Interrelation of Slope, Covariance, and Correlation

Dr. Ratnesh Prasad Srivastava
GGV, Bilaspur

Introduction

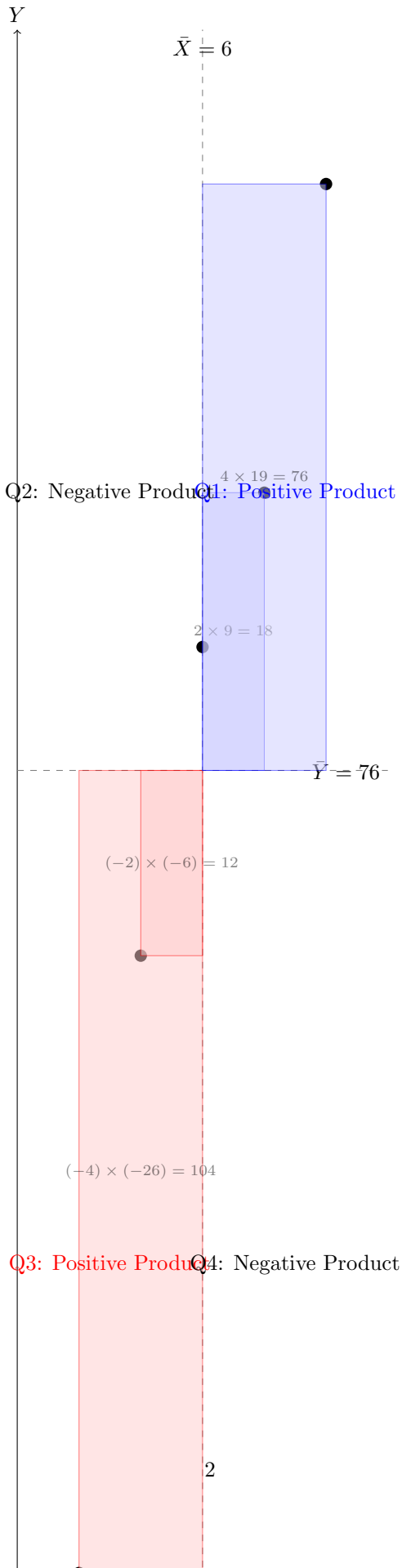
This document explains the interrelation between **slope** (linear regression), **covariance**, and **correlation**. We'll use a practical example with visual intuition for covariance inspired by StatQuest's rectangle method.

Example Dataset: Study Hours vs. Exam Scores

Student	1	2	3	4	5
Hours (X)	2	4	6	8	10
Score (Y)	50	70	80	85	95

Step 1: Visualizing Covariance Intuition (Rectangle Method)

Covariance measures how deviations from the mean in X and Y **co-vary**. Visualize it using rectangles:


Y
$\bar{X} = 6$
Q2: Negative Product
Q1: Positive Product
$4 \times 19 = 76$
$2 \times 9 = 18$
$\bar{Y} = 76$
$(-2) \times (-6) = 12$
$(-4) \times (-26) = 104$
Q3: Positive Product
Q4: Negative Product
2

Covariance Intuition from Rectangles:

- Each rectangle's area = $(X_i - \bar{X}) \times (Y_i - \bar{Y})$
- Red rectangles (Q3): Both negative = Positive product
- Blue rectangles (Q1): Both positive = Positive product
- Quadrants 2 & 4 give negative products (not present here)
- **Covariance** = (Sum of all rectangle areas) / (n-1)
- **Key Insight**: Covariance sign indicates relationship direction

Step 2: Calculate Deviations and Products

Student	1	2	3	4	5	Total
X_i	2	4	6	8	10	
Y_i	50	70	80	85	95	
$X_i - \bar{X}$	-4	-2	0	2	4	0
$Y_i - \bar{Y}$	-26	-6	4	9	19	0
Product $(X_i - \bar{X})(Y_i - \bar{Y})$	104	12	0	18	76	210

Step 3: Covariance Calculation

$$\text{Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y}) = \frac{104 + 12 + 0 + 18 + 76}{4} = \frac{210}{4} = 52.5$$

Why $n - 1$? Corrects bias when estimating population covariance from a sample (degrees of freedom).

Step 4: Variances Calculation

$$\text{Var}(X) = \frac{1}{n-1} \sum (X_i - \bar{X})^2 = \frac{(-4)^2 + (-2)^2 + 0^2 + 2^2 + 4^2}{4} = \frac{16 + 4 + 0 + 4 + 16}{4} = 10$$

$$\text{Var}(Y) = \frac{1}{n-1} \sum (Y_i - \bar{Y})^2 = \frac{(-26)^2 + (-6)^2 + 4^2 + 9^2 + 19^2}{4} = \frac{676 + 36 + 16 + 81 + 361}{4} = \frac{1170}{4} = 292.5$$

Step 5: Correlation Calculation

$$r = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} = \frac{52.5}{\sqrt{10 \times 292.5}} = \frac{52.5}{\sqrt{2925}} \approx \frac{52.5}{54.08} \approx 0.971$$

Correlation Intuition:

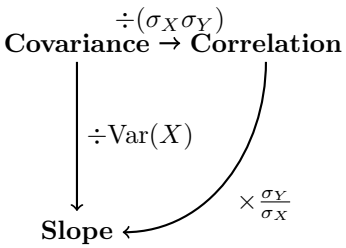
- Standardizes covariance to [-1, 1] range
- Unitless measure of linear relationship strength
- Near +1 indicates strong positive relationship

Step 6: Slope Calculation

$$b_1 = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} = \frac{52.5}{10} = 5.25$$

Slope Interpretation: For every additional hour studied, exam scores increase by 5.25 points on average.

The Fundamental Relationship



$$\boxed{b_1 = \frac{\text{Cov}(X,Y)}{\text{Var}(X)} = r \cdot \frac{\sigma_Y}{\sigma_X} = 0.971 \times \sqrt{\frac{292.5}{10}} \approx 0.971 \times 5.41 \approx 5.25}$$

Teaching Summary

Covariance	Correlation	Slope
• Direction of co-movement • Rectangle areas • Units: X*Y units • Range: $(-\infty, \infty)$	• Standardized covariance • Strength of relationship • Unitless • Range: [-1, 1]	• Practical effect size • $\Delta Y / \Delta X$ • Units: Y/X • Predictive relationship