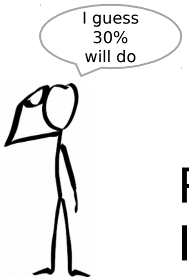
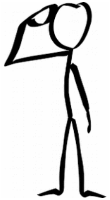
Gradient Descent in Linear Regression

Understanding Gradient Descent

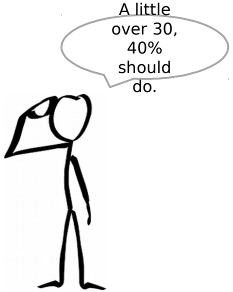


Random Initialisation

Generating Predictions



Calculating cost



Updating parameters

Gradient Descent in Linear Regression

1. Random initialisation

Gradient Descent in Linear Regression

$Y = ßX + b$

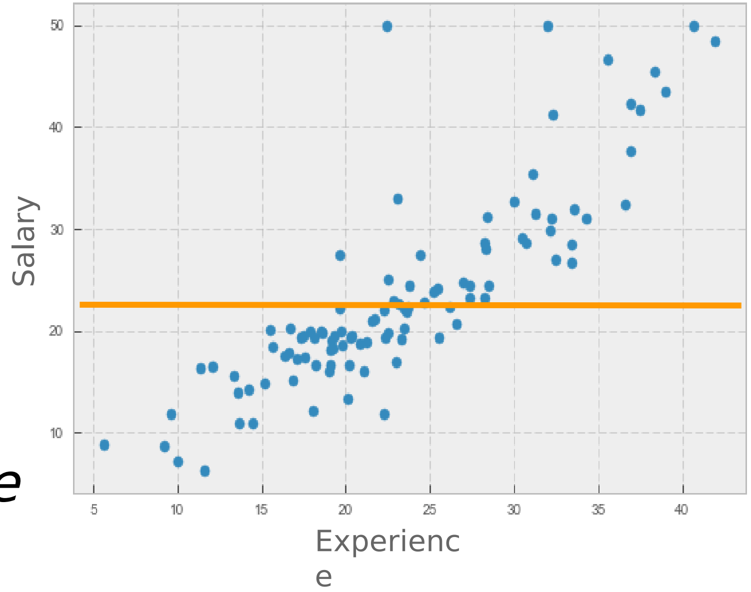
Parameters: *ß and b*

Gradient Descent in Linear Regression

$Y = ßX + b$

Parameters: *ß and b*

- *ß -> 0*
- *b -> Mean of independent Variable*



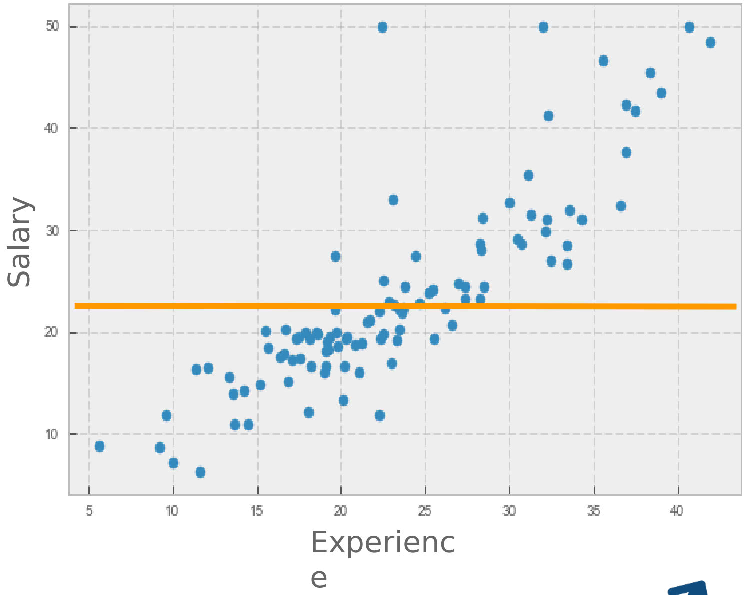
$\hat{Y} = ßX + b$

Gradient Descent in Linear Regression

1. Random initialisation

2. Generating predictions

Gradient Descent in Linear Regression



$\hat{Y} = ßX + b$

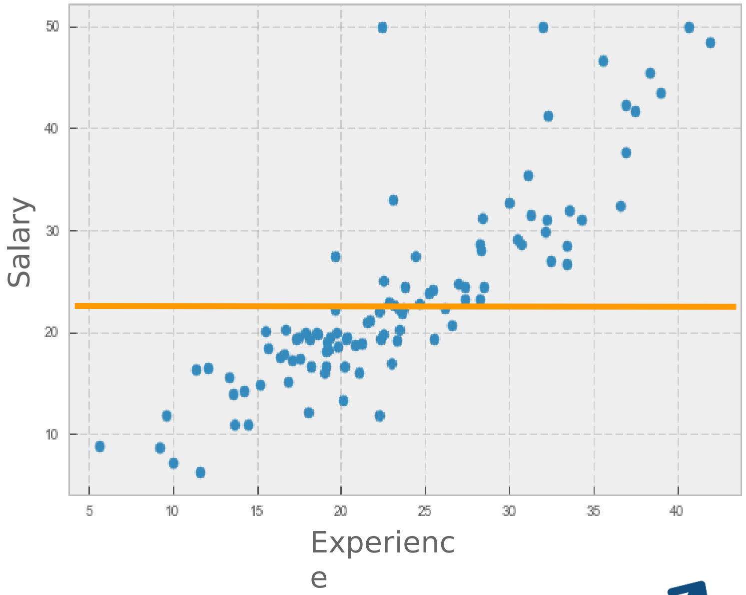
Gradient Descent in Linear Regression

1. Random initialisation

2. Generating predictions

3. Calculating cost

Gradient Descent in Linear Regression



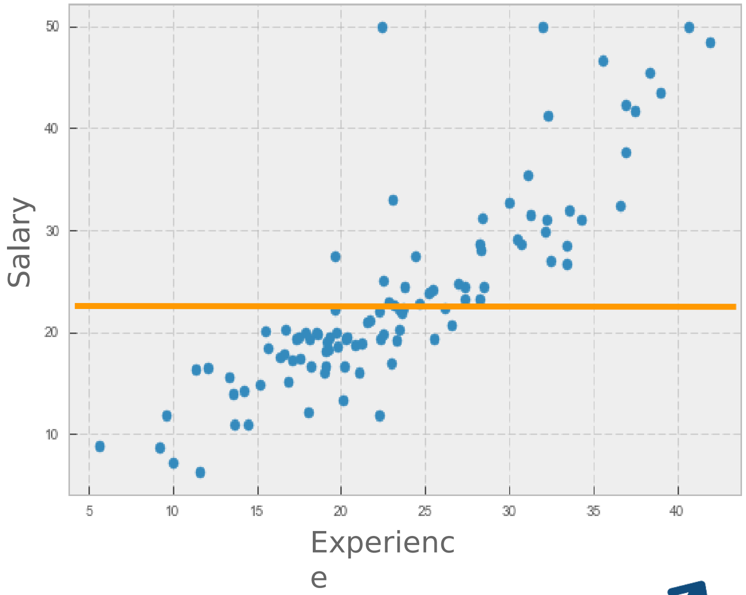
$$\hat{Y} = ßX + b$$

Analytics Vidhya

Learn everything about analytics

Gradient Descent in Linear Regression

J = MSE

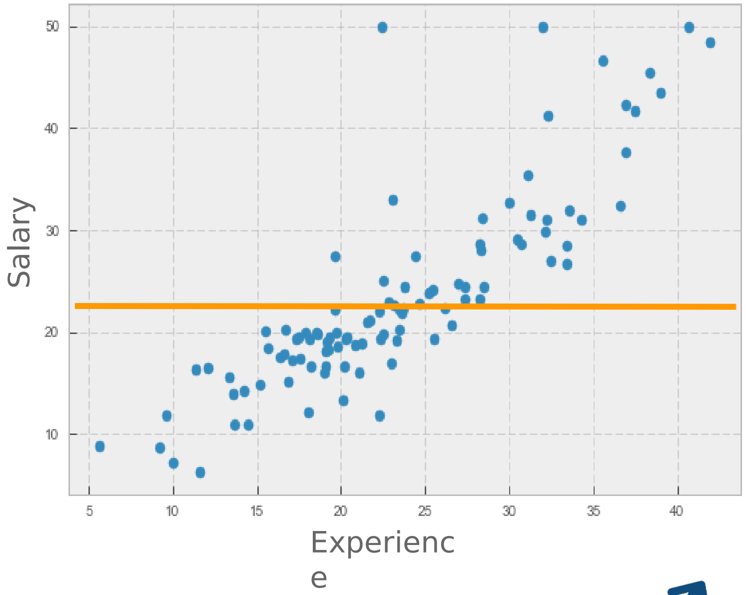


$\hat{Y} = ßX + b$

Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

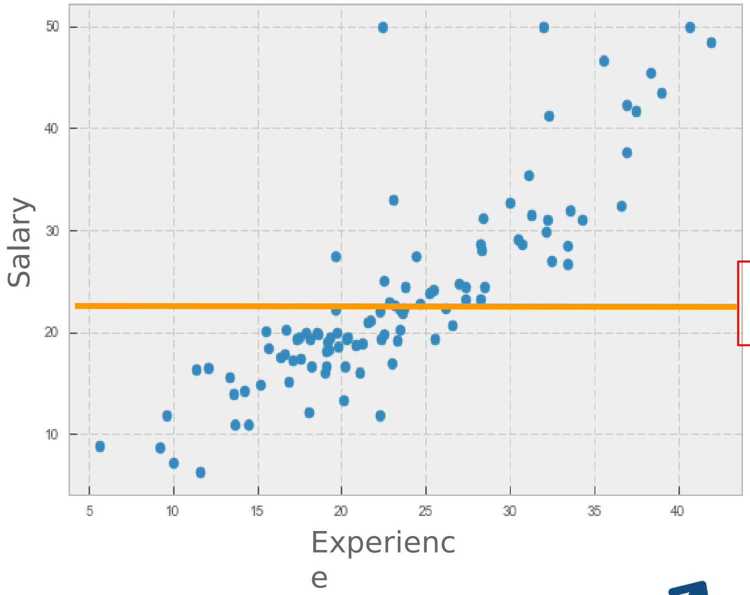
$$J = \frac{\sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2}{n}$$



$\hat{Y} = \beta X + b$

Gradient Descent in Linear Regression

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

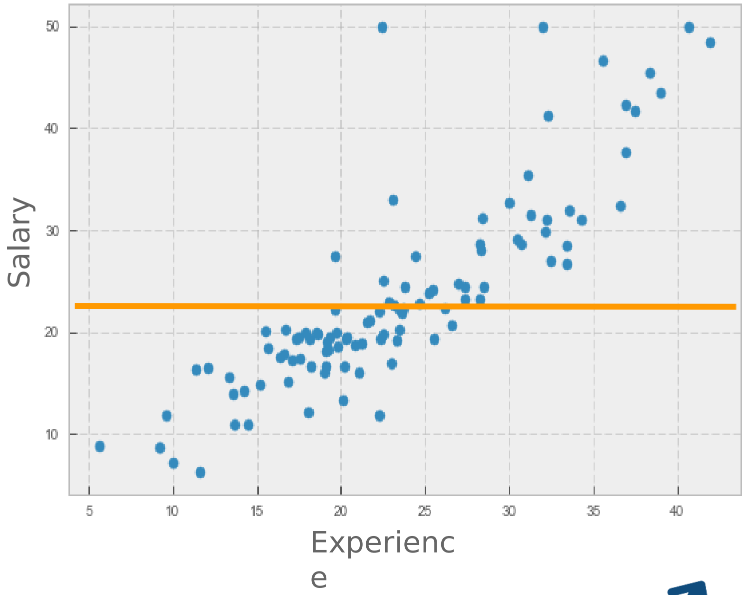


$$\hat{Y} = ßX + b$$

Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$



$\hat{Y} = ßX + b$

Gradient Descent in Linear Regression

1. Random initialisation
2. Generating predictions
3. Calculating cost
4. Updating Parameters

Gradient Descent in Linear Regression

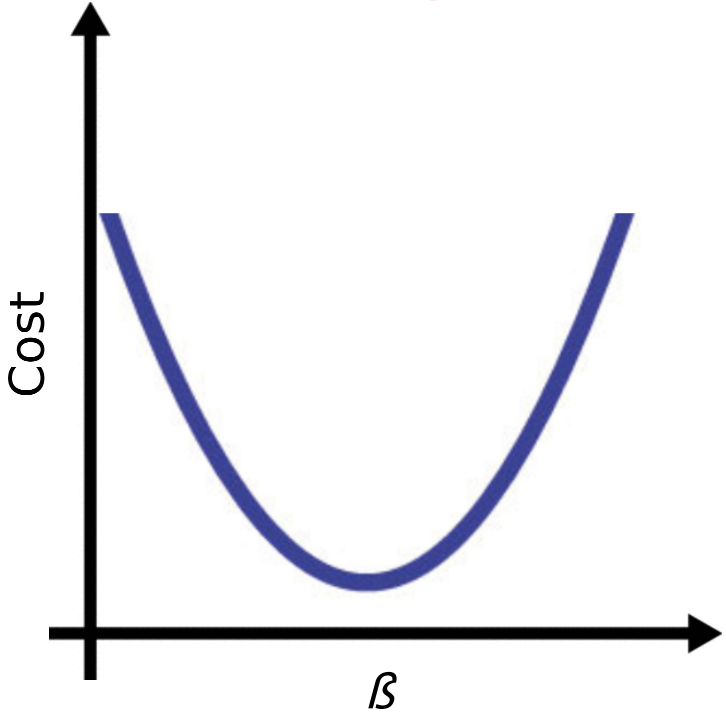
$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

$$ß = ß - Z_1$$

$$b = b - Z_2$$

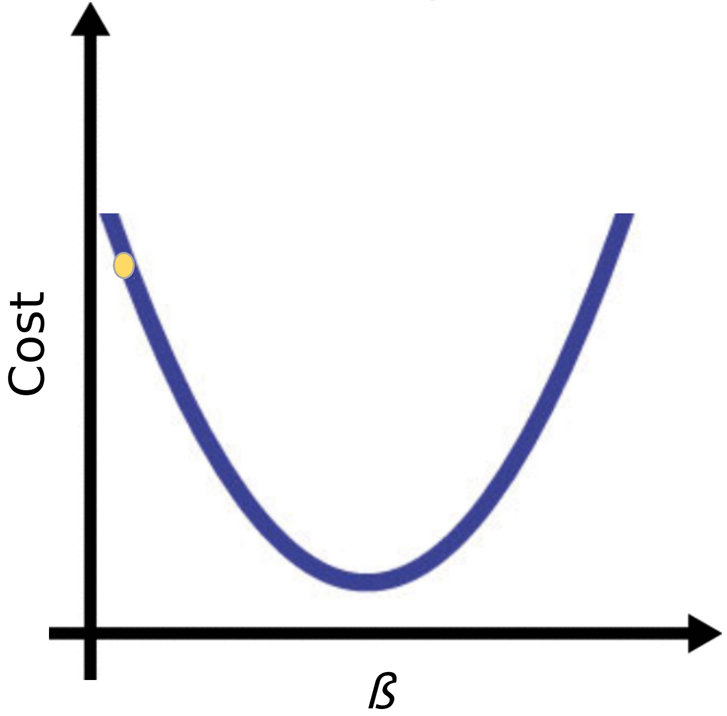
Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$



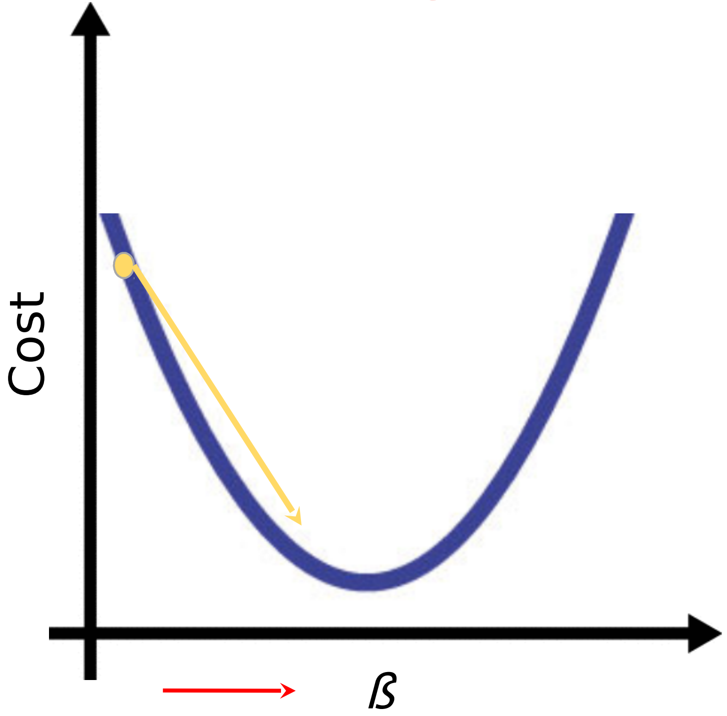
Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$



Gradient Descent in Linear Regression

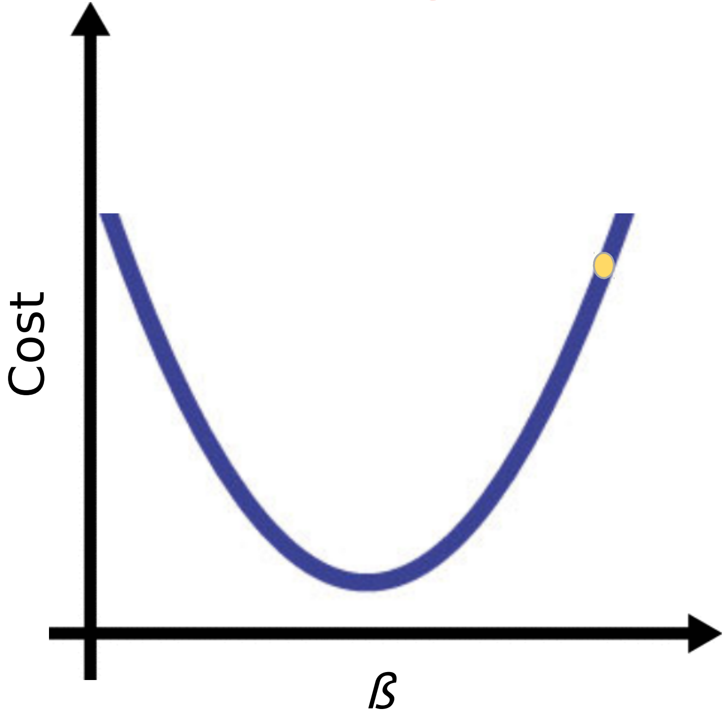
$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$



Analytics Vidhya
Learn everything about analytics

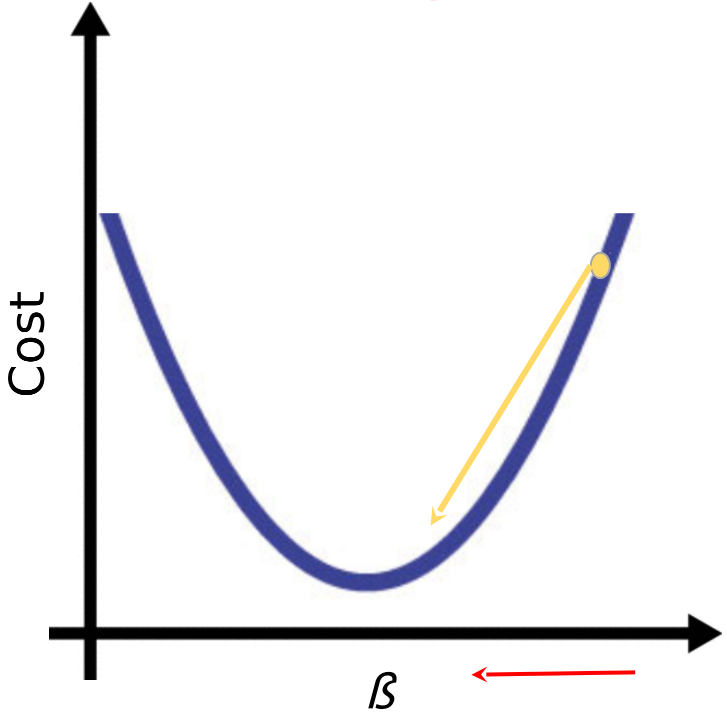
Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$



Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

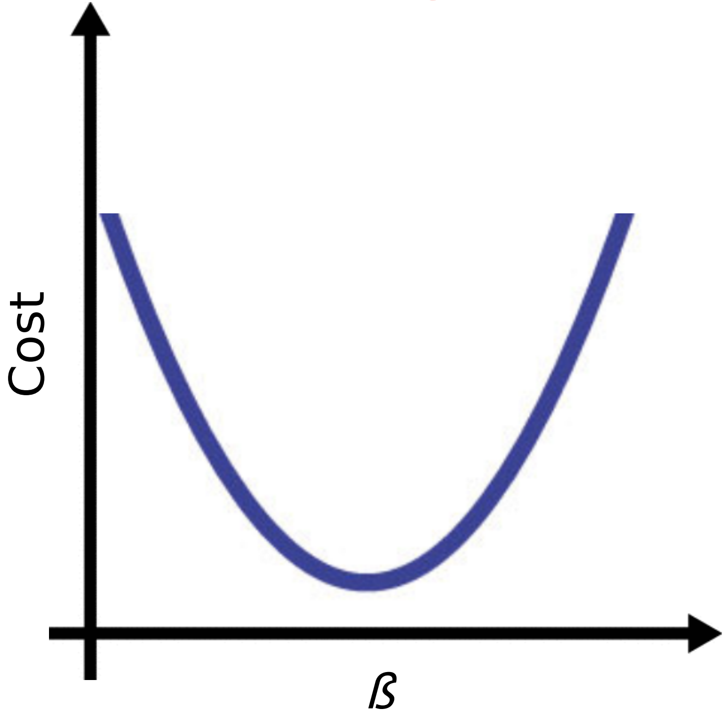


Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

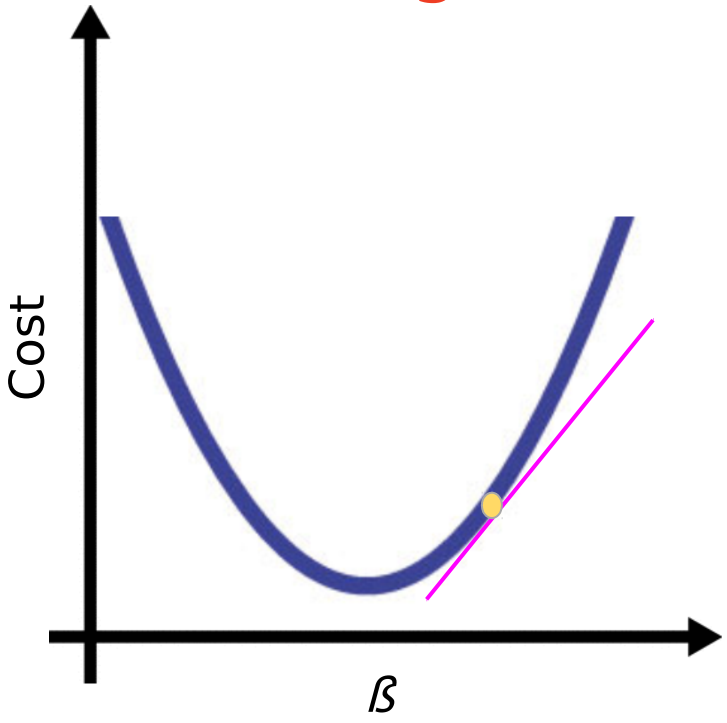
$$\frac{\partial y}{\partial x}$$



Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

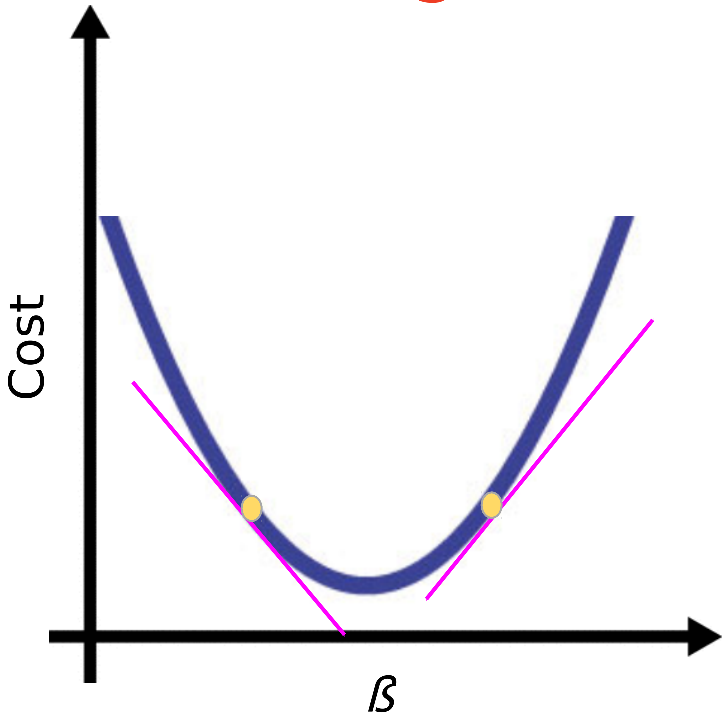
$$\frac{\partial y}{\partial x}$$



Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (\beta X_i + b - Y_i)^2}{n}$$

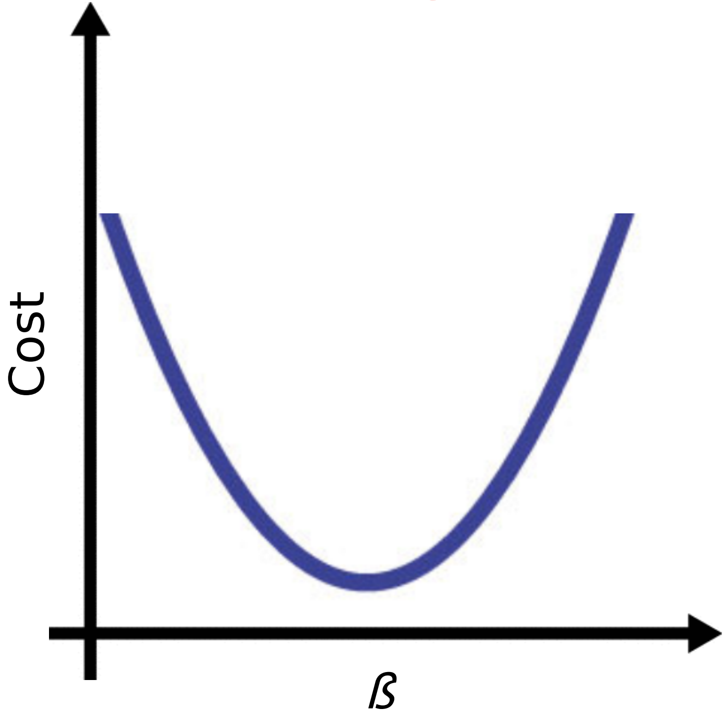
$$\frac{\partial y}{\partial x}$$



Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

$G_ß$

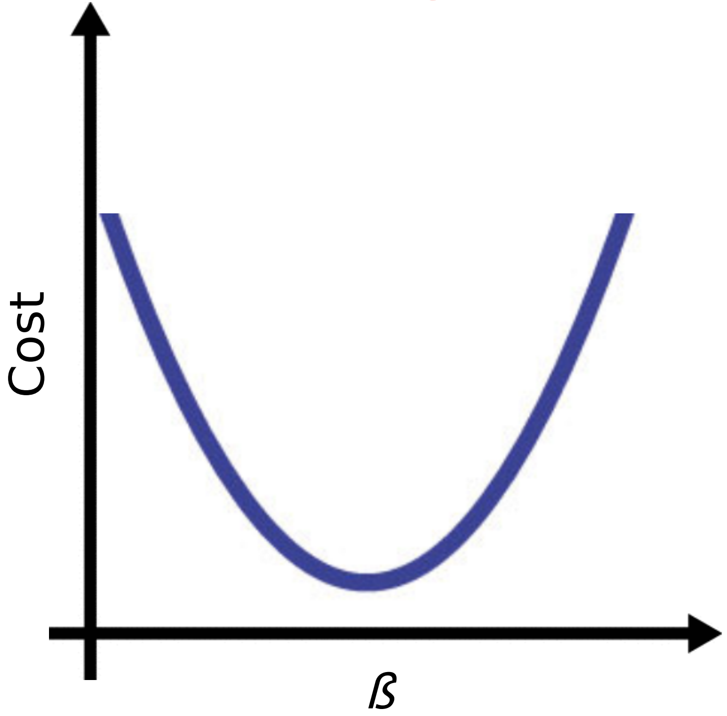


Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

$$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2 \sum_{i=1}^{n} (ßX_i + b - Y_i) X_i}{n}$$



Gradient Descent in Linear Regression

$$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2\sum_{i=1}^{n}(ßX_i + b - Y_i)X_i}{n}$$

$$ß = ß - Z$$

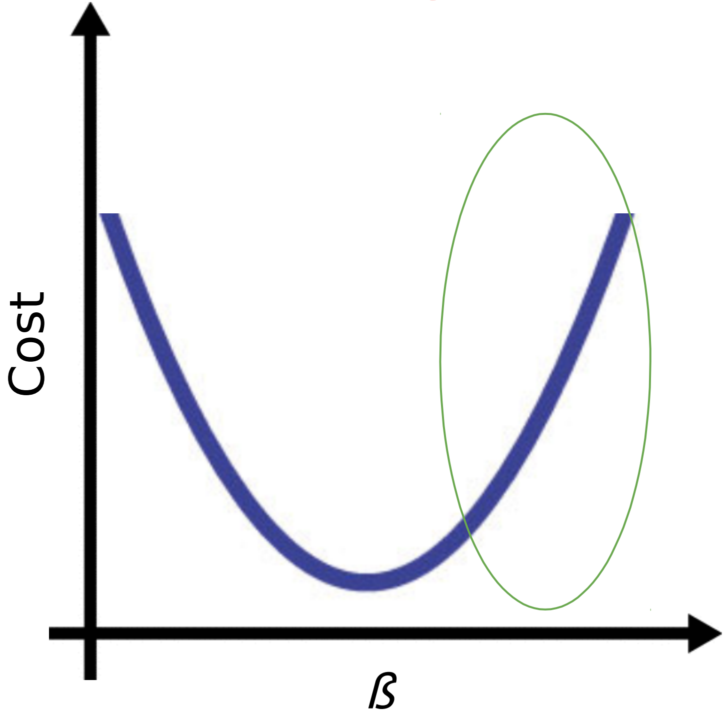
Gradient Descent in Linear Regression

$$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2\sum_{i=1}^{n}(ßX_i + b - Y_i)\,X_i}{n} \qquad ß = ß - \alpha G_ß$$

Analytics Vidhya
Learn everything about analytics

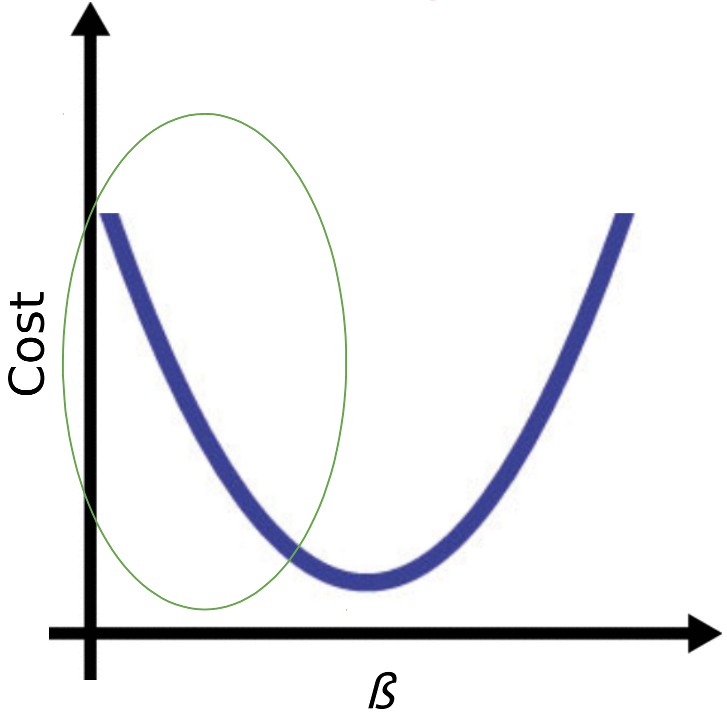
Gradient Descent in Linear Regression

$$ß = ß \boxed{- \alpha G}_{ß}$$



Gradient Descent in Linear Regression

$$ß = ß \boxed{- \alpha} G_{ß}$$



Gradient Descent in Linear Regression

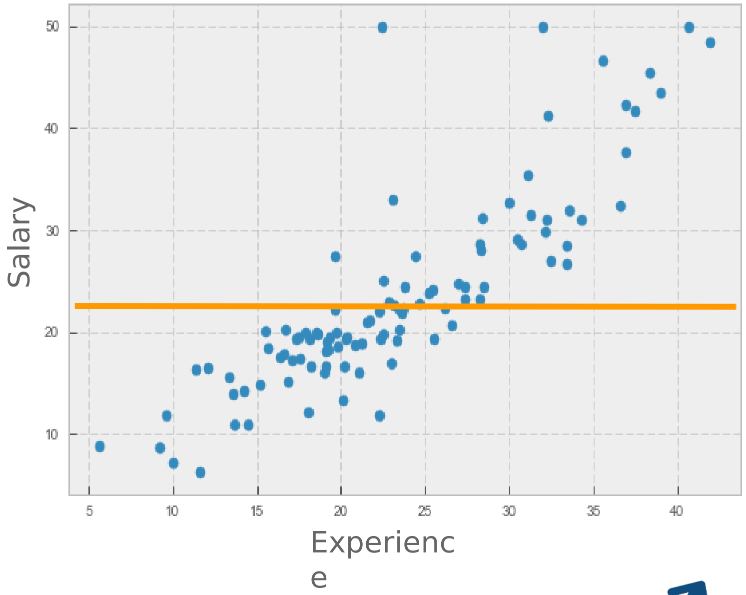
Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*

Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*

2. $\hat{Y}$ = ßX + b



$\hat{Y}$ = ßX + b

Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*

2. $\hat{Y} = ßX + b$

3. $$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*

2. $\hat{Y} = ßX + b$

3. $J = \dfrac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$

4. $G_ß = \dfrac{\partial(J)}{\partial ß} \quad \dfrac{2 \sum_{i=1}^{n} (ßX_i + b - Y_i) X_i}{n}$

Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*
2. $\hat{Y} = ßX + b$
3. $J = \dfrac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$
4. $G_ß = \dfrac{\partial(J)}{\partial ß} = \dfrac{2 \sum_{i=1}^{n} (ßX_i + b - Y_i) X_i}{n}$
5. $ß = ß - \alpha G_ß$

Gradient Descent in Linear Regression

1. *ß -> 0; b -> mean*

2. $\hat{Y} = ßX + b$

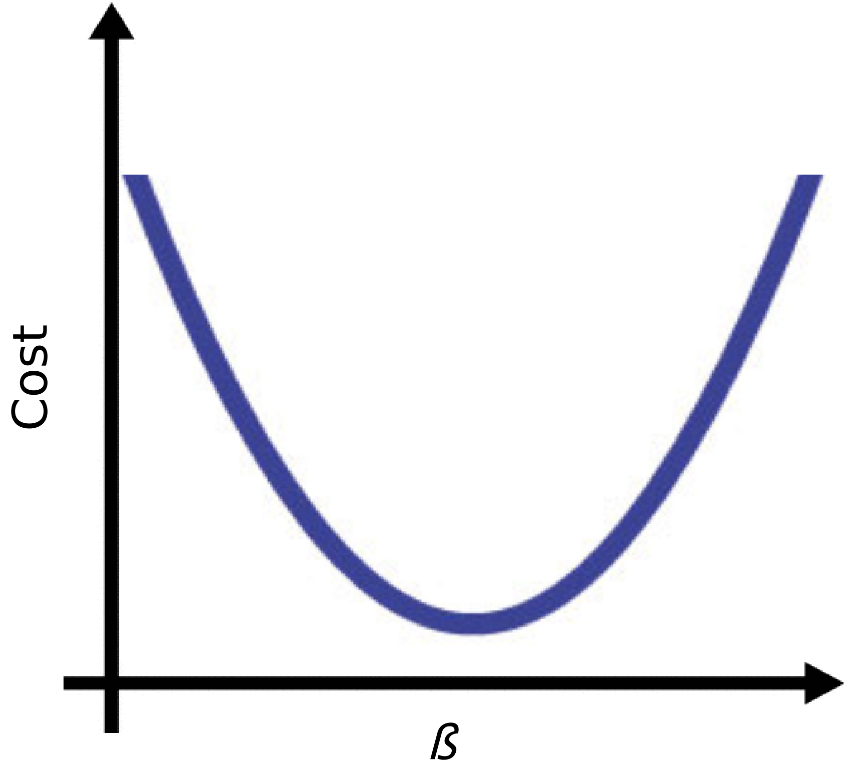
3. $$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

4. $$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2 \sum_{i=1}^{n} (ßX_i + b - Y_i) X_i}{n}$$

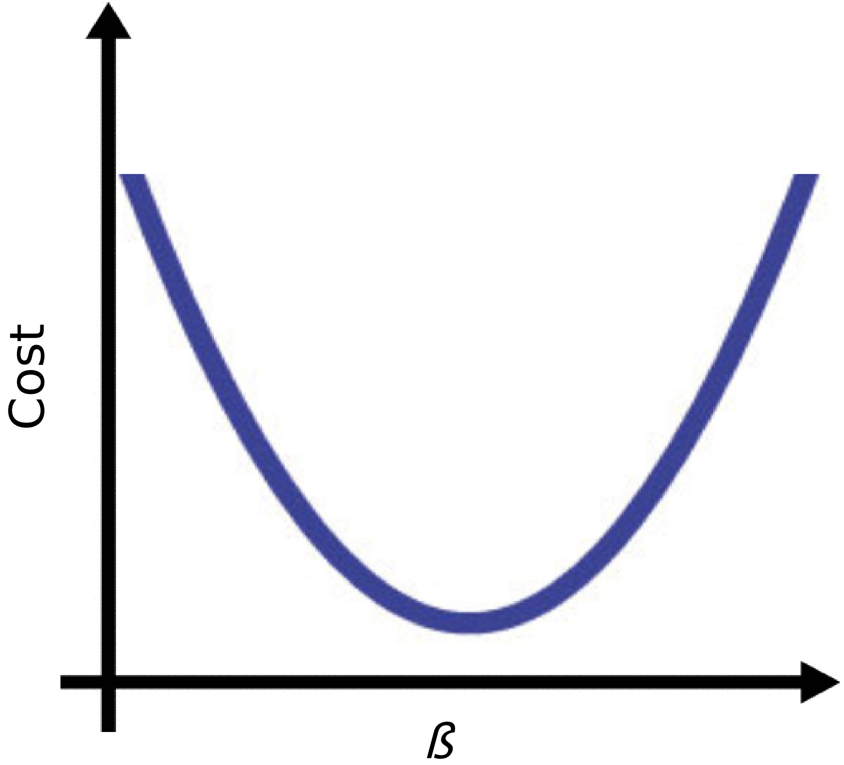
5. $ß = ß - \alpha G_ß$

6. Repeat steps 2 - 5

Gradient Descent in Linear Regression

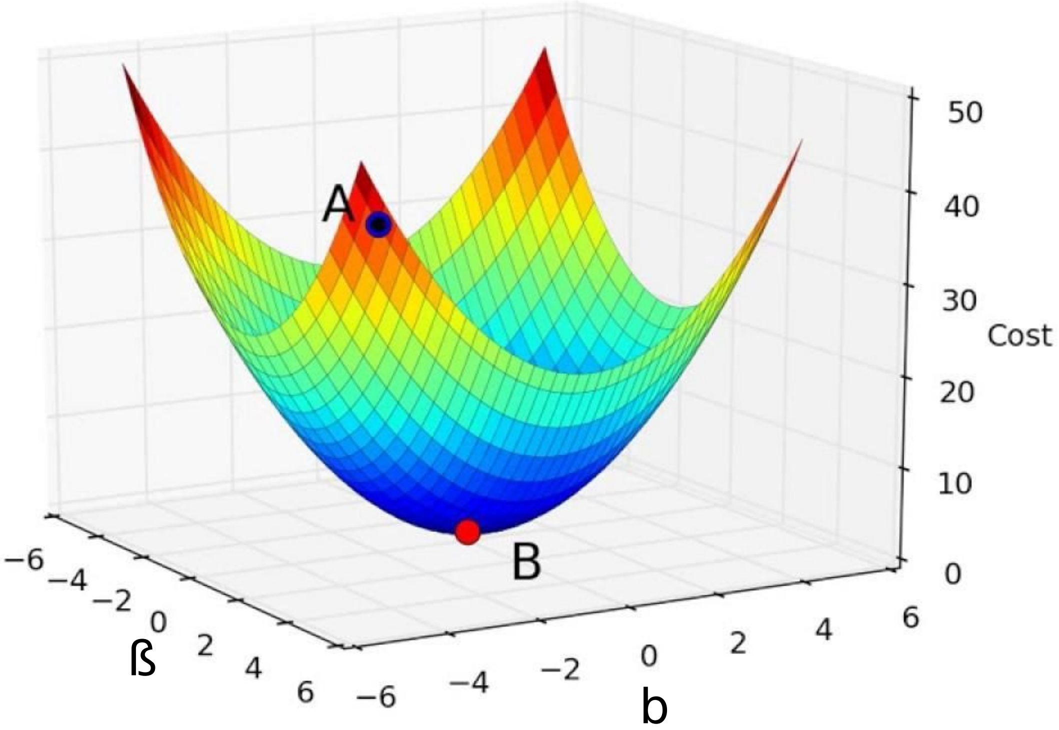


Gradient Descent in Linear Regression



$$ß = ß - \alpha G_{ß}$$

Gradient Descent in Linear Regression



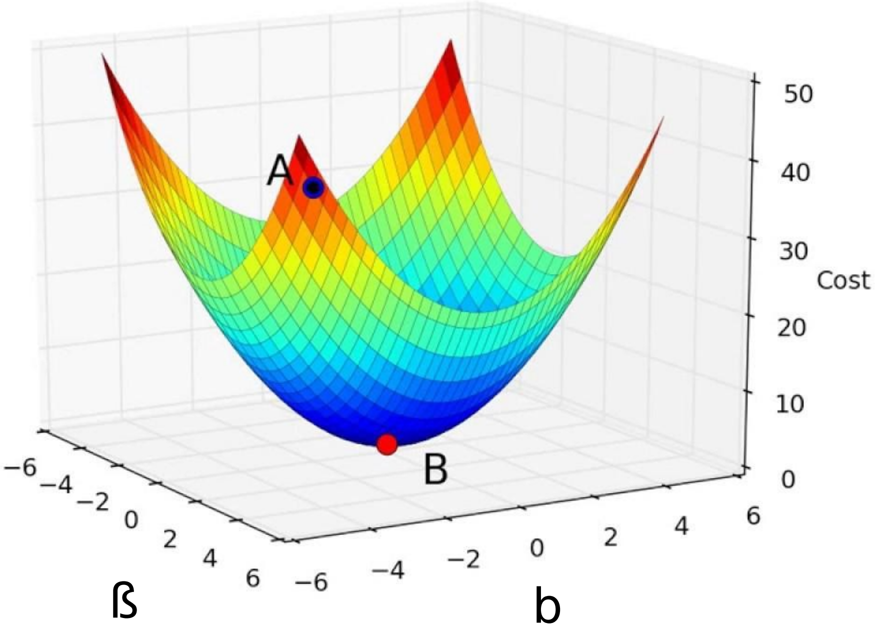
Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

$$J = \frac{\sum_{i=1}^{n} (ßX_i + b - Y_i)^2}{n}$$

$$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2\sum_{i=1}^{n} (ßX_i + b - Y_i) X_i}{n}$$

$$G_b = \frac{\partial(J)}{\partial b} = \frac{2\sum_{i=1}^{n} (ßX_i + b - Y_i)}{n}$$



Analytics Vidhya
Learn everything about analytics

Gradient Descent in Linear Regression

$$G_ß = \frac{\partial(J)}{\partial ß} = \frac{2\sum_{i=1}^{n}(ßX_i + b - Y_i)X_i}{n} \qquad ß = ß - \alpha G_ß$$

$$G_b = \frac{\partial(J)}{\partial b} = \frac{2\sum_{i=1}^{n}(ßX_i + b - Y_i)}{n} \qquad b = b - \alpha G_b$$

Gradient Descent in Linear Regression