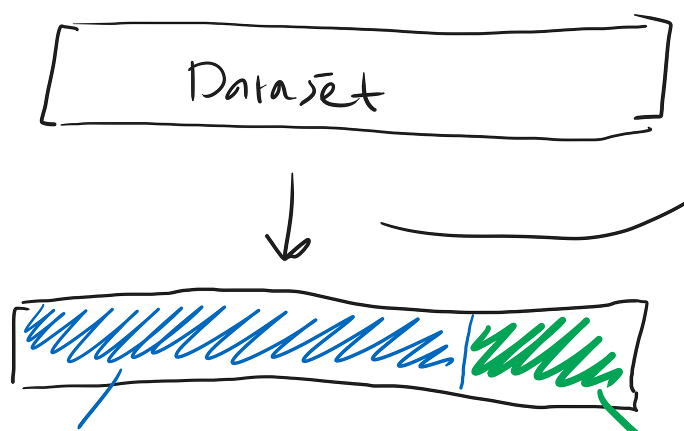
Train test split



Random split data

80 % / 20 %

This ratio is flexible

train data
train set

test data
test set

Unseen data used to estimate actual usefulness

Multiple linear regression

features n st — target

antal train samples $= m$

Antal rum	Boyta (m²)	Byggår	...	Trädgård	Pris
3	82	2004		Ja	4.5Mkr
4	94	1992		Nej	3.2Mkr
3	75	2009		Nej	4.4Mkr

$= (x_1^1, x_2^1, x_3^1, \dots, x_n^1, y^1)$

$\vdots$

$= (x_1^m, x_2^m, x_3^m, \dots, x_n^m, y^m)$

$x_1 \quad x_2 \quad x_3 \quad x_n \quad y$

Multiple linear regression $\rightarrow f_{w,b}(x) = w_1 \cdot x_1 + w_2 x_2 + w_3 \cdot x_3 + \dots + w_n x_n + b$

Simple linear regression $\rightarrow f_{w,b}(x) = w_1 x_1 + b$

Vår uppgift: hitta bra värden på $w_1, w_2, \dots, w_n$ och b

$\bar{x}^{(i)} = [1, x_1^i, x_2^i, x_3^i, \dots, x_n^i]$

$\bar{w} = [w_0, w_1, w_2, w_3, \dots, w_n]$

$\bar{a} \cdot \bar{b} = \sum a_i b_i$

$\bar{x}^{(i)} \cdot \bar{w} = x_1^i w_1 + x_2^i w_2 + x_3^i w_3 + x_n^i w_n + \underbrace{1 \cdot w_0}_{b} = \hat{y}^i \rightarrow$ modellens prediction för training sample i

Låt oss nu skriva om vårt dataset i matrisformat

$X: m \times (n+1) \qquad X^T: (n+1) \times m$

$$\mathbf{X} = \begin{array}{c} n+1 \\ \begin{bmatrix} 1 & x_1^1 & x_2^1 & \dots & x_n^1 \\ 1 & x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^m & x_2^m & \dots & x_n^m \end{bmatrix} \end{array} \Bigg\} m \qquad \mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix}$$

$$Xw = \begin{bmatrix} 1 & x_1^1 & x_2^1 & \dots & x_n^1 \\ 1 & x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^m & x_2^m & \dots & x_n^m \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} \hat{y}^1 \\ \hat{y}^2 \\ \vdots \\ \hat{y}^m \end{bmatrix} \approx \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{bmatrix} = \bar{y}$$

vi vill så gärna

Så, ok! Hur hittar vi optimala värden på våra vikter $w_0, w_1, \dots, w_n$?

Jo, det visar sig att det går att hitta de vikter som ger lägst MSE genom att utföra följande beräkning

Detta kallas för normal ekvationen

$$\underset{(n+1)\times 1}{\text{vektor}} \leftarrow \mathbf{w} = \underbrace{(\underset{(n+1)\times m}{\mathbf{X}^T} \underset{m \times (n+1)}{\mathbf{X}})}_{(n+1)\times(n+1)}{}^{-1} \mathbf{X}^T \mathbf{y}$$