Домашняя работа №2

Огурцов Константин

8 октября 2024 г.

Задача 1

$$\hat{y}(u) = \arg\min_{c \in \mathcal{R}} \sum_{j=1}^{k} w_j \left(y_u^{(j)} - c \right)^2$$

Если в функции, которую мы минимизируем раскрыть все скобки, то получим:

$$L(c) = c^2 \sum_{j=1}^{k} w_j - 2c \sum_{j=1}^{k} w_j y_u^{(j)} + \sum_{j=1}^{k} w_j (y_u^{(j)})^2$$

Это парабола ветвями вверх, точка минимума (тот самый $\hat{y}(u)$) находится по всем знакомой формуле:

$$c^* = \hat{y}(u) = \frac{\sum_{j=1}^{k} w_j y_u^{(j)}}{\sum_{j=1}^{k} w_j}$$

Что и требовалось доказать.

Задача 2

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\sigma(x)' = \frac{0 - (-e^{-x})}{(1 + e^{-x})^2} = \frac{e^{-x}}{(1 + e^{-x})^2}$$

В то же время:

$$\sigma(x)(1 - \sigma(x)) = \frac{1}{1 + e^{-x}} \cdot \frac{e^{-x}}{1 + e^{-x}} = \frac{e^{-x}}{(1 + e^{-x})^2} = \sigma(x)'$$

Что и требовалось доказать.

Задача 3

x_1	y
0	0.1
0.5	0.9
1	2.1
1.5	2.9

$$\hat{y} = w_0 + w_1 x_1$$

$$w^{(0)} = (w_0, w_1) = (1,\ 1)$$

$$\eta = 0.5$$

В данном случае не важно как воспринимать w, как вектор-стобец или как вектор-строка.
Используем функцию потерь MSE:

$$L(w) = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

В данном случае у нас 4 наблюдения, значит n=4. Можем записать функцию в данном виде:

$$L(w_0,\ w_1) = \frac{1}{4}((w_0-0.1)^2+(w_0+0.5w_1-0.9)^2+(w_0+w_1-2.1)^2+(w_0+1.5w_1-2.9)^2)$$

$$\nabla L = \left(\frac{\partial L}{\partial w_0},\ \frac{\partial L}{\partial w_1}\right)$$

$$\frac{\partial L}{\partial w_0} = \frac{1}{2}((w_0-0.1)+(w_0+0.5w_1-0.9)+(w_0+w_1-2.1)+(w_0+1.5w_1-2.9)) =$$

$$= 2w_0 + 1.5w_1 - 3$$

$$\frac{\partial L}{\partial w_1} = \frac{1}{2}(0.5(w_0+0.5w_1-0.9)+(w_0+w_1-2.1)+1.5(w_0+1.5w_1-2.9)) =$$

$$= 1.5w_0 + 1.75w_1 - 3.45$$

На первой итерации $\nabla L = (0.5,\ -0.2)$. Тогда можем произвести эту первую итерацию:

$$w^{(1)} = (1,\ 1) - 0.5(0.5,\ -0.2) = (0.75,\ 1.1)$$

На второй итерации $\nabla L = (0.15,\ -0.4)$. Тогда можем произвести эту вторую итерацию:

$$w^{(2)} = (0.75,\ 1.1) - 0.5(0.15,\ -0.4) = (0.675,\ 1.3)$$

На третьей итерации $\nabla L = (0.3,\ -0.1625)$. Тогда можем произвести эту третью итерацию:

$$w^{(3)} = (0.675,\ 1.3) - 0.5(0.3,\ -0.1625) = (0.525,\ 1.38125)$$

Ответ: после трех итерации получили набор весов $(w_0,\ w_1) = (0.525,\ 1.38125)$

Задача 4

$$\text{X} = \begin{pmatrix} 1 & 0 \\ 1 & 0.5 \\ 1 & 1 \\ 1 & 1.5 \end{pmatrix}$$

$$\text{y} = \begin{pmatrix} 0.1 \\ 0.9 \\ 2.1 \\ 2.9 \end{pmatrix}$$

$$\text{w} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}$$

Найдем вектор весов w по формуле с лекции:

$$w = (X^T \cdot X)^{-1} \cdot X^T \cdot y$$

$$X^T \cdot X = \begin{pmatrix} 4 & 3 \\ 3 & 3.5 \end{pmatrix}$$

$$(X^T \cdot X)^{-1} = \begin{pmatrix} \frac{7}{10} & \frac{-3}{5} \\ \frac{-3}{5} & \frac{4}{5} \end{pmatrix}$$

$$(X^T \cdot X)^{-1} \cdot X^T = \begin{pmatrix} \frac{7}{10} & \frac{2}{5} & \frac{1}{10} & \frac{-1}{5} \\ \frac{-3}{5} & \frac{-1}{5} & \frac{1}{5} & \frac{3}{5} \end{pmatrix}$$

$$(X^T \cdot X)^{-1} \cdot X^T \cdot y = w = \begin{pmatrix} \frac{3}{50} \\ \frac{48}{25} \end{pmatrix}$$

Итого решением задачи является набор весов $(w_0,\ w_1) = (0.06,\ 1,92)$

Задача 5

Функция потерь с L_2 регулизацией задается следующим образом:

$$L_\alpha(w) = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 - \alpha \sum_{j=1}^{d} w_j^2$$

Ее можно записать в матричной форме (используем обозначения с лекции):

$$L_\alpha(w) = (Xw - y)^T (Xw - y) - \alpha w^T w$$

Возьмем производную по вектору w и приравняем ее к 0:

$$L'_\alpha(w) = 2X^T(Xw - y) - 2\alpha w = 0$$

$$(X^T \cdot X - \alpha I)w - X^T \cdot y = 0$$

Выразим w:

$$w^* = (X^T \cdot X - \alpha I)^{-1} X^T \cdot y$$

Что и требовалось доказать.