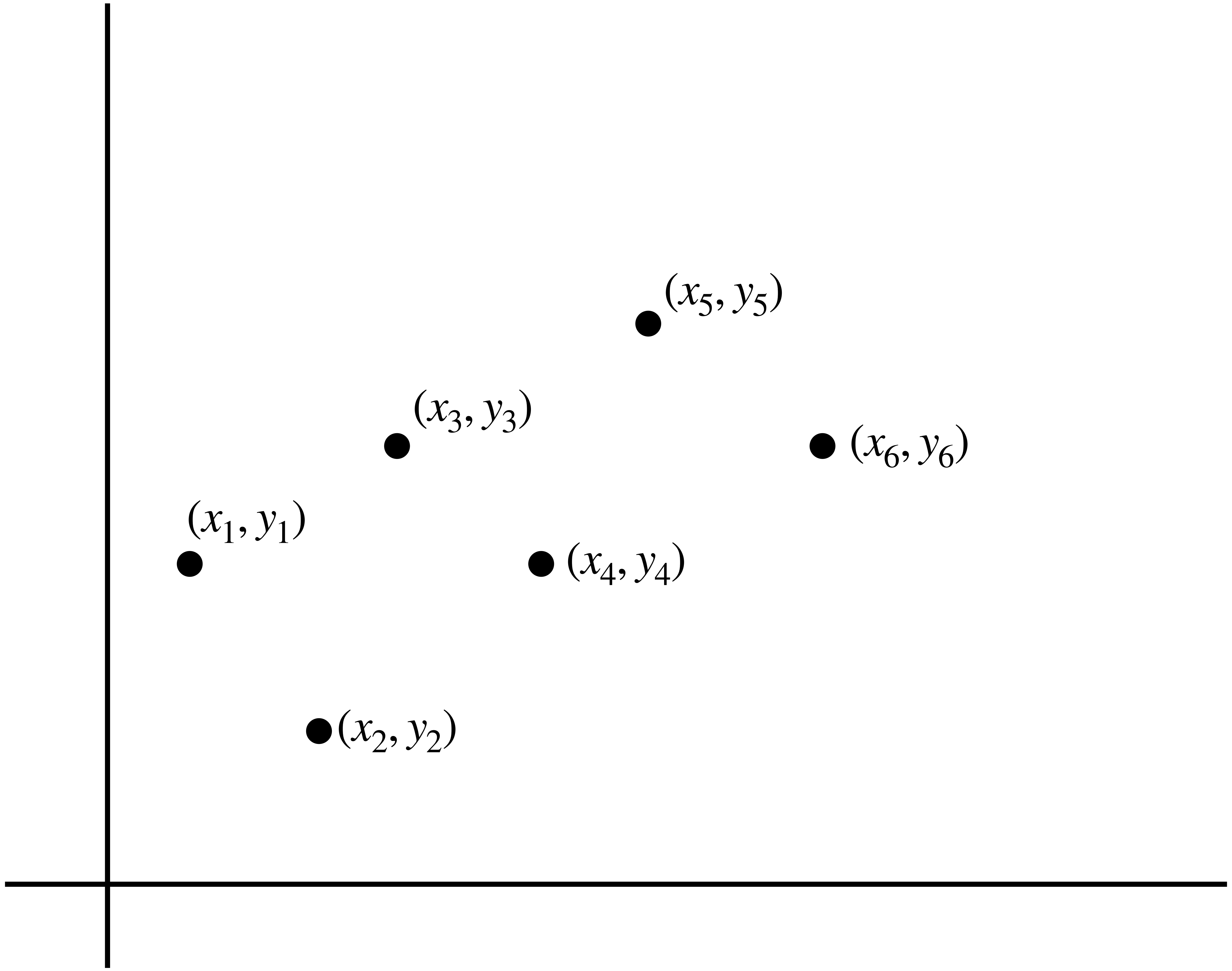
Simple Linear Regression

Proof of the Closed Form Solution

Rahul Singh
rsingh@arrsingh.com

Problem Statement

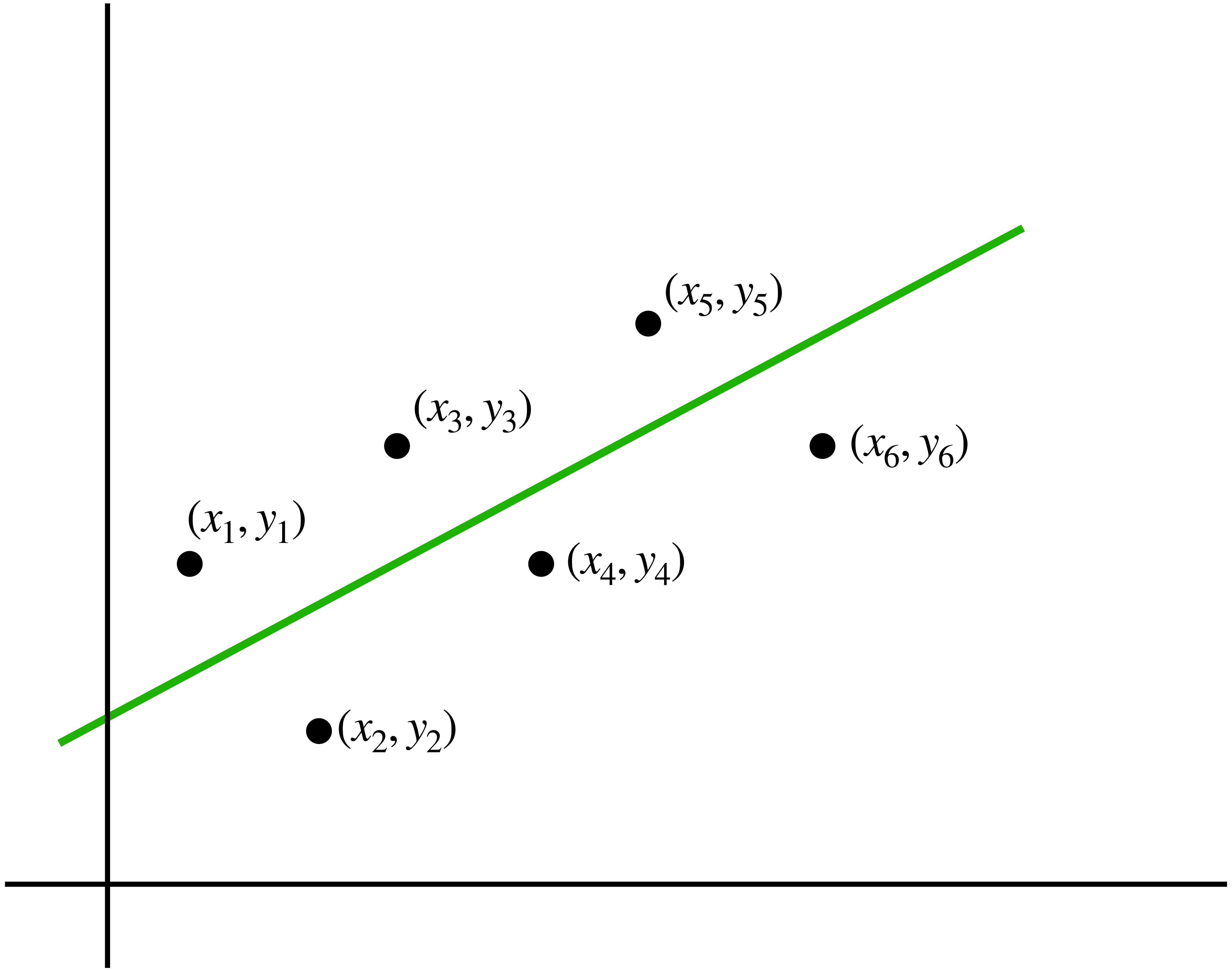
Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data



Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

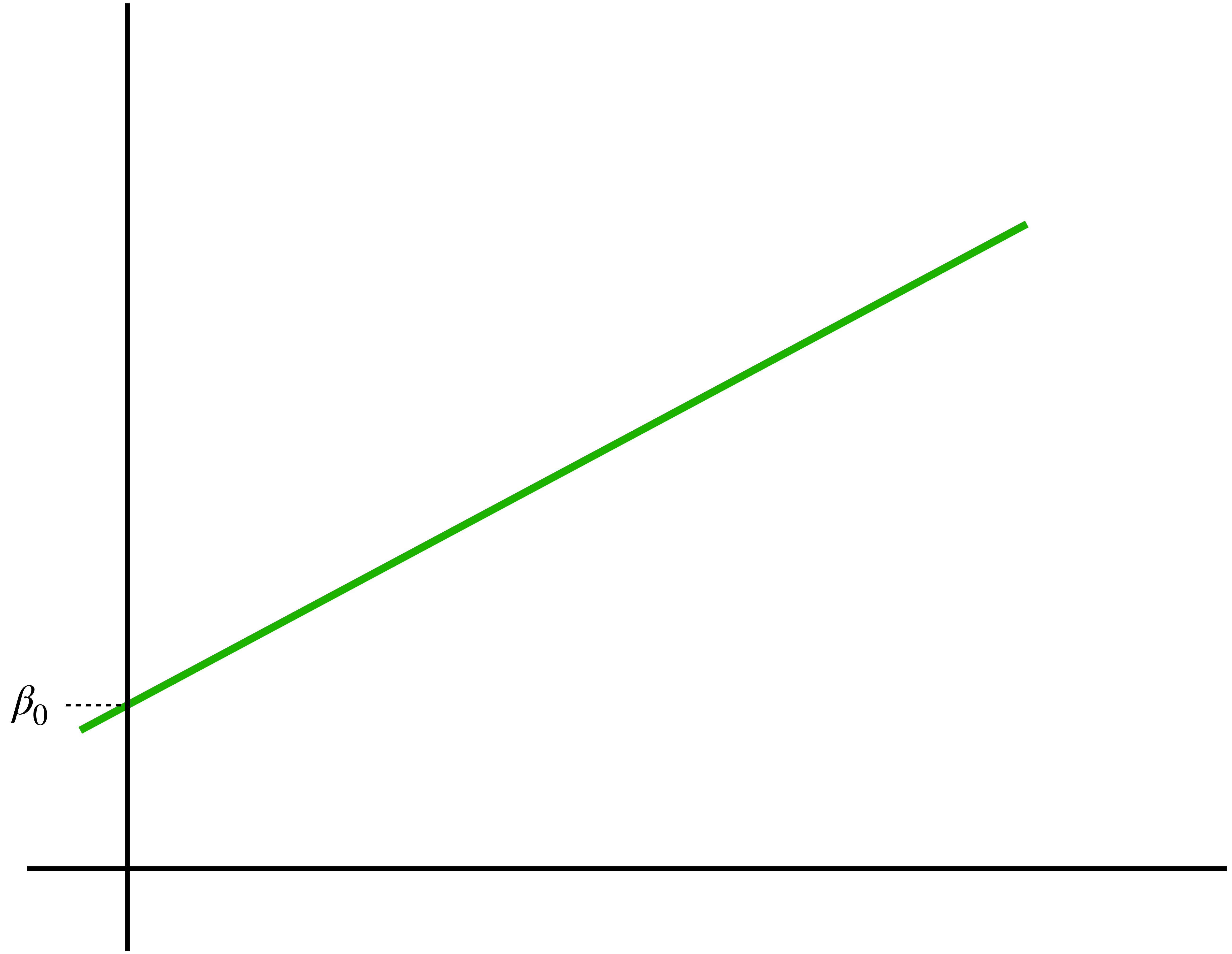


Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

β_0 Is the Y intercept



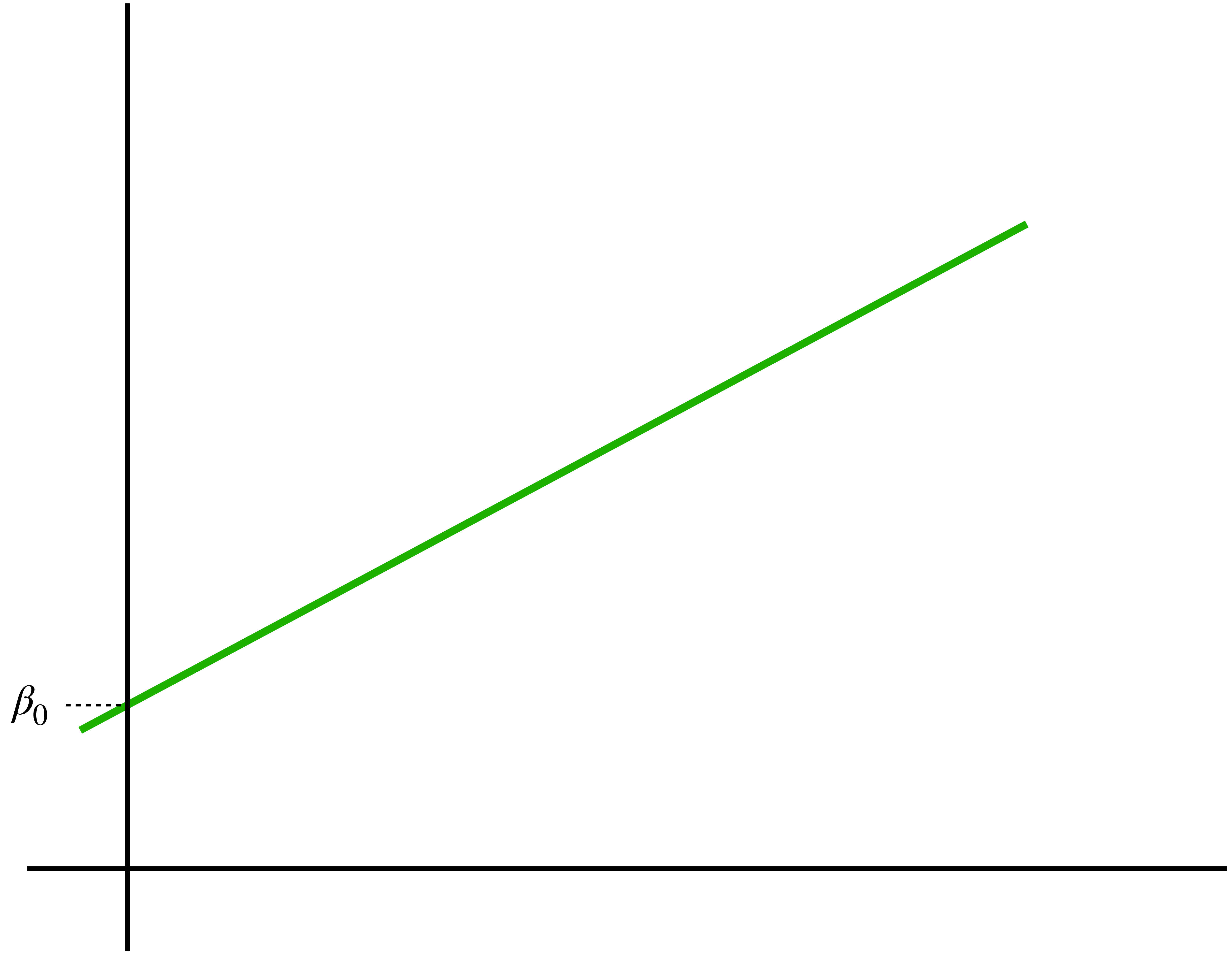
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

β_0 Is the Y intercept

β_1 Is the slope of the line



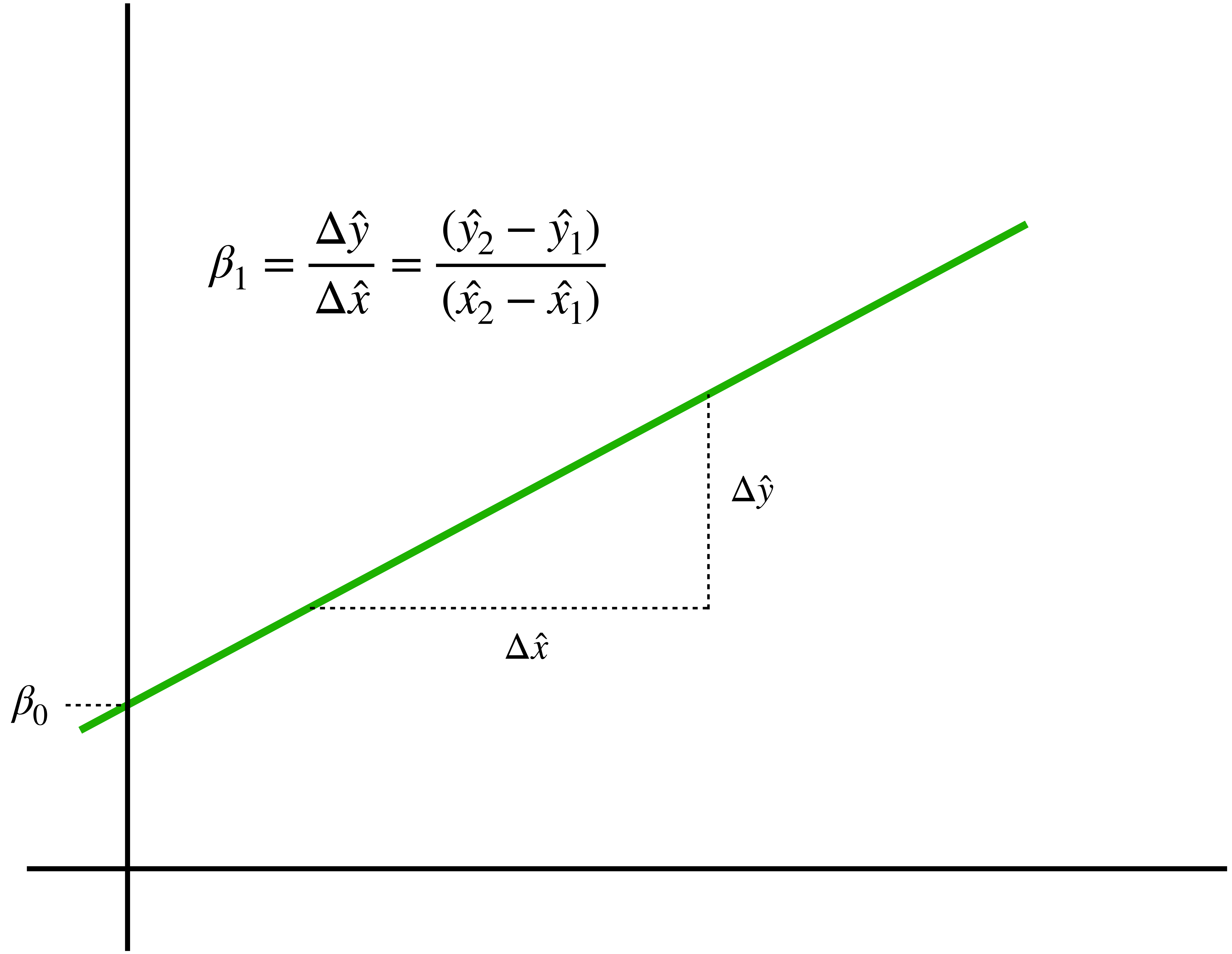
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

β_0 Is the Y intercept

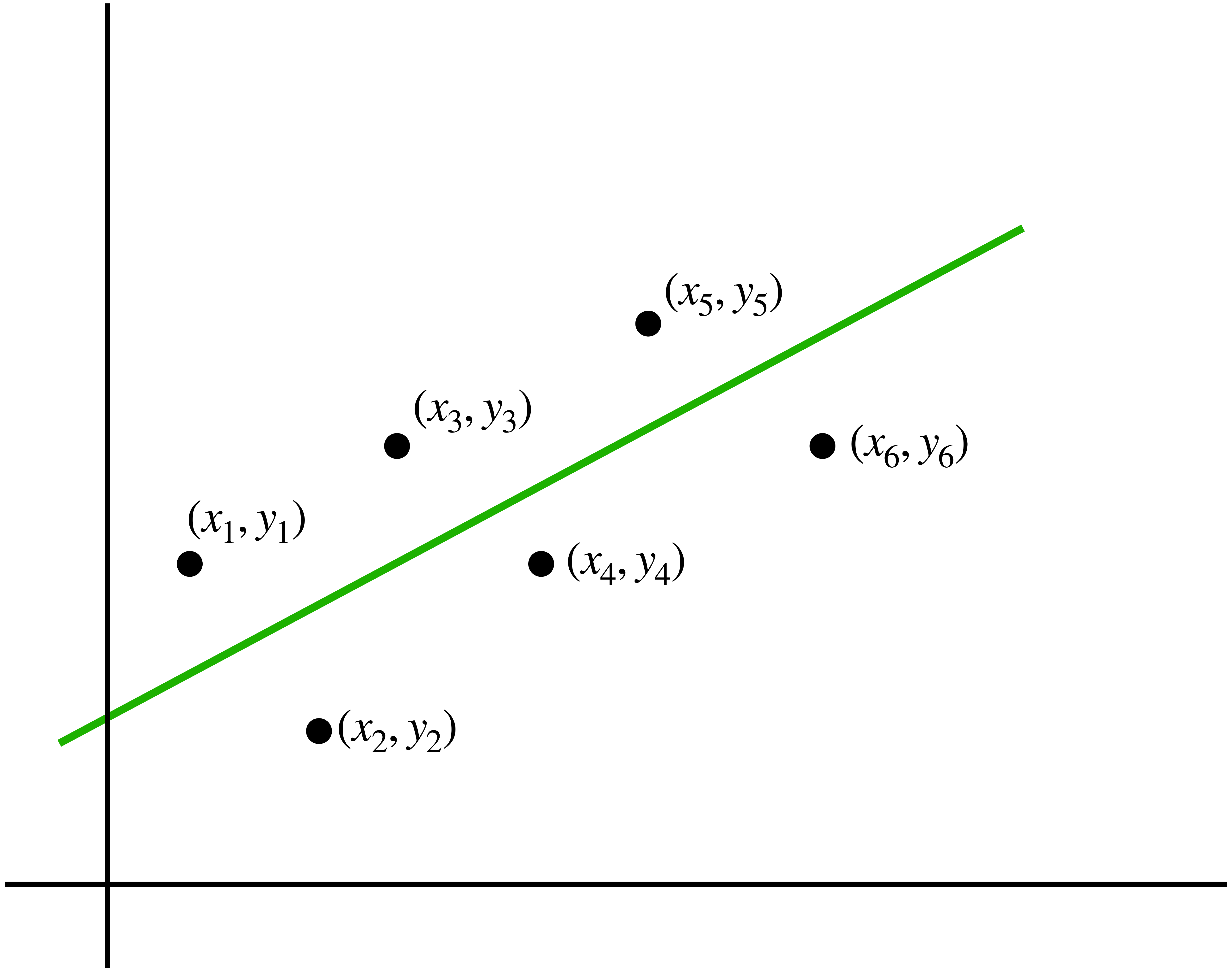
β_1 Is the slope of the line



Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

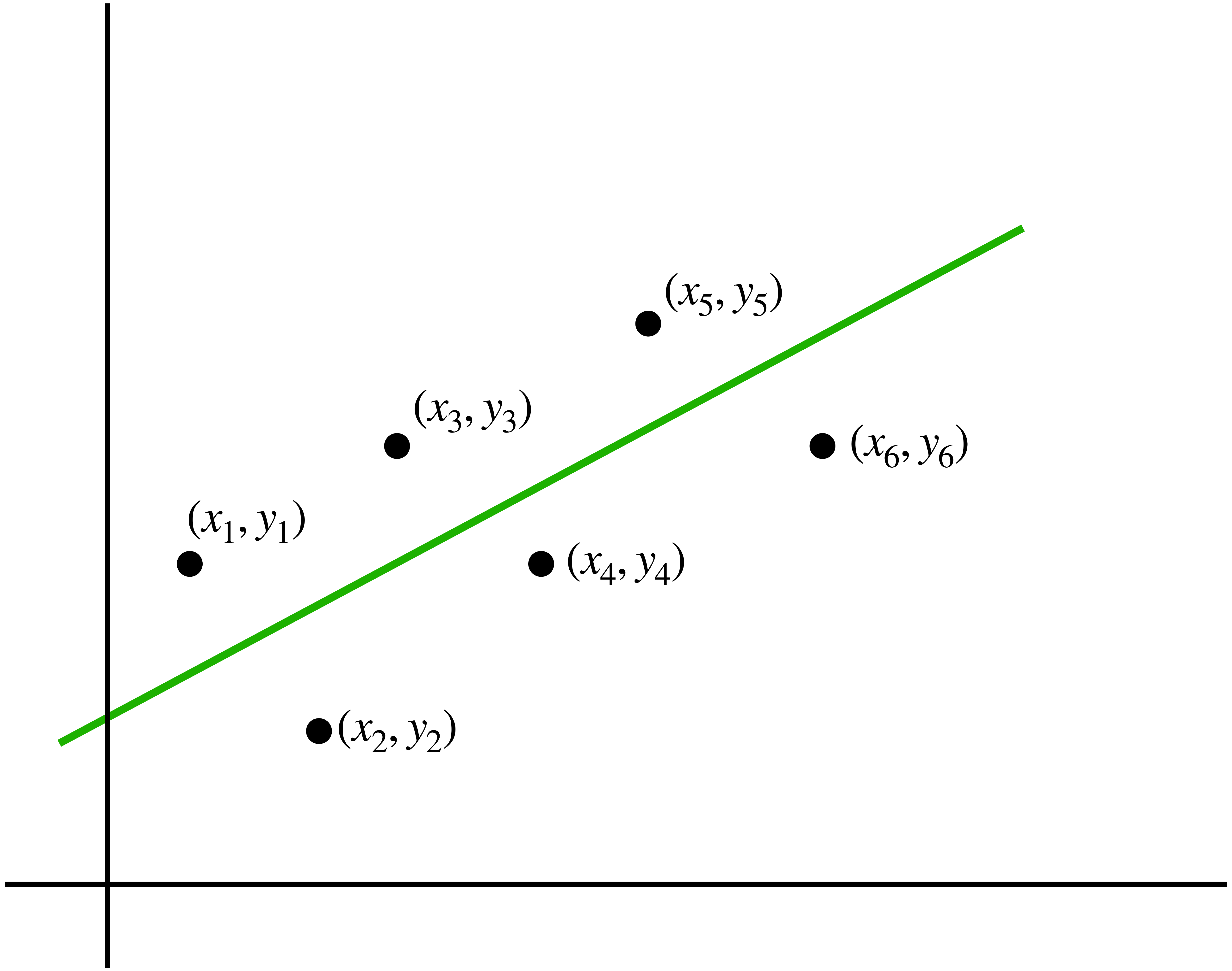


Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):



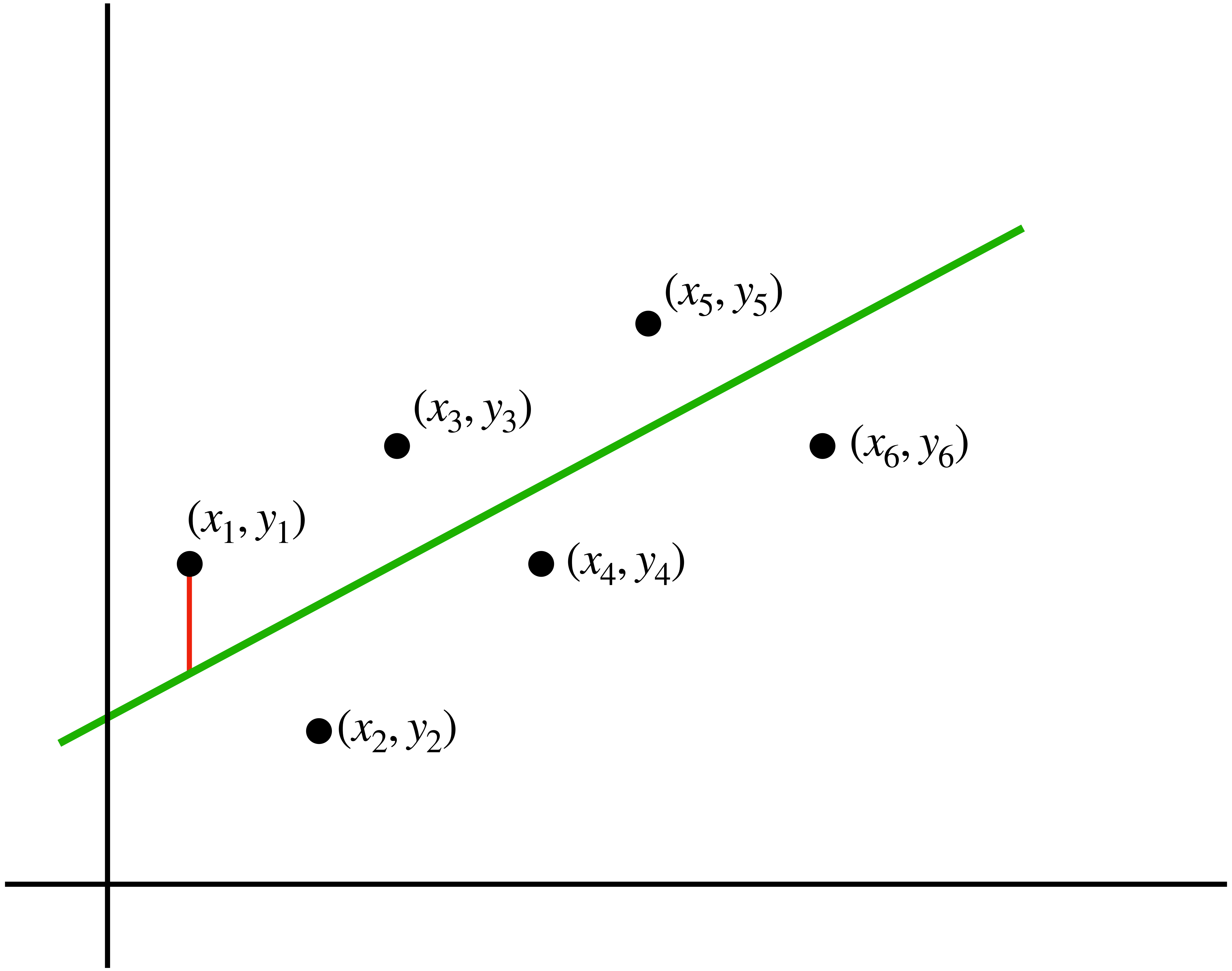
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2$$



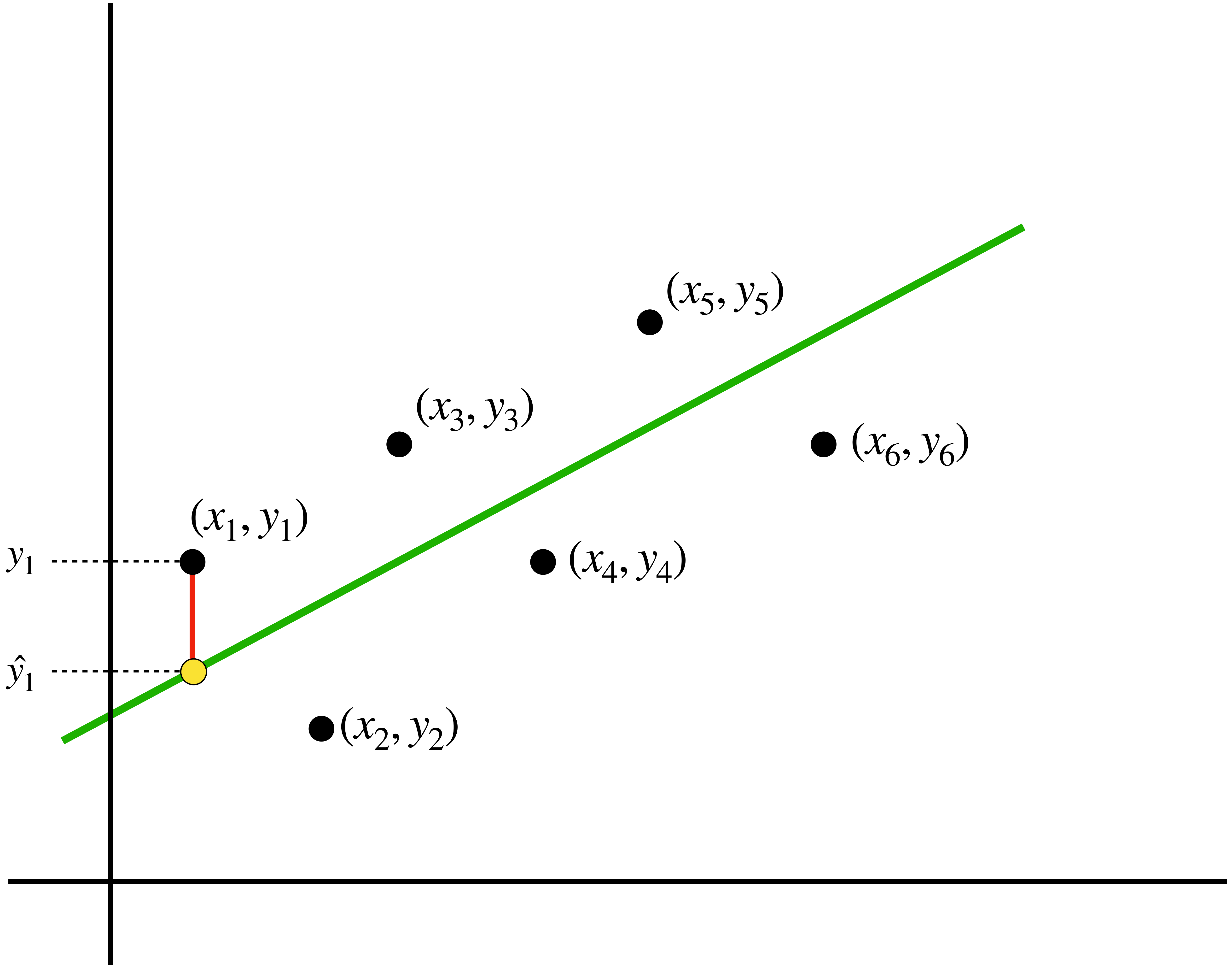
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

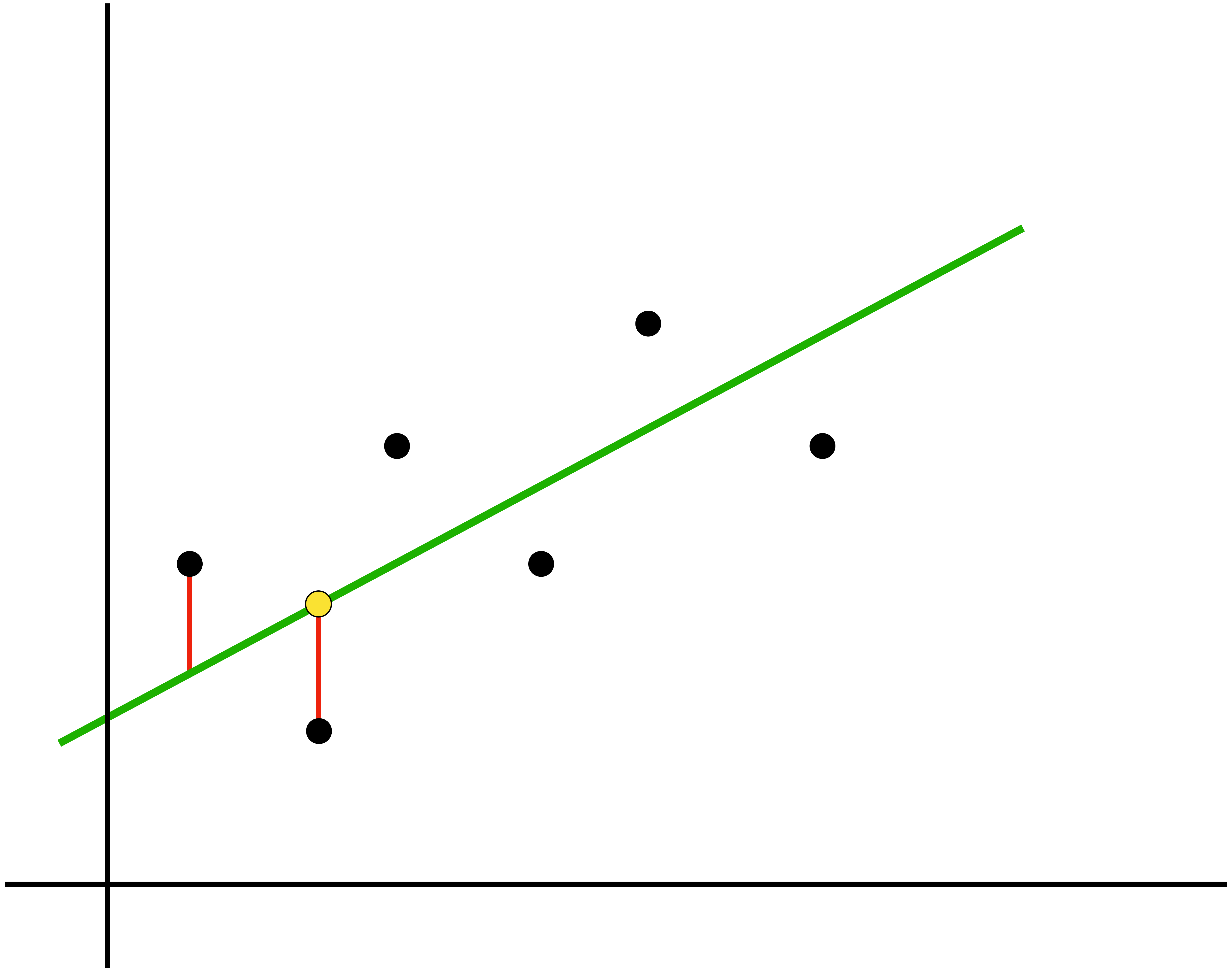
The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2$$



Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data



The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2$$

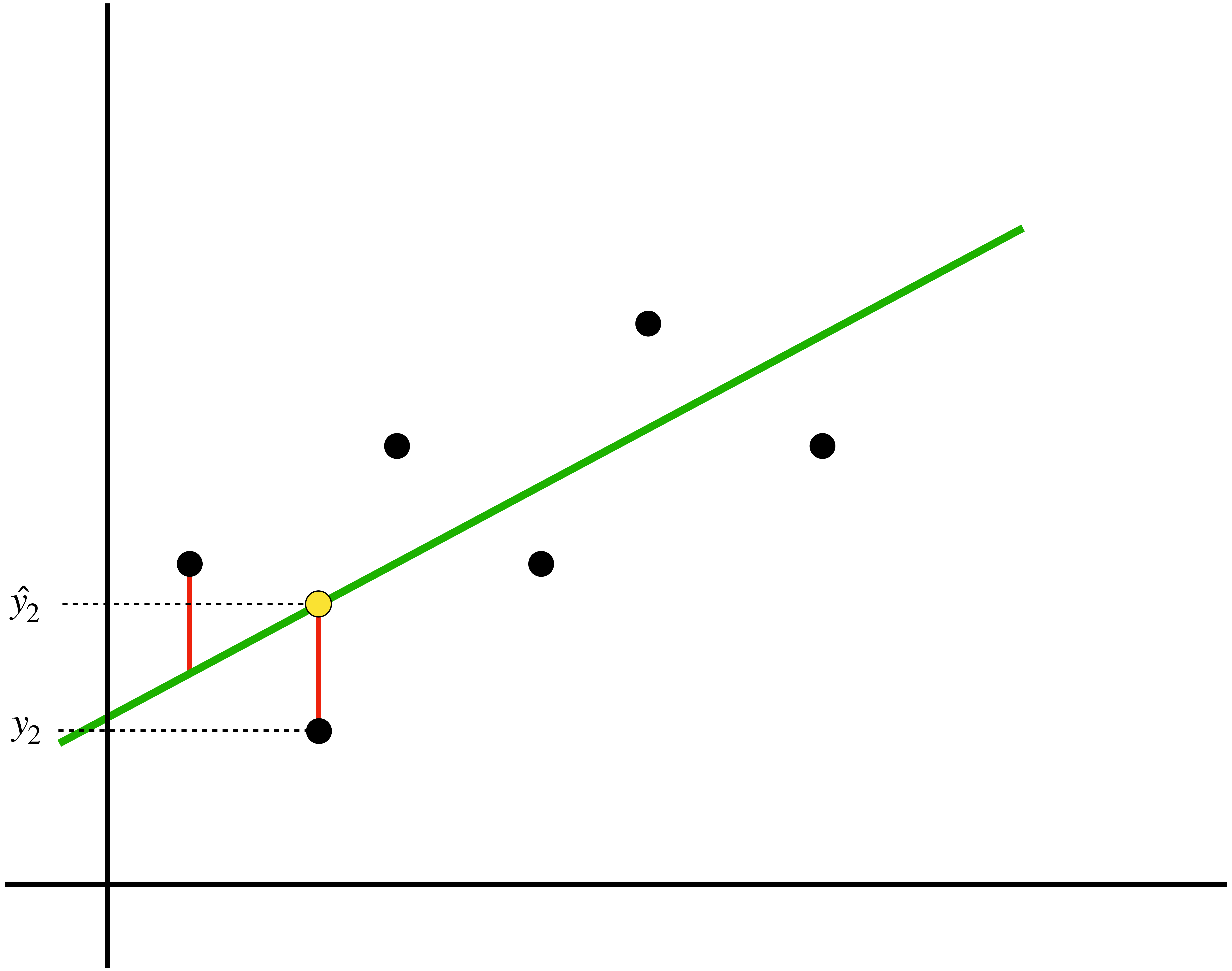
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2$$



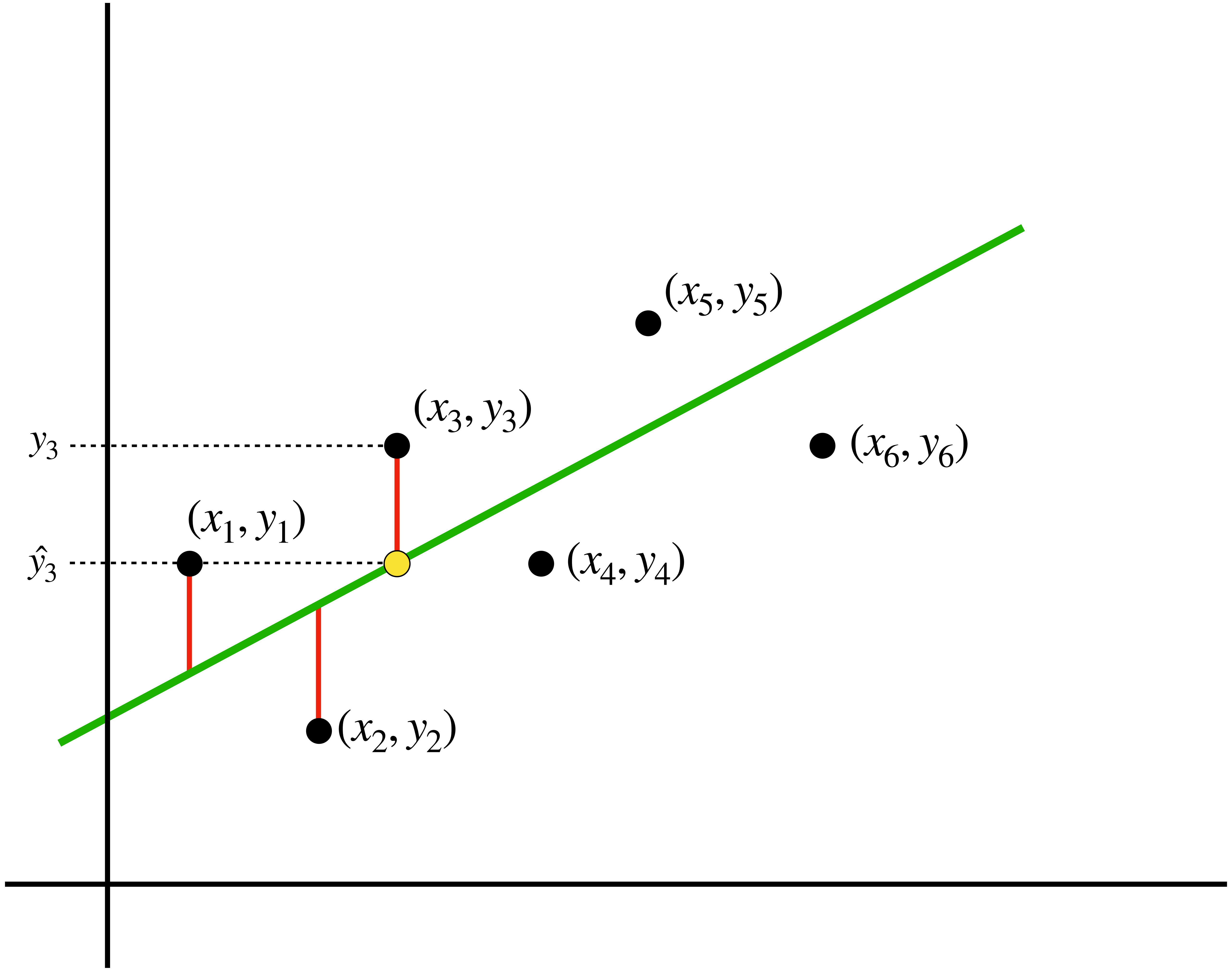
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + (y_3 - \hat{y}_3)^2$$



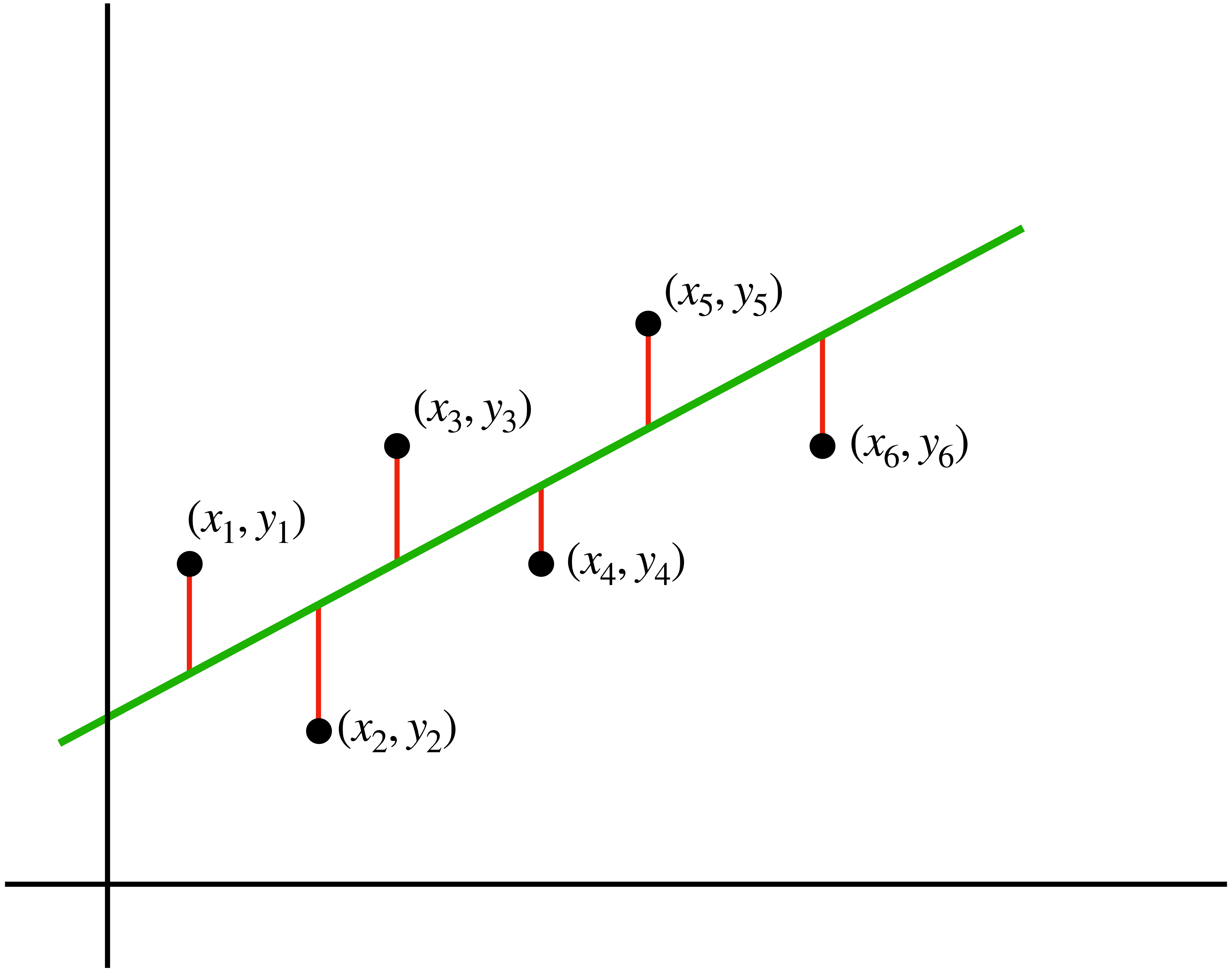
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

The total error (sum of squared errors):

$$(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + (y_3 - \hat{y}_3)^2$$
$$+(y_4 - \hat{y}_4)^2 + (y_5 - \hat{y}_5)^2 + (y_6 - \hat{y}_6)^2$$



Problem Statement

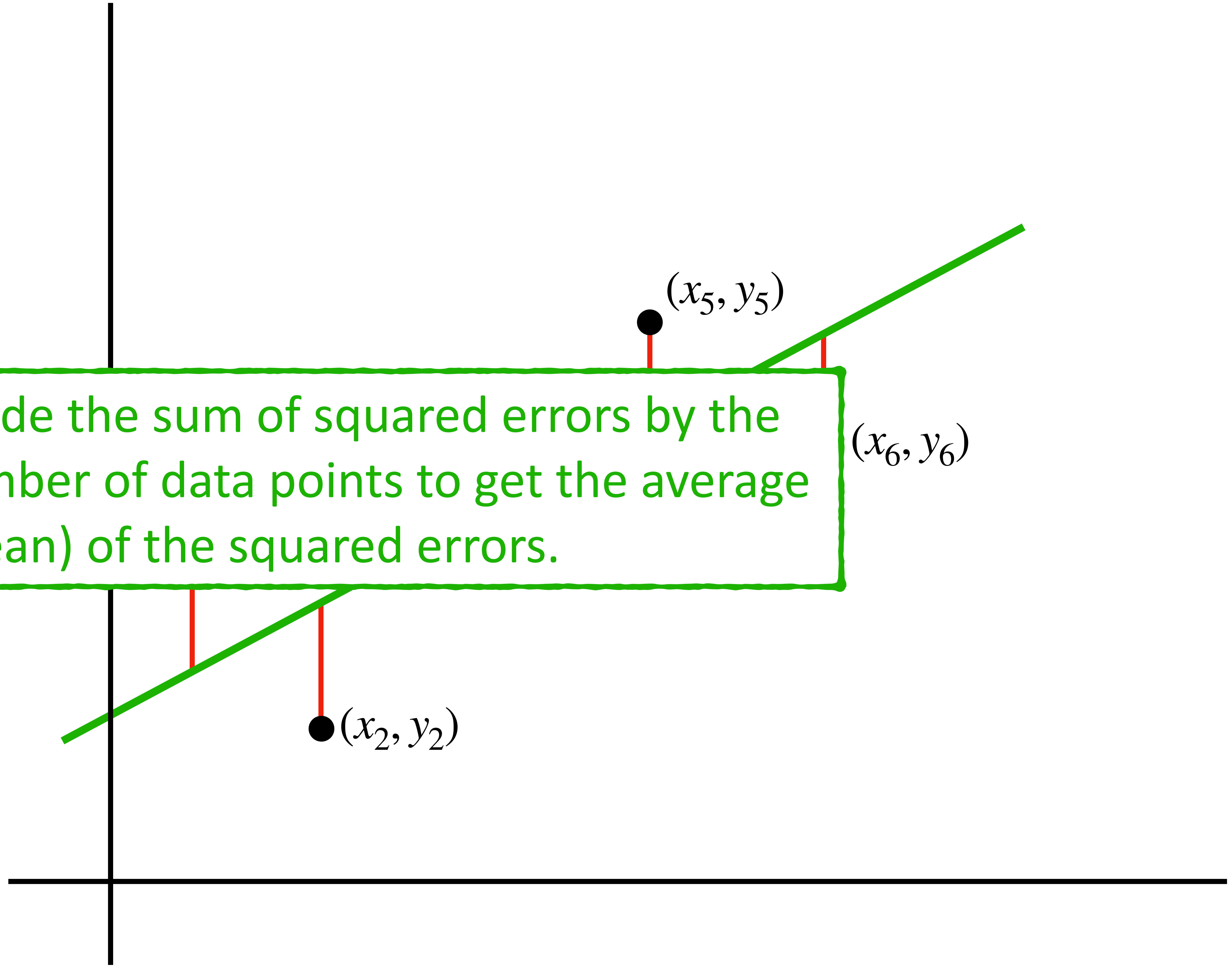
Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{(y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + (y_3 - \hat{y}_3)^2 + (y_4 - \hat{y}_4)^2 + (y_5 - \hat{y}_5)^2 + (y_0 - \hat{y}_0)^2}{n}$$

$$\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$



Divide the sum of squared errors by the number of data points to get the average (mean) of the squared errors.

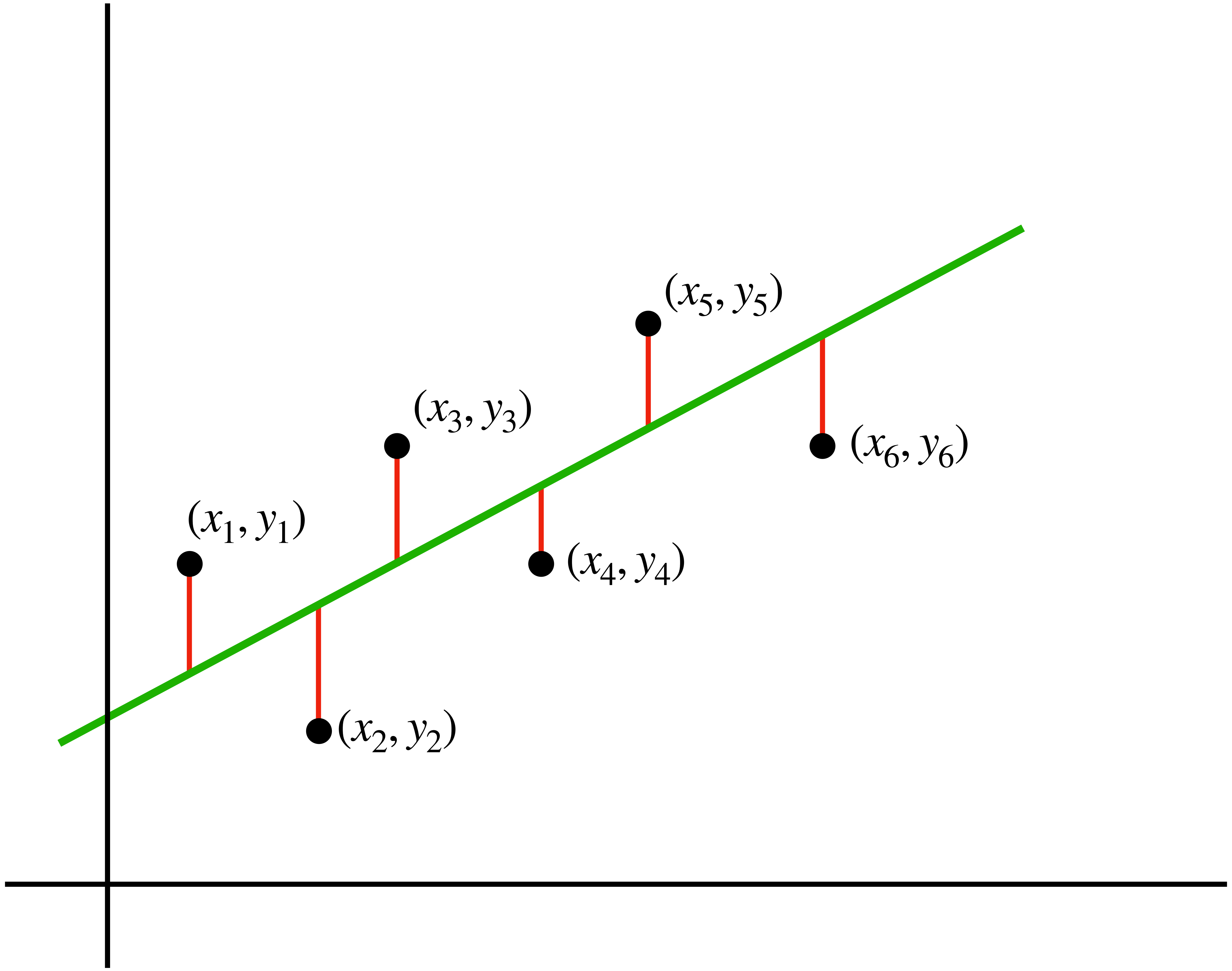
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$



Problem Statement

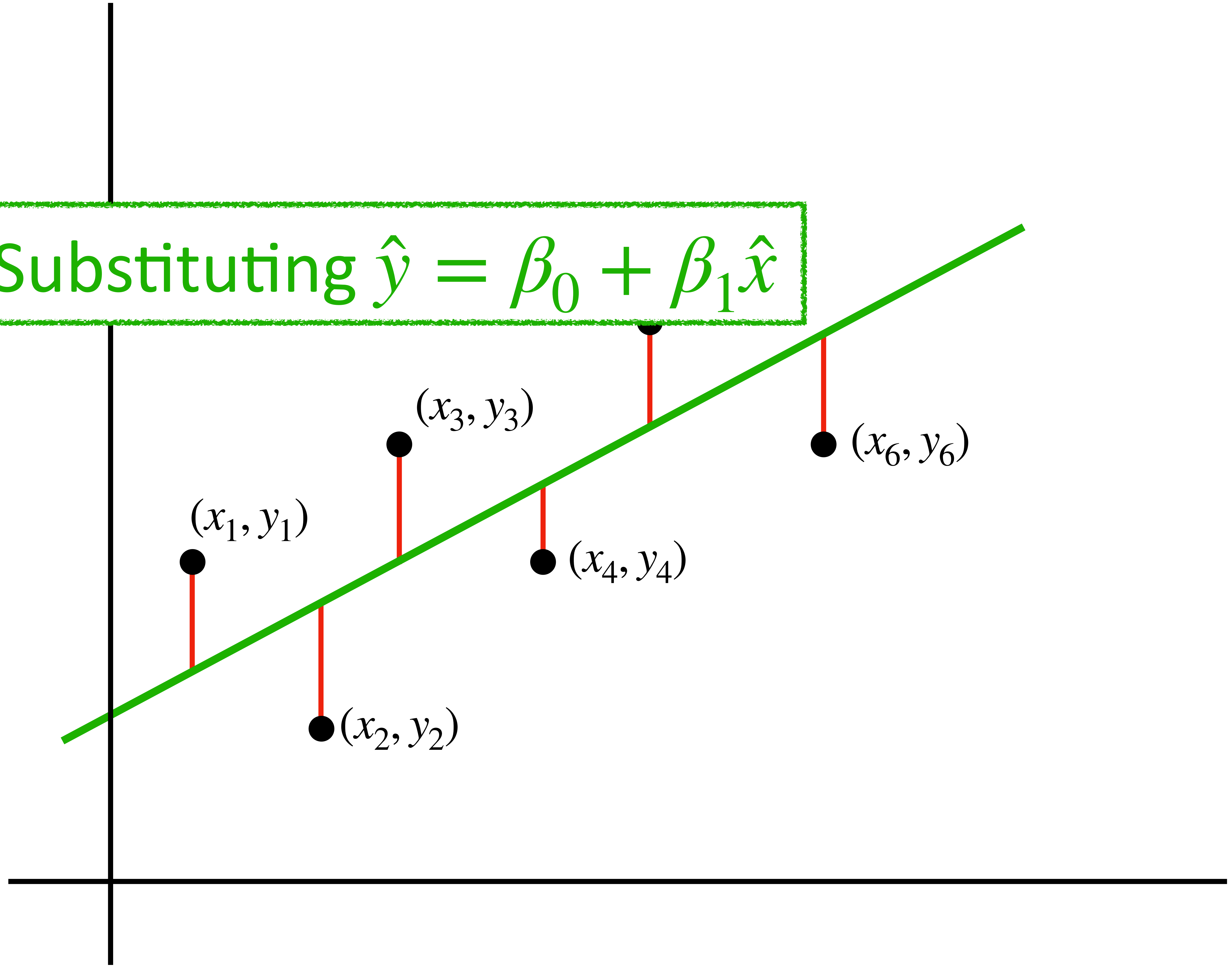
Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1\hat{x}_i)^2$$

Substituting $\hat{y} = \beta_0 + \beta_1 \hat{x}$



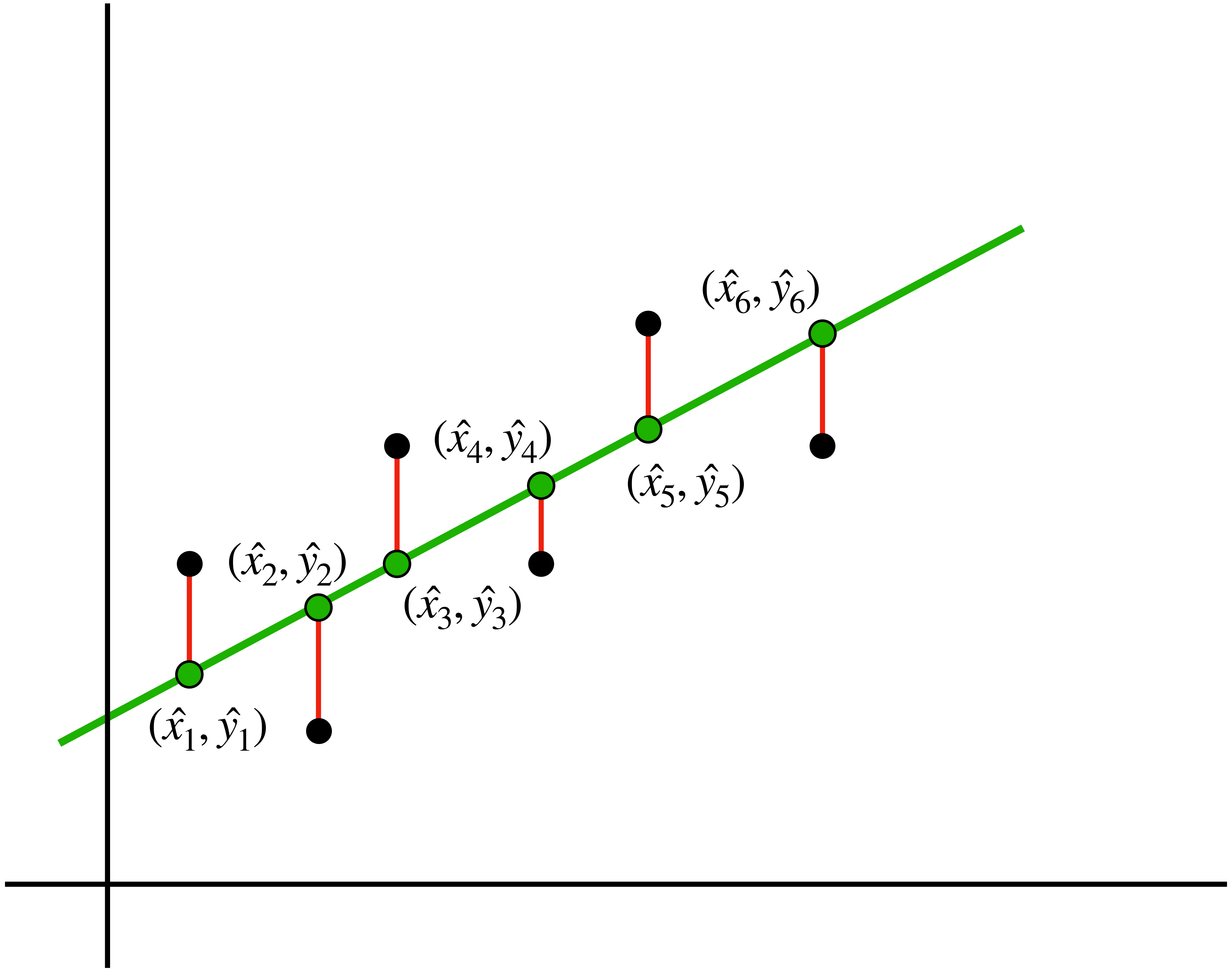
Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 \hat{x}_i)^2$$



Problem Statement

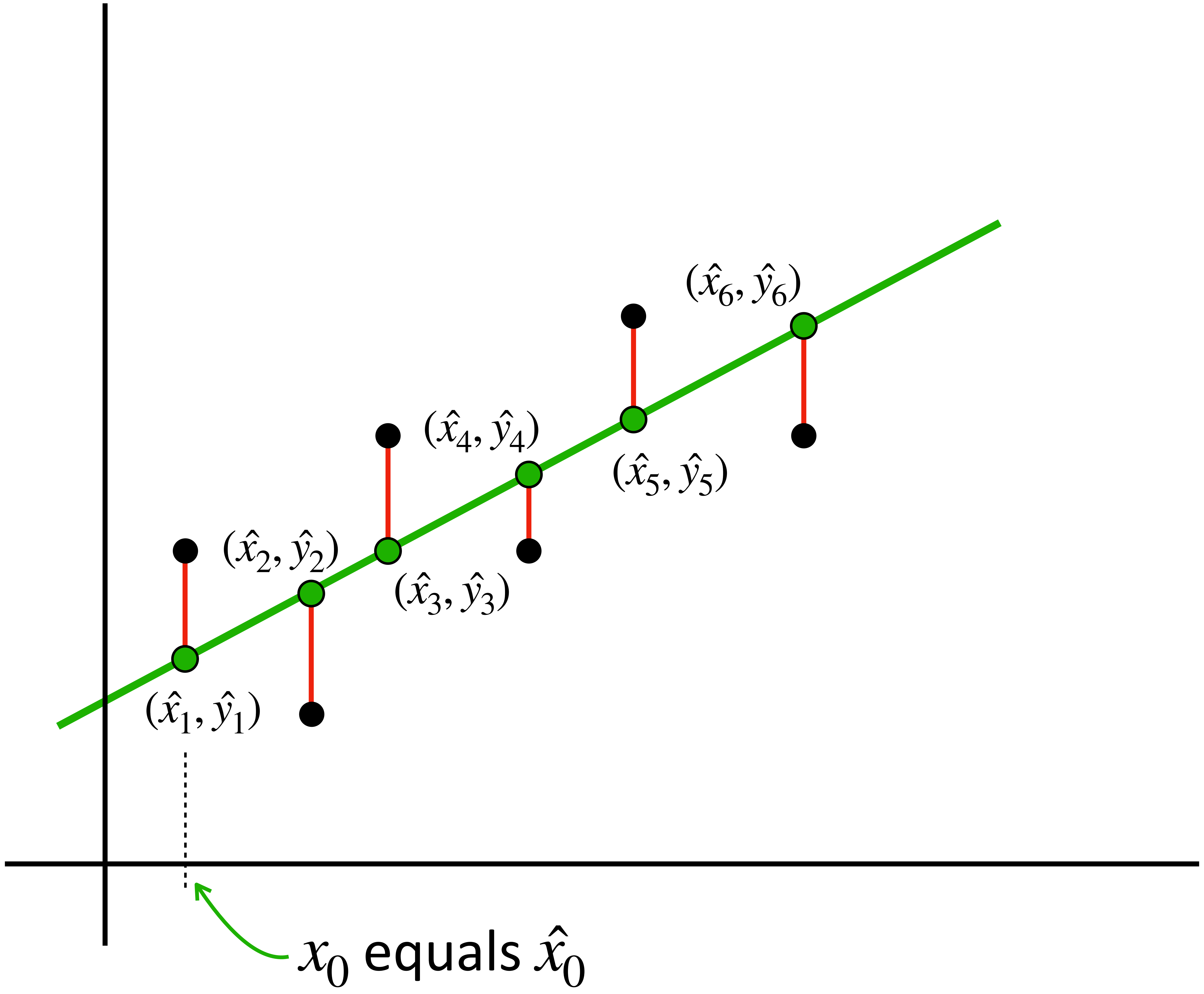
Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 \hat{x}_i)^2$$

For every value of i, $\hat{x}_i$ equals x_i
Substitute x_i for $\hat{x}_i$



Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

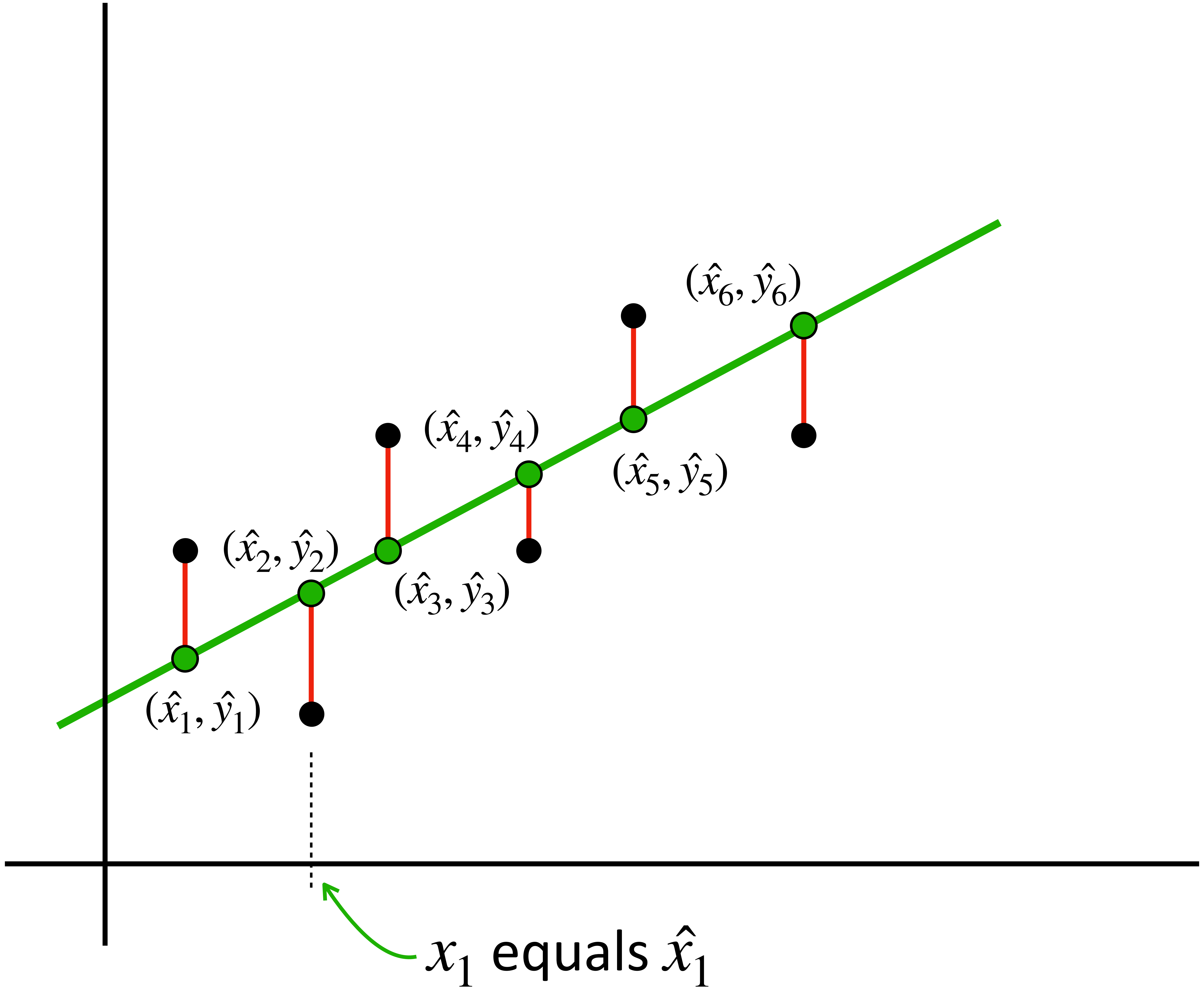
The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 \hat{x}_i)^2$$

For every value of i, $\hat{x}_i$ equals x_i

Substitute x_i for $\hat{x}_i$



Problem Statement

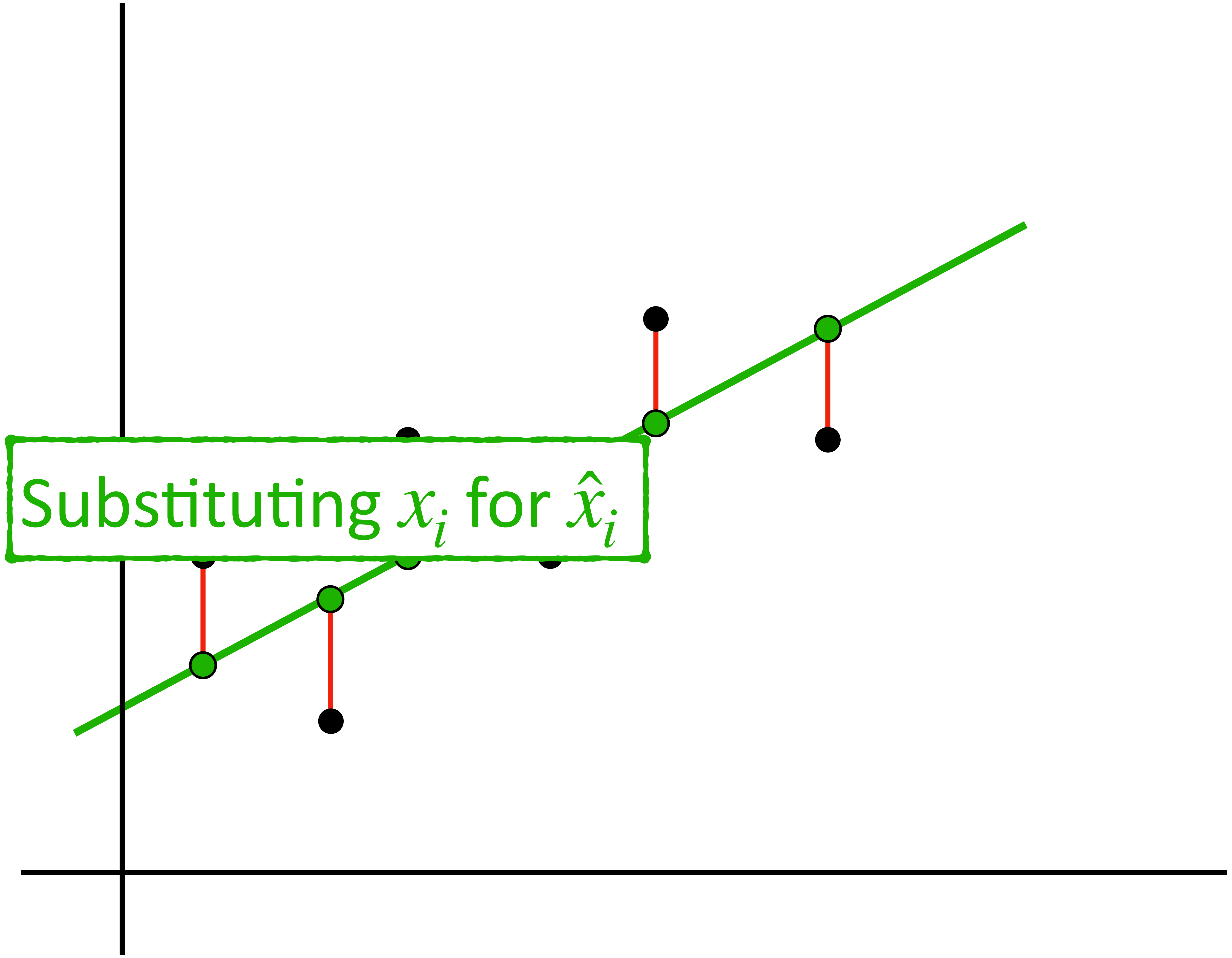
Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data

The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 \hat{x}_i)^2$$

$$= \frac{1}{n}\sum_{i=0}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

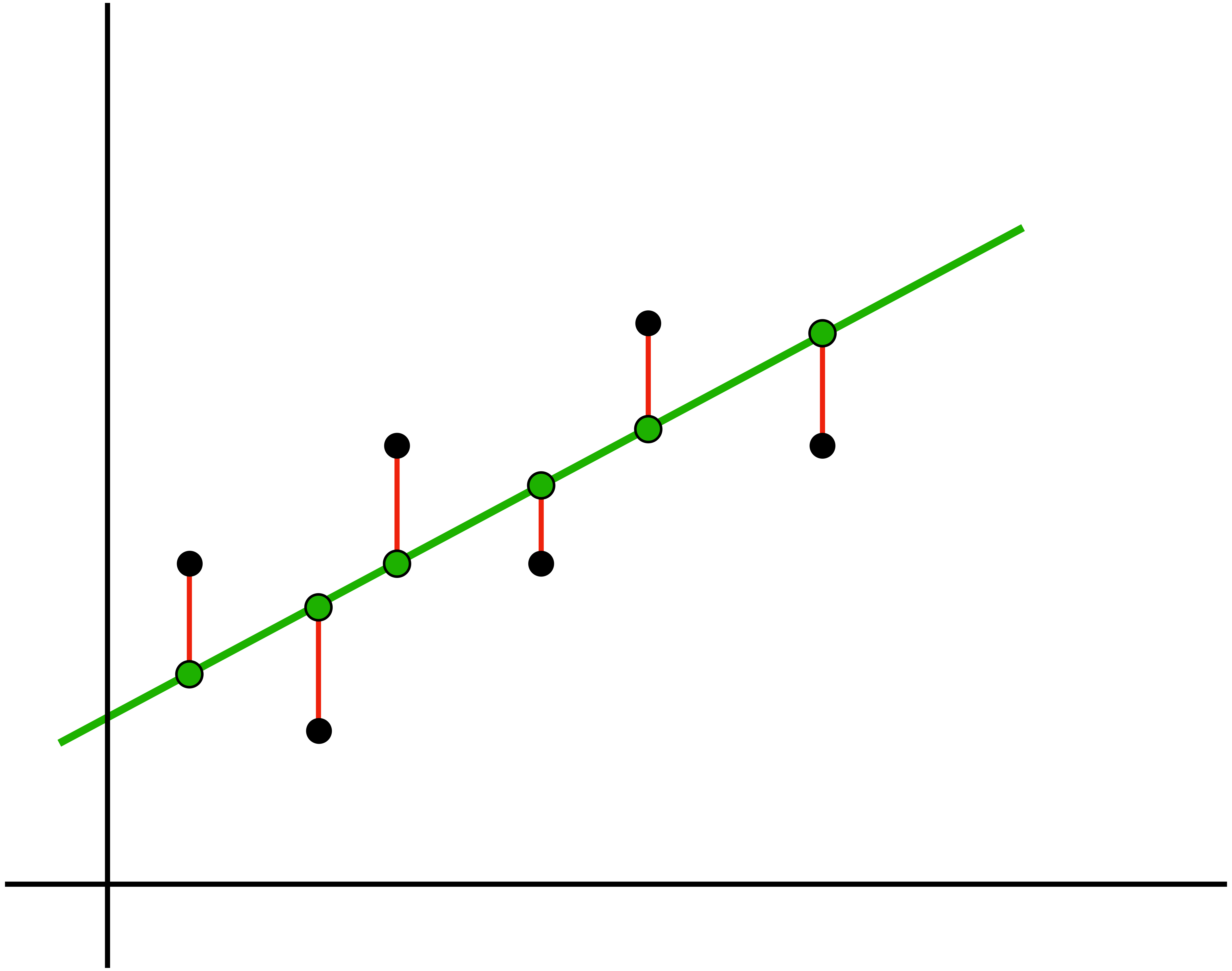


For every value of i, $\hat{x}_i$ equals x_i

Substitute x_i for $\hat{x}_i$

Problem Statement

Given a set of data points in $\mathbf{R}^2$, $(x_1, y_1), (x_2, y_2), (x_3, y_3), \ldots (x_n, y_n)$, find the line that best fits the data



The line of best fit is $\hat{y} = \beta_0 + \beta_1 \hat{x}$

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 \hat{x}_i)^2$$
$$= \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Least Squares Regression: Find the values of β_0 and β_1 such that the Mean Squared Error (MSE) is minimized.

Least Squares Regression: Find the values of β_0 and β_1 such that the Mean Squared Error (MSE) is minimized.

Solution:

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$$

$$\beta_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

Lets walk through the proof…

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

To derive the values of β_0 and β_1, we calculate the partial derivative of the Mean Squared Error (MSE) w.r.t β_0 and β_1 and solve the two equations

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

To derive the values of β_0 and β_1, we calculate the partial derivative of the Mean Squared Error (MSE) w.r.t β_0 and β_1 and solve the two equations

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Partial derivatives w.r.t β_0 and β_1 we get two equations:

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

To derive the values of β_0 and β_1, we calculate the partial derivative of the Mean Squared Error (MSE) w.r.t β_0 and β_1 and solve the two equations

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Partial derivatives w.r.t β_0 and β_1 we get two equations:

$$\frac{\partial}{\partial\beta_0}\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0 \quad \cdots\cdots\cdots\cdots\cdots \; eq(1)$$

$$\frac{\partial}{\partial\beta_1}\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0 \quad \cdots\cdots\cdots\cdots\cdots \; eq(2)$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

To derive the values of β_0 and β_1, we calculate the partial derivative of the Mean Squared Error (MSE) w.r.t β_0 and β_1 and solve the two equations

Mean Squared Error (MSE):

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y})^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

Partial derivatives w.r.t β_0 and β_1 we get two equations:

$$\frac{\partial}{\partial\beta_0}\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0 \quad \cdots\cdots\cdots\cdots\cdots \quad eq(1)$$

$$\frac{\partial}{\partial\beta_1}\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0 \quad \cdots\cdots\cdots\cdots\cdots \quad eq(2)$$

The Mean Squared Error (MSE) is minimized when the partial derivative is zero

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

Chain Rule & Power Rule

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0}\frac{1}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)\frac{\partial}{\partial \beta_0}\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0$$

$\boxed{\text{Chain Rule \& Power Rule}}$

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

Taking the partial derivative of $\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)$

$$\frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = (-1)$$

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Divide both sides by $-\frac{2}{n}$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} x_i = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 1 (take the partial derivative w.r.t β_0):

$$\frac{\partial}{\partial \beta_0} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_0} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} x_i = 0 \qquad \Rightarrow \beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\frac{\partial}{\partial \beta_1} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\frac{\partial}{\partial \beta_1} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

Chain Rule & Power Rule

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\frac{\partial}{\partial \beta_1} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0$$

Taking the partial derivative of $\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)$

$$\frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = (-x_i)$$

See Tutorial on Derivatives

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\frac{\partial}{\partial \beta_1} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(x_i) = 0$$

Divide both sides by $-\frac{2}{n}$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\frac{\partial}{\partial \beta_1} \frac{1}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) \frac{\partial}{\partial \beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i) = 0$$

$$\Rightarrow \frac{2}{n} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(-x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)(x_i) = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

Substitute $\beta_0 = \dfrac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_1 \sum_{i=1}^{n} x_i^2 - \left(\frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} \right) \sum_{i=1}^{n} x_i = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_1 \sum_{i=1}^{n} x_i^2 - \left(\frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} \right) \sum_{i=1}^{n} x_i = 0$$

Divide both sides by n and simplify

$$\Rightarrow n \sum_{i=0}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 \right) = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_0 \sum_{i=1}^{n} x_i - \beta_1 \sum_{i=1}^{n} x_i^2 = 0$$

$$\Rightarrow \sum_{i=1}^{n} x_i y_i - \beta_1 \sum_{i=1}^{n} x_i^2 - \left(\frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n} \right) \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow n \sum_{i=0}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 \right) = 0$$

$$\Rightarrow n \sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow n \sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = 0$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow n\sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = 0$$

$$\Rightarrow \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 - \beta_1 n \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i y_i$$

Add $\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i y_i$ to both sides

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow n\sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = 0$$

$$\Rightarrow \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 - \beta_1 n \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n \sum_{i=1}^{n} x_i y_i$$

$\boxed{\text{Add } \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n \sum_{i=1}^{n} x_i y_i \text{ to both sides}}$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow n\sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = 0$$

$$\Rightarrow \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 - \beta_1 n \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i y_i$$

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

$\boxed{\text{Multiply both sides by } -1}$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow n\sum_{i=1}^{n} x_i y_i - \beta_1 n \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = 0$$

$$\Rightarrow \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 - \beta_1 n \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - n\sum_{i=1}^{n} x_i y_i$$

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i\right)^2 = n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

Multiply both sides by -1

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

Factor out β_1

$$\Rightarrow \beta_1 \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right) = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

$$\Rightarrow \beta_1 \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right) = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

$$\Rightarrow \beta_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

Divide both sides by $n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2$

Least Squares Regression: Find the values of β_0 and β_1 such that the MSE is minimized.

Solving equation 2 (take the partial derivative w.r.t β_1:

$$\Rightarrow \beta_1 n \sum_{i=1}^{n} x_i^2 - \beta_1 \left(\sum_{i=1}^{n} x_i \right)^2 = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

$$\Rightarrow \beta_1 \left(n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \right) = n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i$$

$$\Rightarrow \beta_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

$\boxed{\text{Q.E.D}}$

Least Squares Regression: Find the values of β_0 and β_1 such that the Mean Squared Error (MSE) is minimized.

Solution:

$$\beta_0 = \frac{\sum_{i=1}^{n} y_i - \beta_1 \sum_{i=1}^{n} x_i}{n}$$

$$\beta_1 = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$$

Related Tutorials & Textbooks

Simple Linear Regression

A statistical technique of making predictions from data. The tutorial introduces a linear model in two dimensions and uses that model to predict the value of one dependent variable given one independent variable.

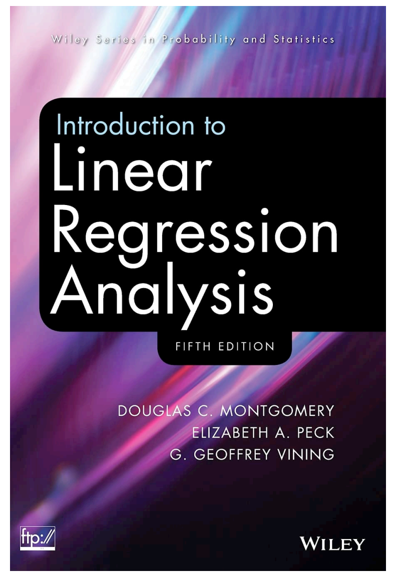
Multiple Regression

Multiple regression extends the two dimensional linear model introduced in Simple Linear Regression to k + 1 dimensions with one dependent variable, k independent variables and k+1 parameters.

Gradient Descent for Simple Linear Regression

An introduction to the Gradient Descent algorithm and a deep dive on how it can be used to optimize the two parameters β_0 and β_1 for Simple Linear Regression.

Recommended Textbooks



Introduction to Linear Regression Analysis

by Douglas C. Montgomery, Elizabeth A. Peck, G. Geoffrey Vining

For a complete list of tutorials see:
https://arrsingh.com/ai-tutorials