Linear Regression Example with 1 feature

Given this data:

Age	Blood Pressure
22	117
32	120
36	122
45	123
52	122
58	129
73	132

Find linear model to predict blood pressure based on the patient's age

Step 1: Set up the model

Let
- X be the independent variable (dataset matrix / vector) with extra column of 1s for bias terms
- y be the dependent variable (labels vector)
- w be the weight vector with extra element β for bias term

We trying to find w s.t.:

$$\| y - Xw \| = \underset{w}{\operatorname{argmin}} \| y - Xw \|$$

Step 2: Find w

Solve using normal equation:

$$X^T X \, w = X^T y$$

- Form $X^T X$:

$$X^T X = \begin{bmatrix} 22 & 32 & 36 & 45 & 52 & 58 & 73 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 22 & 1 \\ 32 & 1 \\ 36 & 1 \\ 45 & 1 \\ 52 & 1 \\ 58 & 1 \\ 73 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 16\,226 & 318 \\ 318 & 7 \end{bmatrix}$$

- Form $X^T y$:

$$X^T y = \begin{bmatrix} 39803 \\ 865 \end{bmatrix}$$

- Solve for w:

$$X^T X w = X^T y$$

$$\Leftrightarrow \begin{bmatrix} 16226 & 318 \\ 318 & 7 \end{bmatrix} \begin{pmatrix} w \\ \beta \end{pmatrix} = \begin{bmatrix} 39803 \\ 865 \end{bmatrix}$$

$$\Leftrightarrow \begin{cases} w = 0.285 \\ \beta = 110.623 \end{cases}$$

So the linear model is $y = 0.285x + 110.623$

Linear Regression with multiple features

Given this data:

Gender	Age	Blood Pressure
M	22	122
M	35	125
M	50	135
M	60	136
F	25	111
F	45	115
F	57	123
F	65	128

Step 1: Set up the model

Let: $X = (x_1 \mid x_2 \mid b)$ be the dataset matrix,

- x_1 is vector of gender
- x_2 is vector of age
- b is vector ones represent bias terms

y be label vector

$w = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \beta \end{pmatrix}$ be weight vector

We try to find w s.t.:

$$\| y - Xw \| = \underset{w}{\operatorname{argmin}} \| y - Xw \|$$

Step 2: Find w

Assume that our dataset matrix X is not invertible, so normal equation is not an option. We either solve using eigendecomposition or gradient descent. Here we try eigendecomposition:

$$\underbrace{A}_{X^TX} \; w = QQ^T \underbrace{b}_{X^Ty}$$

Form X^TX:

$$X^TX = \begin{bmatrix} 4 & 167 & 4 \\ 167 & 17933 & 359 \\ 4 & 359 & 8 \end{bmatrix} \quad \text{where} \quad X = \begin{array}{ccc} \text{Gender} & \text{Age} & \text{bias terms} \\ \downarrow & \downarrow & \downarrow \end{array} \begin{bmatrix} 1 & 22 & 1 \\ 1 & 30 & 1 \\ 0 & 56 & 1 \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Calculate eigenvalues:

$$\det(X^TX - \lambda I) = 0$$

$$\Leftrightarrow \quad \det \begin{bmatrix} 4-\lambda & 167 & 4 \\ 167 & 17933-\lambda & 359 \\ 4 & 359 & 8-\lambda \end{bmatrix} = 0$$

$$\Leftrightarrow \quad (4-\lambda)\left[(17933-\lambda)(8-\lambda) - 359^2\right] - 167\left[167(8-\lambda) - 359 \cdot 4\right]$$
$$+ 4\left[167 \cdot 359 - (17933-\lambda) \cdot 4\right] = 0$$

$$\Rightarrow \begin{cases} \lambda_1 = 17941.74 \\ \lambda_2 = 2.6757 \\ \lambda_3 = 0.58 \end{cases}$$

$\Rightarrow$ This capture almost all variance in the dataset. We could just use this

Calculate eigenvector associated with λ_1:

$$X^TX \; v_1 = \lambda_1 v_1$$

$$\Leftrightarrow \begin{bmatrix} 4 & 167 & 4 \\ 167 & 17933 & 359 \\ 4 & 359 & 8 \end{bmatrix} \cdot v_1 = 17941.74 \; v_1$$

$$\Rightarrow \quad v_1 = \begin{bmatrix} -0.01 \\ -0.1 \\ -0.02 \end{bmatrix}$$

Find projected X:

$$X_{proj} = X \cdot v_1 = \overset{1}{\begin{bmatrix} \vdots \end{bmatrix}} \text{ size of } X$$

Find w:

$$X_{proj}^T X_{proj} \, w = X_{proj}^T y$$

$\Rightarrow$ w = some scalar

On the other hand, if X is invertible, then just solve using normal equation:

$$X^T X \, w = X^T y$$

$$\begin{bmatrix} \text{gender} & \text{age} & \text{intercept} \\ \vdots & \vdots & \vdots \end{bmatrix} \quad \begin{pmatrix} \omega_1 \\ \omega_2 \\ \beta \end{pmatrix} \quad \begin{pmatrix} \text{blood pressure} \\ \vdots \end{pmatrix}$$

$$\Leftrightarrow \begin{bmatrix} 4 & 167 & 4 \\ 167 & 17\,933 & 359 \\ 4 & 359 & 8 \end{bmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \beta \end{pmatrix} = \begin{pmatrix} 518 \\ 45\,250 \\ 995 \end{pmatrix}$$

$$\Rightarrow \begin{cases} \omega_1 = 12.856 \\ \omega_2 = 0.417 \\ \beta = 99.236 \end{cases}$$