

# Simple Linear Regression

Fall 2013

Carola Wenk

# Simple Linear Regression

- Independent variable:  $X=(X[1],\dots,X[n])$   
dependent variable (to be predicted):  $Y=(Y[1],\dots,Y[n])$

- Find a line:

$$Y' = a + bX$$

which minimizes the sum of the squared distances:

$$\sum_{i=1}^n (Y[i] - (a + bX[i]))^2$$

- The solution line  $Y' = a + bX$  can be computed as:

- $$b = \frac{\sum_{i=1}^n X[i]Y[i] - \frac{1}{n} \sum_{i=1}^n X[i] \sum_{i=1}^n Y[i]}{\sum_{i=1}^n X[i]^2 - \frac{1}{n} (\sum_{i=1}^n X[i])^2}$$

- $$a = \frac{1}{n} \sum_{i=1}^n Y[i] - b \frac{1}{n} \sum_{i=1}^n X[i]$$