# Primer in Matrix Algebra

Giulio Rossetti*

giuliorossetti94.github.io

January 16, 2025

* email: giulio.rossetti.1@wbs.ac.uk

# What is a Matrix?

- A matrix is a rectangular set of numbers (i.e., range in Excel).
- Example:

$$A = \underbrace{\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}}_{3\times 2}$$

- Vectors are special cases of matrices with only one column:

$$x = \underbrace{\begin{bmatrix} a \\ c \\ e \end{bmatrix}}_{3\times 1}$$

# Matrix Addition

- Add matrices by adding corresponding elements:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad A + B = \begin{bmatrix} a+e & b+f \\ c+g & d+h \end{bmatrix}$$

- Matrices must have the same shape to add.

# Matrix Multiplication

- Multiply matrices by combining rows of the left matrix with columns of the right matrix.
- Example:

A × B

$\boxed{K \times q \quad q \times P}$

K × Q  Q × P

$$A = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}, \quad B = \begin{bmatrix} g & h \\ i & j \\ k & l \end{bmatrix}$$

2 × 3  3 × 2

$$AB = \begin{bmatrix} ag + bi + ck & ah + bj + cl \\ dg + ei + fk & dh + ej + fl \end{bmatrix}$$

$\boxed{2 \times 2}$

- Matrices must have compatible dimensions.

# Properties of Matrix Multiplication

- Distributive property: $(A + B)C = AC + BC$
- Non-commutative: $AB \neq BA$
- Identity matrix: $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ satisfies $AI = IA = A$.

# Transpose and Symmetric Matrices

- Transpose: Swap rows and columns of a matrix.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad A' = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

$A^T$

- Symmetric matrix: $A = A'$.

$$A = \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}$$

# Matrix Operations

- Element-wise multiplication and division:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad \cancel{A}\,B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

$$A. * B = \begin{bmatrix} ae & bf \\ cg & dh \end{bmatrix}, \quad A./B = \begin{bmatrix} \frac{a}{e} & \frac{b}{f} \\ \frac{c}{g} & \frac{d}{h} \end{bmatrix}$$

# Inner, Outer, and Quadratic Forms

- Inner product: A row vector times a column vector gives a scalar.

$$\underset{1\times 1}{x'y} = \underset{1\times 3}{\begin{bmatrix} a & b & c \end{bmatrix}} \underset{3\times 1}{\begin{bmatrix} d \\ e \\ f \end{bmatrix}} = ad + cf + be$$

scalar
i.e., a number

- Outer product: A column vector times a row vector gives a matrix.

$$\underset{3\times 3}{xy'} = \underset{3\times 1}{\begin{bmatrix} a \\ b \\ c \end{bmatrix}} \underset{1\times 3}{\begin{bmatrix} d & e & f \end{bmatrix}} = \begin{bmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{bmatrix}$$

# Inner, Outer, and Quadratic Forms

*Cont'd*

- Quadratic form: Combines a vector, a symmetric matrix, and another vector.

$$x'Ax = \begin{bmatrix} e & f \end{bmatrix} \begin{bmatrix} a & b \\ b & d \end{bmatrix} \begin{bmatrix} e \\ f \end{bmatrix} = ae^2 + df^2 + 2bef$$

# Matrix Inversion

INVERTING A 2x2 MATRIX

- Inverse of $A$: $A^{-1}$ satisfies $AA^{-1} = A^{-1}A = I$.
- Conditions for inversion:
  - $A$ must be square.
  - $A$ must have full rank.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

① swap elements on main diagonal.
② multiply by -1 the off diagonal matrix
③ multiply by $1/\det(A)$

- Example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\underbrace{ad - bc}_{\text{determinant of } A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

# Applications of Matrices

## *Linear Regression*

Roadmap:

- Linear regression model in matrix form
- Minimization problem
- Derivation of OLS estimate

# Linear Regression (Cont'd)

- The linear regression model is:

$$\boxed{Y = X\beta + \epsilon}$$

- Expanded notation:

$$\overset{N\times 1}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}} = \overset{N\times 2}{\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{N1} & x_{N2} \end{bmatrix}} \overset{2\times 1}{\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}} + \overset{N\times 1}{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_N \end{bmatrix}}$$

$$e'\varepsilon = \varepsilon_1^2 + \varepsilon_2^2 + \dots + \dots = \sum_{i=1}^{n} \varepsilon_i^2$$

in SLR this is a column of 1s

- If there is a constant in the regression, the first column of $X$ is all 1s.
- OLS estimate:

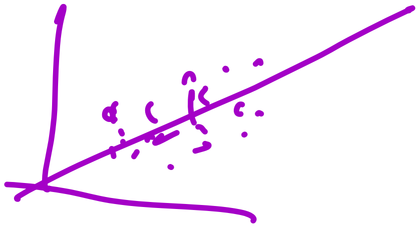
$$\boxed{\hat{\beta} = (X'X)^{-1}X'Y}$$

# Linear Regression (Cont'd)

- Where does $\hat{\beta} = (X'X)^{-1}X'Y$ come from?
- Sum of squared residuals:

$$RSS = \sum_{i}^{N} \epsilon_i^2 = \epsilon'\epsilon = (Y - X\beta)'(Y - X\beta)$$

- Minimize RSS with respect to $\beta$:



$$\begin{aligned}\hat{\beta} &= \arg\min_{\beta} \epsilon'\epsilon \\ &= \arg\min_{\beta} (Y - X\beta)'(Y - X\beta) \\ &= \arg\min_{\beta} Q(\beta)\end{aligned}$$

loss function

# Linear Regression (Cont'd)

$y = X\beta + \varepsilon$

$\hookrightarrow \varepsilon = y - X\beta$

- rewrite $Q(\beta)$:

$\varepsilon \,'\, \varepsilon$

$$Q(\beta) = (y - X\beta)'(y - X\beta),$$

loss funct.

$$= (y' - \beta' X')(y - X\beta), \quad \text{since } (X\beta)' = \beta' X',$$

$$= y'y - \beta' X' y - y' X\beta + \beta' X' X\beta,$$

$$= \boxed{y'y - 2y' X\beta + \beta' X' X\beta,} \quad \text{since } \beta' X' y = y' X\beta.$$

# Linear Regression (Cont'd)

- differentiate[1] $Q(\beta)$ with respect to $\beta$

$\beta_1$
$\beta_2$
$\vdots$
$\beta_K$

$$\left.\frac{\partial Q(\beta)}{\partial \beta'}\right|_{\beta=\hat{\beta}} = \boxed{-2(X'y)' + 2\hat{\beta}'X'X.}$$

- FOC: set the $k$ vector of partial derivatives to zero:

$$-2(X'y)' + 2\hat{\beta}'X'X = 0'_{K+1}$$

a vector if we have K indep variables + intercept
||
K+1

- rewrite the FOC:

$$X'X\hat{\beta} = X'y$$

$$\hat{\beta} = (X'X)^{-1}X'y$$

[1] $\frac{\partial(a'\beta)}{\partial\beta'} = a'$, and $\frac{\partial(\beta' A\beta)}{\partial\beta'} = 2\beta' A$