امین اسدی 810196410

تمرین چهارم یادگیری ماشین

سؤالات پاسخ داده شده: 2 و 3 ، 4 و 6 ، 9 و 10 ، 11 +همزمان

تشریحی: 2 و 3 ، 4 و 6

عملی: 9 و 10 ، 11

2- الف)

$$S_B = \sum_{k=1}^{c} N_k(\mu_k - \mu)(\mu_k - \mu)^T$$

$$= \underbrace{\left[\sqrt{N_1}(\mu_1 - \mu), \sqrt{N_2}(\mu_2 - \mu), \dots, \sqrt{N_c}(\mu_c - \mu)\right]}_{A} \cdot \underbrace{\begin{bmatrix} \sqrt{N_1}(\mu_1 - \mu) \\ \vdots \\ \sqrt{N_c}(\mu_c - \mu) \end{bmatrix}}_{B}$$

کافیست نشان دهیم ستون‌های این ماتریس به صورت خطی به هم وابسته اند.

$$A \cdot \begin{bmatrix} \sqrt{N_1} \\ \vdots \\ \sqrt{N_c} \end{bmatrix} = (N_1\mu_1 + N_2\mu_2 + \dots + N_c\mu_c) - (N_1 + \dots + N_c)\mu$$

$$= N\mu - N\mu = 0 \rightarrow \text{rank}(A) \leq c-1$$

از مفاهیم جبر خطی داریم: $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$

$$\rightarrow \text{rank}(S_B) \leq \text{ran}(K) \leq c-1$$

$$\rightarrow \boxed{\text{ran}(S_B) \leq c-1}$$

2- ب

$$\text{rank}(S_w^{-1} S_B) \leq \min(\text{rank}(S_w^{-1}), \text{rank}(S_B))$$

$$\rightarrow \text{rank}(S_w^{-1} S_B) \leq \text{rank}(S_B) \leq c-1 \quad ✓$$

Scanned by CamScanner

سوال 2)

ت)

$$S_W = \sum_{k=1}^{c} \sum_{x^q \in \omega_k} (x^q - \mu_k)(x^q - \mu_k)^T$$

$$S_B = \sum_{k=1}^{c} N_k (\mu_k - \mu)(\mu_k - \mu)^T$$

$$S_T = \sum_{k=1}^{c} \sum_{x^q \in \omega_k} (x^q - \mu)(x^q - \mu)^T = \sum_{k=1}^{c} \sum_{x^q \in \omega_k} (x^q - \mu_k + \mu_k - \mu)(x^q - \mu_k + \mu_k - \mu)^T$$

$$= \underbrace{\sum_{k=1}^{c} \sum_{x^q \in \omega_k} (x^q - \mu_k)(x^q - \mu_k)^T}_{S_W} + \underbrace{\sum_{k=1}^{c} \underbrace{\sum_{x^q \in \omega_k}}_{N_k} (\mu_k - \mu)(\mu_k - \mu)^T}_{S_B}$$

$$+ \sum_{k=1} \sum_{x^q \in \omega_k} \left[(x^q - \mu_k)(\mu_k - \mu)^T + (\mu_k - \mu)(x^q - \mu_k)^T \right]$$

$$= S_W + S_B + \sum_{k=1} \left[(\mu_k - \mu) \underbrace{\sum_{x^q \in \omega_k} (x^q - \mu_k)}_{0} \right] + \sum_{k=1} \left[(\mu_k - \mu) \underbrace{\sum_{x^q \in \omega_k} (x^q - \mu_k)^T}_{0} \right]$$

مجموع تفاضلات از میانگین صفر است

مجموع تفاضلات از میانگین صفر است

$$= S_W + S_B \rightarrow \boxed{S_T = S_W + S_B}$$

سوال 3

$$y_i - y_j = y_i - y_j + m_1 - m_1 + m_2 - m_2$$

$$= (y_i - m_1) - (y_j - m_2) + m_1 - m_2$$

$$\rightarrow d_1 = \frac{1}{n_1 n_2} \sum_{y_i \in Y_i} \sum_{y_j \in Y_j} (y_i - y_j)^2 = \frac{1}{n_1 n_2} \sum_{y_i \in Y_i} \sum_{y_j \in Y_j} \left[(y_i - m_1) - (y_j - m_2) + (m_1 - m_2)\right]^2$$

مجذور سه جمله ای

$$= \frac{1}{n_1 n_2} \sum_{i \in Y_i} \sum_{j \in Y_j} (y_i - m_1)^2 + (y_j - m_2)^2 + (m_1 - m_2)^2 + 2(y_i - m_1)(y_j - m_2) + 2(y_j - m_2)(m_1 - m_2)$$
$$+ 2(y_i - m_1)(m_1 - m_2)$$

$$= \frac{1}{n_1 n_2} \Bigg(\sum_{i}^{n_1} \sum_{j}^{n_2} (y_i - m_1)^2 + \sum_{i}^{n_1} \sum_{j}^{n_2} (y_j - m_2)^2 + \sum_{i}^{n_1} \sum_{j}^{n_2} (m_1 - m_2)^2$$
$$+ (m_1 - m_2) \sum_{i}^{n_1} \sum_{j}^{n_2} 2(y_j - m_2) + (m_1 - m_2) \sum_{i}^{n_1} \sum_{j}^{n_2} 2(y_i - m_1)$$
$$+ \sum_{i}^{n_1} \sum_{j}^{n_2} (y_i - m_1)(y_j - m_2)$$

$$= \frac{\cancel{n_2}^{\frac{1}{n_1}}}{\cancel{n_1} n_2} S_1^2 + \frac{\cancel{n_1}^{\frac{1}{n_2}}}{\cancel{n_1} n_2} S_2^2 + \frac{\cancel{n_1 n_2}^{1}}{\cancel{n_1 n_2}} (m_1 - m_2)^2$$

$$+ (m_1 - m_2) \underbrace{\cancel{\left(\text{Sum of differences of } y_j \text{ from mean} \right)}}_{0 \;\to\; \sum_j (y_j - m_2)} + (m_1 - m_2) n_2 \underbrace{\cancel{\left(\text{sum of differences of } y_i \text{ from mean} \right)}}_{0 \;\to\; \sum_i (y_i - m_1)}$$

$$+ \sum^{n_1} (y_i - m_1) \underbrace{\cancel{\sum_j^{n_2} (y_j - m_2)}}_{0} = \boxed{(m_1 - m_2)^2 + \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

(از این نکته استفاده شد که مجموع اختلافات از میانگین صفر است)

سؤال ۴

$$P_n(x) = \frac{1}{n} \sum_{n=1}^{n} \frac{1}{V_n} \varphi\left(\frac{x - x_i}{h_n}\right) \qquad d=1 \to V_n = h_n$$

$$\to \overline{P_n(x)} = E(P(x)) = E\left[\frac{1}{n} \sum_{n=1}^{n} \frac{1}{h_n} \varphi\left(\frac{x - x_i}{h_n}\right)\right]$$

$$= \int \frac{1}{h_n} \varphi\left(\frac{x - v}{h_n}\right) p(v)\, dv = \int_{x-v \geq 0} \frac{1}{h_n} e^{-\frac{(x-v)}{h_n}} p(v)\, dv + 0$$

$$= \frac{1}{h_n} \int_{\substack{x \geq v \\ 0 \leq v \leq a}} e^{-\frac{(x-v)}{h_n}} \times \frac{1}{a}\, dv = \frac{1}{h_n} \times \int_{\substack{x \geq v \\ 0 \leq v \leq a}} e^{\frac{-x}{h_n}} \cdot e^{\frac{v}{h_n}} \times \frac{1}{a}\, dv$$

$$= \frac{e^{\frac{-x}{h_n}}}{h_n} \int_{\substack{0 \leq v \leq a \\ x \geq v}} \frac{e^{\frac{v}{h_n}}}{a}\, dv = \frac{e^{\frac{-x}{h_n}}}{h_n} \times \begin{cases} \int_0^a \frac{e^{\frac{v}{h_n}}}{a}\, dv & x \geq a \\ \int_0^x \frac{e^{\frac{v}{h_n}}}{a}\, dv & 0 \leq x \leq a \\ 0 & x < 0 \end{cases}$$

$$= \frac{e^{\frac{-x}{h_n}}}{h_n} \times \begin{cases} \frac{h_n}{a}\left(e^{\frac{a}{h_n}} - 1\right) & x \geq a \\ \frac{h_n}{a}\left(e^{\frac{x}{h_n}} - 1\right) & 0 \leq x \leq a \\ 0 & x < 0 \end{cases}$$

$$= \begin{cases} 0 & x < 0 \\ \frac{e^{\frac{-x}{h_n}}}{\cancel{h_n}} \times \frac{\cancel{h_n}}{a}\left(e^{\frac{x}{h_n}} - 1\right) & 0 \leq x \leq a \\ \frac{e^{\frac{-x}{h_n}}}{\cancel{h_n}} \times \frac{\cancel{h_n}}{a}\left(e^{\frac{a}{h_n}} - 1\right) & x \geq a \end{cases} = \begin{cases} 0 & x < 0 \\ \frac{1}{a}\left(1 - e^{\frac{-x}{h_n}}\right) & 0 \leq x \leq a \\ \frac{1}{a}\left(e^{\frac{a}{h_n}} - 1\right) e^{\frac{-x}{h_n}} & a \leq x \end{cases}$$

سؤال ۶ آ)

adaboost یکی روش‌های ensemble model به روش boosting است.

در روش های boosting به طور کلی چندین مدل پایه به صورت ترتیبی استفاده می‌شود به طوری که هر مدل از اشتباهات انجام شده در مدل قبلی خود یادگیری می‌کند.

در adaboost برای یادگیری از اشتباهات مدل قبلی (نقاط اشتباه طبقه‌بندی)، وزن داده شده به این نقاط را افزایش می دهیم. این روش وقتی از درخت تصمیم استفاده می‌کند عملکرد خوبی دارد.

مزیت روش adaboost نسبت به درخت تصمیم این است که مشکل overfitting را به مقدار خوبی حل می کند. از نقاط ضعف این روش این است که adaboost نسبت به outlier ها حساس است زیرا هر مدل باید اشتباهات مدل قبلی خود را درست کند. از دیگر نقاط ضعف این روش این است که این روش به طور متوالی از مدل ها استفاده می کند که در نتیجه کار هر مدل باید بعد از مدل قبلی انجام شود که این باعث می‌شود scale up کردن غیرممکن باشد.

ب) برای اثبات از دو نکته استفاده می‌کنیم:

۱- همیشه داریم: $e^{-x} \geq 0$

۲- $x \leq 0 \longleftrightarrow e^{-x} \geq 1$

خروجی $H(x)$ 1 یا -1 است پس y_i دو حالت دارد: $y_i = \pm 1$

$f(x)$ هم یا مثبت است یا منفی $\Leftarrow$ ۴ حالت داریم:

الف	$f(x) > 0$	$y_i = 1$
ب	$f(x) > 0$	$y_i = -1$
ج	$f(x) < 0$	$y_i = 1$
د	$f(x) < 0$	$y_i = -1$

الف) $f(x_i) > 0 \rightarrow H(x_i) = 1 \rightarrow H(x_i) \neq y_i = 0$
$\exp(-f(x_i)y_i) \geq 0$

ب) $f(x_i) > 0 \rightarrow H(x_i) = 1 \rightarrow H(x_i) \neq y_i = 1$
$\exp(-f(x_i)y_i) \geq 1$

ج) $f(x_i) < 0 \rightarrow H(x_i) = -1 \rightarrow H(x_i) \neq y_i = 1$
$\exp(-f(x_i)y_i) \geq 1$

د) $f(x_i) < 0 \rightarrow H(x_i) = -1 \rightarrow H(x_i) \neq y_i = 0$
$\exp(-f(x_i)y_i) \geq 0$

همانطور که مشاهده می‌شود، در هر چهار حالت داریم:

$$\overbrace{H(x_i) \neq y_i}^{\text{boolean}} \leq \exp(-f(x_i)y_i)$$

$$\rightarrow \frac{1}{m}\sum_{i=1}^{m}(H(x_i) \neq y_i) \leq \frac{1}{m}\sum_{i=1}^{m}\exp(-f(x_i)y_i) \quad ✓$$

Scanned by CamScanner