# Math for the Social Sciences Module - Young Researchers Fellowship

## Lecture 8 - A Gentle Review of Matrix Algebra

Daniel Sánchez Pazmiño

Laboratorio de Investigación para el Desarrollo del Ecuador

2024

# Why matrices?

- Matrices are a way to organize data in a structured way.

- When dealing with complex mathematical models or operations, using "normal" algebra can be cumbersome.

- For instance, consider a four-variable system of equations:

$$2x + 3y + 4z + 5w = 10$$
$$3x + 4y + 5z + 6w = 20$$
$$4x + 5y + 6z + 7w = 30$$
$$5x + 6y + 7z + 8w = 40$$

# Why matrices?

- If we try to solve this system using algebra, we would have to do a lot of calculations.

- For instance, we may try to solve the first equation for $x$:

$$2x = 10 - 3y - 4z - 5w$$
$$x = \frac{10 - 3y - 4z - 5w}{2}$$

- And then substitute this expression into the second equation:

$$3\left(\frac{10 - 3y - 4z - 5w}{2}\right) + 4y + 5z + 6w = 20$$

- Which would then need to be simplified and solved for $y$:

$$\frac{30 - 9y - 12z - 15w}{2} + 4y + 5z + 6w = 20$$

- And so on...

# Why matrices?

- **Matrices** are rectangular arrays of numbers considered as a single mathematical entity.

- They can help organize a system of equations or complex mathematical operations in a more structured way.

- Through matrix algebra, we can perform operations on matrices quickly and efficiently.

# Matrix notation

- We represent a matrix with a capital letter, usually in boldface, e.g., $\mathbf{A}$.

  - In handwritten text, matrices are often represented with capital letters in block letters, e.g., $\mathbb{A}$, or using a different color.

- Consider, for instance, the following matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

- We usually either use square brackets or round brackets (parentheses) to represent matrices.

# Elements

- The numbers in the matrix are called **elements**. They are usually represented with lowercase letters, e.g., $a_{ij}$.

- We typically define the *dimension* of a matrix as the number of rows and columns it has. Thus, the matrix $A$ above is a $3 \times 4$ matrix.

- To refer to a specific element in a matrix, we use the notation $a_{ij}$, where $i$ is the row number and $j$ is the column number.

    - For instance, in the matrix $A$ above, $a_{23} = 7$.

## Matrix general form

- A matrix can be represented in general form as follows:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

- Where $m$ is the number of rows and $n$ is the number of columns.

- This matrix is said to be an $m \times n$ matrix.

- It is common to see the dimensions of the matrix written as a subscript, e.g., $A_{m \times n}$. or $\underset{m \times n}{A}$

# Vectors

- A vector is a special case of a matrix that has only one column or one row.

- Vectors which have only one column are called **column vectors**.

- Vectors which have only one row are called **row vectors**.

- For instance, consider the following column vector:

$$v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

# Vectors

- We can also represent vectors as row vectors:

$$v = \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$$

- We usually represent vectors with lowercase boldface letters, e.g., $\mathbf{v}$. In harndwritten text, vectors are often represented with lowercase letters in block letters, e.g., $\mathbb{v}$, or using a different color and lowercase letters.

# The main diagonal

- The **main diagonal** of a matrix is the set of elements that run from the top left to the bottom right of the matrix.

- More formally, the main diagonal of a matrix $A$ is the set of elements $a_{ij}$ where $i = j$.

- For instance, consider the following matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

- The main diagonal of this matrix is $\{1, 5, 9\}$.

- Sometimes we refer to the main diagonal as only the "diagonal" of the matrix.

# Special matrices

- There are several types of special matrices that are commonly used, and you should be familiar with them.

- **Square matrix**: A matrix where the number of rows is equal to the number of columns.

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$ is a square matrix.

# Special matrices

- **Diagonal matrix**: A square matrix where all elements off the main diagonal are equal to 0.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \text{ is a diagonal matrix.}$$

- **Note**: Sometimes, everything that is not in the main diagonal is referred to as the "off-diagonals".

# Special matrices

- **Identity matrix**: A square matrix where all elements on the main diagonal are equal to 1 and all other elements are equal to 0.

- This matrix is usually represented by the symbol $I_n$, where $n$ is the dimension of the matrix.

  - For instance, $I_3$ is a $3 \times 3$ identity matrix, as it always a square matrix.

- The identity matrix is analogous to the number 1 in normal algebra.

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

# Special matrices

- **Zero matrix**: A matrix where all elements are equal to 0.

- This matrix is usually represented by the symbol $O_{m \times n}$, where $m$ is the number of rows and $n$ is the number of columns.

## Equality of matrices

- Two matrices are equal if they have the same dimensions and if their corresponding elements are equal.

- Formally, two matrices $A$ and $B$ are equal if $a_{ij} = b_{ij}$ for all $i$ and $j$.

- For instance, consider the following matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

## Matrix operations

- Matrices truly shine when we start performing operations on them.

- Storing large amounts of information will allow large-scale operations to be performed efficiently in the computational sense, as simple algebra can occupy too much computational power.

- In this lecture, we will review the following basic operations:
  - Scalar addition and subtraction
  - Matrix addition and subtraction
  - Scalar multiplication
  - Matrix multiplication
  - Transposition

# Scalars: what are they?

- We've already grown to know and love scalars in our day-to-day algebraic operations.

- These are the numbers that we use to multiply or add to other numbers, such as $2$, $3$, $4$, etc.

- More formally, a scalar is a single number in a field, which is a set of numbers that can be added, subtracted, multiplied, and divided.

    - Unlike vectors and matrices, scalars are not arrays of numbers.

## Scalar addition and subtraction

- We can try to add or subtract a scalar to a matrix. This is done by adding or subtracting the scalar to each element of the matrix.

- For instance, consider the following matrix:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

- If we add the scalar $2$ to this matrix, we get:

$$A + 2 = \begin{bmatrix} 1+2 & 2+2 & 3+2 \\ 4+2 & 5+2 & 6+2 \\ 7+2 & 8+2 & 9+2 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 5 \\ 6 & 7 & 8 \\ 9 & 10 & 11 \end{bmatrix}$$

- Substraction works in the same way.

# Scalar multiplication

- Scalar multiplication is the operation of multiplying a matrix by a scalar.
- This is done by multiplying each element of the matrix by the scalar.
- If we have the matrix $A$ from before and we multiply it by any scalar $k$, we get:

$$kA = \begin{bmatrix} k \cdot 1 & k \cdot 2 & k \cdot 3 \\ k \cdot 4 & k \cdot 5 & k \cdot 6 \\ k \cdot 7 & k \cdot 8 & k \cdot 9 \end{bmatrix}$$

# Matrix addition and subtraction

- We can also add or subtract two matrices, provided that these **have the same dimensions**.
  - When two matrices have the same dimensions, we say that they are of the same order.
- To add or subtract two matrices, we add or subtract the corresponding elements of the matrices, being careful to maintain the same position of each element.

# Example: matrix addition

■ For instance, we have matrices $A$ and $B$. We can compute
$C = A + B$, by defining each element of $C$ as $c_{ij} = a_{ij} + b_{ij}$. See
below:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad B = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix}$$

$$C = \begin{bmatrix} 1+9 & 2+8 & 3+7 \\ 4+6 & 5+5 & 6+4 \\ 7+3 & 8+2 & 9+1 \end{bmatrix} = \begin{bmatrix} 10 & 10 & 10 \\ 10 & 10 & 10 \\ 10 & 10 & 10 \end{bmatrix}$$

## Matrix addition and subtraction

- Substraction works in the same way, but with the operation $c_{ij} = a_{ij} - b_{ij}$ if we were to compute $C = A - B$.
- Notice that if the matrices do not have the same dimensions, we cannot add or subtract them.
- Key difference between matrix addition and scalar addition: matrices must have the same dimensions, and thus matrix addition and subtraction may not always be possible for all pairs of matrices.
- Notice that because vectors are just special cases of matrices, the same rules apply to them!

## Properties of matrix addition

- Matrix addition has several properties that are similar to addition of two scalars (numbers).
- **Commutative property**: $A + B = B + A$
- **Associative property**: $(A + B) + C = A + (B + C)$
- **Identity property**: $A + O = A$
- **Inverse property**: $A + (-A) = O$
- **Distributive property**: $k(A + B) = kA + kB$, where $k$ is a scalar.
- **Distribution of scalar over matrix addition**: $(k + l)A = kA + lA$, where $k$ and $l$ are scalars.

# Why do we care about these properties?

- Why even care about these? It turns out that in many social science applications, we will need to work with matrices that we don't really observe, but we can infer from the data we have. i.e. they are *variable* matrices.

- For instance, the matrix $X$ usually represents the matrix of all observed explanatory variables in a regression model. The vector $y$ represents the explained variable.

- You will need to learn the properties above (and some others) by heart to follow along with the math in some statistical models, since there is no easy way to work with them without knowing these properties, as there are just too many elements to work with.

# The big one: matrix multiplication

- Matrix multiplication is a bit more complex than the previous operations.

- First of all, it requires the condition of **conformability**. No two matrices can be multiplied unless they are conformable with each other.

- Conformable means that the number of columns in the first matrix must be equal to the number of rows in the second matrix. See below:

$$A_{m \times n} \times B_{n \times p} = C_{m \times p}$$

- $A$ is conformable with $B$, since it has $n$ columns and $B$ has $n$ rows.

- The resulting matrix $C$ will have $m$ rows and $p$ columns (the "outer" dimensions of the matrices being multiplied).

# The big one: matrix multiplication

- How do we actually compute the elements of the resulting matrix $C$?

- This is easiest seen with an example. Consider the following matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix}$$

- The resulting elements of C are computed as follows:

$$C = \begin{bmatrix} 1 \cdot 7 + 2 \cdot 9 + 3 \cdot 11 & 1 \cdot 8 + 2 \cdot 10 + 3 \cdot 12 \\ 4 \cdot 7 + 5 \cdot 9 + 6 \cdot 11 & 4 \cdot 8 + 5 \cdot 10 + 6 \cdot 12 \end{bmatrix}$$

# The big one: matrix multiplication

- This is defined as the **dot product** of the rows of the first matrix and the columns of the second matrix. Think of this as the SUMPRODUCT function in Excel, or multiplying the rows of the first matrix by the columns of the second matrix, and summing the results.

- More generally, for any two matrices $A$ and $B$ with dimensions $m \times n$ and $n \times p$, respectively, the element $c_{ij}$ of the resulting matrix $C$ is given by:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} \cdot b_{kj}$$

where $i$ is the row number and $j$ is the column number, and $k$ is the index of the sum, which runs from 1 to $n$ (the number of columns in the first matrix).

# The big one: matrix multiplication

- Note that matrix multiplication is **not commutative**. That is, $A \times B \neq B \times A$.

- This is because the number of columns in the first matrix must be equal to the number of rows in the second matrix. So, for our example above, $A \times B$ is defined, but $B \times A$ is not, since the number of columns in $B$ is not equal to the number of rows in $A$.

- This is a common mistake that students make when first learning about matrix multiplication.

# Premultiplication and postmultiplication

- Because of the non-commutative nature of matrix multiplication, we have to always write down the order of the matrices when multiplying them, either in the equation or by specifying whether we are pre- or postmultiplying.

- If we say premultiply $A$ by $B$, we write $BA$. This means that $B$ is on the left side of $A$.

- If we say postmultiply $A$ by $B$, we write $AB$. This means that $B$ is on the right side of $A$.

- Note that the only case where both premultiplication and postmultiplication are defined is when $A$ and $B$ are square matrices, since the number of rows and columns are equal.

## Properties of matrix multiplication

- Matrix multiplication has several properties that are similar to multiplication of two scalars (numbers), but we need to adjust them to the matrix context.

- **Associative property**: $(AB)C = A(BC)$

- **Distributive property**, left: $A(B + C) = AB + AC$

- **Distributive property**, right: $(A + B)C = AC + BC$

- **Distributive property**, scalar: $k(AB) = (kA)B = A(kB)$

- **Identity property**: $AI = IA = A$ (where $I$ is the identity matrix).

- **Zero property**: $O_n A = AO_n = O_n$

- **Transpose property**: $(AB)^T = B^T A^T$

- **In general, matrix multiplication is not commutative**: $AB \neq BA$

## Powers of matrices

- We can also raise a matrix to a power, just like we can with a scalar.

- To raise a matrix to a power, we multiply the matrix by itself the number of times indicated by the power.

- For instance, consider the matrix $A$:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

- If we raise this matrix to the power of 2, we get:

$$A^2 = A \times A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + 2 \cdot 3 & 1 \cdot 2 + 2 \cdot 4 \\ 3 \cdot 1 + 4 \cdot 3 & 3 \cdot 2 + 4 \cdot 4 \end{bmatrix}$$

- Do NOT confuse this with the element-wise power of a matrix, which is not a standard operation in matrix algebra.

Daniel Sánchez Pazmiño
Math for the Social Sciences Module - Young Researchers Fellowship
Laboratorio de Investigación para el Desarrollo del Ecuador

# Transposition

- The **transpose** of a matrix is a new matrix that is obtained by exchanging the rows and columns of the original matrix.

- The transpose of a matrix $A$ is denoted by $A'$ or $A^T$.

- For instance, consider the matrix $A$:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

- The transpose of this matrix is:

$$A' = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

## Transposition properties

- The transpose of a matrix has several properties that are useful in matrix algebra.
- **Transposition of a sum**: $(A + B)' = A' + B'$
- **Transposition of a product**: $(AB)' = B'A'$
- **Transposition of a transposition**:
    - $(A')' = A$
- **Transposition of a scalar multiple**: $(kA)' = kA'$
- **Transposition of an identity matrix**: $(I_n)' = I_n$
- **Transposition of a zero matrix**: $(O_{m \times n})' = O_{n \times m}$
- **Transposition of a diagonal matrix**: $(D)' = D$

# Inverse of a matrix

- The **inverse** of a matrix is a matrix that, when multiplied by the original matrix, results in the identity matrix.

- The inverse of a matrix $A$ is denoted by $A^{-1}$.

- It must hold that $AA^{-1} = A^{-1}A = I$, where $I$ is the identity matrix.

- Not all matrices have an inverse. A matrix that has an inverse is called **invertible** or **nonsingular**.

- A matrix that does not have an inverse is called **noninvertible** or **singular**.

# Inverse of a matrix

- In order for a matrix to have an inverse, it must be a square matrix and its determinant must be different from zero.

- The determinant of a matrix is a scalar value that can be computed from the elements of the matrix. It is denoted by $\det(A)$.

- The values of the inverse itself are a bit more complex to compute, and we will not cover them in this lecture; a computer can easily compute the inverse of a matrix.

# Properties of the inverse of a matrix

- The inverse of a matrix has several properties that are useful in matrix algebra.
- **Inverse of a product**: $(AB)^{-1} = B^{-1}A^{-1}$
- **Inverse of a transpose**: $(A')^{-1} = (A^{-1})'$
- **Inverse of an inverse**: $(A^{-1})^{-1} = A$
- **Inverse of an identity matrix**: $(I_n)^{-1} = I_n$
- **Inverse of a scalar multiple**: $(kA)^{-1} = \frac{1}{k}A^{-1}$
- **Inverse of a diagonal matrix**: $(D)^{-1} = D^{-1}$
- **Inverse of a zero matrix**: $(O_{m \times n})^{-1}$ is undefined

## Division of matrices

- Division of matrices is not a standard operation in matrix algebra.

- We understand division as the inverse of multiplication, but in matrix algebra, we do not have a standard operation for division.

- We do the inverse of multiplication by multiplying by the inverse of the matrix. This is the closest thing we have to division in matrix algebra.

- For instance, consider the equation $AX = B$. To solve for $X$, we premultiply both sides by the inverse of $A$:

$$A^{-1}AX = A^{-1}B$$

- This will give us $X = A^{-1}B$. This becomes useful in defining the solution to a system of linear equations.

## Linear systems of equations in matrix form

- One of the most common applications of matrix algebra is in solving systems of linear equations.

- Consider a three variable system of equations:

$$2x + 3y + 4z = 10$$
$$3x + 4y + 5z = 20$$
$$4x + 5y + 6z = 30$$

- We write systems of linear equations in the following general form:

$$Ax = B$$

- Where $A$ is a matrix of coefficients, $x$ is a column vector of variables, and $B$ is a column vector of constants.

## Linear systems of equations in matrix form

- We write the $A$ matrix by taking the coefficients of the variables in the equations, except what is on the right side of the equal sign (the constants).

$$A = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix}$$

- The $x$ vector is the vector of variables we are trying to solve for:

$$x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

- The $B$ vector is the vector of constants on the right side of the equal sign:

# Linear systems of equations in matrix form

- We can then write the system of equations in matrix form as:

$$\begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 10 \\ 20 \\ 30 \end{bmatrix}$$

- This is equivalent to the system of equations we had before if we are to perform the matrix multiplication.

- Using our knowledge of matrix multiplication, we can solve this system of equations by finding the inverse of the matrix $A$ and multiplying it by the vector $B$.

## Solving linear systems of equations

- To solve a system of linear equations in matrix form, we should go back to the general form of the equation:

$$Ax = B$$

- To solve for $x$, we multiply both sides by the inverse of $A$:

$$A^{-1}Ax = A^{-1}B$$

- Since $A^{-1}A = I$, we get:

$$Ix = A^{-1}B$$

# Solving linear systems of equations

- Since $Ix = x$, we get:

$$x = A^{-1}B$$

- This is the solution to the system of linear equations.

# Solving linear systems of equations

- So, to solve a system of linear equations in matrix form, we need to:
  - Find the matrix $A$ of coefficients.
  - Find the vector $B$ of constants.
  - Find the inverse of the matrix $A$.
  - Multiply the inverse of $A$ by the vector $B$.
- This is a very efficient way to solve systems of linear equations, especially when the number of variables is large.

# Solving linear systems of equations

- The conditions for the matrix $A$ to have an inverse are the same as for any matrix to have an inverse: it must be a square matrix and its determinant must be different from zero.

- In terms of the system of equations, this means that the system must have a unique solution; if the determinant of the matrix is zero, the system has no solution or an infinite number of solutions.

- Notice that, for the two-variable system of equations, being able to solve the system of equations is equivalent to the determinant of the matrix being different from zero. This means having a unique solution: not having the two lines be parallel.

## The dot product

- Because vectors are just special cases of matrices, we can perform the same operations on vectors as we can on matrices.

- It is worth noting that the dot product (or inner or scalar product) are names for multiplications among vectors, which always end up in a scalar.

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + ... + a_n b_n = \sum_{i=1}^{n} a_i \cdot b_i$$

where $\mathbf{a}$ and $\mathbf{b}$ are vectors of the same length $n$.

- Notice that the dot product holds no regard with whether the vectors are row or column vectors, it always "assumes" conformability. However, these necessarily have to be vectors of the same length $n$.

# Dot product as matrix multiplication

- The dot product can also be seen as a special case of matrix multiplication, when we multiply a row vector by a column vector.

- For instance, consider the following row and column vectors of length $n$

$$\mathbf{a} = \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix}$$

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

## Dot product as matrix multiplication

- The dot product of these two vectors is given by:

$$\mathbf{a} \cdot \mathbf{b} = \begin{bmatrix} a_1 & a_2 & ... & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1 b_1 + a_2 b_2 + ... + a_n b_n$$

- This is the same as the dot product we defined before.

# Properties of the dot product

- The dot product has several properties that are useful in matrix algebra.
- **Commutative property**: $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
- **Distributive property**: $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$
- **Distributive property**: $(k\mathbf{a}) \cdot \mathbf{b} = k(\mathbf{a} \cdot \mathbf{b}) = \mathbf{a} \cdot (k\mathbf{b})$
- **Orthogonality**: If $\mathbf{a} \cdot \mathbf{b} = 0$, then $\mathbf{a}$ and $\mathbf{b}$ are orthogonal.
- **Distribution of scalar over dot product**: $(k\mathbf{a}) \cdot (l\mathbf{b}) = kl(\mathbf{a} \cdot \mathbf{b})$