Derivation

找到这里了啊...

最小二乘法

首先定义残差 ϵ，以及前文定义的方程组 $b = Ax + \epsilon$

$$\varepsilon = b - \hat{b}$$

根据 $b = Ax + \varepsilon$ 的变形，我们可以得到 $\varepsilon = b - Ax$，所以就是 $\varepsilon = b - \hat{b} = b - Ax$。
现在我们需要得到残差平方和公式，一般情况我们默认把 ε 视为列向量，所以我们需要将向量 ε 先转置为行向量，也就是这样得出残差平方和 S 的等式。在转置后，$\varepsilon^\top$ 是 $1 \times n$ 的行向量，而没转置的 ε 是 $n \times 1$ 的列向量。

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix} \quad \varepsilon^\top = [\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n]$$

$$S = \varepsilon^\top \varepsilon$$

将 $\varepsilon = b - Ax$ 代入 S，我们得到如下式子，并对其展开：

$$\begin{aligned} S &= (b - Ax)^\top (b - Ax) \\ &= (b^\top - (x^\top A^\top))(b - Ax) \\ &= b^\top b - 2x^\top A^\top b + x^\top A^\top Ax \end{aligned}$$

我们要对上公式中的 x 求导，将导数置为0，记为 $\frac{dS}{dx} = 0$，由于第一项 $b^\top b$ 是常数向量，所以求导为0，我们有

$$\begin{aligned} \frac{dS}{dx} &= b^\top b - 2x^\top A^\top b + x^\top A^\top Ax \\ &= 0 - 2x^\top A^\top b + x^\top A^\top Ax \end{aligned}$$

进行移项和推导，得到

$$\begin{aligned} -2A^\top b + 2A^\top Ax &= 0 \\ 2A^\top Ax &= 2A^\top b \\ A^\top Ax &= A^\top b \\ x &= (A^\top A)^{-1} A^\top b \end{aligned}$$