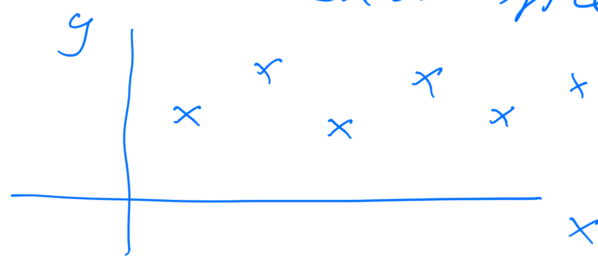
Machine Learning Basics:

- Data sets (Regression example)

$y(x)$



- Make a model

straight line

unknown parameters

$$y(x) \simeq a_0 + a_1 x = \tilde{y}(x)$$

- Cost/Loss function

Measure how well the model is doing

absolute error $= |y(x) - \tilde{y}(x)|$

$y_i = y(x_i)$ $\quad \mathbf{x} \in [x_0, x_1 \ldots x_{n-1}]$

$\mathbf{y} \in [y_0, y_1 \ldots y_{n-1}]$

$i = 0, 1, \ldots n-1$

Mean-square error

$$MSE(\mathbf{a}) = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \tilde{y}_i)^2$$

$$\mathbf{a} = [a_0, a_1]$$

$$\boxed{\frac{\partial MSE}{\partial a}} = 0$$

in linear regression this leads to a "plain" matrix inversion problem to find the unknown parameters a.

The math of Linear Reg

$$y = [y_0 \; y_1 \; \dots \; y_{n-1}]$$

Polynomial fit:

$$\boxed{y_i \simeq \tilde{y}_i = f(x_i) = \sum_{j=0}^{n-1} \beta_j x_i^j}$$

$$\tilde{y}_0 = \beta_0 + \beta_1 x_0 + \beta_2 x_0^2 + \dots \beta_{n-1} x_0^{n-1}$$

$$\tilde{y}_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + \dots \beta_{n-1} x_1^{n}$$

$$\vdots$$

$y \simeq f(x) = \beta_0 + \beta_1 x$ - for all x values.

$$\vdots$$

$$\tilde{y}_{n-1} = \beta_0 + \beta_1 x_{n-1} + \beta_2 x_{n-1}^2 + \dots$$

$$\tilde{y} = \begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \end{bmatrix}$$

$$y = \begin{bmatrix} \vdots \\ y_{n-1} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & x_0^1 & \cdots & x_0^{n-1} \\ 1 & x_1^1 & & \vdots \\ \vdots & \vdots & & \vdots \\ 1 & x_{n-1}^1 & & x_{n-1}^{n-1} \end{bmatrix}$$

$$X \in \mathbb{R}^{n \times n}$$

parameters $\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \end{bmatrix}$

$\beta \in \mathbb{R}^n \implies$

$$\tilde{y} = \underline{X \cdot \beta}$$

Design/feature matrix

β defines the features of the model.

$\beta \in \mathbb{R}^n \rightarrow \beta \in \mathbb{R}^p$

p = number of features/predictors