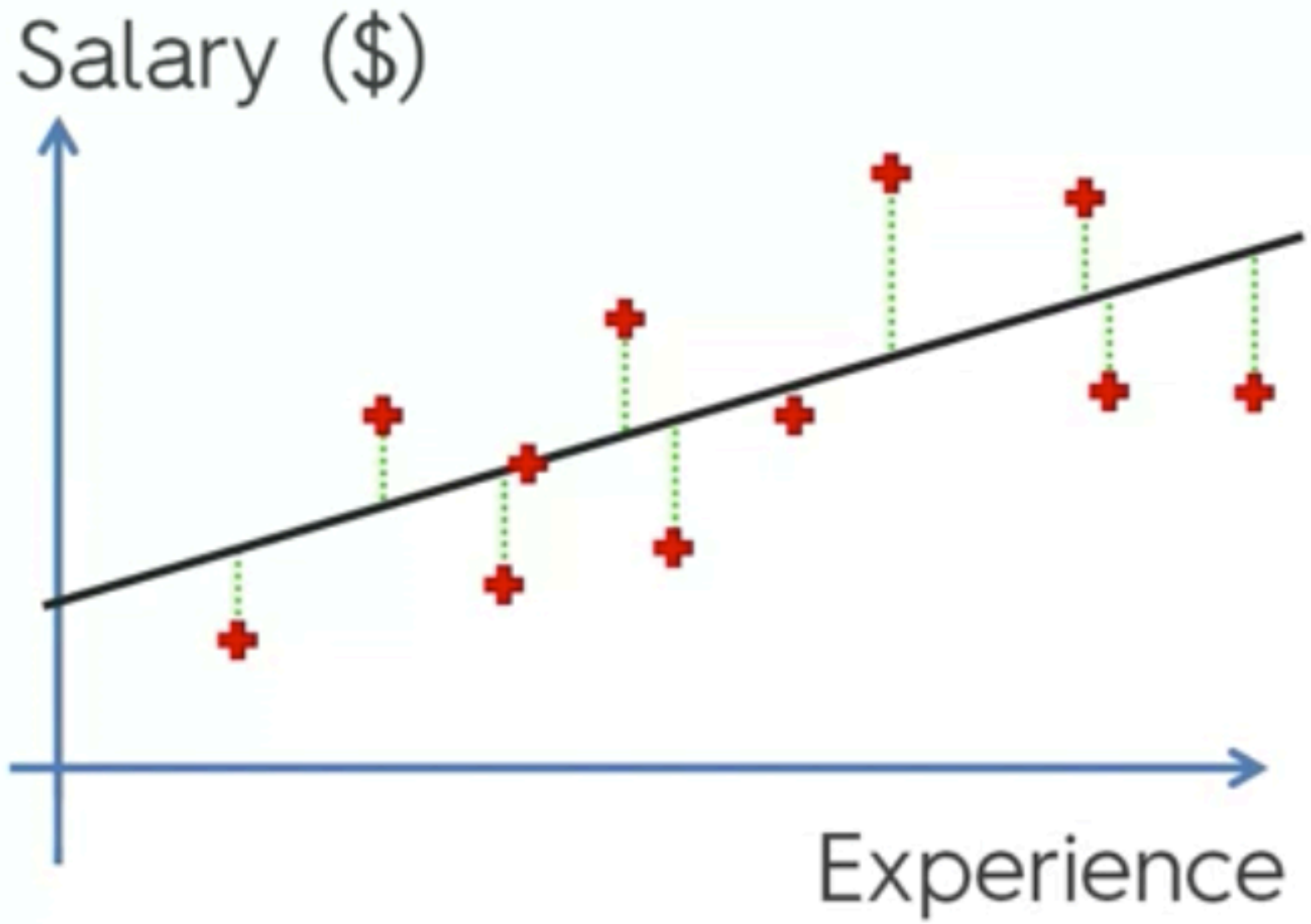
R Squared

Simple Linear Regression:



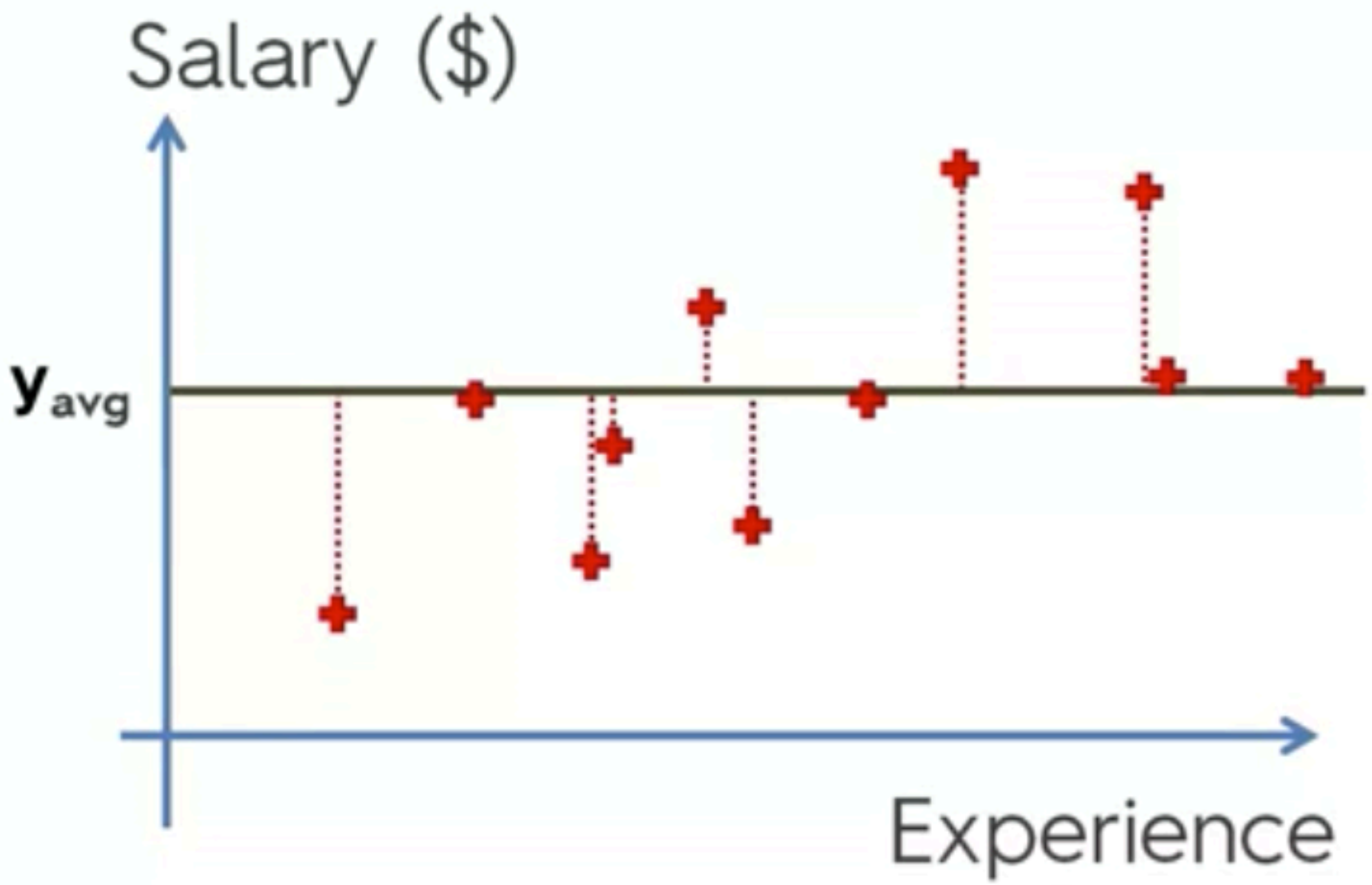
Sum of Squares of Residuals

$$SS_{res} = SUM\ (y_i - \hat{y}_i)^2$$

Predicted Value

R Squared

Simple Linear Regression:



$$SS_{res} = SUM\ (y_i - \hat{y}_i)^2$$

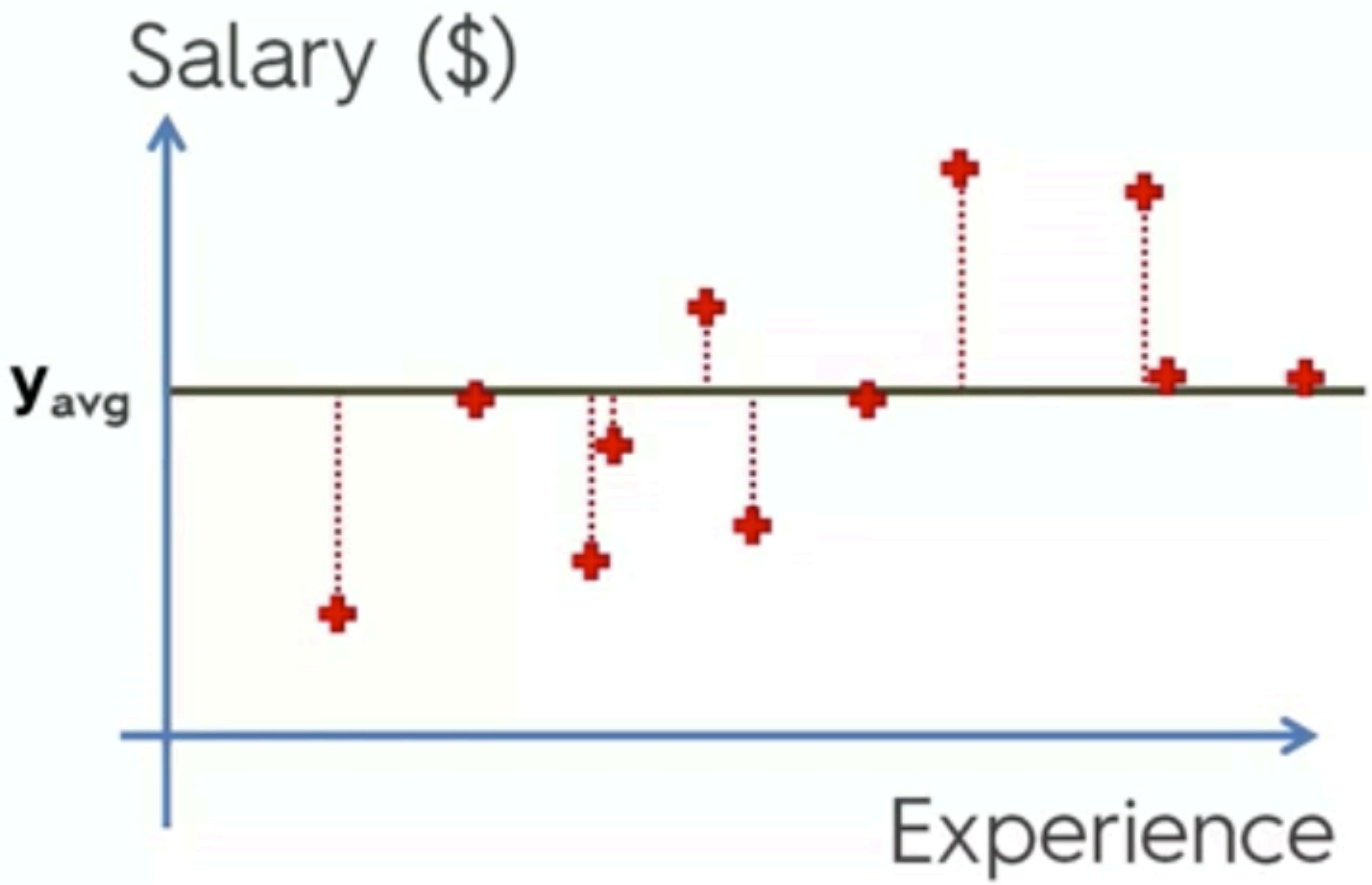
$$SS_{tot} = SUM\ (y_i - y_{avg})^2$$

Total Sum of Squares

Sample Mean Value

R Squared

Simple Linear Regression:



$$SS_{res} = SUM\ (y_i - \hat{y_i})^2$$

$$SS_{tot} = SUM\ (y_i - y_{avg})^2$$

R2 means how good is the sloped line as compared to the horizontal line at the average.

$$\boxed{R^2 = 1 - \frac{SS_{res}}{SS_{tot}}}$$

While fitting, intention is to minimise SSres
Usually, 0 <= R2 <= 1, R2 tending to 1 is better
R2 can also be negative, we then go with avg.

Adjusted R^2

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

R^2 – Goodness of fit
(greater is better)

$y = b_0 + b_1 * x_1$

$y = b_0 + b_1 * x_1 + b_2 * x_2$

Problem:

$\boxed{+ b_3 * x_3}$

SS_{res} -> Min

R^2 will never decrease

Adjusted R^2

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

$$Adj\ R^2 = 1 - (1 - R^2)\frac{n - 1}{n - p - 1}$$

p - number of regressors

n - sample size