Math Econ

YY

Thursday 5th February, 2026

1 Math Notes

These notes build intuition for core tools used in economics, data science, and optimization.

Topics (in progress)

- Linear Algebra
- Calculus
- Difference & Differential Equations

Note

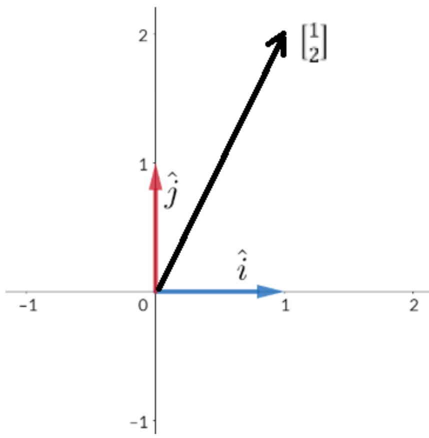
This is a living document. I will revise and expand as the course evolves.

2 Linear Algebra: Basics

2.1 Vector

A vector is a mathematical quantity that has both **magnitude** (or size) and **direction**.

Geometrically, a vector is represented as a directed line segment, like an arrow, where the length signifies the magnitude and the arrowhead indicates the direction.



More conveniently, you may write a vector as $\begin{bmatrix}1\\2\end{bmatrix}$ or simply $\vec{v}$. Can you tell what the magnitude (or length or norm) of the above vector is?

Another way to represent a vector is by using basis vectors, i.e. $\hat{\imath}$ and $\hat{\jmath}$, where $\hat{\imath} = \begin{bmatrix}1\\0\end{bmatrix}$ and $\hat{\jmath} = \begin{bmatrix}0\\1\end{bmatrix}$. Hence the vector $\begin{bmatrix}1\\2\end{bmatrix}$ can be written as $\vec{v} = 1\hat{\imath} + 2\hat{\jmath}$.

Make sure you know how to add, subtract vectors, and also multiply/scale vectors by some scalar.

2.2 Matrix Basics

We have already introduced the basis vectors $\hat{\imath}$ and $\hat{\jmath}$, which we can put into a matrix. That is, the $\hat{\imath}$ and $\hat{\jmath}$ vectors placed side-by-side in a 2×2 matrix:

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} \tag{1}$$

Now suppose we want to rotate $\vec{v} = \begin{bmatrix}1\\2\end{bmatrix}$ anticlockwise by 90°. Where will it go?

We can rotate the 2-D space, such that $\hat{\imath}$ will go to $\begin{bmatrix}0\\1\end{bmatrix}$ and $\hat{\jmath}$ will go to $\begin{bmatrix}-1\\0\end{bmatrix}$. Combining these into a matrix gives:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{2}$$

This can easily be done by pre-multiplying the vector by the transformation matrix, as follows:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Basically, a matrix can be viewed as a way to transform/change a vector!

Matrix and vector multiplication

The proper way to multiply a matrix and a vector is:

$$1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

Matrix addition and subtraction: If $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ are two $m \times n$ matrices of same dimension, then $\mathbf{A} + \mathbf{B}$ is defined as $(a_{ij} + b_{ij})$. That is, we add element by element the two matrices.
Clearly $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$. The rule applies to matrix subtraction.

Transpose of a matrix: If $\mathbf{A}$ is an $m \times n$ matrix, then $\mathbf{A}'$ is the $n \times m$ matrix whose rows are the columns of $\mathbf{A}$. So $\mathbf{A}' = (a_{ji})$. For example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \tag{3}$$

- Note: $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$, however $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

Null matrix: Has all elements as 0. Clearly $\mathbf{A} + \mathbf{0}_{m,n} = \mathbf{0}_{m,n} + \mathbf{A} = \mathbf{A}$ for all $m \times n$ matrices.

Scalar multiplication: If $\mathbf{A} = (a_{ij})$ then for any constant k, define $k\mathbf{A} = (ka_{ij})$. That is, multiply each element in $\mathbf{A}$ by k.

Matrix multiplication: Say $\mathbf{A}$ is $m \times n$, and $\mathbf{B}$ is $n \times p$, then the $m \times p$ matrix $\mathbf{AB}$ is the product of $\mathbf{A}$ and $\mathbf{B}$. For example:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \tag{4}$$

and the matrices $\mathbf{A}$ and $\mathbf{B}$ are conformable.

Fro example, given the matrices:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 0 \\ -6 & 7 \end{bmatrix}$$

To multiply A and B, we apply the row-by-column rule:

$$= \begin{bmatrix} 5-12 & 0+14 \\ 15+24 & 0-28 \end{bmatrix} = \begin{bmatrix} -7 & 14 \\ 39 & -28 \end{bmatrix}$$

Long story!!

Matrix Multiplication: The Column-Wise (Basis Vector) Method

Instead of dot products, we can view AB as taking the columns of A (the transformed $\hat{\imath}$ and $\hat{\jmath}$) and scaling them by the components in each column of B.

From the sam eexample above, let the columns of A be:

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$$

Finding Column 1 of the Result:

We use the first column of B $\begin{bmatrix} 5 \\ -6 \end{bmatrix}$ as scalars:

$$5\begin{bmatrix} 1 \\ 3 \end{bmatrix} + (-6)\begin{bmatrix} 2 \\ -4 \end{bmatrix} = \begin{bmatrix} 5 \\ 15 \end{bmatrix} + \begin{bmatrix} -12 \\ 24 \end{bmatrix} = \begin{bmatrix} -7 \\ 39 \end{bmatrix}$$

Finding Column 2 of the Result:

We use the second column of B $\begin{bmatrix} 0 \\ 7 \end{bmatrix}$ as scalars:

$$0\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 7\begin{bmatrix} 2 \\ -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 14 \\ -28 \end{bmatrix} = \begin{bmatrix} 14 \\ -28 \end{bmatrix}$$

Final Result:

$$AB = \begin{bmatrix} -7 & 14 \\ 39 & -28 \end{bmatrix}$$

as before!

Diagonal matrix: A square $n \times n$ matrix $\mathbf{A}$ is diagonal if all entries off the 'main diagonal' are zero, i.e.

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \tag{5}$$

Note: The elements in the diagonal of $\mathbf{A}$ are the same as those in $\mathbf{A}'$.

Trace of a square matrix: If $\mathbf{A}$ is a square matrix, the trace of $\mathbf{A}$, denoted $\text{tr}(\mathbf{A})$, is the sum of the elements on the main/leading diagonal of $\mathbf{A}$.

Identity matrix: Denoted by $\mathbf{I}_n$, the $n \times n$ diagonal matrix with $a_{ii} = 1$ for all i. That is:

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{6}$$

Symmetric matrix: A square matrix $\mathbf{A}$ is symmetric if $\mathbf{A} = \mathbf{A}'$. The identity matrix and the square null matrix are symmetric.

Idempotent matrix: A square matrix is said to be idempotent if $\mathbf{A}^n = \cdots = \mathbf{A}^2 = \mathbf{A}$. Below is an example (try squaring the matrix and see what you get).

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

Singular Matrix: A matrix that is linearly dependent (in the columns or rows) is singular and has determinant of zero. Note that any two (or three) vectors $\mathbf{x}$ and $\mathbf{y}$ (and $\mathbf{z}$) are said to be linearly dependent iff one can be written as a scalar multiple of the other. Two vectors $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$ are said to be **linearly independent** iff the ONLY solution to $a\mathbf{x} + b\mathbf{y} = \mathbf{0}$ is $a = 0$ AND $b = 0$. And $\mathbf{x}$ and $\mathbf{y}$ are said to be **orthogonal**.

2.2.1 Some Notes on Matrices

1. Usually $\mathbf{AB} \neq \mathbf{BA}$. For example, try $\mathbf{A} = \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

2. We can have $\mathbf{AB} = \mathbf{0}$ even if $\mathbf{A}$ or $\mathbf{B}$ are not zero. For example, try $\mathbf{A} = \begin{bmatrix} 6 & -12 \\ -3 & 6 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 12 & 6 \\ 6 & 3 \end{bmatrix}$.

3. Say, $\mathbf{A} = \begin{bmatrix} 4 & 8 \\ 1 & 2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}$, and $\mathbf{C} = \begin{bmatrix} -2 & 1 \\ 4 & 2 \end{bmatrix}$. We find that $\mathbf{AB} = \mathbf{AC}$ even though $\mathbf{B} \neq \mathbf{C}$.

4. Say $\mathbf{A}$ and $\mathbf{B}$ are singular matrices. Then $\mathbf{AB}$ will not be zero, although the product will be a singular matrix.

2.3 Systems of Equations in Matrix Form

One of the uses of linear algebra in economics is to represent and then solve systems of equations. For instance, consider the system of two equations:

$$2x_1 + x_2 = 5 - x_1 + x_2 = 2 \tag{7}$$

Here, we can define:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 2 \end{bmatrix} \tag{8}$$

Then we see that:

$$\mathbf{Ax} = \begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2x_1 + x_2 \\ -x_1 + x_2 \end{bmatrix}$$

The original system is in fact equivalent to the matrix form:

$$\mathbf{Ax} = \mathbf{b} \tag{9}$$

2.4 Matrix Row Operations

Matrix row operations, also known as elementary row operations, are three basic actions performed on a matrix: **swapping two rows**, **multiplying a row by a non-zero constant**, and **adding a multiple of one row to another row**.

1. **Interchanging two rows** (Row Swapping): You can swap the positions of any two rows in a matrix. As we will see later, this operation is useful for changing the order of equations in a system without altering the solution.
2. **Multiplying a row by a non-zero constant** (Scalar Multiplication): You can multiply every element in a specific row by any non-zero number. This is equivalent to multiplying both sides of an equation by a constant.
3. **Adding a multiple of one row to another row** (Row Addition): You can multiply one row by a constant and then add the result to another row. The original row and the row being multiplied remain unchanged. This operation is often the most powerful for simplifying systems, as it corresponds to adding a modified version of one equation to another.

Definition: REF (Row Echelon Form)
A matrix is in **Row Echelon Form** if...

1. Every non-zero row begins with a leading one.
2. A leading one in a lower row is further to the right.
3. Zero rows are at the bottom of the matrix.

Note: In some books, leading by one is not required.

Definition: RREF Reduced Row Echelon Form
A matrix is in **Reduced Row Echelon Form** if...

1. Every non-zero row begins with a leading one.
2. A leading one in a lower row is further to the right.
3. Zero rows are at the bottom of the matrix.
4. Every **column** with a leading one has zeros elsewhere.

While there can exist several row echelon forms for a matrix, there is only one (a unique) reduced row echelon form.

We can solve system of equation by row operations.

To solve for x_1 and x_2, we can use an augmented matrix and perform elementary row operations—swapping rows, scalar multiplication, and row addition—to reach **Reduced Row Echelon Form (RREF)**.

1. Set up the augmented matrix:

$$\begin{bmatrix} 2 & 1 & \mid & 5 \\ -1 & 1 & \mid & 2 \end{bmatrix}$$

2. Create a leading one in Row 1:

We can swap R_1 and R_2:

$$\begin{bmatrix} -1 & 1 & \mid & 2 \\ 2 & 1 & \mid & 5 \end{bmatrix}$$

Then multiply R_1 by -1 to get a leading one ($R_1 \leftarrow -1 \cdot R_1$):

$$\begin{bmatrix} 1 & -1 & \mid & -2 \\ 2 & 1 & \mid & 5 \end{bmatrix}$$

3. Create a zero below the leading one:

Add -2 times R_1 to R_2 ($R_2 \leftarrow R_2 - 2R_1$):

$$\begin{bmatrix} 1 & -1 & \mid & -2 \\ 0 & 3 & \mid & 9 \end{bmatrix}$$

4. Create a leading one in Row 2:

Multiply R_2 by 1/3 ($R_2 \leftarrow \frac{1}{3}R_2$):

$$\begin{bmatrix} 1 & -1 & \mid & -2 \\ 0 & 1 & \mid & 3 \end{bmatrix}$$

5. Create a zero above the leading one (RREF):

To reach reduced row echelon form, every column with a leading one must have zeros elsewhere. Add R_2 to R_1 ($R_1 \leftarrow R_1 + R_2$):

$$\begin{bmatrix} 1 & 0 & \mid & 1 \\ 0 & 1 & \mid & 3 \end{bmatrix}$$

Basically, a matrix can be viewed as a way to transform or change a vector to solve these systems!

Final Result

The system provides the unique solution:

$$x_1 = 1, \quad x_2 = 3$$

Basically, a matrix can be viewed as a way to transform or change a vector to solve these systems!

- Note: We need not reach RREF, and could stop at **Step 4** (or even **Step 3**). At Step 4, the last row tells us that $x_2 = 3$. We can then use **back-substitution** into the first row ($x_1 - x_2 = -2$). HEnce $x_1 - 3 = -2 \implies x_1 = 1$.
- This confirms the same result without performing the final row addition.

2.5 The Determinant of a Matrix

The determinant is a scalar value uniquely associated with a square non-singular matrix, and is usually denoted as $|\mathbf{A}|$. The question of whether or not a matrix is nonsingular and therefore invertible is linked to the value of its determinant.

We can write a generic 2×2 matrix as:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and its determinant is defined as:

$$|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21}$$

For a 3×3 matrix, say:

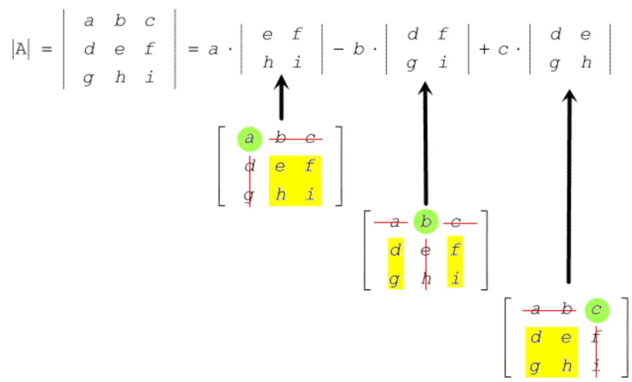
$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

its determinant is:

$$|\mathbf{A}| = a(ei - fh) - b(di - fg) + c(dh - ef)$$

Note the sign in front of b is begative, because we imppose (multiply by) the sign matrix $\begin{bmatrix} + & - & + \\ - & + & - \\ + & - & + \end{bmatrix}$.

Visually:



The 'yellow' bits called **minors** are 'smaller' determinants, now 2 by 2 and easier to handle. Note that the elements of the minor come from remaining elements after deleting the rows and columns of the corresponding elements in the selected row (or column).

Let's take a numerical example.

To find the determinant of matrix A using **Cofactor Expansion** along the first row (technically we could pick any row to work with) involves multiplying each element of the first row by its corresponding 2×2 minor, following the sign pattern: $(+, -, +)$ for the first row.

Given:

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

The formula for expansion along the first row is:

$$\det(A) = a(M_{11}) - b(M_{12}) + c(M_{13})$$

1. First Element $(a = 2)$**:**

$$+2 \cdot \begin{vmatrix} -1 & 2 \\ 1 & -1 \end{vmatrix} = 2 \cdot [(-1)(-1) - (2)(1)] = 2(1 - 2) = -2$$

2. Second Element $(b = -3)$**:** Note the negative sign from the checkerboard pattern:

$$-(-3) \cdot \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix} = 3 \cdot [(1)(-1) - (2)(3)] = 3(-1 - 6) = -21$$

3. Third Element $(c = 1)$**:**

$$+1 \cdot \begin{vmatrix} 1 & -1 \\ 3 & 1 \end{vmatrix} = 1 \cdot [(1)(1) - (-1)(3)] = 1(1+3) = 4$$

Final Result
Summing the results from each step gives $\det(A) = -2 - 21 + 4 = -19$.

2.6 Properties of Determinants

Property 1: $|\mathbf{A}| = |\mathbf{A}'|$ and $|\mathbf{A}| \cdot |\mathbf{B}| = |\mathbf{B}| \cdot |\mathbf{A}|$.

Property 2: Interchanging any two rows or columns will affect the sign of the determinant, but not its absolute value.

Property 3: Multiplying a single row or column by a scalar will cause the value of the determinant to be multiplied by the scalar.

Property 4: Addition or subtraction of a nonzero multiple of any row or column to or from another row or column does not change the value of the determinant.

Property 5: The determinant of a triangular matrix is equal to the product of the elements along the principal diagonal.

Property 6: If any of the rows or columns equal zero, the determinant is also zero.

Property 7: If two rows or columns are identical or proportional, i.e. linearly dependent, then the determinant is zero.

> Note: **Properties 2,3** and **4** are standard row operations already discussed above.

2.7 Matrix Inversion

Consider,the following.

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

Then pre-multiplying by:

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

Gives:

$$\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} = 10\mathbf{I}$$

Where does the 10 in the matrix after the = sign come from? It is the determinant of **A**.

Then, (pre-)multiplying both sides of the equation by $\frac{1}{10}$ gives:

$$\frac{1}{10} \begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} = \mathbf{I}$$

The matrix $\frac{1}{10}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix}$ is in fact the inverse of $\mathbf{A}$.

Hence we have the relationship $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$. In fact, $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$.

2.8 Finding the Inverse of a Matrix by Row Operations

We can find the inverse of a matrix using row operations or **Gauss-Jordan Elimination** method. The objective is to transform the matrix A into the Identity matrix I. By applying the same sequence of elementary row operations to an Identity matrix simultaneously, that Identity matrix transforms into A^{-1}.

Given the matrix A from the example above

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

We set up an **augmented matrix** $[A|I]$ by placing the 2×2 Identity matrix to the right of A:

$$[A|I] = \begin{bmatrix} 1 & -2 & | & 1 & 0 \\ 3 & 4 & | & 0 & 1 \end{bmatrix}$$

Step-by-Step Row Operations

1. Create a zero below the first leading one:

Subtract 3 times the first row from the second row ($R_2 \leftarrow R_2 - 3R_1$):

$$\begin{bmatrix} 1 & -2 & | & 1 & 0 \\ 0 & 10 & | & -3 & 1 \end{bmatrix}$$

2. Create a leading one in the second row:

Multiply the second row by 1/10 ($R_2 \leftarrow \frac{1}{10}R_2$):

$$\begin{bmatrix} 1 & -2 & | & 1 & 0 \\ 0 & 1 & | & -3/10 & 1/10 \end{bmatrix}$$

3. Create a zero above the second leading one:

To reach **Reduced Row Echelon Form (RREF)**, we add 2 times the second row to the first row ($R_1 \leftarrow R_1 + 2R_2$):

$$\begin{bmatrix} 1 & 0 & | & 1 + 2(-3/10) & 0 + 2(1/10) \\ 0 & 1 & | & -3/10 & 1/10 \end{bmatrix}$$

Simplifying the arithmetic in the first row:

$$\begin{bmatrix} 1 & 0 & | & 4/10 & 2/10 \\ 0 & 1 & | & -3/10 & 1/10 \end{bmatrix}$$

Final Result

The left side of the augmented matrix has been transformed into the Identity matrix. Therefore, the right side is now the inverse, A^{-1}:

$$A^{-1} = \begin{bmatrix} 0.4 & 0.2 \\ -0.3 & 0.1 \end{bmatrix}$$

Verification

A matrix multiplied by its inverse must result in the Identity matrix:

$$\begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 0.4 & 0.2 \\ -0.3 & 0.1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2.9 Properties of Inverse

Here are some properties of the inverse of a matrix worth knowing:

Property 1: For any nonsingular matrix $\mathbf{A}$, $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$.

Property 2: The determinant of the inverse of a matrix is equal to the reciprocal of the determinant of the matrix. That is $|\mathbf{A}^{-1}| = \frac{1}{|\mathbf{A}|}$.

Property 3: The inverse of matrix $\mathbf{A}$ is unique.

Property 4: For any nonsingular matrix $\mathbf{A}$, the inverse of the transpose of a matrix is equal to the transpose of the inverse of the matrix. $(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'$.

Property 5: If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular and of the same dimension, then $\mathbf{AB}$ is also nonsingular.

Property 6: The inverse of the product of two matrices is equal to the product of their inverses in reverse order. $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.

2.10 Finding the Inverse of a Matrix (Standard Approach)

The inverse of a matrix is defined as:

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \operatorname{adj}(\mathbf{A}) \tag{10}$$

Let's say we have the following linear system:

$$2x_1 - 3x_2 + x_3 = -1x_1 - x_2 + 2x_3 = -33x_1 + x_2 - x_3 = 9$$

Writing the above system in the form $\mathbf{Ax} = \mathbf{b}$, where:

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & -1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -1 \\ -3 \\ 9 \end{bmatrix}$$

We can therefore solve the system by $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$, which means that we need the inverse of $\mathbf{A}$.

Step 1. The minor $|M_{ij}|$ is the determinant of the submatrix formed by deleting the i-th row and j-th column of the original matrix. So we have:

$$\begin{bmatrix} -1 & -7 & 4 \\ 2 & -5 & 11 \\ -5 & 3 & 1 \end{bmatrix}$$

Step 2: The cofactor C_{ij} is a minor with the prescribed sign that follows the rule:

$$C_{ij} = (-1)^{i+j}|M_{ij}| \tag{11}$$

Hence, we have:

$$\begin{bmatrix} -1 & 7 & 4 \\ -2 & -5 & -11 \\ -5 & -3 & 1 \end{bmatrix}$$

Step 3: An adjoint matrix is simply the transpose of a cofactor matrix. Continuing the example above, we have:

$$\begin{bmatrix} -1 & -2 & -5 \\ 7 & -5 & -3 \\ 4 & -11 & 1 \end{bmatrix}$$

Since $|\mathbf{A}| = -19$, we then have:

$$\mathbf{A}^{-1} = -\frac{1}{19}\begin{bmatrix} -1 & -2 & -5 \\ 7 & -5 & -3 \\ 4 & -11 & 1 \end{bmatrix}$$

and

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = -\frac{1}{19}\begin{bmatrix} -1 & -2 & -5 \\ 7 & -5 & -3 \\ 4 & -11 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ -3 \\ 9 \end{bmatrix}$$

which gives $x_1 = 2, x_2 = 1$ and $x_3 = -2$.

2.11 Cramer's Rule

Cramer came up with a simplified method for solving a system of linear equations through the use of determinants.

Basically, given $\mathbf{Ax} = \mathbf{b}$, we have:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Define:

$$\Delta = |\mathbf{A}| \tag{12}$$

and

$$\Delta_1 = \begin{vmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} a_{11} & b_1 & a_{13} \\ a_{21} & b_2 & a_{23} \\ a_{31} & b_3 & a_{33} \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ a_{31} & a_{32} & b_3 \end{vmatrix} \tag{13}$$

where the columns in the **A** matrix are replaced one by one by the **b** vector as shown above.

Then:

$$x_i = \frac{\Delta_i}{\Delta} \tag{14}$$

So, for the previous example:

$$\mathbf{A} = \begin{bmatrix} 2 & -3 & 1 \\ 1 & -1 & 2 \\ 3 & 1 & -1 \end{bmatrix} \quad \text{and} \quad \det(A) = -19, \quad \mathbf{b} = \begin{bmatrix} -1 \\ -3 \\ 9 \end{bmatrix}$$

It follows that

$$\Delta_1 = \begin{vmatrix} -1 & -3 & 1 \\ -3 & -1 & 2 \\ 9 & 1 & -1 \end{vmatrix}, \quad \Delta_2 = \begin{vmatrix} 2 & -1 & 1 \\ 1 & -3 & 2 \\ 3 & 9 & -1 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} 2 & -3 & -1 \\ 1 & -1 & -3 \\ 3 & 1 & 9 \end{vmatrix}$$

where $\Delta_1 = -38, \Delta_2 = -19$, and $\Delta_3 = 38$.

So dividing through, you can easily verify that you get the soluitons $\{2, 1, -2\}$ as before!

3 Linear Algebra: Special Matrices

Let's look at some special matrices and determinants that are useful in economic analysis.

3.1 The Jacobian

The Jacobian (named after German mathematician Karl Gustav Jacobi, 1804–1851) is generally used in conjunction with partial derivatives to provide an easy test for the existence of functional dependence (linear and nonlinear). A Jacobian determinant $|\mathbf{J}|$ is composed of all the first-order partial derivatives of a system of equations, arranged in order sequence. Given

$$y_1 = f_1(x_1, x_2, x_3) y_2 = f_2(x_1, x_2, x_3) y_3 = f_3(x_1, x_2, x_3)$$

Then

$$|\mathbf{J}| = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \\ \frac{\partial y_3}{\partial x_1} & \frac{\partial y_3}{\partial x_2} & \frac{\partial y_3}{\partial x_3} \end{vmatrix} \tag{15}$$

For example, say you have

$$y_1 = 5x_1 + 3x_2 y_2 = 25x_1^2 + 30x_1x_2 + 9x_2^2$$

Then taking the first-order partials gives

$$\partial y_1 \partial x_1 = 5, \qquad \frac{\partial y_1}{\partial x_2} = 3, \qquad \frac{\partial y_2}{\partial x_1} = 50x_1 + 30x_2, \qquad \frac{\partial y_2}{\partial x_2} = 30x_1 + 18x_2$$

And the Jacobian is

$$|\mathbf{J}| = \begin{vmatrix} 5 & 3 \\ 50x_1 + 30x_2 & 30x_1 + 18x_2 \end{vmatrix}$$

The determinant of the Jacobian matrix is

$$|\mathbf{J}| = 5(30x_1 + 18x_2) - 3(50x_1 + 30x_2) = 0$$

Since $|\mathbf{J}| = 0$, there is functional dependence between the equations.

Note: $(5x_1 + 3x_2)^2 = 25x_1^2 + 30x_1x_2 + 9x_2^2$.

3.2 The Hessian

Given that the first-order conditions $z_x = z_y = 0$ are met, a sufficient condition for a multivariate function $z = f(x, y)$ to be an optimum is

1. $z_{xx}, z_{yy} > 0$ for a minimum; $z_{xx}, z_{yy} < 0$ for a maximum.
2. $z_{xx} z_{yy} > (z_{xy})^2$.

A convenient test for this second-order condition is provided by the Hessian. A Hessian $|\mathbf{H}|$ is a determinant composed of all the second-order partial derivatives, with the second-order direct partials on the principal diagonal and the second-order cross partials off the principal diagonal. That is,

$$|\mathbf{H}| = \begin{vmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{vmatrix} \tag{16}$$

where (by Young's Theorem) $z_{xy} = z_{yx}$.

If the first element on the principal diagonal, a.k.a. the first principal minor, $|H_1| = z_{xx}$ is positive, and the second principal minor also

$$|H_2| = \begin{vmatrix} z_{xx} & z_{xy} \\ z_{yx} & z_{yy} \end{vmatrix} = z_{xx} z_{yy} - (z_{xy})^2 > 0, \tag{17}$$

then the second-order conditions for a minimum are set. That is, when $|H_1| > 0$ and $|H_2| > 0$, the Hessian $|\mathbf{H}|$ is said to be **positive definite**. And a positive definite Hessian fulfills the second-order conditions for a **minimum**.

If however the first principal minor $|H_1| < 0$, while the second principal minor remains positive, then the second-order conditions for a maximum are met. That is, when $|H_1| < 0$ and $|H_2| > 0$, the Hessian $|\mathbf{H}|$ is said to be **negative definite** and fulfills the second-order conditions for a **maximum**.

Let's take a numerical example:

Given

$$z = f(x, y) = 2x^2 - xy + 2y^2 - 5x - 6y + 20,$$

then

$$z_x = 4x - y - 5, \quad z_y = -x + 4y - 6, \quad z_{xx} = 4, \quad z_{yy} = 4, \quad z_{xy} = z_{yx} = -1.$$

The Hessian is therefore

$$|\mathbf{H}| = \begin{vmatrix} 4 & -1 \\ -1 & 4 \end{vmatrix}$$

We have $|H_1| = 4 > 0$ and $|H_2| = (4)(4) - (-1)^2 = 16 - 1 = 15 > 0$, i.e. both princpal minnores are positive and hence the Hessian is said to be **positive definite** and the function z is characterized by a **minimu**m at the critical values (can you find these?).

3.3 The Discriminant

Determinants can be used to test for positive and negative definiteness of any quadratic form. The determinant of a quadratic form is called a **discriminant** $|\mathbf{D}|$. Given the quadratic form

$$z = ax^2 + bxy + cy^2, \tag{18}$$

the determinant is formed by placing the coefficients of the squared terms on the principal diagonal and dividing the coefficients of the non-squared terms equally between the off-diagonal positions. Hence, we have

$$|\mathbf{D}| = \begin{vmatrix} a & b/2 \\ b/2 & c \end{vmatrix} \tag{19}$$

We then evaluate the principal minors like we did for the Hessian test, where

$$|D_1| = a \quad \text{and} \quad |D_2| = \begin{vmatrix} a & b/2 \\ b/2 & c \end{vmatrix} = ac - \frac{b^2}{4}. \tag{20}$$

If $|D_1|, |D_2| > 0$, $|\mathbf{D}|$ is positive definite and z is positive for all nonzero values of x and y. If $|D_1| < 0$ and $|D_2| > 0$, $|\mathbf{D}|$ is negative definite and z is negative for all nonzero values of x and y. If $|D_2| \neq 0$, z is not sign definite and z may assume both positive and negative values.

Let's take an example to test for sign definiteness of the following quadratic form

$$z = 2x^2 + 5xy + 8y^2.$$

We can easily form the discriminant

$$|\mathbf{D}| = \begin{vmatrix} 2 & 2.5 \\ 2.5 & 8 \end{vmatrix}$$

Then evaluating the principal minors gives

$$|D_1| = 2 > 0, \qquad |D_2| = \begin{vmatrix} 2 & 2.5 \\ 2.5 & 8 \end{vmatrix} = 16 - 6.25 = 9.75 > 0.$$

Hence, z is positive definite, meaning that it will be greater than zero for all nonzero values of x and y.

3.4 The Quadratic Form

A quadratic form is defined as a polynomial expression in which each component term has a uniform degree.

Here are some examples:

$6x^2 - 2xy + 3y^2 \quad$ is a quadratic form in 2 variables.$x^2 + 2xy + 4xz + 2yz + y^2 + z^2 \quad$ is a quadratic form in 3 variabl

It would be useful to determine the sign definiteness so we can make statements concerning the optimum value of the function as to whether it is a minimum or a maximum.

More generally, a quadratic form in n variables $(x_1, x_2, \ldots, x_n)$ can be written as $\mathbf{x}'\mathbf{A}\mathbf{x}$ where

$\mathbf{x}'$ is a row vector $[x_1, x_2, \ldots, x_n]$ and
$\mathbf{A}$ is an $n \times n$ matrix of scalar elements.

To see this more clearly, let

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

Then the quadratic form is:

$$\begin{aligned} Q(x,y) &= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= \begin{bmatrix} 2x+y & x+3y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \\ &= (2x+y)x + (x+3y)y \\ &= 2x^2 + 2xy + 3y^2 \end{aligned}$$

Furthermore, completing the square gives:

$$\begin{aligned} Q(x,y) &= 2x^2 + 2xy + 3y^2 = 2(x^2 + xy) + 3y^2 \\ &= 2\left(x + \frac{y}{2}\right)^2 - \frac{1}{2}y^2 + 3y^2 = 2\left(x + \frac{y}{2}\right)^2 + \frac{5}{2}y^2 \end{aligned}$$

Since both squared terms have positive coefficients, $Q(x, y) > 0$ for all $(x, y) \neq (0, 0)$, the quadratic form is **positive definite**.

Sign Definiteness of Quadratic Form

A quadratic form is said to be:

- **Positive definite** if $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$.
- **Positive semi-definite** if $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ and $\exists \mathbf{x} \neq \mathbf{0}$ such that $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$.
- **Negative definite** if $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$ for all $\mathbf{x} \neq \mathbf{0}$.
- **Negative semi-definite** if $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0$ and $\exists \mathbf{x} \neq \mathbf{0}$ such that $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$.
- **Sign indefinite** if it takes both positive and negative values.

Quadratic forms are used in many areas such as:

- Optimization: To determine if a critical point is a minimum, maximum, or saddle point (via the Hessian matrix).
- Geometry: Describing conic sections (ellipses, hyperbolas) and quadric surfaces.
- Physics: Representing energy functions (e.g., kinetic energy in mechanics).

Quadratic forms are particularly useful in determining the concavity or convexity of a differentiable function.

For a function $y = f(x_1, x_2, \ldots, x_n)$, we can form the Hessian consisting of second-order partial derivatives and construct a quadratic form as $\mathbf{x}'\mathbf{H}\mathbf{x}$. Then

a) The function y is strictly convex if the quadratic form is positive (implying that the Hessian is positive definite, i.e., the determinants of the principal minors are all positive).

b) The function y is strictly concave if the quadratic form is negative (that is, the Hessian is negative definite, i.e. the determinants of the principal minors alternate in sign).

Let's take an example to see this more clearly.

Say we have

$$z = x^2 + y^2.$$

Then

$$z_x = 2x, \quad z_{xx} = 2, \quad z_{xy} = 0, z_y = 2y, \quad z_{yx} = 0, \quad z_{yy} = 2.$$

But

$$\mathbf{x}'\mathbf{H}\mathbf{x} = z_{xx}x^2 + z_{yy}y^2 + z_{xy}z_{yx}xy = 2x^2 + 2y^2 > 0.$$

Hence the function $z = x^2 + y^2$ is characterized by a minimum at its optimal value and is therefore a strictly convex function. And it is easy to see that the Hessian is positive definite (i.e. the value of the function is always positive for any non-zero values of x and y).

Here's yet another example:

Say we have a quadratic form in 2 variables

$$z = 8x^2 + 6xy + 2y^2.$$

This can be rearranged as

$$z = 8x^2 + 3xy + 3xy + 2y^2.$$

We can write this as

$$\begin{aligned}\mathbf{x}'\mathbf{H}\mathbf{x} &= z_{xx}x^2 + z_{yy}y^2 + (z_{xy} + z_{yx})xy \\ &= 8x^2 + 2y^2 + (3+3)xy \\ &= 8x^2 + 2y^2 + 6xy\end{aligned}$$

The Hessian is

$$\mathbf{H} = \begin{bmatrix} 8 & 3 \\ 3 & 2 \end{bmatrix}$$

The principal minors are $|H_1| = 8$ and $|H_2| = \begin{vmatrix} 8 & 3 \\ 3 & 2 \end{vmatrix} = 16 - 9 = 7$. Both are positive so we have the Hessian and the quadratic form as positive definite!

3.5 Higher Order Hessian

Given $y = f(x_1, x_2, x_3)$, the third-order Hessian is

$$|\mathbf{H}| = \begin{vmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{vmatrix} \tag{21}$$

where the elements are the various second-order partial derivatives of y:

$$y_{11} = \frac{\partial^2 y}{\partial x_1^2}, \quad y_{12} = \frac{\partial^2 y}{\partial x_2 \partial x_1}, \quad y_{23} = \frac{\partial^2 y}{\partial x_3 \partial x_2}, \quad \text{etc.} \tag{22}$$

Conditions for a relative minimum or maximum depend on the signs of the first, second, and third principal minors, respectively. If

$$|H_1| > 0, \quad |H_2| = \begin{vmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{vmatrix} > 0, \quad \text{and} \quad |H_3| = |\mathbf{H}| > 0, \tag{23}$$

then $|\mathbf{H}|$ is positive definite and fulfills the second-order conditions for a minimum.

If

$$|H_1| < 0, \quad |H_2| = \begin{vmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{vmatrix} > 0, \quad \text{and} \quad |H_3| = |\mathbf{H}| < 0, \tag{24}$$

then $|\mathbf{H}|$ is negative definite and fulfills the second-order conditions for a maximum.

Let's take an example. Given the function:

$$y = -5x_1^2 + 10x_1 + x_1x_3 - 2x_2^2 + 4x_2 + 2x_2x_3 - 4x_3^2.$$

The first order conditions (F.O.C) are

$$\frac{\partial y}{\partial x_1} = y_1 = -10x_1 + 10 + x_3 = 0, \frac{\partial y}{\partial x_2} = y_2 = -4x_2 + 2x_3 + 4 = 0, \frac{\partial y}{\partial x_3} = y_3 = x_1 + 2x_2 - 8x_3 = 0.$$

which can be expressed in matrix form as

$$\begin{bmatrix} -10 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -10 \\ -4 \\ 0 \end{bmatrix}$$

Using Cramer's rule we get $x_1 \approx 1.04$, $x_2 \approx 1.22$ and $x_3 \approx 0.43$ (please verify this).

Taking the second partial derivatives from the first-order conditions to create the Hessian,

$$y_{11} = -10, \quad y_{12} = 0, \quad y_{13} = 1, y_{21} = 0, \quad y_{22} = -4, \quad y_{23} = 2, y_{31} = 1, \quad y_{32} = 2, \quad y_{33} = -8.$$

Thus,

$$\mathbf{H} = \begin{bmatrix} -10 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -8 \end{bmatrix}$$

Finally, applying the Hessian test, we have

$$|H_1| = -10 < 0,$$

$$|H_2| = \begin{vmatrix} -10 & 0 \\ 0 & -4 \end{vmatrix} = 40 > 0,$$

$$|H_3| = |\mathbf{H}| = \begin{vmatrix} -10 & 0 & 1 \\ 0 & -4 & 2 \\ 1 & 2 & -8 \end{vmatrix} = -276 < 0.$$

Since the principal minors alternate correctly in sign, the Hessian is **negative definite** and the function is **maximized** at $x_1 \approx 1.04$, $x_2 \approx 1.22$ and $x_3 \approx 0.43$.

3.6 The Bordered Hessian

To optimize a function $f(x_1, x_2)$ subject to a constraint $g(x_1, x_2)$ the first-order conditions can be found by setting up what is known as the Lagrangian function $F(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda(g(x_1, x_2) - c)$.

The second-order conditions can be expressed in terms of a bordered Hessian $|\bar{\mathbf{H}}|$ as

$$|\bar{\mathbf{H}}| = \begin{vmatrix} 0 & g_1 & g_2 \\ g_1 & F_{11} & F_{12} \\ g_2 & F_{21} & F_{22} \end{vmatrix} \tag{25}$$

or

$$|\bar{\mathbf{H}}| = \begin{vmatrix} 0 & g_1 & g_2 \\ g_1 & f_{11} & f_{12} \\ g_2 & f_{21} & f_{22} \end{vmatrix} \tag{26}$$

which is the usual Hessian bordered by the first derivatives of the constraint with zero on the principal diagonal.

The order of a bordered principal minor is determined by the order of the principal minor being bordered. Hence $|\bar{H}_2|$ above represents a second bordered principal minor, because the principal minor being bordered has dimensions 2×2.

For the bivariate case with a single constraint, we simply look at $|\bar{H}_2|$. If this is negative, the bordered Hessian is said to be positive definite and satisfies the second-order condition for a minimum. However, if it is positive, the bordered Hessian is said to be negative definite, and meets the sufficient conditions for a maximum.

Let's try optimizing the following objective function

$$f(x_1, x_2) = 4x_1^2 + 3x_2^2 - 2x_1x_2$$

subject to

$$x_1 + x_2 = 56.$$

Setting up the Lagrangian function, taking the first-order partials and solving gives $x_1^* = 36$, $x_2^* = 20$ (and $\lambda = 348$). (You should verify this.)

The bordered Hessian for this optimization problem is

$$|\bar{\mathbf{H}}| = \begin{vmatrix} 0 & 1 & 1 \\ 1 & 8 & -2 \\ 1 & -2 & 6 \end{vmatrix}$$

Starting with the second principal minor, we have

$$\begin{aligned} |\bar{H}_2| = |\bar{\mathbf{H}}| &= 0 \cdot \begin{vmatrix} 8 & -2 \\ -2 & 6 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & -2 \\ 1 & 6 \end{vmatrix} + 1 \cdot \begin{vmatrix} 1 & 8 \\ 1 & -2 \end{vmatrix} \\ &= -(6+2) + (-2-8) = -8 - 10 = -18 \end{aligned}$$

which is negative, hence $|\bar{\mathbf{H}}|$ is positive definite and we have met sufficient conditions for a minimum.

For the more general case in which the objective function has say n variables, i.e., $f(x_1, \dots, x_n)$ which is subject to some constraint $g(x_1, \dots, x_n)$, we can set up the bordered Hessian as

$$|\bar{\mathbf{H}}| = \begin{bmatrix} 0 & g_1 & g_2 & \dots & g_n \\ g_1 & F_{11} & F_{12} & \dots & F_{1n} \\ g_2 & F_{21} & F_{22} & \dots & F_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_n & F_{n1} & F_{n2} & \dots & F_{nn} \end{bmatrix}$$

where $|\bar{\mathbf{H}}| = |\bar{H}_n|$ because the $n \times n$ principal minor is bordered.

In this case, if all the principal minors are negative, i.e. $|\bar{H}_2|, |\bar{H}_3|, \dots, |\bar{H}_n| < 0$, the bordered Hessian is said to be positive definite and satisfies the second-order condition for a minimum.

On the other hand, if the principal minors alternate consistently in sign from positive to negative, i.e. $|\bar{H}_2| > 0, |\bar{H}_3| < 0, |\bar{H}_4| > 0$ etc., the bordered Hessian is negative definite, and meets the sufficient conditions for a maximum.

3.7 Input-Output Analysis

If a_{ij} is a technical coefficient representing the value of input i required to produce one dollar's worth of product j, the total demand for good i can be expressed as

$$x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + b_i \tag{27}$$

where b_i is the final demand for product i. What is important to realize here is that the total demand for a product consists of that product being the final demand plus that product being an intermediate good required for the production of other products.

In matrix form we have

$$\mathbf{X} = \mathbf{AX} + \mathbf{B} \tag{28}$$

where, for an n sector economy,

$$\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \tag{29}$$

A is called the matrix of technical coefficients.

To find the output (intermediate and final goods) needed to satisfy demand, all we have to do is to solve for **X**:

$$\mathbf{X} - \mathbf{AX} = \mathbf{B} \quad \Rightarrow \quad (\mathbf{I} - \mathbf{A})\mathbf{X} = \mathbf{B} \quad \Rightarrow \quad \mathbf{X} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}. \tag{30}$$

where $(I - A)$ is kowns as the **Leontief matrix**.

Thus for a 3-sector economy, we have

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & 1 - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & 1 - a_{33} \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \\ b_n \end{bmatrix}$$

Say we are asked to determine total output for three sectors/industries given A and B as below:

$$A = \begin{bmatrix} 0.3 & 0.4 & 0.1 \\ 0.5 & 0.2 & 0.6 \\ 0.1 & 0.3 & 0.1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 20 \\ 10 \\ 30 \end{bmatrix}$$

Since $X = (I - A)^{-1}B$

$$I - A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.3 & 0.4 & 0.1 \\ 0.5 & 0.2 & 0.6 \\ 0.1 & 0.3 & 0.1 \end{bmatrix} = \begin{bmatrix} 0.7 & -0.4 & -0.1 \\ -0.5 & 0.8 & -0.6 \\ -0.1 & -0.3 & 0.9 \end{bmatrix}$$

And taking the inverse

$$(I - A)^{-1} = \frac{1}{0.151} \begin{bmatrix} 0.54 & 0.39 & 0.32 \\ 0.51 & 0.62 & 0.47 \\ 0.23 & 0.25 & 0.36 \end{bmatrix}$$

Hence,

$$\begin{aligned} X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} &= \frac{1}{0.151} \begin{bmatrix} 0.54 & 0.39 & 0.32 \\ 0.51 & 0.62 & 0.47 \\ 0.23 & 0.25 & 0.36 \end{bmatrix} \begin{bmatrix} 20 \\ 10 \\ 30 \end{bmatrix} \\ &= \frac{1}{0.151} \begin{bmatrix} 24.3 \\ 30.5 \\ 17.9 \end{bmatrix} = \begin{bmatrix} 160.93 \\ 201.99 \\ 118.54 \end{bmatrix} \end{aligned}$$

3.8 Characteristic Roots and Vectors (Eigenvalues and Eigenvectors)

The sign and definiteness of a Hessian and a quadratic form has been tested by using the principal minors. Sign definiteness can also be tested by using the characteristic roots of a matrix. Given a square matrix $\mathbf{A}$, is possible to find a vector $\mathbf{V} \neq 0$ and a scalar λ such that

$$\mathbf{AV} = \lambda \mathbf{V} \tag{31}$$

the scalar λ is called the *characteristic root*, *latent value* or *eigenvalue*; and the vector $\mathbf{V}$ is called the *characteristic vector*, *latent vector* or *eigenvector*. The above can be written as

$$\mathbf{AV} - \lambda \mathbf{V} = 0 \tag{32}$$

which can be rearranged so that

$$\begin{aligned} \mathbf{AV} - \lambda \mathbf{IV} &= 0 \\ (\mathbf{A} - \lambda \mathbf{I})\mathbf{V} &= 0 \end{aligned} \tag{33}$$

where $\mathbf{A} - \lambda \mathbf{I}$ is called the *characteristic matrix* of $\mathbf{A}$. Since we have $\mathbf{V} \neq 0$, the characteristic matrix $\mathbf{A} - \lambda \mathbf{I}$ must be singular and thus its determinant is zero.

If $\mathbf{A}$ is a 3×3 matrix, then

$$|\mathbf{A} - \lambda \mathbf{I}| = \begin{vmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ a_{21} & a_{22} - \lambda & a_{23} \\ a_{31} & a_{32} & a_{33} - \lambda \end{vmatrix} = 0 \tag{34}$$

With $|\mathbf{A} - \lambda \mathbf{I}| = 0$, there will be an infinite number of solutions for $\mathbf{V}$. To force a unique solution, the solution may be *normalized* by requiring of the elements v_i of $\mathbf{V}$ such that $\sum v_i^2 = 1$.

Sign Definiteness and Characteristc Roots

For a square matrix $\mathbf{A}$ if

- All characteristic roots λ are positive $\Rightarrow$ $\mathbf{A}$ is positive definite.
- All λ's are negative $\Rightarrow$ $\mathbf{A}$ is negative definite.
- All λ's are nonnegative and at least one $\lambda = 0 \Rightarrow$ $\mathbf{A}$ is positive semi-definite.
- All λ's are nonpositive and at least one $\lambda = 0 \Rightarrow$ $\mathbf{A}$ is negative semi-definite.
- Some λ's are positive and some negative $\Rightarrow$ $\mathbf{A}$ is sign indefinite.

Let's take an example. Given a square matrix

$$\mathbf{A} = \begin{bmatrix} -6 & 3 \\ 3 & -6 \end{bmatrix}$$

To find the characteristic roots (eigenvalues) of $\mathbf{A}$, we simply set $|\mathbf{A}-\lambda\mathbf{I}| = 0$:

$$|\mathbf{A} - \lambda\mathbf{I}| = \begin{vmatrix} -6-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix} = 0$$

This means

$$(-6-\lambda)(-6-\lambda) - 9 = 0\lambda^2 + 12\lambda + 27 = 0(\lambda+9)(\lambda+3) = 0\lambda = -9, -3.$$

Since both characteristic roots λ are negative, we say $\mathbf{A}$ is negative definite.

Note:

$$\text{(i)} \sum \lambda_i = \operatorname{tr}(\mathbf{A}), \qquad \text{(ii)} \prod \lambda_i = |\mathbf{A}|. \tag{35}$$

Let's continue with the example above to find the characteristic vector.

We know one of the roots $\lambda = -9$, so substituting in the characteristic matrix gives

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0} \quad \Rightarrow \quad \begin{bmatrix} -6-(-9) & 3 \\ 3 & -6-(-9) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since the coefficient matrix is linearly dependent, there are infinite number of solutions. The product of the matrices gives two equations which are identical:

$$3v_1 + 3v_2 = 0 \quad \Rightarrow \quad v_2 = -v_1.$$

By normalizing we have

$$v_1^2 + v_2^2 = 1.$$

Substituting $v_2 = -v_1$ gives

$$v_1^2 + (-v_1)^2 = 1 \quad \Rightarrow \quad 2v_1^2 = 1 \quad \Rightarrow \quad v_1^2 = \frac{1}{2}.$$

Taking the positive square root gives $v_1 = \sqrt{1/2} = \frac{\sqrt{2}}{2}$ and substituting into $v_2 = -v_1$ gives $v_2 = -\frac{\sqrt{2}}{2}$. That is

$$\mathbf{v}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

Using the second characteristic root $\lambda = -3$:

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \begin{bmatrix} -6-(-3) & 3 \\ 3 & -6-(-3) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ 3 & -3 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Multiplying out gives $-3v_1 + 3v_2 = 0$ and $3v_1 - 3v_2 = 0$, so $v_1 = v_2$. Normalizing as before:

$$v_1^2 + v_2^2 = 1 \quad \Rightarrow \quad 2v_1^2 = 1 \quad \Rightarrow \quad v_1 = \frac{\sqrt{2}}{2}.$$

Hence,

$$\mathbf{v}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}.$$

3.9 Diagonalization

A square matrix A is diagonalizable if it can be written as:

$$\mathbf{A} = \mathbf{T}\mathbf{D}\mathbf{T}^{-1}. \qquad (36)$$

where

- **D** is a diagonal matrix (entries only on the main diagonal) consisting of eignevalues of **A**,
- **T** is an invertible matrix whose columns are eigenvectors of **A**.

Note that not all matrices are diagonalizable, however.
For example, given

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

Let us start by finding its eigenvalues.
The characteristic polynomial is given by

$$\det(A - \lambda I) = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0.$$

Hence,

$$\lambda = 1, \qquad \lambda = 3.$$

- **For** $\lambda = 1$:

$$(A - I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \mathbf{v} = \mathbf{0}.$$

A corresponding eigenvector is

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

- **For** $\lambda = 3$:

$$(A - 3I)\mathbf{v} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \mathbf{v} = \mathbf{0}.$$

A corresponding eigenvector is

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then,

$$T = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}.$$

The inverse of T is

$$T^{-1} = \tfrac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

We can verify that $A = TDT^{-1}$.

$$TDT^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \cdot \tfrac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = A.$$

Note (1):

- All **symmetric matrices** are diagonalizable (even with repeated eigenvalues).
- Diagonalizable $\neq$ invertible.
  For example,

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

  is diagonalizable but not invertible.
- If A is diagonalizable, then A^n and e^A are easy to compute.

Note (2):
This is closely related to the transformation form of diagonalization, which states that if A is diagonalizable, then there exists an invertible matrix T and a diagonal matrix D such that

$$T^{-1}AT = D.$$

Exercise. Can you show this?

4 Basic Differentiation

Before we study the differentiation of single-variable functions, we briefly review several foundational mathematical concepts.

4.1 Functions

A function f from a set X to a set Y, written $f : X \to Y$, is a rule that assigns **exactly one** element of Y to each element of X.

- X is called the **domain**
- The **range** is the set of values in Y that are actually attained

Using $x \in X$ and $y \in Y$, a function is written as $y = f(x)$, where x is the independent variable and y the dependent variable.

4.2 Graphs

If X and Y are sets of real numbers, the **graph** of a function f is the set of points (x, y) such that $y = f(x)$.

Economic convention.
Economists often draw demand curves with quantity on the horizontal axis and price on the vertical axis, even when the function is written as $q = f(p)$.

4.3 Slope

The slope of a line through points (x, y) and (x', y') is

$$m = \frac{y' - y}{x' - x}.$$

Differentiation is the method of finding the slope of a function and is denoted by $f'(x)$.

4.4 Limits

We say that a function f has **limit** L as $x \to a$ if, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that[1]

[1] A function need not be defined at the point a in order to have a limit as $x \to a$. For example,

$$f(x) = \frac{x^2 - 1}{x - 1}$$

is not defined at $x = 1$, but

$$\lim_{x \to 1} f(x) = 2.$$

$$|f(x) - L| < \varepsilon \quad \text{whenever} \quad 0 < |x - a| < \delta.$$

When this condition holds, we write

$$\lim_{x \to a} f(x) = L.$$

4.5 Continuity

A function f is **continuous at** a if:[2]

1. $f(a)$ is defined
2. $\lim_{x \to a} f(x)$ exists
3. $\lim_{x \to a} f(x) = f(a)$

4.6 Derivative at a Point

Let $y = f(x)$. When x changes by Δx, the change in y is

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

4.7 Derivative as a Function

The **derivative of** f **at** x is defined as

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

If the derivative exists for every x in the domain of f, then the derivative itself defines a new function, denoted $f'(x)$.

Geometrically, $f'(x)$ is the slope of the tangent line to the graph of f at $(x, f(x))$.

Common notations include:

- $f'(x)$
- $\dfrac{dy}{dx}$
- $Df(x)$

[2] The two functions discussed above are not continuous. The first is not continuous because $f(a)$ is not defined. The second is not continuous because f does not converge to a limit as $x \to a$. For example, if

$$f(x) = \begin{cases} -1, & x < 0, \\ 1, & x > 0, \end{cases}$$

then f has no limit as $x \to 0$, since the right-hand limit equals 1 while the left-hand limit equals -1.

Note that what we usually think of as a variable x is held constant while Δx varies and converges to zero. It is useful to keep in mind that the derivative of a function f at x is the slope of a line tangent to the graph of the function f at the point $(x, f(x))$. It is crucial to understand the implications of the existence of the derivative at a point x. The function must be smooth—meaning it is both continuous and differentiable—at the point x. The tangent line provides a high-quality linear approximation to the graph of the function near x. In general, if we know that the function f is differentiable at a, then the tangent line approximation to f at a is:

$$y = f(a) + f'(a)(x - a)$$

where $a, f(a)$, and $f'(a)$ are constants, x is the independent variable, and y is the dependent variable. We will see this point again with Taylor series expansions. Many important concepts in economics—such as marginal cost or marginal utility—are based on this derivative function.

4.8 Second Derivative

The **second derivative** is the derivative of the derivative and is written as

$$f''(x) = \frac{d^2 f(x)}{dx^2}.$$

Economic interpretation.
If $\ln p(t)$ describes log prices over time, then:

- the first derivative is inflation
- the second derivative is the change in inflation

4.9 Basic Rules of Differentiation

Let $y = f(x)$.

4.9.1 Constant-function Rule

The derivative of a contsant function $y = f(x) = k$ is zero, for all values of x-it has zero slope!

$$\frac{d}{dx}(k) = 0.$$

4.9.2 Power-function Rule

The derivative of a power function $f(x) = x^n$ is:

$$\frac{d}{dx}(x^n) = nx^{n-1}.$$

4.9.3 Generalized Power-function Rule

When a multiplicaytive constant k appears in the power fuction, so that $f(x) = kx^n$, then:

$$\frac{d}{dx}(kx^n) = knx^{n-1}.$$

4.9.4 Logarithmic Rule

The derivatice of the log-function $f(x) = lnx$ is:

$$\frac{d}{dx}(\ln x) = \frac{1}{x}.$$

4.9.5 Exponential Rule

For some exponential function $f(x) = a^x$, where a is some constant, then:

$$\frac{d}{dx}(a^x) = a^x \ln a.$$

Note that a particular case of the above is

$$\frac{d}{dx}e^x = e^x$$

While

$$\frac{d}{dx}\ln x = \frac{1}{x}$$

Now, let's consider some further useful rules of differentiation involving two or more functions of the same variable. Specifically, suppose $f(x)$ and $g(x)$ are two different functions of x and that $f'(x)$ and $g'(x)$ exist. That is, let $f(x)$ and $g(x)$ be differentiable, then:

4.9.6 Sum-difference Rules

The derivative of a sum (difference) of two functions is the sum (difference) of the derivatives of the two functions.

$$\frac{d}{dx}[f(x) \pm g(x)] = f'(x) \pm g'(x).$$

4.9.7 Product Rule

The derivative of the product of two (differentiable) functions is equal to the first function times the derivative of the second function plus the second function times the derivative of the first function.

$$\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x).$$

4.9.8 Quotient Rule

The derivative of the quotient of two (differentiable) functions, $f(x)/g(x)$, is

$$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$$

provided that $g(x) \neq 0$. Note that $[g(x)]^2 = g^2(x)$.

4.9.9 Chain Rule

If $z = f(y)$ and $y = g(x)$, then

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}.$$

The chain rule provides a convenient way to study how one variable (say, x) affects another variable (z) through its influence on some intermediate variable (y).

Sometimes, we can write for a composite function $y = f(g(x))$:

$$\frac{dy}{dx} = f'(g(x)) \cdot g'(x)$$

4.9.10 Chain Rule for Exponential and Logarithmic Functions

The general exponential function rule

$$\frac{d}{dx}e^{g(x)} = e^{g(x)}g'(x)$$

For example:

$$\frac{d}{dx}e^{ax} = \frac{d}{d(ax)}e^{ax}\frac{d}{dx}(ax) = e^{ax}a = ae^{ax}$$

If we are using a base other than e:

$$\frac{d}{dx}(a^{g(x)}) = a^{g(x)}g'(x)\ln a, \text{ where } a > 0, a \neq 0$$

The general natural logarithmic function rule

$$\frac{d}{dx}\ln(g(x)) = \frac{g'(x)}{g(x)}$$

Interestingly:

$$\frac{d}{dx}\ln(ax) = \frac{d}{d(ax)}\ln(ax)\frac{d}{dx}(ax) = \frac{1}{ax}a = 1/x$$

while

$$\frac{d}{dx}\ln(x^2) = \frac{d}{d(x^2)}\ln(x^2)\frac{d}{dx}(x^2) = \frac{1}{x^2}2x = 2/x$$

Note also when considered base other than e. Because

$$\log_b(x) = \frac{\ln(x)}{\ln(b)}$$

we have

$$\frac{d}{dx}\log_b(x) = \frac{1}{x}\frac{1}{\ln(b)}$$

Or more generally:

$$\begin{aligned}\frac{d}{dx}\log_b g(x) &= \frac{g'(x)}{g(x)}\frac{1}{\ln b}, \text{ where } b > 0, b \neq 1 \\ &= \frac{g'(x)}{g(x)}\log_b e\end{aligned}$$

Note that $\log_b e = \dfrac{1}{\ln b}$.

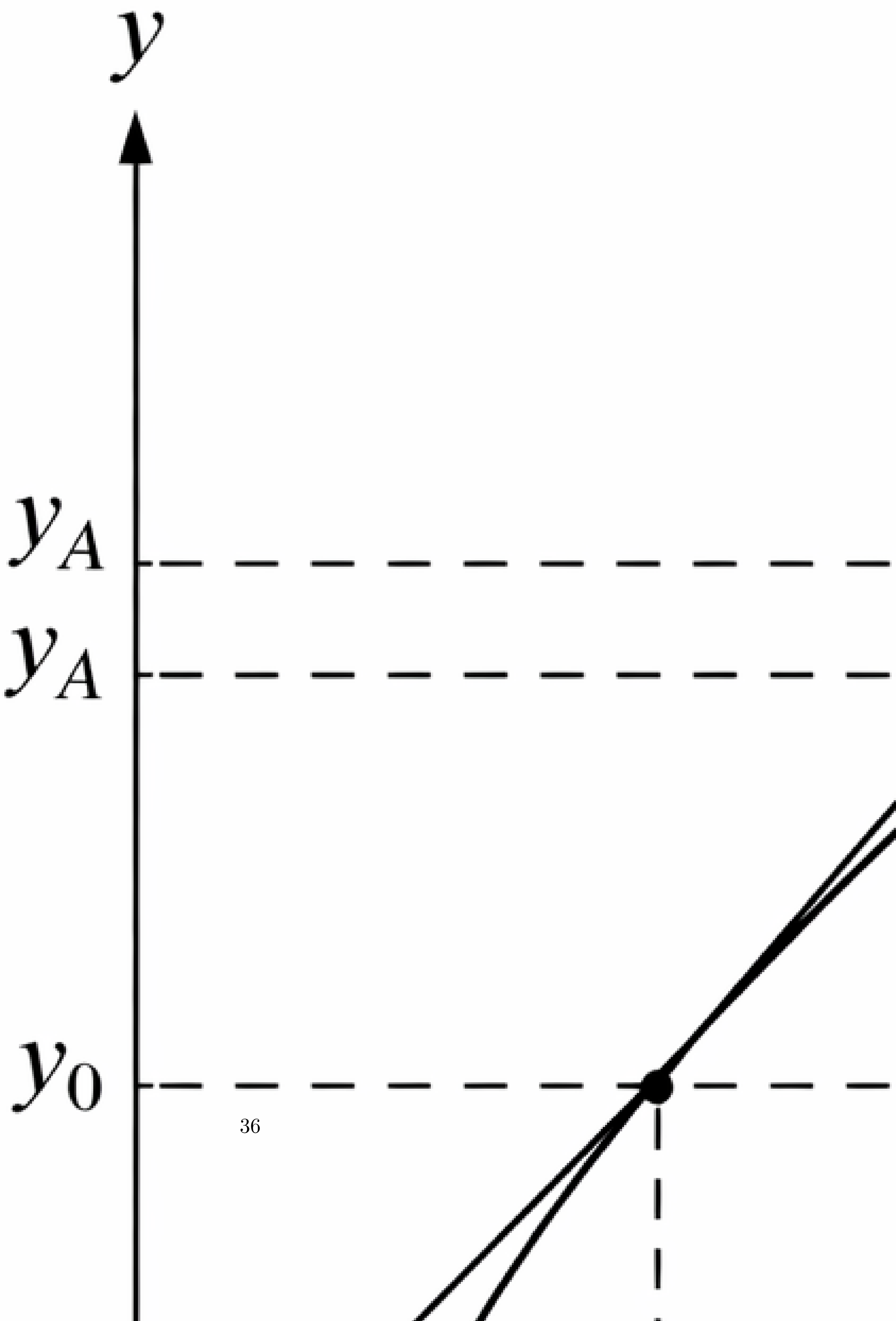
4.10 The Differential

Define dx as an arbitrary change in x from its initial value x_0 and dy as the resulting change in y **along the tangent line** from the initial value of the function $y_0 = f(x_0)$.

The **differential** of $y = f(x_0)$ evaluated at x_0 is

$$dy = f'(x_0)\,dx.$$

This represents the change in y along the tangent line at x_0. Graphically, we have:


y
y_A
y_A
y_0

4.11 Taylor Series (Preview)

A smooth complex function $z(x)$ can be approximated around $x = a$ by

$$f(x) = z(a) + z'(a)(x - a) + \frac{1}{2}z''(a)(x - a)^2 + \frac{1}{6}f'''(a)(x - a)^3 + \cdots$$

This idea underlies many approximation methods in economics.


y

The image provided illustrates a function $z(x)$ being approximated by three different Taylor polynomials (or Taylor series expansions) centered around the point $x = a$.

The simplest approximation perhaps would simply be $g(x) = a$. This constant-valued function does not work well, especially if we move away from the point a.

A better approximation would be a linear function of the form $h(x) = z(a) + b(x - a)$, where b is some slope. But what would be a good value of b? We saw above that the differential is an equation for the tangent line (or slope) at the point $x = a$. So, we could argue that the best **linear approximation** to the function around this point would be

$$h(x) = z(a) + z'(a)(x - a)$$

where $z'(a)$ is the derivative of the function evaluated at $x = a$.

But why stop here? We could improve on this. A better approximation could allow for some curvature. The general form would then be, say, $f(x) = z(a) + z'(a).(x - a) + c.(x - a)^2$. Again, we ask, "What would be the best value for c?" The rate of change of the slope of the quadratic approximation should be equal to the rate of change of change of the function at the a. And since the second derivative of $z(x)$ is $2c$, then for $f''(x)$ to equal $z''(x)$ at $x = a$, we need $c = 1/2z''(a)$. Hence the **quadratic approximation** to the function aound $x = a$ is:

$$f(x) = z(a) + z'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

Exteding the above argument for cubic and higher-degree approximations, we could find the nth-degree approximation to the function $z(x)$, which we could call $m(x)$, around the point $x = a$ is

$$m(x) = \frac{z(a)}{0!} + \frac{z'(a)}{1!}(x - a) + \frac{z''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x - a)^n$$

where $f^{(n)}(a)$ is the n the derivative of $z(x)$ evaluated at $x = a$. The function $m(x)$ above is called the **n-th degree Taylor expansion series** of $z(x)$ evaluated at $x = a$.

To sum, $z(x)$ is the original function being approximated (the solid curve). $g(x)$ represents a constant function. $h(x)$ represents the first-order Taylor polynomial, i.e. a straight line that has the same value and slope as $z(x)$ at $x = a$ (or a tangent to $z(x)$ at $x = a$). The formula for $f(x)$ is $f(x) = z(a) + z'(a)(x - a) + \frac{1}{2}z''(a)(x - a)^2$ representing a second-order (quadratic) polynomial-the dashed curve, which matches better the function's value, slope, and concavity (curvature) at $x = a$. It is a better approximation of $z(x)$ near $x = a$ than the linear approximation $h(x)$ and of course the constant function $g(a)$. The graph demonstrates that as more terms are included in the Taylor polynomial, the approximation of the original function becomes more accurate over a larger range around the center point $x = a$.

Example
For example, consider the function

$$y = e^{x/2} - e^{-x/2},$$

expanded around the point $x = 2$.

Linear approximation
The linear approximation to this function is

$$h(x) = \left(e^1 - e^{-1}\right) + \left(\tfrac{1}{2}\left(e^1 + e^{-1}\right)\right)(x - 2).$$

Quadratic approximation
The quadratic approximation is

$$j(x) = \left(e^1 - e^{-1}\right) + \left(\tfrac{1}{2}\left(e^1 + e^{-1}\right)\right)(x - 2) + \left(\tfrac{1}{8}\left(e^1 - e^{-1}\right)\right)(x - 2)^2.$$

4.12 Implicit Differentiation

Let's consider a very simple function,

$$xy = 7.$$

Here, possible solutions include $(x, y) = (1, 7)$, $(7, 1)$, and so on. If we want to find the slope of this function, we can differentiate it.

Finding y'

To find y', we proceed as follows.

(a) We make the main assumption that y is a function of x, i.e. $y = f(x)$. We then differentiate both sides of the equation with respect to x.

Hence, we obtain

$$\frac{d}{dx}[xf(x)] = 0.$$

Using the product rule, this gives

$$1 \cdot f(x) + xf'(x) = 0.$$

Equivalently,

$$y + xy' = 0.$$

(b) Solving the resulting equation for y' gives

$$y' = -\frac{y}{x}.$$

So, if we substitute, for example, $x = 1$ and $y = 5$, we obtain the slope of the function at that point:

$$y' = -5.$$

4.13 Inverse Function Rule for Implicit Functions

We can show that

$$\frac{dy}{dx} = -\frac{f_x}{f_y}.$$

That is, if we have an implicit function written as

$$f(x, y) = 0,$$

then the derivative of y with respect to x can be obtained by:

1. differentiating f with respect to x to obtain f_x,
2. differentiating f with respect to y to obtain f_y,
3. taking the ratio $-\frac{f_x}{f_y}$.

This gives the derivative of the implicit function y with respect to x.

It often feels like magic — but it is simply a consequence of the chain rule.

Why this works
Since $f(x, y) = 0$ holds along the curve, differentiating both sides with respect to x and solving for y' naturally leads to the ratio $-\frac{f_x}{f_y}$.

4.14 Some Uses of Differentiation in Economics

Some common applications of differentiation include:

- Increasing and decreasing functions
- Relative extrema (maximum or minimum)
- Inflection points
- Optimization of functions

etc.

4.14.1 A CES production function example

Given the CES production function

$$Q = A\big[\alpha K^{-\beta} + (1-\alpha)L^{-\beta}\big]^{-1/\beta},$$

we can show that the elasticity of substitution is constant, as follows.

First-order conditions

The first-order conditions require that

$$\frac{\partial Q/\partial L}{\partial Q/\partial K} = \frac{P_L}{P_K}.$$

Using the generalized power function rule, we take the first-order partial derivatives.

For labor,

$$\frac{\partial Q}{\partial L} = -\frac{1}{\beta}A\big[\alpha K^{-\beta} + (1-\alpha)L^{-\beta}\big]^{-(1/\beta+1)}(-\beta)(1-\alpha)L^{-\beta-1}.$$

Canceling the $-\beta$ terms, rearranging $(1-\alpha)$, and adding the exponents $-(1/\beta)-1$, we obtain

$$\frac{\partial Q}{\partial L} = (1-\alpha)A\big[\alpha K^{-\beta} + (1-\alpha)L^{-\beta}\big]^{-(1+\beta)/\beta}L^{-(1+\beta)}.$$

Substituting $A^{1+\beta}/A^{\beta} = A$, we can write

$$\frac{\partial Q}{\partial L} = (1-\alpha)\frac{A^{1+\beta}}{A^{\beta}}\big[\alpha K^{-\beta} + (1-\alpha)L^{-\beta}\big]^{-(1+\beta)/\beta}L^{-(1+\beta)}.$$

From the CES production function,

$$A^{1+\beta}\big[\alpha K^{-\beta} + (1-\alpha)L^{-\beta}\big]^{-(1+\beta)/\beta} = Q^{1+\beta},$$

and

$$L^{-(1+\beta)} = \frac{1}{L^{1+\beta}}.$$

Thus,

$$\frac{\partial Q}{\partial L} = \frac{1-\alpha}{A^{\beta}}\left(\frac{Q}{L}\right)^{1+\beta}.$$

The marginal product of capital

Similarly,

$$\frac{\partial Q}{\partial K} = \frac{\alpha}{A^{\beta}}\left(\frac{Q}{K}\right)^{1+\beta}.$$

Dividing the two equations and equating the result to P_L/P_K (from the FOC) leads to the cancellation of A^{β} and Q:

$$\frac{1-\alpha}{\alpha}\left(\frac{K}{L}\right)^{1+\beta} = \frac{P_L}{P_K}.$$

Rearranging,

$$\left(\frac{K}{L}\right)^{1+\beta} = \frac{\alpha}{1-\alpha}\frac{P_L}{P_K},$$

and therefore,

$$\frac{K}{L} = \left(\frac{\alpha}{1-\alpha}\right)^{1/(1+\beta)}\left(\frac{P_L}{P_K}\right)^{1/(1+\beta)}.$$

Elasticity of substitution
Since α and β are constants, we can treat K/L as a function of P_L/P_K.
Let

$$h = \left(\frac{\alpha}{1-\alpha}\right)^{1/(1+\beta)}.$$

Then

$$\frac{K}{L} = h\left(\frac{P_L}{P_K}\right)^{1/(1+\beta)}.$$

The marginal function is

$$\frac{d(K/L)}{d(P_L/P_K)} = \frac{h}{1+\beta}\left(\frac{P_L}{P_K}\right)^{1/(1+\beta)-1}.$$

The average function is

$$\frac{K/L}{P_L/P_K} = h\left(\frac{P_L}{P_K}\right)^{1/(1+\beta)-1}.$$

Dividing the marginal function by the average function, we obtain the elasticity of substitution:

$$\text{MRS} = \frac{d(K/L)}{d(P_L/P_K)}\bigg/\frac{K/L}{P_L/P_K} = \frac{1}{1+\beta}.$$

This is constant, hence the CES production function exhibits **constant elasticity of substitution**.

Interpretation

- If $-1 < \beta < 0$, then MRS > 1.

- If $\beta = 0$, then MRS $= 1$ (Cobb–Douglas case).
- If $0 < \beta < \infty$, then MRS < 1.

Intuition: elasticity of substitution

The elasticity of substitution measures how easily a firm can substitute labor for capital when their relative prices change. In a CES production function, this elasticity is constant: it does not depend on the levels of K, L, or output. When the elasticity is high, firms can adjust input combinations easily in response to wage or rental-rate changes; when it is low, substitution is difficult and input proportions are relatively rigid. The parameter β governs this flexibility: values of β close to zero imply unit elasticity (the Cobb–Douglas case), while larger values of β imply more limited substitution.