CS 129.18

Linear Regression

Linear Regression

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon$$

A function $\mathbf{y} = X\beta + \varepsilon$

$X \rightarrow$ data

$\epsilon \rightarrow$ error in data

$\beta \rightarrow$ coefficients

$y \rightarrow$ target

$$\mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ \downarrow \\ y_n \end{bmatrix}^T$$

$\mathbf{y}$ is the target variable vector

$$X = \begin{bmatrix} x_0 \\ x_1 \\ \downarrow \\ x_n \end{bmatrix}^T$$

$\mathbf{X}$ is a vector of features you want to predict with

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \downarrow \\ \beta_n \end{bmatrix}^T$$

β is the coefficient vector

$$\varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \downarrow \\ \varepsilon_n \end{bmatrix}^T$$

ε is the error term or noise vector

Linear regression is basically $y = mx + b$

Let’s make it more interesting

Multivariate Linear Regression

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon$$

Linear method for modelling the relationship between a dependent or target variable, and one or more or independent variables.

A function $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \varepsilon$

X $\longrightarrow$ data

$\epsilon \longrightarrow$ error in data

β $\longrightarrow$ coefficients

y $\longrightarrow$ target

$$\mathbf{y} = \begin{bmatrix} y_0 \\ y_1 \\ \downarrow \\ y_n \end{bmatrix}^T$$

$\mathbf{y}$ is the target variable vector

$$\mathbf{X} = \begin{bmatrix} x_{0,0} & x_{0,1} & x_{0,2} & \cdots\rightarrow & x_{0,n} \\ x_{1,0} & x_{1,1} & x_{1,2} & \cdots\rightarrow & x_{1,n} \\ & \downarrow & & & \downarrow \\ x_{n,0} & x_{n,1} & x_{n,2} & \cdots\rightarrow & x_{n,n} \end{bmatrix}$$

$\mathbf{X}$ are the features or values you want to predict against

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \downarrow \\ \beta_n \end{bmatrix}^T$$

β is the coefficient vector

$$\varepsilon = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \downarrow \\ \varepsilon_n \end{bmatrix}^T$$

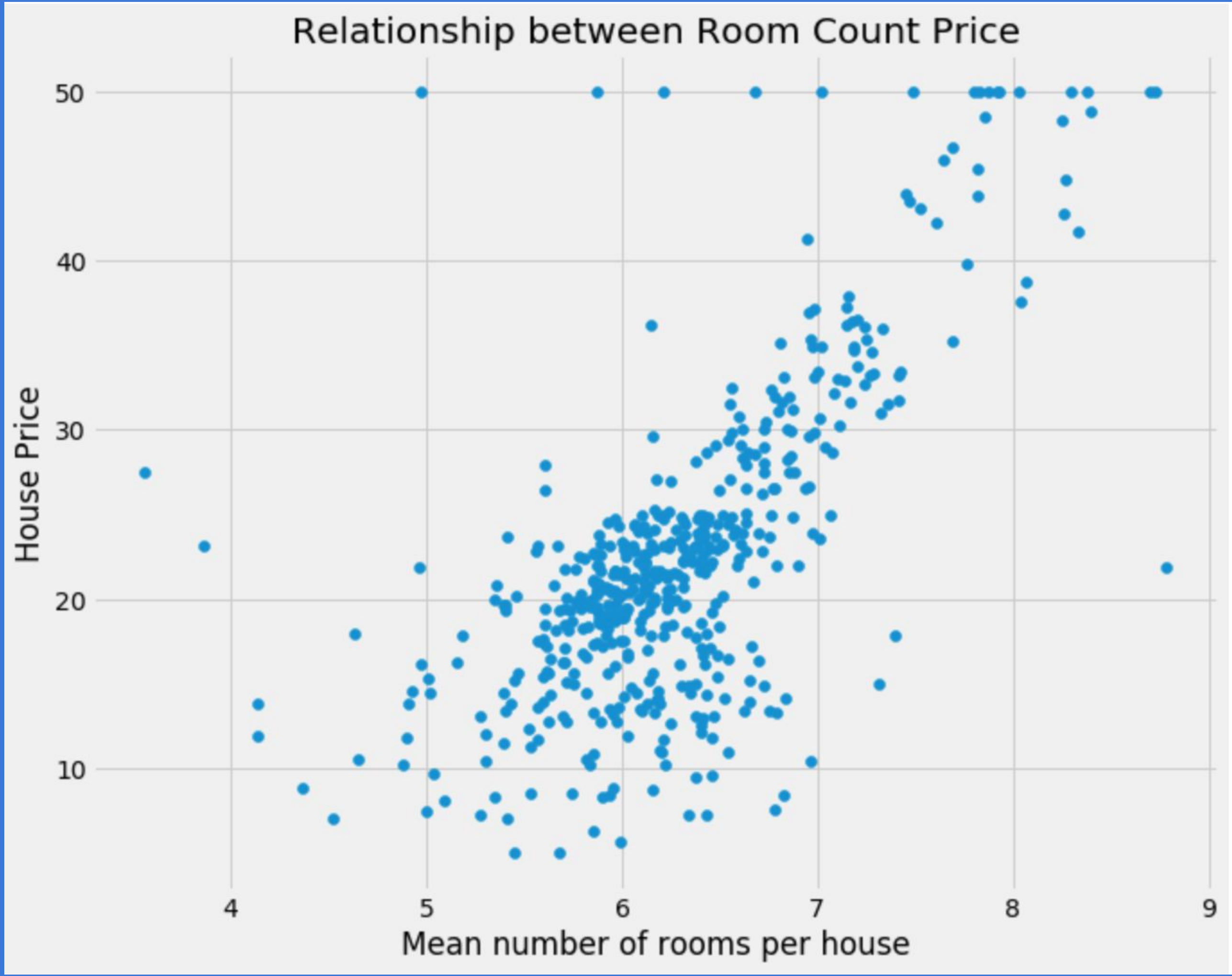
ε is the error term or noise vector

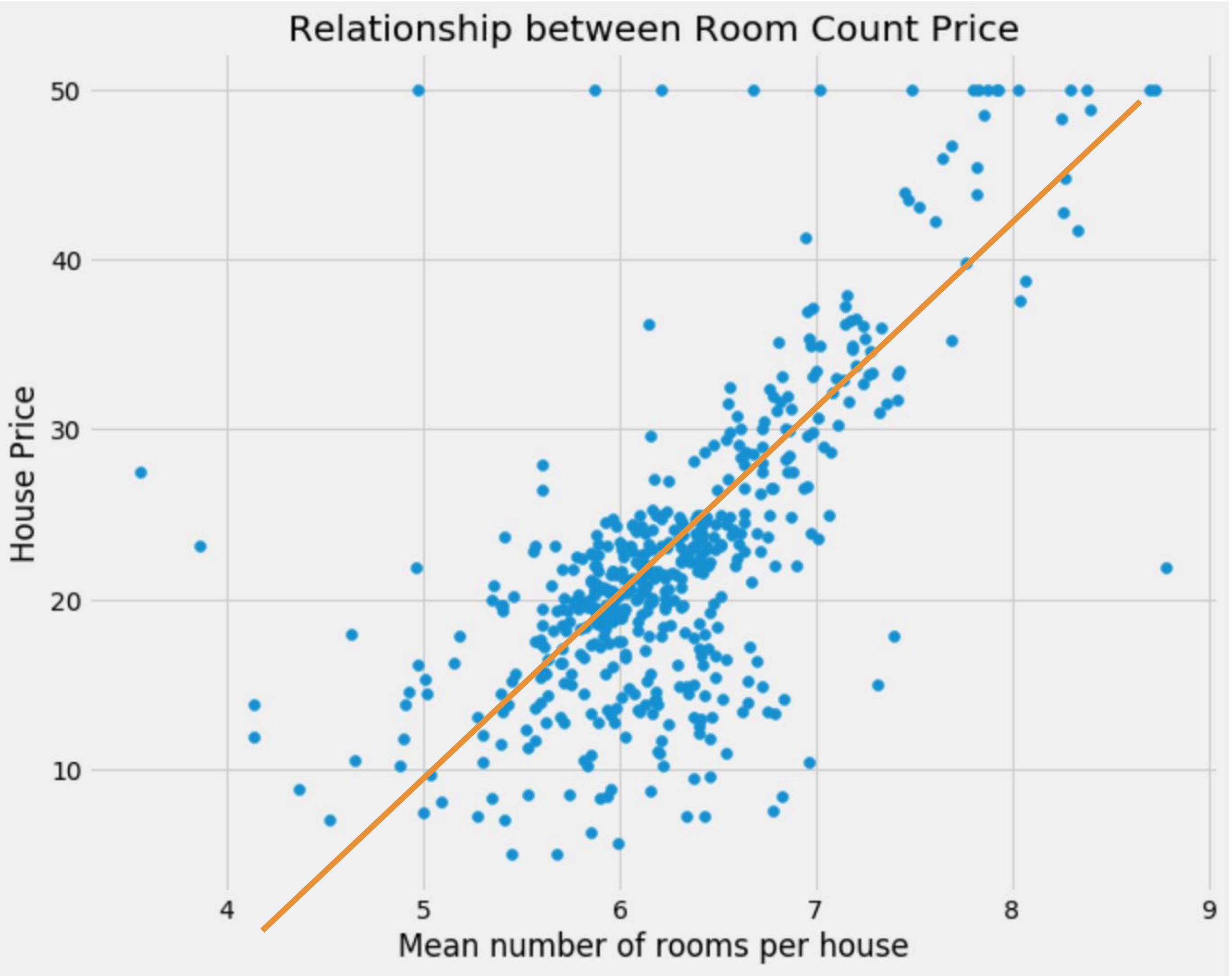
Mean Squared Error

The MSE is a measure of the quality of an estimator—it is always non-negative, and values closer to zero are better.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

But what does that all mean?


Relationship between Room Count Price
House Price
Mean number of rooms per house
10
20
30
40
50
4
5
6
7
8
9


Relationship between Room Count Price
50
40
30
20
10
House Price
4
5
6
7
8
9
Mean number of rooms per house

You get to predict new values based off existing data

Thank you