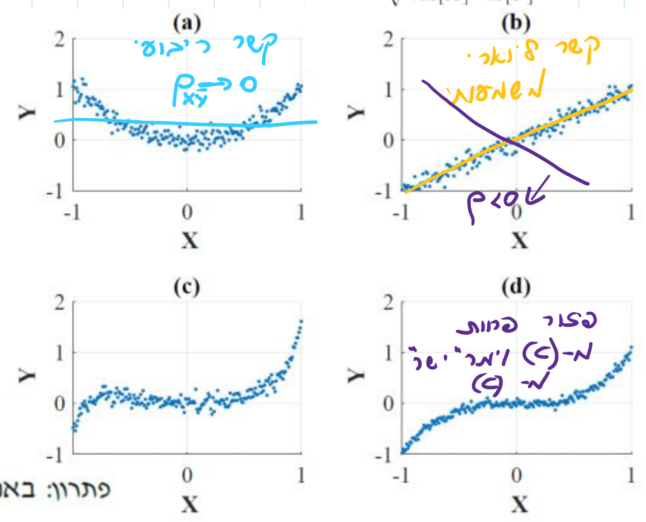
התוצאה זהה לחיזוי ע"י b בלבד

התוצאה כזה לחיצוני ע״י b בלבד

$\rho_{xy} = \pm 1$ קשר ליניארי מובהק ⇐ חיצוני <u>מושלם</u>
סימן = סימן של שיפוע

$$\rho_{XY} = \frac{\mathrm{Cov}[X,Y]}{\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}}$$

שאלה: יש למיין גרפים לפי $|\rho_{xy}|$



הערה: לא ניתן לקבל ערך מספרי של ρ_{xy} ללא חישוב מסודר

1. (b) 0.98621
2. (d) 0.90839
3. (c) 0.6233
4. (a) −0.037787

פתרון: באופן כללי,

- ☐ עבור קו "ישר" יותר $|\rho_{XY}|$ יותר קרוב ל-1.
- ☐ עבור קו ישר, עבור פיזור "קטן" יותר $|\rho_{XY}|$ יותר קרוב ל-1.

נתונים זוג משתנים אקראיים בלתי תלויים $W_1, W_2 \sim U[0,1]$ עבורם מתקיים

קשר ליניארי

$$X = W_1$$
$$Y = 0.9W_1 + 0.1W_2$$

תוספת רעש

$? = E[X], E[Y], \mathrm{Var}[X], \mathrm{Var}[Y], E[XY], \mathrm{Cov}[X,Y], \rho_{XY}$

עבור $X \sim U[a,b]$ מתקיים (איור 2.1א)

$a=0$ $b=1$

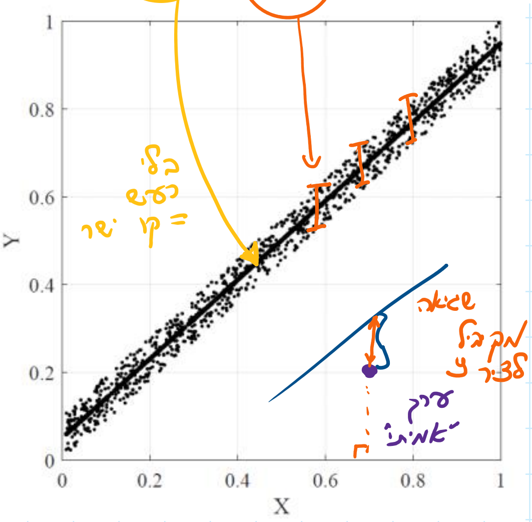
תכונות הרצאה מס׳ 2 התפלגות אחידה

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a \leqslant x \leqslant b \\ 0 & \text{אחרת} \end{cases}$$

$$E[X] = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\cdot\frac{x^2}{2}\Big|_a^b = \frac{a+b}{2}$$

$$E\left[X^2\right] = \int_a^b x^2\frac{1}{b-a}dx = \frac{1}{b-a}\cdot\frac{x^3}{3}\Big|_a^b = \frac{1}{3}\frac{b^3-a^3}{b-a} = \frac{b^2+ab+a^2}{3}$$

$$\mathrm{Var}[X] = E\left[X^2\right] - E^2[X] = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} = \frac{(b-a)^2}{12}$$



$$E[X] = E[W_1] = \frac{1}{2} = E[W_2]$$
$$E[Y] = 0.9E[W_1] + 0.1E[W_2] = \frac{1}{2}$$
$$\mathrm{Var}[X] = \mathrm{Var}[W_1] = \frac{1}{12}$$

הצבה

שונות של זוג משתנים בלתי תלויים

$$\mathrm{Var}[cX + dY] = c^2\mathrm{Var}[X] + d^2\mathrm{Var}[Y]$$

$$X = W_1$$
$$Y = 0.9W_1 + 0.1W_2$$

$c=0.9$ $d=0.1$

$$\mathrm{Var}[Y] = \left(\frac{9}{10}\right)^2\mathrm{Var}[W_1] + \frac{1}{10^2}\mathrm{Var}[W_2] = \frac{82}{100}\cdot\frac{1}{12} = \frac{41}{600} \approx 0.06833$$

סיכום:

$$E[XY] = 0.9E\left[W_1^2\right] + 0.1E[W_1W_2]$$

$$0.1E[W_1]\,E[W_2]$$

הצבה

שונות משותפת

$$\mathrm{Cov}[X,Y] = E[XY] - E[X]\,E[Y] = \left(\frac{3}{10}+\frac{1}{40}\right) - \frac{1}{2}\cdot\frac{1}{2} = \frac{3}{40} = 0.075$$

הצבות

$$\rho_{XY} = \frac{3/40}{\sqrt{\frac{1}{12}\cdot\frac{41}{600}}} = \frac{9}{\sqrt{82}} \approx 0.99388$$

מאוד קרוב ל-1

$$\hat{Y} = \frac{1}{2} + \frac{3/40}{1/12}\left(x-\frac{1}{2}\right) = \frac{1}{2} + \frac{9}{10}\left(x-\frac{1}{2}\right)$$

$$mse = \mathrm{Var}[Y]\left(1-\rho_{XY}^2\right) = \frac{41}{600}\left(1-\frac{81}{82}\right) = \frac{1}{1200} \approx 8.33\times10^{-4}$$

הערה: X,Y <u>לא</u> בלתי תלויים
W_1, W_2 כן בלתי תלויים

דוגמה: $X_1, X_2, \ldots X_n$ n משתנים ? הצבה תוחלת ליניארי הצבה

600 (82) 1200

דוגמה: $X_1, X_2, \ldots, X_n$ n משתנים

נתון
$$E[X_i] = m \qquad \forall i$$
$$Var[X_i] = \sigma^2$$

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$Y_j = X_j - \bar{X}, \quad j = 1, \ldots, n$$

צריך להוכיח:

* $E[Y_j] = 0$
* $E\left[\bar{X}Y_j\right] = 0$

הצבה, תוחלת לינארית, הצבה, ?

$$0 \stackrel{?}{=} E[Y_j] = E\left[X_j - \bar{X}\right] = E[X_j] - E\left[\bar{X}\right] = m - E\left[\bar{X}\right]$$

$$E\left[\bar{X}\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\underbrace{\sum_{i=1}^{n} m}_{n \text{ פעמים } m} = m = \frac{1}{n}\cdot n\cdot m$$

מתקיים תמיד
$$\boxed{E[X_1 + \ldots + X_n] = E[X_1] + \ldots + E[X_n]}$$

סיכום
$$E[Y_j] = m - m = 0$$

$$E\left[\bar{X}Y_j\right] = E\left[\bar{X}(X_j - \bar{X})\right] = \overset{①}{E\left[\bar{X}X_j\right]} - \overset{②}{E\left[\bar{X}\bar{X}\right]}$$

② $$E\left[\bar{X}\bar{X}\right] = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E[X_iX_j]$$

הצבה + תכונה

$$\boxed{E\left[X^2\right] = \underbrace{\mathrm{Var}[X]}_{\sigma^2} + \underbrace{E^2[X]}_{m^2}}$$

א) $i=j$ $\quad E[x_i x_j] = E[x_i^2] = \sigma^2 + m^2$

2) $i \neq j$ $\quad E[x_i]E[x_j] = m^2$

$i=j$ ישנם n פעם לגבי

$i \neq j$ $\quad n(n-1) = n^2 - n$

סה"כ אפשרויות (i,j)

$$= \frac{1}{n^2}\left(n(n-1)m^2 + n(\sigma^2 + m^2)\right)$$
$$= \frac{\sigma^2}{n} + m^2$$

①

הצבה
תוחלת ← x_j

$$E\left[\bar{X}X_j\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_iX_j]$$
$$= \frac{1}{n}\sum_{i \neq j}^{n} E[X_iX_j] + \frac{1}{n}\underbrace{E[X_jX_j]}_{i=j}$$
$$= \frac{1}{n}\left[(n-1)m^2 + (\sigma^2 + m^2)\right] = \frac{\sigma^2}{n} + m^2$$

X_i, X_j בלתי תלויים

$E[x_i x_j] = E[x_i]E[x_j]$ $\quad i \neq j$ מותר

$E[x_i x_i] = E[x_i^2]$ $\quad i=j$ לא ניתן להפריד ריבוע של משתנה.