Linear Regression: The Objective Function

- Parameter $\mathbf{w}$ that satisfies $y_i = \mathbf{w}^T\mathbf{x}_i$ *exactly* for each i may not exist
- So we look for the closest approximation
- Specifically, $\mathbf{w}$ that minimizes the following sum-of-squared-differences between the truth (y_i) and the predictions ($\mathbf{w}^T\mathbf{x}_i$), just as we did for the one-dimensional case:

$$E(\mathbf{w}) = \frac{1}{2}\sum_{i=1}^{N}(y_i - \mathbf{w}^T\mathbf{x}_i)^2$$