Primer in Matrix Algebra

Giulio Rossetti*
giuliorossetti94.github.io

January 23, 2026

* 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:

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

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

$$x = \begin{bmatrix} a \\ c \\ e \end{bmatrix}$$

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:

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

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

- 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.

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

- 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:

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

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

Inner, Outer, and Quadratic Forms

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

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

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

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

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

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

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

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:

$$Y = X\beta + \epsilon$$

- Expanded notation:

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

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

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

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 β:

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

Linear Regression (Cont'd)

- rewrite $Q(\beta)$:

$$
\begin{aligned}
Q(\beta) &= (y - X\beta)'(y - X\beta), \\
&= (y' - \beta' X')(y - X\beta), \quad \text{since } (X\beta)' = \beta' X', \\
&= y'y - \beta' X' y - y' X\beta + \beta' X' X\beta, \\
&= y'y - 2y' X\beta + \beta' X' X\beta, \quad \text{since } \beta' X' y = y' X\beta.
\end{aligned}
$$

Linear Regression (Cont'd)

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

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

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

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

- 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 $\frac{\partial(\beta' A\beta)}{\partial\beta'} = 2\beta' A$

Roadmap

Appendix

Appendix A: Matrix Products as Sums

- Let X be an $N \times K$ matrix with rows x_i'.
- Then:

$$X'X = \sum_{i=1}^{N} x_i x_i'$$

- Each observation contributes a $K \times K$ outer product.
- Econometric interpretation:
  - $X'X$ collects information about regressor variation
  - $\frac{1}{N} X'X \to \mathbb{E}[x_i x_i']$

Appendix B: The Meaning of $X'y$

- Let y be an $N \times 1$ vector.
- Then:

$$X'y = \sum_{i=1}^{N} x_i y_i$$

- This is a vector of sample covariances.
- OLS estimator:

$$\hat{\beta} = \left(\sum x_i x_i'\right)^{-1} \sum x_i y_i$$

- OLS is a ratio of sample moments.

Appendix C: OLS as Orthogonality

- OLS first-order condition:

$$X'(y - X\hat{\beta}) = 0$$

- Equivalent to:

$$\sum_{i=1}^{N} x_i \hat{\varepsilon}_i = 0$$

- Interpretation:
  - Residuals are orthogonal to regressors
  - Finite-sample analogue of $\mathbb{E}[x_i \varepsilon_i] = 0$

Appendix D: Numerical Example — Computing $X'X$

- Suppose:

$$X = \begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 1 & 5 \end{bmatrix}$$

- Rows:

$$x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

- Then:

$$X'X = \sum_{i=1}^{3} x_i x_i'$$

Appendix E: Try It Yourself

- Compute each outer product:

$$x_1 x_1' = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad x_2 x_2' = \begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix}, \quad x_3 x_3' = \begin{bmatrix} 1 & 5 \\ 5 & 25 \end{bmatrix}$$

- Sum them to obtain $X'X$.

Appendix F: Numerical Example — OLS by Hand

- Suppose:

$$y = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}$$

- Compute:

$$X'y = \sum x_i y_i$$

- Then compute:

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

- Compare with Excel or Python.

Appendix G: Quadratic Forms and RSS

- Residual sum of squares:

$$RSS = (y - X\beta)'(y - X\beta)$$

- This is a quadratic form.
- Expanded:

$$RSS = \sum_{i=1}^{N} \varepsilon_i^2$$

- Matrix notation is compact — the object is still a sum.

Appendix H: Key Takeaways

- Matrix products in econometrics are sums of observations.
- OLS is a ratio of sample moments.
- Regression = projection.
- Inference depends on second moments of regressors.
- If you can expand the matrix, you understand the estimator.

Appendix J: The Trace Operator

- The trace of a square matrix is the sum of its diagonal elements:

$$\mathrm{tr}(A) = \sum_i a_{ii}$$

- Key properties:
  - $\mathrm{tr}(A) = \mathrm{tr}(A')$
  - $\mathrm{tr}(AB) = \mathrm{tr}(BA)$
- Trace allows us to rewrite quadratic forms:

$$x'Ax = \mathrm{tr}(Axx')$$

Appendix K: Why Derivatives Are Row Vectors

- Let $f(x)$ be a scalar and x a vector.
- The derivative is defined as:

$$\frac{\partial f(x)}{\partial x'} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_K} \end{bmatrix}$$

- Gradients are row vectors by definition.
- This ensures dimensional consistency in optimization.