1. **Linear Regression**

A linear model makes a prediction by computing the weighted sum of the input features plus a bias term/intercept term

$$\hat{y} = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$

$\hat{y}$: predicted value

x_n: n^{th} feature value

n = number of features

$$\hat{y} = h_\theta(\bar{x}) = \bar{\theta} \cdot \bar{x}$$

← vectorized form

Mean Squared Error (MSE)

Most well suited measure of linear regression model performance is RMSE. Therefore we need to find a theta that minimizes the RMSE. Alternatively, in practice MSE is used.

$$MSE(\bar{x}, h_\theta) = \frac{1}{m} \sum^{m} \left(\bar{\theta}^T \bar{x}^{(i)} - y^{(i)} \right)^2$$

$$RMSE = \sqrt{MSE}$$

The Normal Equation

There is a closed-form solution (a mathematical way) to compute theta without having to train the model

$$\hat{\theta} = (X^T X)^{-1} X^T y$$

Questions for interview