1

Assume that we have some data $x_1, \dots x_n \in \mathbb{R}$. Our goal is to find a constant b such that $\sum_i (y_i - b)^2$ is minimized.

1.1

Find an analytic solution for the optimal value of b.

Consider the function

$$\begin{aligned} f : \mathbb{R}^n \times \mathbb{R} &\to [0, \infty) \\ (y, b) &\mapsto \sum_{i=1}^{n} (y_i - b)^2. \end{aligned}$$

For fixed $y \in \mathbb{R}^n$ we want to solve the following minimization problem, find b_y^* such that

$$\min_{b \in \mathbb{R}} f(y, b) = f(y, b_y^*)$$

$$\sum_{i=1}^{n} (y_i - b)^2 = \sum_{i=1}^{n} (y_i^2 - 2y_i b + b^2) = \sum_{i=1}^{n} y_i^2 \underbrace{-2b\underline{y} + b^2}_{=:g(b)}$$

As such we are looking for the minimum of $g(b)$. Obviously g is continuously differentiable and we can calculate

$$g'(b) = -2\sum_{i=1}^{n} y_i + 2b = 2(b - \sum_{i=1}^{n} y_i) = 0$$

meaning that $b^* = \sum_{i=1}^{n} y_i$ is the (unique!) solution to our minimization problem.

1.2

How does this problem and its solution relate to the normal distribution?

1.3

What if we change the loss from $\sum_i (y_i - b)^2$ to $\sum_i |y_i - b|$? Can you find the optimal solution for b?

Consider $b' := \overline{y}$, our new "linear" error function can be written as follows:

$$\begin{aligned} \sum_{i=1}^{n} |y_i - b| &= \sum_{i=1, y_i < \overline{y}}^{n} |y_i - \overline{y}| + \sum_{i=1, y_i > \overline{y}}^{n} |y_i - \overline{y}| \\ &= \sum_{i=1, y_i < \overline{y}}^{n} \overline{y} - y_i + \sum_{i=1, y_i > \overline{y}}^{n} y_i - \overline{y} \end{aligned}$$

2

Prove that the affine functions that can be expressed by $y = Wx + b$ are equivalent to linear functions on $(x, 1)$.

Consider $C := (w, 1)$ then $C^t(x, b) = (w^t|1)$ and

$$\begin{pmatrix} w & 1 \end{pmatrix} \begin{pmatrix} x \\ b \end{pmatrix} = wx + b$$

3

4

4.1

In that case there exists $v \in \mathbb{R}^n \backslash \{0\}$ such that $X^t X v = 0$.

4.2

The set of invertible matrices is dense in the set of all matrices so we'd almost certainly get an invertible matrix.

5