Курасбедиани Г.К.
3824М1ПМки

№35

* задачу восстановления регрессии с квадратичной функцией потерь $L(y', y) = (y' - y)^2$.
Д-ть, что если $f^*(x) = \arg\min\limits_c E((Y-c)^2 \mid X = x)$, то $f^*(x) = E(Y \mid X = x)$.
Чему равен средний риск $R(f^*)$?

$* \ g(c) = E((Y-c)^2 \mid X = x)$

$g'(c) = E(-2(Y-c) \mid X = x) = -2E((Y-c) \mid X = x)$

$g''(c) = 2 > 0 \Rightarrow$ т. min

$g'(c) = 0 \Leftrightarrow -2E((Y-c) \mid X = x) = 0$

$E(Y - c \mid X = x) = 0$

$E(Y \mid X = x) - c = 0 \Rightarrow c = E(Y \mid X = x)$

$\Rightarrow f^*(x) = E(Y \mid X = x)$

$R(f^*) = \int E((Y - E(Y \mid x))^2 \mid x)\, p(x)\, dx =$

$= \int Var(Y \mid x)\, p(x)\, dx = E(Var(Y \mid x))$

№36

* задачу восстановления регрессии с функцией потерь $L(y', y) = |y' - y|$.
Д-ть, что минимум среднего риска достигается при этом условии медиана $f(x) = median(Y \mid X = x)$

$R(f) = \int E(|f(x) - Y| \mid X = x)\, p(x)\, dx$

$f^*(x) = \arg\min\limits_c E(|c - Y| \mid x)$

$* \ g(c) = E(|c - Y| \mid x)$

$g(c) = \int\limits_{-\infty}^{+\infty} |c - y|\, p(y \mid x)\, dy$

$g(c) = \int\limits_{-\infty}^{c} (c - y)\, p(y \mid x)\, dy + \int\limits_{c}^{+\infty} (y - c)\, p(y \mid x)\, dy$

$g'(c) = \int\limits_{-\infty}^{c} p(y \mid x)\, dy - \int\limits_{c}^{+\infty} p(y \mid x)\, dy = F_{y|x}(c) - (1 - F_{y|x}(c)) =$

$= 2F_{y|x}(c) - 1$

Курбанова Г.К
3824 МиПМкИ

$y'(c) = 0 \Leftrightarrow 2F_{y|x}(c) - 1 = 0$

$F_{y|x}(c) = 1/2 \Rightarrow c = \mathrm{median}(Y \mid X = x)$

$\Rightarrow f^*(x) = \mathrm{median}(Y \mid X = x)$

№37

Как должна выглядеть функция потерь, чтобы минимум средниму риску был условной модой

Условная мода: $\mathrm{mode}(X_\cdot) = \arg\max_x p_{X|Y}(x)$

$L(y', y) = -\delta(y - y')$

$R(f) = -\int_X \int_Y \delta(y - f(x))\, p(y|x)\, p(x)\, dy\, dx = -\int_X p(f(x) \mid x)\, p(x)\, dx$

$f^*(x) = \arg\min_c -p(c|x) = \arg\max_c p(c|x)$

№38

∃ в задаче классификации с двумя классами $\{0,1\}$ используется функция потерь $L(y', y)$, такая, что $L(0,0) = L(1,1) = 0$, $L(1,0) = L_1$, $L(0,1) = L_0$.

Д.т., что в этом случае байесов классификатор удовлетворяет условию

$f(x) = \arg\max_{y \in \{0,1\}} L_y \Pr(y|x)$

$R(f) = \int_X \left(\sum_{y=0}^{1} L(f(x), y) \Pr(y|x) \right) p(x)\, dx$

$y' = 0: \quad R(f) = \int_X L_0 \Pr(1|x)\, p(x)\, dx$

$y' = 1 \quad R(f) = \int_X L_1 \Pr(0|x)\, p(x)\, dx$

$\Rightarrow f^*(x) = \arg\min_{y \in \{0,1\}} L_y \Pr(\bar{y}|x)$

$f^*(x) = \arg\min_{y \in \{0,1\}} L_y \Pr(\bar{y}|x) = \arg\min_{y \in \{0,1\}} L_y (1 - \Pr(y|x)) = \arg\max_{y \in \{0,1\}} L_y \Pr(y|x)$

№39

Выразить байесов классификатор $f^*(x)$ для задачи классификации с K классами, если ф-я потерь $L(y', y) = L_{y'y}$ $(y', y = \overline{1,K})$

$R(f) = \int_X \left(\sum_{y=1}^{K} L_{f(x), y} \Pr(Y = y \mid X = x) \right) p(x)\, dx$

$f^*(x) = \arg\min_f R(f) = \arg\min_{y' \in \{1,K\}} \sum_{y=1}^{K} L_{y'y} \Pr(y|x)$