Problem Set 2

Sept. 23, 2021

1. (Optional) Search on the Internet or in the literature about the history of people's study to show smoking is harmful to human health. Write a short summary on it and discuss why it took so long to reach the conclusion.

2. For the two-class problem, FLD can be obtained as a special case of MSE. Assume that class C_1 has n_1 samples $\boldsymbol{x}_i, i = 1, \cdots, n_1$ and class C_2 has n_2 samples $\boldsymbol{x}_i, i = 1, \cdots, n_2$. In MSE, we take targets for samples class C_1 and C_2 to be $b_1 = \frac{n}{n_1}$ and $b_2 = -\frac{n}{n_2}$, respectively. n is the total number of samples i.e. $n = n_1 + n_2$. So the sum-of-squares error function can be written as:

$$\mathrm{E} = \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{w}^T\boldsymbol{x}_i + w_0 - t_i)^2, \qquad t_i = \begin{cases} \frac{n}{n_1}, & \text{if } \boldsymbol{x}_i \in C_1 \\ -\frac{n}{n_2}, & \text{if } \boldsymbol{x}_i \in C_2 \end{cases}$$

Our goal is to get the MSE solution of $\boldsymbol{w}$ and w_0. Tasks:

(1) Prove that the optimal $w_0 = -\boldsymbol{w}^T\boldsymbol{m}$, where $\boldsymbol{m}$ is the mean vector of all samples

(2) Derive that the optimal $\boldsymbol{w}$ satisfy the following formula:

$$\left(\mathbf{S}_\mathbf{w} + \frac{n_1 n_2}{n}\boldsymbol{S_B}\right)\boldsymbol{w} = n(\boldsymbol{m}_1 - \boldsymbol{m}_2)$$

where $\mathbf{S}_\mathbf{w}$ and $\boldsymbol{S_B}$ are the within-class scatter matrix and between-class scatter matrix, and $\boldsymbol{m}_1$, $\boldsymbol{m}_2$ are the means of the two classes, respectively.

(3) Following the formula in (2), derive that $\boldsymbol{w} \propto \mathbf{S}_\boldsymbol{w}^{-1}(\boldsymbol{m}_1 - \boldsymbol{m}_2)$.

3. Find or work out a proof of the Perceptron convergence theorem. Show that the algorithm will not converge when the samples are not linearly separable.

4. Study the basic formulas of the logistic/sigmoid function.

(1) For

$$\theta(s) = \frac{1}{1 + e^{-s}}$$

show that

$$\theta(s) = \frac{1}{1 + e^{-s}} = \frac{e^s}{e^s + 1},$$
$$\theta(-s) = 1 - \theta(s),$$
$$\theta'(s) = \theta(s)\big(1 - \theta(s)\big).$$

(2) Study the relationship of the hyperbolic tangent function

$$f(\mathrm{s}) = \tanh s$$

to the $\theta(s)$, and derive the derivative of $f(\mathrm{s}) = \tanh s$.

Due date: Sept. 29 (Wed.) 23:00 Beijing time