

Training Data

We have N input, output pairs $\mathcal{D} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$ where $\mathbf{x}_n \in \mathbb{R}^D$ where D is the dimension and $y_n \in \mathbb{R}$.

Assumptions

Underlying function f is linear, so that

$$f(\mathbf{x}_i) = \theta^T \mathbf{x}_i$$

Observation y is a noisy version of f .

$$f(\mathbf{x}_i) \approx y$$

Linear Regression

Linear regression means linear in the parameters, not in the input data.