Machine Learning 1 - Week 3

Linh Nguyen

September 2022

1 Problems

1. Re-transform linear regression from

$$t = y(x, w) + noise- > w = (X^T X)^{-1} X^T t$$

2. Using numpy, find linear regression model for house's price prediction problem, dataset
   (a) Plot the prediction model (line) and data (point - scatter).
   (b) Predict the price of houses with an area of 50, 100, 150.
3. Using numpy, find linear regression model for house's price prediction problem, dataset
4. Prove $X^T X$ is invertible when X is full rank.

2 Solutions

2.1 Re-transform Linear Regression

We have a data set of observations $x = (x_1, x_2, ..., x_N)^T$, representing N observations of the scalar variable x and their corresponding target values $t = (t_1, t_2, ..., t_N)^T$.

$$t = y(x, w) + noise$$

One method to find the parameters w is making y(x, w) close to y. We define the cost function:

$$L = \frac{1}{2} \sum_{n=1}^{n} (y(x_n, w) - t_n)^2$$

Using the fact that for a vector z, we have that $z^t z = \sum_i z_i^2$:

$$L = \frac{1}{2} \sum_{n=1}^{n} (y(x_n, w) - t_n)^2 = \frac{1}{2} (y(x, w) - t)^T (y(x, w) - t)$$

To minimize L, let's find its derivatives with respect to w:

$$\begin{aligned}\frac{\partial L}{\partial w} &= \frac{\partial L}{\partial y}\frac{\partial y}{\partial w} \\ &= (y(x,w))^T X \\ &= (Xw - t)^T X \\ &= X^T(Xw - t) \\ &= X^T Xw - X^T t\end{aligned}$$

We set the derivatives to zero and obtain:

$$X^T Xw = X^T t$$

Thus, the value of w that minimizes L is:

$$w = (X^T X)^{-1} X^T t$$

2.2 Prove $X^T X$ is invertible when X is full rank

Let X is an $m \times n$ matrix, v $\in \mathbb{R}^n$.

X is full rank $=> rank(X) = min(m, n)$

If $Xv = 0$ then $X^T Xv = 0$

If $X^T Xv = 0$ then $v^T X^T Xv = 0$, that is $(Xv)^T Xv = 0$, which implies that $Xv = 0$

$\longrightarrow null(X) = null(X^T X)$

Case 1: $m < n$

When $m < n$, $rank(X) = m$. According to the rank-nullity theorem:

$$\begin{aligned}&rank(X) + null(X) = n \\ &null(X) = n - rank(X) \\ &null(X) = n - m \neq 0\end{aligned}$$

We have proved that $null(X) = null(X^T X) \Rightarrow null(X^T X) \neq 0$
$\longrightarrow X^T X$ can not be invertible when X is full row rank

Case 2: $m \geq n$

When $m \geq n$, $rank(A) = n$. According to the rank-nullity theorem:

$$\begin{aligned}&rank(X) + null(X) = n \\ &null(X) = n - rank(X) \\ &null(X) = n - n = 0\end{aligned}$$

We have proved that $null(X) = null(X^T X) \Rightarrow null(X^T X) = 0$
$\longrightarrow X^T X$ is invertible when X is full column rank