

LINEAR REGRESSION

FINDING PARAMETERS

$$\underset{\substack{\uparrow \\ \text{Parameter} \\ \text{vector}}}{W} = \left(\underset{\substack{\uparrow \\ \text{Feature matrix}}}{X}^T \underset{\substack{\uparrow \\ \text{Feature matrix}}}{X} \right)^{-1} \underset{\substack{\uparrow \\ \text{Transposed feature matrix}}}{X}^T \underset{\substack{\uparrow \\ \text{Target vector}}}{y}$$

Analytical solutions are most common. However, the cost of computing pseudo-inverse of $X^T X$ means that gradient descent can be better in large data.

ChrisAlbon