# Part II, Day 2: Comparative Statics and Linear Algebra

Jacob Bradt[1]
jbradt@g.harvard.edu
Harvard Kennedy School

Harvard Public Policy, Health Policy, and Business Administration PhD Math Camp
18 August 2022

[1]These notes are a slight adaptation of previous versions by Robert French, Rebecca Sachs, Todd Gerarden, Wonbin Kang, Sam Richardson, and Jonathan Borck. I am deeply indebted to previous Math Camp TF's for their work on these notes.

# Outline

# Outline

# Total differentiation

- We reviewed how to compute how much a function changes when one variable changes – that's a partial derivative

- How does one compute how much a function changes when multiple variables change at the same time?

- To answer this question, we compute the **total differential**

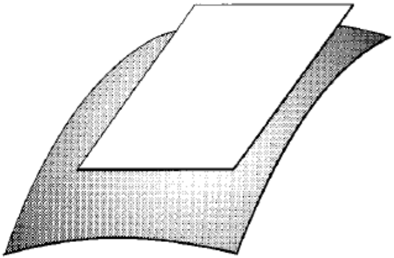
# Total differentiation

- To understand the total differential, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$
- Consider the tangent plane to $f$ at some point in the domain of $f$, $(x^*, y^*)$ (example at right)
- The change on the function $f$ is approximately the change $df$ *on the tangent plane*, given $\Delta x$ and $\Delta y$, which we call the **total differential** of $f$ at $(x^*, y^*)$

$$df = \frac{\partial f}{\partial x}(x^*, y^*)dx + \frac{\partial f}{\partial y}(x^*, y^*)dy$$

- This of course generalizes to functions of more than two variables; consider $f : \mathbb{R}^n \to \mathbb{R}$:

$$df = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} dx_j$$



# Total differentiation: Example #1

*Example:* Compute the total differential of the function $f(x, y) = x^2 + xy + y^2$.

Using the formula for the total differential:

$$\begin{aligned} df &= \frac{\partial f}{\partial x}(x, y)dx + \frac{\partial f}{\partial y}(x, y)dy \\ &= (2x + y)dx + (x + 2y)dy \end{aligned}$$

# Total differentiation: Example #2

*Example:* Compute the total differential of the savings function $S = S(y, i)$ where $y$ is income, $i$ is the interest rate, and $S$ is savings.
Using the formula for the total differential:

$$dS = \frac{\partial S}{\partial y}(y, i)dy + \frac{\partial S}{\partial i}(y, i)di$$

# Total differentiation: Example #3

*Example:* Compute the total differential of the utility function $U = U(x_1, x_2, \ldots, x_n)$.

Using the formula for the total differential:

$$\begin{aligned} dU &= \frac{\partial U}{\partial x_1}(\mathbf{x})dx_1 + ... + \frac{\partial U}{\partial x_1}(\mathbf{x})dx_n \\ &= \sum_{j=1}^{n} \frac{\partial U}{\partial x_j}(\mathbf{x})dx_j \end{aligned}$$

# Total derivatives

- How does one compute the derivative of a function when the arguments of that function are related?

- To answer this question, compute the **total derivative**

- To understand the total derivative, consider a function $f(x_1(t), \ldots, x_n(t)) : \mathbb{R}^n \to \mathbb{R}$
  - As the notation suggests, let us assume that the variables $x_1, \ldots, x_n$ themselves are related through $t$
  - To see how the function changes with $t$, we compute

$$\frac{df}{dt} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \frac{dx_i}{dt}.$$

  - This is what is known as the **total derivative** of $f$ with respect to $t$.

# Total differentiation: Example #1

*Example:* Compute the total derivative of the function $f(x, y) = x^2 + xy + y^2$ with respect to $x$.

There's actually nothing new to this: note that there is no explicit relationship between $x$ and $y$ specified in this setup. Standard partial differentiation gives:

$$\frac{df}{dx} = \frac{\partial f}{\partial x} = 2x + y$$

# Total differentiation: Example #2

*Example:* Compute the total derivative, $\frac{dS}{di}$, of the function $S = S(y, i)$, where $y = f(i)$.
Note now that $y$ is a function of $i$. So using our formula for the total derivative (i.e., the chain rule):

$$\begin{aligned}\frac{dS}{di} &= \frac{\partial S}{\partial y}\frac{\partial y}{\partial i} + \frac{\partial S}{\partial i}\frac{\partial i}{\partial i} \\ &= \frac{\partial S}{\partial y} f'(i) + \frac{\partial S}{\partial i}\end{aligned}$$

# Total differentiation: Example #3

*Example:* Consider the production function $F = F(K, L, t)$, where $L$ denotes labor, $K$ denotes capital, and $t$ denotes time. Assume that capital and labor both vary with time. Compute the rate of change of output with respect to time.

Note that output is a function of $K(t)$, $L(t)$, *and* $t$ itself. Using our formula for the total derivative (i.e., the chain rule):

$$
\begin{aligned}
\frac{dF}{dt} &= \frac{\partial F}{\partial K}\frac{\partial K}{\partial t} + \frac{\partial F}{\partial L}\frac{\partial L}{\partial t} + \frac{\partial F}{\partial t}\frac{\partial t}{\partial t} \\
&= \frac{\partial F}{\partial K}\frac{\partial K}{\partial t} + \frac{\partial F}{\partial L}\frac{\partial L}{\partial t} + \frac{\partial F}{\partial t}
\end{aligned}
$$

# Total differentiation: Example #4

*Example:* Given the production function $F(K, L)$, show that, along an isoquant,[2] the marginal rate of technical substitution is

$$MRS_K^L = \frac{dL}{dK} = -\frac{\frac{\partial F}{\partial K}}{\frac{\partial F}{\partial L}}$$

First, take the total differential of $F$:

$$dF = \frac{\partial F}{\partial K} dK + \frac{\partial F}{\partial L} dL$$

We know that along an isoquant, Firms' output does not change, so that $dF = 0$. Then, setting $\frac{\partial F}{\partial K} dK + \frac{\partial F}{\partial L} dL$ equal to 0, we can re-arrange to get the desired result.

[2] An isoquant is the set of production inputs that produce the same level of output.

# Implicit functions: Some definitions

- An **explicit function** is a function of the general form $y = F(x_1, ..., x_n)$ where the endogenous (dependent) variables (e.g., $y$) are solved in terms of the exogenous (independent) variables (e.g. $x$).
  - For each set of exogenous variables, there exists a *unique* value for each endogenous variable, e.g.

$$y = +\sqrt{1 - x^2} \qquad\qquad y = -\sqrt{1 - x^2}$$

- An **implicit equation** is of the form $F(y, x_1, \ldots, x_n) = 0$
  - For each set of $x$, there may exist multiple values for $y$ that solve the equation. e.g. consider the unit circle:

$$x^2 + y^2 - 1 = 0.$$

# Implicit functions: Some definitions

- If for each exogenous variable $(x_1, ..., x_n)$ an implicit equation determines a corresponding value $y$, then we say that the implicit equation defines the endogenous variable $y$ as an **implicit function** of the exogenous variables, e.g.:

$$y^2 + f(x)^2 - 1 = 0$$

- Just because we can write down an implicit equation, say $F(x, y) = c$ does not mean that this equation automatically defines $y$ as a function of $x$
  - Consider the implicit equation defining the unit circle

$$x^2 + y^2 - 1 = 0$$

  - When $x > 1$, there is no $y$ satisfying the above implicit equation
  - However, typically think of defining implicit functions at a specific solution to an implicit equation (e.g., $x = 0, y = 1$)

# Implicit functions: Some examples

- $y = 3x^4$ is an explicit function

- $y = f(x)$ is an explicit function

- The implicit equation $4x + 2y - 5 = 0$ defines $y$ as an implicit function of $x$

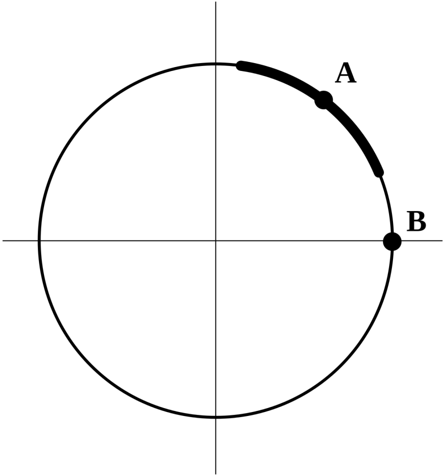
- $y^2 - 5xy + 4x^2 = 0$ defines $y$ as multi-valued implicit function of $x$

- $y^5 - 5xy + 4x^2 = 0$ is an implicit equation—which cannot be solved into an explicit function—that defines $y$ as a function of $x$ *for certain values of* $x$

- $F(y, x_1, \ldots, x_n) = 0$ is an implicit equation that may or may not implicitly define an implicit equation

# Implicit functions: Some examples

*Question*: Is $F(y, x) = x^2 + y^2 - 1 = 0$ an implicit function?



# Why do we care about implicit functions?

- Often in economics, we cannot easily express exogenous variables as a function of independent variables
- Frequently equations which arise natrually have the exogenous variables mixed in with the endogenous variables
  - For example, first order conditions in a maximization problem (e.g., profit maximization)

# Implicit Function Theorem

- When dealing with implicit functions, say $G(x, y) = c$, we may want to know how changes in $x$ affect $y$ in the neighborhood of a given solution, $(x_0, y_0)$

- **Implicit Function Theorem:** Given an *implicit equation* $F(y, x_1, \dots, x_n) = 0$, if
  1. $F$ has continuous partial derivatives, and
  2. at a point $(y_0, x_{10}, \dots, x_{n0})$ satisfying the equation $F(y, x_1, \dots, x_n) = 0$,

  $$\frac{\partial F}{\partial y}(y_0, x_{10}, \dots, x_{n0}) \neq 0,$$

  then there exists an $n$-dimensional neighborhood of $(x_{10}, \dots, x_{n0})$ in which $y$ is an implicitly defined function of the variables $(x_1, \dots, x_n)$.

# Implicit Function Theorem

Now, test whether the conditions of the Implicit Function Theorem hold for $F(y, x) = x^2 + y^2 - 9 = 0$. When does the implicit equation (implicitly) define $y$ as a function of $x$?

- The two conditions for us to use the implicit function rule is that the derivatives of $F$ with respect to $x$ and $y$ are continuous at $(0, 3)$ and that $\frac{\partial F}{\partial y}\Big|_{(0,30)} \neq 0$.
- Seeing that $\frac{\partial F}{\partial x} = 2x$ and $\frac{\partial F}{\partial y} = 2y$ are continuous for all values of $x$ and $y$, the partial derivatives are continuous at $(0, 3)$.
- We also have that $\frac{\partial F}{\partial y}\Big|_{(0,3)} = 6 \neq 0$

# Implicit function rule

- Now we know when an implicit equation defines a one-to-one (or many-to-one) relationship between its independent and dependent arguments

- With such a one-to-one or many-to-one mapping, it now makes sense to consider how the function's values (equivalently, its output $y$) changes when its arguments change

- The **implicit function rule** states:
  - Given an *implicit equation* $F(y, x_1, \ldots, x_n) = 0$, if an implicit function exists, the the partial derivatives are given by:

$$\frac{\partial y}{\partial x_i} = -\frac{\frac{\partial F}{\partial x_i}}{\frac{\partial F}{\partial y}}$$

# Implicit function rule: Example

*Example:* Compute the partial derivative $\frac{\partial y}{\partial x}$ of the implicit equation $F(y, x) = x^2 + y^2 - 9 = 0$ and evaluate it at $(0, 3)$.

We know that this defines an implicit function. We can apply our inverse function rule:

$$\begin{aligned}\frac{\partial y}{\partial x} &= -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} \\ &= \frac{-x}{y}\end{aligned}$$

which evaluated at $(0, 3)$ is 0.

# Applying the Implicit Function Theorem: Comparative statics

- As is hinted by the name, **comparative statics** typically concern how an optimum or an equilibrium condition changes when an underlying, exogenous parameter changes
- To see how equilibrium conditions or optimal choices change when such underlying, exogenous parameters change, we can employ the tools of the Implicit Function Theorem/implicit function rule
- We'll consider two examples to illustrate the concept of comparative statics

# Comparative statics: Tax incidence example

Market equilibrium is given by:

$$Q^d = D(p+t)$$
$$Q^s = S(p)$$
$$Q = Q^d = Q^s$$

where $Q^d$ is quantity, $Q^s$ is quantity supplied, $p$ is price, and $t$ is a tax levied on consumers. Determine how a change in the tax, $t$, affects the equilibrium price $p$.

- Assume consumers respond identically to changes in prices and taxes
- Hint: Try rewriting your solution in terms of elasticities of demand and supply
- If you use any theorems, note what assumptions you have made

# Comparative statics: Tax incidence example

We are looking for $\frac{dp}{dt}$. In equilibrium, we have that $D(p+t) - S(p) = 0$, so we can define $F(p,t) \equiv D(p+t) - S(p)$ to form an implicit equation. We can then go ahead an use the implicit function rule:

$$\begin{aligned}\frac{dp}{dt} &= -\frac{\frac{\partial F}{\partial t}}{\frac{\partial F}{\partial p}} \\ &= -\frac{D'(p+t)}{D'(p+t) - S'(p)}\end{aligned}$$

Since we have used the implicit function rule, we have assumed that $\frac{\partial F}{\partial p} = D'(p+t) - S'(p) \neq 0$, and that $F$ has continuous partial derivatives.

# Comparative statics: Tax incidence example

We can write this in terms of elasticities for a small enough tax rate by noticing that in equilibrium $Q = Q^d = Q^s$, so we can multiply the top and bottom of the quotient by $\frac{p}{Q}$, to get

$$\frac{dp}{dt} = -\frac{\varepsilon_D}{\varepsilon_D - \varepsilon_S}$$

where $\varepsilon_D = D'(p + t)\frac{p}{Q}$ and $\varepsilon_S = S'(p)\frac{p}{Q}$.

# Comparative statics: The optimal forest rotation problem

Imagine we have a forest whose value at time $T$ is given by the function $V(T)$, where $V''(T) < 0$. The value function $V(T)$ is maximized when $V'(T) = 0$, but if you can cut down and sell the trees, put your money in the bank, and earn interest rate $\delta$, then you wouldn't necessarily want to let the trees grow until they reach their maximum value. Assume the value function is globally nonnegative. Your goal is to maximize the net present value of the forest:

$$\max_T f(T) = e^{-\delta T} V(T)$$

1. Derive an expression for the optimal harvest time $T^*$, taking $\delta$ as fixed. Note any assumptions you make.
2. Calculate $\frac{dT^*}{d\delta}$, which shows how the optimal harvest time changes if the interest rate changes.

# Comparative statics: The optimal forest rotation problem

1. Derive an expression for the optimal harvest time $T^*$, taking $\delta$ as fixed. Note any assumptions you make.
   Here we just need to find the optimal value of $T$:

$$\frac{\partial f(T)}{\partial T} = -\delta e^{-\delta T} V(T) + e^{-\delta T} V'(T) \overset{\text{Set}}{=} 0$$

   which gives us the optimality condition

$$\delta V(T^*) = V'(T^*)$$

# Comparative statics: The optimal forest rotation problem

2 Calculate $\frac{dT^*}{d\delta}$, which shows how the optimal harvest time changes if the interest rate changes.
We can define an implicit equation as follows,

$$F(\delta, T) \equiv \delta V(T^*) - V'(T^*) = 0$$

and so by the implicit function rule,

$$\frac{\partial T^*}{\partial \delta} = -\frac{\frac{\partial F}{\partial \delta}}{\frac{\partial F}{\partial T^*}} = \frac{-V(T^*)}{\delta V'(T^*) - V''(T^*)}$$

where we have again assumed that $\frac{\partial F}{\partial T^*} \neq 0$ and that the partial derivatives are continuous.

# Key takeaways

- Know how to calculate a total derivative, total differential
- Be familiar with explicit/implicit function distinction
- Be familiar with the Implicit Function Theorem and applying the implicit function rule in comparative statics

# Additional resources

- Chiang and Wainwright, chapters 8.1-8.7
- Simon and Blume, chapters 14.4, 15.1 - 15.4

# Outline

# What is a matrix?

- A matrix is a rectangular array of numbers, variables, or functions. Matrices are used to efficiently store and manipulate information. Examples:

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{33} \\ a_{41} & a_{43} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

- Key matrix concepts:
  - **Elements:** Individual terms in a matrix (e.g., $a_{11}$).
  - **Equality:** Two matrices are equal if and only if they have the same dimension and identical elements: $A = B \iff a_{ij} = b_{ij}, \forall i, j$.
  - **Dimensions of a matrix:** (number of rows) x (number of columns)

# Important classes of matrices

- **Scalar:** a matrix with one row and one column

$$\begin{bmatrix} 7 \end{bmatrix}$$

- **Vector:** a matrix with either one row or one column. Here is a column and a row vector. When left unspecified, a vector typically refers to a column vector.

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \qquad B = \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}$$

- **Square matrix:** a matrix with the same number of rows as columns ($n = m$)

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

# Important classes of matrices

- **Symmetric matrix:** a matrix that equals its own transpose

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

- **Identity matrix:** a square diagonal matrix with all diagonal elements equal to one

$$I_1 = \begin{bmatrix} 1 \end{bmatrix} \qquad I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

- **Null matrix:** a matrix of zeros; often denoted $\mathbf{0}_{(n \times m)}$

$$\mathbf{0}_{(2\times 3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

# Important classes of matrices

- **Idempotent matrix:** a matrix that, when multiplied by itself, equals itself

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

- **Positive-definite matrix:** an $n \times n$ matrix is **positive definite** if:

$$v'Av > 0, \quad \forall v \in \mathbb{R}^n, \quad v \neq \mathbf{0}_{n\times 1}$$

Similarly, an $n \times n$ matrix is **positive semidefinite** if:

$$v'Av \geq 0, \quad \forall v \in \mathbb{R}^n$$

Flip inequalities for **negative definite** and **negative semidefinite**.

# Aside: Matrix definiteness applications

- Statistics: All variance-covariance matrices are positive semidefinite
- Optimization: The definiteness of a symmetric matrix is important for multivariate optimization.
    - These properties are roughly analogous to concavity and convexity; the Second Order Conditions of an optimization problem can be checked by determining whether the matrix of second derivatives (Hessian) is positive or negative definite[3]

[3] See Simon and Blume Chapters 16-18 for more information.

# Basic operations: Transposition

**Transposition:** a matrix with its rows and columns switched. The transpose of a matrix $A$ is denoted by $A'$ or $A^T$.

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

$$A' = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \dots & \dots & \dots & \dots \\ a_{1m} & a_{n2} & \dots & a_{nm} \end{bmatrix}$$

# Basic operations: Addition (and subtraction)

**Addition and Subtraction:** matrices must be *conformable for addition*—they must have equivalent dimensions. If the matrices are conformable, simply add or subtract the corresponding elements of the matrices.

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2m} + b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \dots & a_{nm} + b_{nm} \end{bmatrix}$$

# Basic operations: Scalar multiplication and dot product

**Scalar multiplication:** multiply each of the elements of the matrix by the scalar.
For $c \in \mathbb{R}$, we have,

$$cA = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1m} \\ ca_{21} & ca_{22} & \dots & ca_{2m} \\ \dots & \dots & \dots & \dots \\ ca_{n1} & ca_{n2} & \dots & ca_{nm} \end{bmatrix}$$

**Dot (scalar) product of two vectors:** to multiply a $1 \times n$ row vector by a $n \times 1$ column vector, multiply the corresponding elements, and add up the products. Note the length of the two vectors must be the same for the dot product to be defined.
Let $v_{n\times 1}$ and $u_{n\times 1}$, then

$$v \cdot u = v'u = \sum_{i=1}^{n} v_i u_i$$

The dot product is a special case of the inner product, specific to Euclidean space.

# Basic operations: Matrix multiplication

1. Are the matrices *conformable for multiplication*? The number of columns of the first matrix must equal the number of rows of the second matrix.
   - *Exception:* a scalar ($1 \times 1$ matrix) and matrix of any size can be multiplied together.

2. Determine the dimensions of the product. If $A$ is $n \times m$ and $B$ is $m \times q$, then $C = AB$ is $n \times q$.

3. Multiply! The element $c_{ij}$ of the product matrix $C$ is equal to the dot product of row $i$ of matrix $A$ and column $j$ of matrix $B$.
   - That is, $AB = [c_{ij}]$, where $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}$, for every $i$th row and every $j$th column.

# Basic operations: Practice problems

- Addition:

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} =$$

- Scalar multiplication:

$$4 \begin{bmatrix} 1 & 3 & 5 \\ -1 & -8 & 10 \\ -7 & -5 & 13 \end{bmatrix} =$$

- Matrix multiplication:

$$\begin{bmatrix} 2 & 0 & -1 & 1 \\ 1 & 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & -7 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} =$$

# Basic operations: Practice problems

- Addition:

$$\begin{bmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 7 & 5 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 8 & 5 \\ 3 & 3 \end{bmatrix}$$

- Scalar multiplication:

$$4\begin{bmatrix} 1 & 3 & 5 \\ -1 & -8 & 10 \\ -7 & -5 & 13 \end{bmatrix} = \begin{bmatrix} 4 & 12 & 20 \\ -4 & -32 & 40 \\ -28 & -20 & 52 \end{bmatrix}$$

- Matrix multiplication:

$$\begin{bmatrix} 2 & 0 & -1 & 1 \\ 1 & 2 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & -7 \\ 1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 11 & -15 \\ 5 & 7 & -7 \end{bmatrix}$$

## Rules/properties of matrix operations

- Commutative Law of Addition

$$A + B = B + A$$

- (No) Commutative Law of Multiplication

$$AB \neq BA$$

- Associative Law of Addition

$$(A + B) + C = A + (B + C)$$

- Associative Law of Multiplication

$$(AB)C = A(BC)$$

- Additive Inverse: For any matrix $A$, define $-A = [-a_{ij}]\ \forall i, j$. Then

$$A + (-A) = \mathbf{0}$$

- Distributive Law

$$A(B + C) = AB + AC$$

# Rules/properties of matrix operations

- **No "Zero Product Rule"**

$$AB = \mathbf{0} \not\Rightarrow A = \mathbf{0} \text{ or } B = \mathbf{0}$$

- **Properties of the Transpose**

$$(A^T)^T = A$$
$$(A + B)^T = A^T + B^T$$
$$(ABC)^T = C^T B^T A^T$$

- **Properties of the Identity Matrix:** For identity matrix $I_n$ and any $n \times n$ matrix $A$,

$$AI = IA = A$$
$$I(I) = I$$

- **Properties of the Null Matrix:** For any matrix $A$ which is $m \times n$,

$$A + \mathbf{0}_{(m\times n)} = A$$
$$A\mathbf{0}_{(n\times p)} = \mathbf{0}_{(m\times p)}$$

# Matrix rank: The fundamental criterion

- **Row (column) rank:** the number of linearly independent row (column) vectors, where a row (column) vector is linearly independent if it cannot be computed as a linear function of the other row (column) vectors

- **Rank:** the number of linearly independent rows or columns, or the row rank or the column rank (the two are always equal). The rank of a matrix is denoted $\text{rank}(A)$.

# Determining matrix rank

- For small matrices, it is often possible to determine the rank of a matrix by inspection, comparing rows or columns to test for linear dependence. For instance, find

$$\operatorname{rank}\begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} =$$

- Matrix rank can be computed using **Gauss-Jordan elimination**
  - Gauss-Jordan elimination: elementary row operations (addition, subtraction, and scalar multiplication)
  - GJ elimination reduces a matrix to **row echelon form**, in which each row has more leading zeros than the row preceding it

$$\begin{bmatrix} 2 & -4 \\ 0 & 2 \end{bmatrix}$$

  - Row echelon form can be used to determine the matrix rank, which is also defined as the number of nonzero rows of a matrix in row echelon form

# Gauss-Jordan elimination operations

- Operations:
  - Swapping two rows
  - Multiplying a row by a nonzero number
  - Adding a multiple of one row to another

- Example: Use Gauss-Jordan elimination to verify your solution by inspection for

$$\operatorname{rank}\begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} =$$

Adding 2 times the second row to the first row gives:

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

Swapping the rows gives

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

There is one row with non-zero elements $\Longrightarrow$ rank$=1$!

# Properties of the matrix rank

For an $m \times n$ matrix $A$,

- $\mathrm{rank}(A) \leq \min\{M, N\}$. If $\mathrm{rank}(A) = \min\{M, N\}$, matrix $A$ is of **full rank**.
- If $M = N$, $A$ is **invertible** if and only if $A$ has full rank.
- If the product $AB$ is well–defined, $\mathrm{rank}(AB) \leq \min\{\mathrm{rank}(A), \mathrm{rank}(B)\}$.
- Subadditivity: If the sum $A + B$ is well–defined, $\mathrm{rank}(A + B) \leq \mathrm{rank}(A) + \mathrm{rank}(B)$.

# Matrix rank: Examples

*Examples:* Find the maximum possible rank of each matrix based solely on dimensions.

$$\operatorname{rank}\begin{bmatrix} 2 & -4 & 3 \\ -1 & 2 & 0 \end{bmatrix} \leq$$

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

$$\operatorname{rank}\left(\begin{bmatrix} a & b \\ d & e \end{bmatrix}\begin{bmatrix} 2 & -4 & 3 \\ -1 & 2 & 0 \end{bmatrix}\right) \leq$$

# Matrix rank: Examples

*Examples:* Find the maximum possible rank of each matrix based solely on dimensions.

$$\operatorname{rank}\begin{bmatrix} 2 & -4 & 3 \\ -1 & 2 & 0 \end{bmatrix} \leq 2$$

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

$$\operatorname{rank}\left(\begin{bmatrix} a & b \\ d & e \end{bmatrix}\begin{bmatrix} 2 & -4 & 3 \\ -1 & 2 & 0 \end{bmatrix}\right) \leq 2$$

# Matrix rank: Examples

*Examples:* Use Gauss-Jordan elimination to compute

$$\operatorname{rank}\begin{bmatrix} 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \\ 1 & 3 & -8 & 4 \end{bmatrix} =$$

Does this matrix have full rank?

Is this matrix invertible?

# Matrix rank: Examples

*Examples:* Use Gauss-Jordan elimination to compute

$$\operatorname{rank}\begin{bmatrix} 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \\ 1 & 3 & -8 & 4 \end{bmatrix} =$$

Switch rows around: move row 3 to row 1, move row 1 to row 2 and row 2 to row 3

$$\begin{bmatrix} 1 & 3 & -8 & 4 \\ 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \end{bmatrix}$$

Subtract 1 times row 1 from row 2

$$\begin{bmatrix} 1 & 3 & -8 & 4 \\ 0 & 3 & 1 & -1 \\ 1 & 9 & -6 & 4 \end{bmatrix}$$

# Matrix rank: Examples

*Examples:* Use Gauss-Jordan elimination to compute

$$\operatorname{rank}\begin{bmatrix} 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \\ 1 & 3 & -8 & 4 \end{bmatrix} =$$

Subtract 1 times row 1 from row 3

$$\begin{bmatrix} 1 & 3 & -8 & 4 \\ 0 & 3 & 1 & -1 \\ 0 & 6 & 2 & 0 \end{bmatrix}$$

Subtract 2 times row 2 from row 3

$$\begin{bmatrix} 1 & 3 & -8 & 4 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

# Matrix rank: Examples

*Examples:* Use Gauss-Jordan elimination to compute

$$\operatorname{rank}\begin{bmatrix} 1 & 6 & -7 & 3 \\ 1 & 9 & -6 & 4 \\ 1 & 3 & -8 & 4 \end{bmatrix} = 3$$

Does this matrix have full rank? YES!

Is this matrix invertible? No, it's not square, but what's "invertible" and how would I know this?

# Matrix determinant

- Like rank, the determinant is a real-valued function of matrices
- The determinant of the $n \times n$ matrix $A$ is denoted $\det(A)$
- The key takeaway is the determinant is used to test whether a matrix is of full rank (i.e., $\det(A) \neq 0 \Leftrightarrow \mathrm{rank}(A) = n$), which also serves as a test of invertibility (more on this soon)
- In practice, you can use a computer to calculate the determinant of large matrices.

# Calculating the determinant

- **Calculating the determinant of a 2 x 2 matrix**

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$$

- **Calculating the determinant of a 3 x 3 matrix**

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

Simplify the determinant using the formula for the determinant of a $2 \times 2$ matrix

- **Calculating the determinant of a triangular matrix**: The determinant of a lower-triangular, upper-triangular, or diagonal matrix is simply the product of its diagonal entries: $\det(A) = \prod_{i=1}^{N} a_{ii}$. Verify this using the formula for a $2 \times 2$ matrix:

$$\det\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} = \qquad \det\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \qquad \det\begin{pmatrix} a & 0 \\ c & d \end{pmatrix} =$$

# Properties of the determinant

Given an $N \times N$ matrix $A$, matrix determinants satisfy the following properties:

- $\det(A) \neq 0 \Leftrightarrow \mathrm{rank}(A) = N$.
- If $\det(A) \neq 0$, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$.
- For any scalar $\alpha \neq 0$, $\det(\alpha A) = \alpha^N \det(A)$.
- If $A$ is diagonal or triangular, $\det(A) = \prod_{n=1}^{N} a_{nn}$.
- $\det(A^T) = \det(A)$.

# Matrix trace

- The trace is another real–valued function that can be defined only on square matrices, and corresponds to the sum of the terms along the main diagonal

- Given an $N \times N$ matrix $A$, the trace of $A$ is given by

$$\mathrm{tr}(A) = \sum_{n=1}^{N} a_{nn}.$$

- Given an $N \times N$ matrix $A$, matrix trace satisfy the following properties:
    - Linearity: $\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B)$ and $\mathrm{tr}(cA) = c\mathrm{tr}(A)$.
    - Invariance under Transposition: $\mathrm{tr}(A) = \mathrm{tr}(A^T)$.

# Matrix trace: Examples

*Examples:* Compute the trace of the following matrices:

$$\operatorname{tr}\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) =$$

$$\operatorname{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) =$$

$$\operatorname{tr}\left(3\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\right) =$$

# Matrix inverse

- Loosely speaking, the inverse of a matrix is a generalization of the scalar inverse.
- Formally, the inverse of a matrix $A$ is a matrix $A^{-1}$ such that $AA^{-1} = A^{-1}A = I$. A few notes:
  - Only square matrices have inverses.
  - A matrix possessing an inverse is also referred to as nonsingular, nondegenerate, or invertible.
  - If an inverse exists, then it is unique.
- The following definitions are equivalent to an $N \times N$ matrix $A$ being invertible:
  - $\mathrm{rank}(A) = N$.
  - The rows (or columns) of $A$ are linearly independent.
  - $\det(A) \neq 0$.
  - $A^T$ is invertible.

# Matrix inverse

- **Common use:** Solve systems of linear equations

- **Properties of Inverses:**

$$(A^{-1})^{-1} = A$$
$$(A^T)^{-1} = (A^{-1})^T$$
$$(AB)^{-1} = B^{-1}A^{-1}$$
$$(kA)^{-1} = k^{-1}A^{-1} \text{ for nonzero scalar } k$$

- **Handy fact:** The inverse of a diagonal matrix is a matrix of the (scalar) inverses of the diagonal elements

# Inverting matrices

Three procedures:

1. Special case of $2 \times 2$ matrices
2. Gauss-Jordan Elimination
3. Matrix inversion using minors, cofactors, and adjoint

# Inverting matrices: Special case of $2 \times 2$ matrices

**Formula for 2 x 2 Matrix**

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

*Example:* Compute

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

# Inverting matrices: Gauss-Jordan Elimination

1. Does the inverse exist? Check by verifying the matrix is full rank (e.g., is the determinant nonzero?).
2. Write the matrix in augmented form with the identity matrix.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad \left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right]$$

3. Perform row operations to reduce the original matrix to the identity matrix.
   This works for any square matrix, regardless of size.

*Example:* Compute

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

# Inverting matrices: Minors, cofactors, and adjoint

1. Does the inverse exist? Check by calculating the determinant.
2. Create the matrix of minors: for each element of the matrix, ignore values in the current row and column, and calculate the determinant of the remaining values.
3. Create the cofactor matrix: change the sign of every other cell in the matrix of minors according to the rule:
$$\sigma(i+j) = \begin{cases} 1, & \text{if } i+j \text{ is even} \\ -1, & \text{if } i+j \text{ is odd,} \end{cases}$$
4. Create the adjoint matrix, which is just the transpose of the cofactor matrix.
5. Divide the adjoint matrix by the determinant of the original matrix to find the inverse matrix.

*Example:* Use this method to verify

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

# Inverting matrices: Examples

*Example:* Determine whether the following matrix is invertible by computing its determinant. If it is, use the formula for a $2 \times 2$ matrix to compute:

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

*Example:* Determine whether the following matrix is invertible by computing its determinant. If it is, use Gauss-Jordan elimination to compute:

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

*Example:* Determine whether the following matrix is invertible by computing its determinant. If it is, use Gauss-Jordan elimination to compute:

$$\begin{bmatrix} 4 & 5 \\ 2 & 4 \end{bmatrix}^{-1} =$$

# Inverting matrices to solve systems of linear equations

- Suppose we've been given a system of linear equations:

$$\begin{aligned} 4x + 3y &= 28 \\ 2x + 5y &= 42 \end{aligned}$$

- Solutions methods:
  1. Regular old algebra
  2. Linear algebra

# Inverting matrices to solve systems of linear equations

1. Rewrite the system in matrix form: $Ax = b$.
2. Invert the matrix $A$. *Note: If the matrix is invertible, the system of equations has a unique solution.*
3. Use matrix multiplication to solve for $x$ using the formula $x = A^{-1}b$.
   - $\rightarrow$ To see this, pre-multiply the original equation by $A^{-1}$: $A^{-1}Ax = A^{-1}b \Rightarrow Ix = A^{-1}b \Rightarrow x = A^{-1}b$.

# Solving systems of linear equations: Example

*Example:* Invert the coefficient matrix ($A$) to solve the following system of equations:

$$\begin{aligned} 2x_1 + x_2 &= 5 \\ x_1 + x_2 &= 3. \end{aligned}$$

We can write this out in matrix form:

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

We can manipulate these set of matrices to give,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

which we can evaluate to,

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \quad \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

and so $x_1 = 2$ and $x_1 = 1$.

# Matrix calculus

- The gradient of a scalar function $f : \mathbb{R}^N \to \mathbb{R}$ is given by

$$\nabla f \equiv \frac{\partial f}{\partial x} = \left[ \begin{array}{ccc} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_N} \end{array} \right]'$$

  *Note:* This is a special case of the Jacobian, which generalizes the gradient of scalar functions to vector-valued functions.

# Matrix calculus

- The tangent vector to a curve $x : \mathbb{R} \to \mathbb{R}^N$ is given by

$$\frac{\partial x(t)}{\partial t} = \begin{bmatrix} \frac{\partial x_1(t)}{\partial t} \\ \vdots \\ \frac{\partial x_N(t)}{\partial t} \end{bmatrix}.$$

- Given the linear function $z = c^T x = x^T c$,

$$\frac{\partial z}{\partial x} = c$$

# Matrix calculus

- Given a matrix $A$ and the linear function $y = Ax$,

$$\begin{aligned}\frac{\partial y}{\partial x} &= A^T \\ \frac{\partial (x^T A x)}{\partial x} &= (A + A^T)x \\ \frac{\partial (x^T A x)}{\partial A} &= xx^T\end{aligned}$$

- Given $x, v \in \mathbb{R}^k$,

$$\frac{\partial x^T v}{\partial v} = \frac{\partial v^T x}{\partial v} = x$$

# Common Matrices in Economics and Econometrics

**Budget constraint:** $p \cdot x = w$ or $p'x = w$

**Quadratic matrix:** $X'X$—analogous to squaring a scalar; produces a symmetric matrix.

**Jacobian:** the matrix of all first-order partial derivatives.

# Common Matrices in Economics and Econometrics

**Hessian:** the matrix of all second-order partial derivatives. Note: this is a symmetric matrix (recall the order of partial differentiation does not matter).

**Sample mean:** If $X$ is a $n \times 1$ random matrix, the sample mean equals $\frac{1}{n}X'\mathbf{1}_n$.

**Sample variance-covariance matrix:** If $X$ is a $n \times m$ random matrix, the sample variance-covariance matrix is given by:

$$S = \frac{1}{n-1}X'\left(I_n - \frac{1}{n}\mathbf{1}_{n\times n}\right)X$$

# Common Matrices in Economics and Econometrics: Examples

*Example:* What are the Jacobian and Hessian for the function $f(x, y) = xy$

*Example:* What are the Jacobian and Hessian for the function $f(x, y) = 4x^2y - 3xy^3 + 6x$

# Key takeaways

- Be comfortable with the basic matrix concepts and important classes of matrices
- Know how to do basic matrix operations
- Be familiar with key properties of matrices (rank and determinant)
- Be able to compute a matrix inverse (and solve a system of linear equations)

# Additional resources

- Chiang and Wainwright, chapters 4-5
- Simon and Blume, chapters 8-11
- The Matrix Cookbook
- MIT OpenCourseWare (great video lectures)