

Simple Case: $\{(x_i, y_i)\}_{i=1}^N$ $y_i \approx m x_i + b$


Multivariate Case: features of our data (K such)
 samples: each sample consists of (N samples)
 a measurement of the K features
 target variable that we are trying to predict

$mpg \approx$ displacement of engine, weight of car, acceleration of car.

$$y \approx m_1 \cdot \text{disp} + m_2 \cdot \text{weight} + m_3 \cdot \text{accel} + b$$

"TIDY DATA CONVENTION"

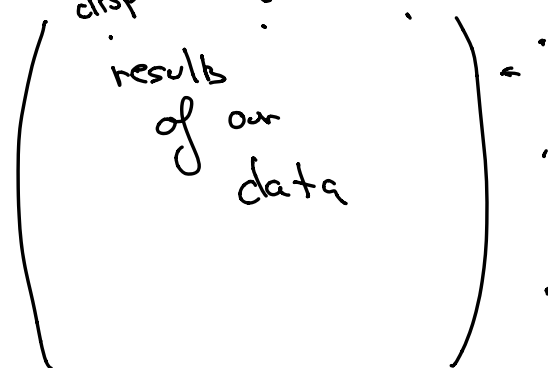
N rows
 each is a sample



K columns
 each is a feature

Car model₁
 Car model₂
 ⋮
 Car model_N

disp weight accel
 results of our data



X

N entries

Y

y₁ mpg

y_N

Y



$$y_i \approx m_1 x_{i1} + m_2 x_{i2} + m_3 x_{i3} + b$$

$$\begin{array}{c} \text{observed} \rightarrow Y \approx X M + b \\ \text{"target" values} \\ N \times 1 \text{ vector} \end{array} \quad \begin{array}{c} \text{tidy} \\ \text{data} \\ \text{matrix} \\ N \times K \end{array} \quad \begin{array}{c} \uparrow \\ K \times 1 \\ \text{entries "} \\ \text{are the slopes"} \end{array} \quad \begin{array}{c} = b \\ \uparrow \\ N \times 1 \\ \text{matrix} \end{array} \quad \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

add an extra column to our data matrix

$$X = \begin{pmatrix} x_{i1} & x_{i2} & \dots & x_{iK} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ x_{N1} & x_{N2} & \dots & x_{NK} & 1 \end{pmatrix} \quad \begin{array}{c} \uparrow \\ N \times (K+1) \\ \text{\# of} \\ \text{samples} \end{array} \quad \begin{array}{c} \uparrow \\ \text{\# features} \end{array}$$

$$Y \approx X M$$

M is $K+1 \times 1$ column vector

$$m_1 x_{i1} + m_2 x_{i2} + \dots + m_K x_{iK} + m_{K+1}$$

$$\begin{pmatrix} x_{i1} & x_{i2} & \dots & x_{iK} \end{pmatrix} \begin{pmatrix} m_1 \\ \vdots \\ m_K \\ m_{K+1} \end{pmatrix} =$$

$$\|Y - X M\|^2$$

$$X M = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1K} & 1 \\ x_{21} & x_{22} & \dots & x_{2K} & 1 \\ \vdots & \vdots & & \vdots & \vdots \end{pmatrix} \begin{pmatrix} m_1 \\ \vdots \\ m_K \\ m_{K+1} \end{pmatrix}$$

$$= \begin{pmatrix} x_{11} m_1 + x_{12} m_2 + \dots + x_{1K} m_K + m_{K+1} \\ x_{21} m_1 + x_{22} m_2 + \dots + x_{2K} m_K + m_{K+1} \\ \vdots \end{pmatrix}$$

$$\|Y - X M\|^2$$

$$= \sum_{i=1}^N (y_i - x_{i1} m_1 - x_{i2} m_2 - \dots - x_{iK} m_K - m_{K+1})^2$$

$= \text{MSE}(m_1, \dots, m_{K+1})$

multivariate optimization problem:

Find m_1, \dots, m_{x+1} so that

$\|Y - XM\|^2$ is minimized.