Train / test split

Dataset

Random split

70% / 30%

This ratio is flexible

train set — we train on this

test set — not trained on. Unseen data which we use to evaluate performance

---

Multiple linear regression

features n:st — target

antal 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

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

$\rightarrow (x_1^1, x_2^1, x_3^1, \ldots, x_n^1, y^1)$

$\vdots$

$\rightarrow (x_1^m, x_2^m, x_3^m, \ldots, x_n^m, y^m)$

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

multiple linear regression $\rightarrow f_{\bar{w}}(x) = w_n \cdot x_n + w_{n-1} \cdot x_{n-1} + \ldots + w_2 \cdot x_2 + w_1 \cdot x_1 + \underbrace{w_0}_{=b}$

Vår uppgift blir nu att hitta bra värden på $\bar{w} = (w_n, w_{n-1}, \ldots, w_1, w_0)$

Anta att vi, på något sätt, har hittat bra värden på **w**.

Då kan jag ta en sample, vilken som helst

$\bar{x}^i = (x_n^i, x_{n-1}^i, \ldots, x_1^i, 1)$

$\bar{w} = (w_n, w_{n-1}, \ldots, w_1, w_0)$

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

$\bar{x}^i \bar{w} = x_n^i w_n + x_{n-1}^i w_{n-1} + \ldots + x_1^i w_1 + w_0 = \hat{y}^i$ Detta är nu vår prediction för den i:te samplen

Låt oss nu skriva om hela vår dataset i matris format

$X = m \times (n+1)$, $\bar{w} = (n+1) \times 1$

$$\mathbf{X} = \overbrace{\begin{bmatrix} 1 & x_1^1 & x_2^1 & \ldots & x_n^1 \\ 1 & x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^m & x_2^m & \ldots & x_n^m \end{bmatrix}}^{n+1} \Big\} m \qquad \mathbf{w} = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_n \end{bmatrix}$$

hoppas!

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

OK! MEN HUR HITTAR JAG VÄRDEN på $w_n, w_{n-1}, \ldots, w_1, w_0$??

Jo, just i detta fall visar det vis att vi kan hitta optimala värden som minimerar MSE med följande beräkning

Detta kallas för normalekvationen

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

vektor $(n+1) \times 1$