

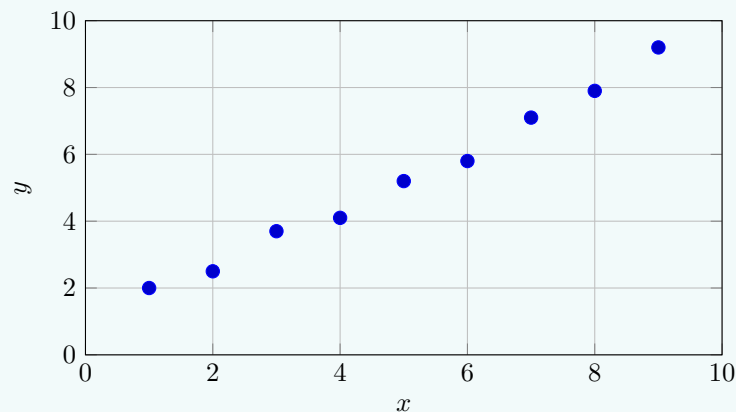
# STAT 151A Lecture 4

Henry Liev

5 September 2025

## Remark 0.1 (Graphical interpretation of Simple Regression)

Scatter Plot of  $y$  vs  $x$



Line of best fit is the regression line

## Remark 0.2 (Optimization approach to Simple Regression)

$$\min_c \sum_{i=1}^n (y_i - c)^2$$

$$\min_{a,b} \sum_{i=1}^n (y_i - a - bx_i)^2$$

$$\frac{\partial}{\partial a} \sum_{i=1}^n (y_i - a - bx_i)^2 = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0$$

$$\frac{\partial}{\partial b} \sum_{i=1}^n (y_i - a - bx_i)^2 = -2 \sum_{i=1}^n x_i (y_i - a - bx_i) = 0$$

Above equations are known as the normal equations

$$\hat{b} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = r_{xy} \frac{\sigma_y}{\sigma_x}$$

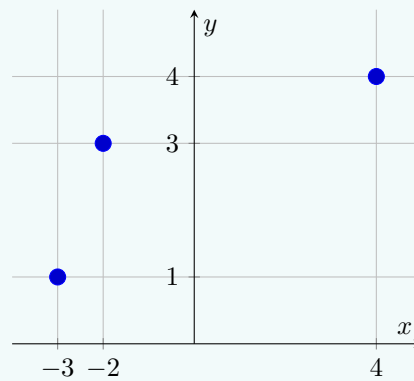
$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

---

**Remark 0.3** (Probabilistic approach to Simple Regression) $(x_1, y_1), \dots$ Model:  $y_i = \alpha + \beta x_i + \varepsilon_i$ ,  $\mathbb{E}(\varepsilon_i) = 0$ ,  $\text{Var}(\varepsilon_i) = \sigma^2$  $\mathbb{E}(y_i | x_i) = \alpha + \beta x_i$  $y_i, x_i, \varepsilon_i$  are random variables,  $\alpha, \beta, \sigma$  are parameters**Remark 0.4** (Linear Algebra approach to Simple Regression)

$y$	$x$
3	-2
1	-3
4	4

Scatter Plot of Table Data

**Remark 0.5** (Linear Algebra Review)

Vector: Length and a direction, can find angles between vectors

1. They can be scaled,  $2\vec{v}$  has double length
2. They can be added,  $\vec{v} + \vec{u}$

Can represent vectors in Cartesian coordinates  $\vec{v} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ ,  $\vec{u} = \begin{pmatrix} -3 \\ 0 \end{pmatrix}$ If you write them this way,  $2\vec{v} = \begin{pmatrix} 2v_1 \\ 2v_2 \end{pmatrix}$ ,  $\vec{v} + \vec{u} = \begin{pmatrix} v_1 + u_1 \\ v_2 + u_2 \end{pmatrix}$

---

**Remark 0.6** (Geometric concepts)

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}, \text{length}(\vec{v}) = \sqrt{a^2 + b^2}$$

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{pmatrix}, \|\vec{v}\| = \text{length}(\vec{v}) = \sqrt{v_1^2 + v_2^2 + \cdots + v_p^2} = \sqrt{\vec{v} \cdot \vec{v}}$$

$$\vec{u} \cdot \vec{v} = u_1 \cdot v_1 + u_2 \cdot v_2 + \cdots + u_p \cdot v_p$$

Distance between two vectors?  $\|\vec{u} - \vec{v}\|$

Alternate definition of dot product:  $\vec{u} \cdot \vec{v} = \|\vec{u}\| \cdot \|\vec{v}\| \cos \theta$

Orthogonality

$$\vec{u} \perp \vec{v} \Leftrightarrow \vec{u} \cdot \vec{v} = 0$$

If  $\vec{u} \perp \vec{v}$ :  $\|\vec{u} + \vec{v}\|^2 = (\vec{u} + \vec{v}) \cdot (\vec{u} + \vec{v}) = \vec{u} \cdot \vec{u} + \vec{v} \cdot \vec{u} + \vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v} = \|\vec{u}\|^2 + \|\vec{v}\|^2$  by orthogonality

**Definition 0.7** (Vector spaces)

$$\vec{v}_1, \dots, \vec{v}_p \in \mathbb{R}^n$$

Subspace generated by  $\vec{v}_1, \dots, \vec{v}_p$  is the set of all vectors of the form:  $c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_p \vec{v}_p$  where  $c_i \in \mathbb{R}$

Also called  $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$

**Definition 0.8** (Vector Space Basis)

We call  $\{\vec{v}_1, \dots, \vec{v}_p\}$  if  $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} = V$