Today we've learned linear algebra formulas to solve least square regression. Partial differentiation of least square formula will give you the solution formula. Check the transformation yourself and comment on it.

$(x_1, y_1), \cdots, (x_n, y_n)$ というデータが得られた時に、
Xを説明変数とし、係数aと切片bを用いて $Y \sim aX + b$ という線形回帰モデルで説明したい。

最小二乗法で

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

を最小にするようなa,bの値を求めたい。

説明変数が複数ある場合、

$$X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} & 1 \\ x_{12} & x_{22} & \cdots & x_{k2} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{1n} & x_{2n} & \cdots & x_{kn} & 1 \end{pmatrix}$$

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \\ b \end{pmatrix}$$

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

とおくと、

$$\vec{y} \sim X\vec{a}$$

であり、

$$\|\vec{y} - X\vec{\mathbf{a}}\|^2$$

を最小にするようなベクトルaを求める。

この式を整理すると、

$$\begin{aligned}\|y - Xa\|^2 &= (Xa - y)^\top (Xa - y) \\ &= (a^\top X^\top - y^\top)(Xa - y) \\ &= a^\top X^\top Xa - 2a^\top X^\top y + y^\top y\end{aligned}$$

これをaで微分（各要素で偏微分）してゼロになるところが最小値なので、

$$2A^\top Xa - 2X^\top y = 0$$

$$X^\top Xa = X^\top y$$

$X^\top X$が正則であれば

$$a = (X^\top X)^{-1} X^\top y$$

が解となり、このベクトルaが最小値を与える。