BASICS OF MATHEMATICAL ECONOMICS

by Alex Nikolsko-Rzhevskyy, Ph.D.

University of Memphis

August, 2011

TABLE OF CONTENTS

WEEKLY SCHEDULE

Week 1: Quiz. Review of Basic Concepts

Week 2: Linear Models and Matrix algebra

- Matrices and vectors
- Matrix and vector operations
- Identity and null matrices
- Determinants and Nonsingularity

Week 3: Linear Models and Matrix algebra (cont.)

- Finding the Inverse Matrix
- Transposes, inverses, and operations with them
- Solving systems of linear equations

Week 4: Differentiation

- The Concept of Limit and Continuity
- The Derivative
- Rules of Differentiation of a Function of One Variable
- Rules of Differentiation of Two or more Functions of One Variable

Week 5: Differentiation (cont.)

- Partial Derivatives and Functions of Several Variables
- Differentials and Total Derivatives
- Implicit and Inverse Function Rules
- Comparative Statics

Week 6: Exam #1

Week 7: Optimization

- Increasing and Decreasing Functions , Concavity and Convexity
- Relative and Absolute (global) Extrema, Unconstraint Optimization
- Taylor and Maclaurin Series

Week 8: Optimization (cont.)

- Extreme Values of a Function of Two Variables
- Extreme Values of a Function of n Variables

Week 9: Optimization (cont.)

- Optimization with Bounding Constrains
- Optimization with Inequality Constrains

Week 10: Integration

- Indefinite Integrals
- Rules of Integration
- Integration by Substitution

Integration by Parts
Definite Integrals and Area Under the Curve
Improper Integrals
L'Hopital's Rule

Week 11: Exam #2

Week 12: Differential and Difference Equations

First Order Differential Equations
Higher Order Differential Equations

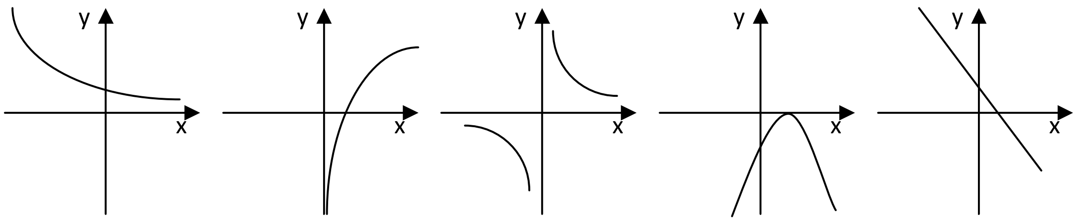
Week 13: Differential and Difference Equations (cont.)

Difference Equations in Discrete Time
Optimal Control Theory and the Hamiltonian

Week 14: Exam #3

QUIZ

1. Assign each graph 1 through 5 to a corresponding function $y(x)$ (write a letter below each graph):



   a. $y(x) = 5\ln(x)$
   b. $y(x) = \frac{5}{x}$
   c. $y(x) = -x^2 + 2x - 1$
   d. $y(x) = -2x + 3$
   e. $y(x) = \mathrm{e}^{-2\mathrm{x}}$

2. Find $\frac{dy}{dx}$ for $y(x) = x \ln x$

3. Find $\frac{dF}{dL}$ for $F(K, L) = AK^{\alpha}L^{1-\alpha}$

4. Find $\int xe^x dx$

5. Find $\lim_{\mathrm{x}\to 0} \frac{(1+\mathrm{x})^2 - 1}{\mathrm{x}}$

6. Is $y(x) = 2e^x$ convex, concave, or neither?

7. What is the slope of $y(x) = 2x^2 - x + 4$ at $x = 2$?

8. If vectors $A = (1, 1, 2)'$ and $B = (-1, 0, 3)'$, what is $A'B$?

9. Find $\det A$ if $A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & 3 & 4 \end{pmatrix}$

10. Find x and y that solve $\begin{cases} x = 3 - y \\ y = 2x + 9 \end{cases}$

11. Solve the differential equation $\dot{y} = 2y - 1$ given $y(0) = 2$.

12. $A = 2 + 3i$, $B = 1 - i$. Find AB.

13. Which trigonometric function is this: $\frac{1}{2}(e^{ix} + e^{-ix})$?

REVIEW OF BASIC CONCEPTS

1. Quantifiers are used to simplify the notation: $\exists =$ *exists*. $\nexists =$ *does not exist*. $\exists! =$ *exists only one*. $\in =$ *belongs*. $\cup =$ *unification*. $\cap =$ *intersection*. $\emptyset =$ *empty set*. $\forall =$ *any*. $\infty =$ *infinity*. $R =$ *real number*. $Q =$ *rational number*. $I =$ *irrational number*. $\setminus =$ *excluded*.

> *For example:* $I = R \setminus Q$.
>
> *It's also useful to formulate statements. For instance, for any linear function* $f(x) = kx + b$, *there exists one and only one real value of* $x = x^*$ *such as* $y(x^*) = 0$.

2. Exponents, x^n . Here, n is called an exponent. Also, you can read this as x raised to the power of n. Here are the basic rules of working with exponents:

> $x^0 = 1$ *(unless* $x = 0$, *for which this expression is undefined)*
>
> $x^a x^b = x^{a+b}$
>
> $\sqrt[n]{x} = x^{1/n}$
>
> $\frac{y^a}{y^b} = y^{a-b}$ *(regardless of whether* a, b *are positive of negative)*

3. Polynomials. Expressions consisting of a real number multiplied by a variable raised to the power of a positive integer are called monomials. Monomials can be added or subtracted to form polynomials. Monomials that have the same variables and exponents are called like terms.

> $(5x)(14y^2) = 65x^5$
>
> $(2x + 3y)(8x - 5y - 7z) = 16x^2 + 14xy - 14xz - 21yz - 15y^2$

4. Equations: Linear and Quadratic. If you have a mathematical statement that sets 2 algebraic expressions equal to each other, you have an equation. If all the variables are raised to the first power, you have a linear equation. If all the variables are raised either to the first or second power, you have a quadratic equation.

> $\frac{x}{4} - 3 = \frac{x}{5} + 1$, *Answer:* $x = 80$
>
> $5x^2 - 55x + 140 = 0$, *Answer:* $x = \frac{55 \pm \sqrt{55^2 - 4(5)(140)}}{2(5)} = \frac{55 \pm \sqrt{225}}{10} = (7; 4)$

5. Functions. A function is a rule that assigns to each value of a variable (e.g. x), called an argument, one and only one value $f(x)$, referred to as the value of the function at x. The domain of a function is the set of all possible values of x. The range is the set of all possible values of $f(x)$. Function can be Linear (the highest power of x is "1"), Quadratic (the highest power is "2"), Polynomial, Rational (ratio of 2 polynomials), power, etc. Note that any letter could be an argument:

> This is still a linear function: $y(e) = x^2 + e\,ln(x) + 7$.
>
> $f(x) = \frac{x^2-9}{x+4}$. *For this function, the domain is* $x \in R/\{2\}$ *(or simply* $x \neq 2$*), the range is* R*, or simply* $-\infty < x < +\infty$.

Other functions that you need to know: $ln(x)$ and $\exp(x) = e^x$, where $e = 2.718$. Note that $e^{\ln(x)} = x$, meaning that $ln(x)$ shows to which power the number e needs to be raised to obtain x.

6. Graphs, Slopes, and Intercepts. x is usually plotted on the horizontal axis and is called the independent variable; y is plotted on the vertical axis, and is called the dependent variable. The graph of a linear function $y = kx + b$ is a straight line. The slope k is that change in y, divided by the change in x; the greater – the steeper. The intercept b is the value of y when $x = 0$; it's also called the y-intercept. To easiest way to plot the line is to find points when $x = 0$ and $y = 0$. This is easily extended to non-linear functions.

> $2y - 6x = 12$, *Answer:* $y = 3x + 6$, *so points are* $(0; 6)$ *and* $(-2; 0)$.
>
> *Alternatively, you could have been given the 2 datapoints (say:* $(1;\ 2)$ *and* $(-2;\ 5)$*) and asked to find the parameters of original equation. So, first you write it as* $y = kx + b$ *and then take into account that you know 2 pairs of* (x, y)*:* $2 = 1k + b$ *and* $4 = -2k + b$ *and solve the system.*

Note that a function can have only one y-intercept (as there should be one and only one f(x) for each x), but it can have multiple x-intercepts.

> $y(x) = x^2 - x - 2$ *has only 1 y-intercept at* $y = -2$ *but 2 x-intercepts at* $x = 2$ *and* $x = -1$

LINEAR MODELS AND MATRIX ALGEBRA

WHY DO WE NEED TO KNOW LINEAR ALGEBRA

1. Linear algebra (1) Permits expression of a complicated system of equations in a simplified way (2) Allows to determine whether a solution exists before it is attempted (3) Provides the tools to solve the linear systems of equations.

> *Econometrics. From basic economics, we know that $GDP_t = \alpha + \beta \cdot G_t + \epsilon_t$. During recessions, GDP is down. To bring it up by ΔGDP, we need to increase G by $\frac{\Delta GDP}{\beta}$. Thus, need to "estimate" α and β from the data, and to do this, we need to be able to compute $(X'X)^{-1}(X'GDP)$, where GDP is the vector of historical GDP values and the matrix $X = (1\ G)$.*

> *Input-Output analysis. Usually, the production of one good requires the input of many other goods as "intermediate goods" (steel would require coal, electricity, iron ore, etc.). Also, steel itself is an intermediate good in producing some other goods. When a manager decides how much steel, x_i, to produce, he should take into account the final demand for steel, b_i, and intermediate demand for steel, required to produce other goods: $a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n$, where a_{ij} are "technical coefficients" expressing the value of input i, required to produce one dollar worth of product j. Thus, the total demand for product i would be $x_i = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + b_i$, or in matrix notation, $X = AX + B$. The goal is to find the value of "total output" X which has to be produced to satisfy total demand.*

MATRICES AND VECTORS

1. A matrix is a rectangular array of numbers, referred as elements. $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, elements of a matrix are referred to as a_{rc}. When the matrix is a square matrix, elements a_{ii} are called the principle or main diagonal.

2. The number of rows r and the number of columns c are the dimensions of the matrix (r x c). For a square matrix, $r = c$.

3. If a matrix consists of a single column, (r x 1), it's a vector, also referred to as a column vector. You can think of a matrix consisting of c column-vectors (1 x r), or of r row-vectors (1 x c).

4. If in a matrix, one of the rows (columns) is a linear combination of other rows (columns), that matrix is called singular.

MATRIX AND VECTOR OPERATIONS

1. Transposition. A matrix which converts rows of matrix A into columns, and columns into rows, is called the transpose of A and is designated by A' or A^T. A matrix for which $A = A'$ is called symmetric. $A' = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}' = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$. Similarly, if $v = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, then $v' = [3\ 5]$.

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

2. Addition and Subtraction. You can add and subtract only matrices of equal dimensions. Addition and subtraction is done element by element. Same is true for vectors, as vector is just a special case of matrices. Of course, $A + B = B + A$ and $A + (B + C) = (A + B) + C$.

3. Scalar multiplication. If you multiply a matrix A by a scalar k, each element of A gets multiplied by k. $kA = \begin{bmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{bmatrix}$. It's also obvious that $kA = Ak$.

4. Multiplication of row and column vectors. If a row vector $v = (1 \times n)$ and a column vector $u = (n \times 1)$ have the same number of elements, n, their product will be a scalar $C = v_{11}u_{11} + v_{22}u_{22} + \cdots + v_{nn}u_{nn}$.

$$v = [1 \quad 0 \quad -2] \ \textit{and}\ u = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix} \ \textit{then}\ C = vu = 1(1) + 2(0) + 4(-2) = -7.$$

Note that this formula works only if you multiply a column vector by a row vector, not otherwise. This will be explained later.

5. Multiplication of Matrices. Suppose you have $A = (n \times m)$, $B = (l \times k)$, and you need to find $C = AB$. To multiply A by B, they need to be conformable, i.e. $m = l$. Moreover, $(n \times m) \cdot (m \times k) = (n \times k)$, so C is a matrix which has the dimensions equal the first and the last dimension of the product matrices. In this matrix $C = (n \times k)$, each element c_{nk} (a scalar) is the product of n-th row of the matrix A onto k-th column of the matrix B. Thus, if $A = \begin{bmatrix} 1 & 3 \\ 2 & 8 \\ 4 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 5 \\ 9 \end{bmatrix}$, then $C = AB$ should be $(3 \times 2)(2 \times 1) = (3 \times 1)$, thus a column vector with the following elements: $C = \begin{bmatrix} 1(5) + 3(9) \\ 2(5) + 8(0) \\ 4(5) + 0(9) \end{bmatrix} = \begin{bmatrix} 32 \\ 82 \\ 20 \end{bmatrix}$.

Another example:

If $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ *and* $B = \begin{bmatrix} 2 & 0 \\ 3 & 0 \\ 0 & 4 \end{bmatrix}$, *then*

$$C = AB = \begin{bmatrix} 1(2) + 2(3) + 0(0) & 1(0) + 2(0) + 0(4) \\ (-1)2 + 0(3) + 1(0) & (-1)0 + 0(0) + 1(4) \end{bmatrix}.$$

A matrix for which $A \times A = A$ is called an idempotent matrix. Note that non-trivial matrices can have this property.

$\begin{bmatrix} 0 & 0 \\ \alpha & 1 \end{bmatrix}$ *is an example of an idempotent matrix.*

6. More on multiplication of vectors. In real life, you will be working with column vectors rather than row vectors. Suppose as before, $v = (n \times 1)$. As vectors is just a special case of matrices, this means you cannot multiply vectors by themselves as $(n \times 1)(n \times 1)$ is undefined. Instead, you can multiply $v' = (1 \times n)$ by $v = (n \times 1)$ and the result will be a (1×1) matrix, which is a scalar:

   If $= \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, v' *would be* $v' = [1\ 2\ 3]$ *and* $v'v = 1^2 + 2^2 + 4^2 = 21.$

   This explains that a product of a column vector and a row vector is a scalar.
   You can also multiply v by v' and the result will be an $(n \times n)$ matrix as $(n \times 1)(n \times 1) = (n \times n)$.

   If $= \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}$, v' *would be* $v' = [1\ 2\ 3]$ *and* $vv' = \begin{bmatrix} 1(1) & 1(2) & 1(3) \\ 2(1) & 2(2) & 2(3) \\ 4(1) & 4(2) & 3(3) \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 4 & 8 & 9 \end{bmatrix}$.

   Note that $v'v$ would decrease the dimension to (1×1), and vv' will increase the dimensions to $(n \times n)$.

7. Commutative, Associative, and Distributive Laws in Matrix Algebra.
   Matrix addition is commutative, i.e. $A + B = B + A$ and $A - B = -B + A$. Also, it's associative, i.e. $A + (B + C) = (A + B) + C$.
   Obviously, matrix multiplication is (generally) not commutative, i.e. $AB \neq BA$, but it's still associative, i.e. $A(BC) = (AB)C$. Also, the distributive law holds: $A(B + C) = AB + AC$.

IDENTITY AND NULL MATRICES

1. An identity matrix I is a square matrix that has 1 for every element on the principal diagonal and 0 everywhere else. Sometimes, I_n is written with a subscript n to specify its dimensions. For example, $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. The identity matrix is somewhat similar to the number 1, since $AI = IA = A$ (you can check that yourself). Also, $I \times I = I^2 = I$. The identity matrix is both symmetric and idempotent.

2. A null matrix is composed entirely of 0s and can be of any dimensions. It's not necessary a square. A product of 2 non-zero matrices can produce a null matrix (which wasn't the case in algebra).

   For example: $A = \begin{bmatrix} 6 & -12 \\ -3 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 12 & 6 \\ 6 & 3 \end{bmatrix}$, *and* $C = AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

DETERMINANTS AND NONSINGULARITY

1. The determinant of a matrix A is a scalar and it can be computed for any square matrix

2. The determinant $|A|$ of a 2×2 matrix is calculated as $\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{12}$. For a singular matrix (linearly dependent rows or columns), $\det A = 0$. The opposite is also true – if a matrix is nonsingular, its determinant is different from zero. Note that the determinant is a scalar and can be computed only for square matrices.

3. For a 3×3 matrix, the determinant can be computed in several ways. One of them is the following: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{21}a_{32}a_{13}) - (a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{23}a_{32}a_{11})$

4. A minor determinant or subdeterminant $|M_{jk}|$ is the determinant of a submatrix A obtained after deleting its jth row and kth column. For $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, $|M_{21}| = \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$.

5. A cofactor $|C_{jk}|$ is the product of $(-1)^{j+k}$ and the minor $|M_{21}|$: $|C_{jk}| = (-1)^{j+k}|M_{jk}|$

6. Laplace expansion and higher order determinants. Laplace expansion can be used to compute a determinant of any size by reducing its dimensionality to a 2×2 determinant. For any row or column (I will choose the first row as an example), it's true that: $|A| = a_{11}(-1)^{1+1}|M_{11}| + a_{12}(-1)^{1+2}|M_{12}| + a_{13}(-1)^{1+3}|M_{13}|$. As the elements of that row of column are multipliers before each minor, it usually makes sense to choose a row or a column with the highest number of zeros.

   Here I choose the last column: $\begin{vmatrix} 12 & 7 & 0 \\ 5 & 8 & 3 \\ 6 & 7 & 0 \end{vmatrix} = 0(-1)^4 \begin{vmatrix} 5 & 8 \\ 6 & 7 \end{vmatrix} + 3(-1)^5 \begin{vmatrix} 12 & 7 \\ 6 & 7 \end{vmatrix} + 0(-1)^6 \begin{vmatrix} 12 & 7 \\ 5 & 8 \end{vmatrix} = 3(-1)^5 \begin{vmatrix} 12 & 7 \\ 6 & 7 \end{vmatrix} = -3(84 - 42) = -126$.

   Using recursion, it's possible to reduce any dimension determinant to a 2×2 determinant.

7. Properties of a determinant.
   a. Adding or subtracting any linear combination of rows (columns) from one of the rows (columns) does not change the value of the determinant.
   b. Interchanging any two rows or columns of a matrix will change the sign of the determinant.
   c. The determinant of a matrix equals the determinant of its transpose.
   d. If all the elements in a row or column are zero, the determinant is zero.
   e. The determinant of a triangular matrix which has zeros under the principal diagonal equals to the product of values on the principal diagonal.
   f. The $\det kA = k^n \det A$.

FINDING THE INVERSE MATRIX

1. A cofactor matrix C is a matrix in which each element is replaced with its cofactor: $C = \begin{bmatrix} |C_{11}| & |C_{12}| & |C_{13}| \\ |C_{21}| & |C_{22}| & |C_{23}| \\ |C_{31}| & |C_{32}| & |C_{33}| \end{bmatrix}$.

If $A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 5 & 3 & 4 \end{bmatrix}$ *then* $C = \begin{bmatrix} \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} & -\begin{vmatrix} 4 & 2 \\ 5 & 4 \end{vmatrix} & \begin{vmatrix} 4 & 1 \\ 5 & 3 \end{vmatrix} \\ -\begin{vmatrix} 3 & 1 \\ 3 & 4 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ 5 & 4 \end{vmatrix} & -\begin{vmatrix} 2 & 3 \\ 5 & 3 \end{vmatrix} \\ \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 4 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 4 & 1 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -2 & -6 & 7 \\ -9 & 3 & 9 \\ 5 & 0 & -10 \end{bmatrix}$.

2. An adjoint matrix is the transpose of a cofactor matix.

$$Adj\ A = C' = \begin{bmatrix} |C_{11}| & |C_{21}| & |C_{31}| \\ |C_{12}| & |C_{22}| & |C_{32}| \\ |C_{13}| & |C_{23}| & |C_{33}| \end{bmatrix} = \begin{bmatrix} -2 & -9 & 5 \\ -6 & 3 & 0 \\ 7 & 9 & -10 \end{bmatrix}$$

3. An inverse matrix A^{-1} which can be found only for a square, non-singular matrix A, is a unique matrix such as $AA^{-1} = I = A^{-1}A$. The formula to find A^{-1} is the following: $A^{-1} = \frac{1}{|A|}C' = \frac{1}{|A|}Adj\ A$.

For the above setup $A^{-1} = \frac{1}{|A|}\begin{bmatrix} -2 & -9 & 5 \\ -6 & 3 & 0 \\ 7 & 9 & -10 \end{bmatrix} = \frac{1}{(-15)}\begin{bmatrix} -2 & -9 & 5 \\ -6 & 3 & 0 \\ 7 & 9 & -10 \end{bmatrix} = \begin{bmatrix} 2/15 & 3/5 & -1/3 \\ 2/5 & -1/5 & 0 \\ -7/15 & -3/5 & 2/3 \end{bmatrix}$.

Another example: $A = \begin{bmatrix} 4 & 1 & -5 \\ -2 & 3 & 1 \\ 3 & -1 & 4 \end{bmatrix}$ *then* $|A| = 98$, $C = \begin{bmatrix} 13 & 11 & -7 \\ 1 & 31 & 7 \\ 16 & 6 & 14 \end{bmatrix}$, *and thus* $A^{-1} = \frac{1}{|A|}C' = \begin{bmatrix} 13/98 & 1/98 & 16/98 \\ 11/98 & 31/98 & 6/98 \\ -7/98 & 7/98 & 14/98 \end{bmatrix}$.

Obviously, an inverse does not exists for singular matrices as their $\det A = 0$.

TRANSPOSES, INVERSES, AND OPERATIONS WITH THEM

1. Note the following rule of inverting a product of 2 matrices: $(AB)^{-1} = B^{-1}A^{-1}$. This can be used recursively for more complex products, e.g. $(ABC)^{-1} = C^{-1}(AB)^{-1} = C^{-1}B^{-1}A^{-1}$.
2. Similar rule holds for transposing a product of 2 matrices: $(AB)' = B'A'$.

Note that you can easily apply the same rule when working with triplets or quadruplets of matrices. For example, $(ABC)' = \left((AB)C\right)' = C'(AB)' = C'B'A'$

SOLVING SYSTEMS OF LINEAR EQUATIONS

1. Matrix representation of a system of linear equations. Suppose you have the following system of equations: $\begin{cases} 4x_1 + x_2 - 5x_3 = 8 \\ -2x_1 + 3x_2 + x_3 = 12 \\ 3x_1 - x_2 + 4x_3 = 5 \end{cases}$. It can be expressed in matrix form as $AX = B$, where $A = \begin{bmatrix} 4 & 1 & -5 \\ -2 & 3 & 1 \\ 3 & -1 & 4 \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, and $B = \begin{bmatrix} 8 \\ 12 \\ 5 \end{bmatrix}$. Here, A is the coefficient matrix, X is the unknown or solution vector, and B is the vector of constant terms or simply constants. Regardless of the dimensionality of the system, X and B will always be column vectors, and A will be a square matrix.

2. Solving the system by substitution. Often, this is the easiest method for small systems. Take the first equation, express, say, $x_2 = 8 - 4x_1 + 5x_3$ and plug it into the 2nd and 3rd equations, reducing dimensionality of the system. Then do the same with either x_1 or x_3.

3. Solving the system by elimination or triangulation. For this method, you take into account that the solution to the system stays unchanged if you subtract from each equation a linear combination of other equations or another row, premultiplied by a constant. You can do these operations to achieve zeros under the principle diagonal of A, keeping in mind that you also need to change B.
   The easiest way to do this is to write the combined $A|B$ matrix in the following form: $\begin{bmatrix} 4 & 1 & -5 & \vdots & 8 \\ -2 & 3 & 1 & \vdots & 12 \\ 3 & -1 & 4 & \vdots & 5 \end{bmatrix}$, which will be transformed into $\begin{bmatrix} 4 & 1 & -5 & \vdots & 8 \\ 0 & 7 & -3 & \vdots & 32 \\ 0 & 7/3 & -31/3 & \vdots & 4/3 \end{bmatrix}$, and then into $\begin{bmatrix} 4 & 1 & -5 & \vdots & 8 \\ 0 & 7 & -3 & \vdots & 32 \\ 0 & 0 & 28 & \vdots & 28 \end{bmatrix}$. If you now recall that this is just another representation of the system, it becomes clear that $x_3 = 1$ because the last equation says that $28\, x_3 = 28$. The rest of unknowns are determined by moving from the last equation to the first.

4. Solving the system with the inverse. This is the most useful method when working with real-life data and estimating regression using a computer. The original system you have is $AX = B$. Premultiply both sides by A^{-1} to obtain $X = A^{-1}B$. This is already the answer.

   Using the above example, $|A| = 98$, $C = \begin{bmatrix} 13 & 11 & -7 \\ 1 & 31 & 7 \\ 16 & 6 & 14 \end{bmatrix}$, *and thus* $A^{-1} = \frac{1}{|A|}C' = \begin{bmatrix} 13/98 & 1/98 & 16/98 \\ 11/98 & 31/98 & 6/98 \\ -7/98 & 7/98 & 14/98 \end{bmatrix}$. *The solution* $X = \begin{bmatrix} 13/98 & 1/98 & 16/98 \\ 11/98 & 31/98 & 6/98 \\ -7/98 & 7/98 & 14/98 \end{bmatrix} \begin{bmatrix} 8 \\ 12 \\ 5 \end{bmatrix} = \begin{bmatrix} \frac{104}{98} + \frac{12}{98} + \frac{16}{98} \\ \frac{88}{98} + \frac{372}{98} + \frac{30}{98} \\ -\frac{56}{98} + \frac{84}{98} + \frac{70}{98} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$.

Another example on solving the systems using linear algebra. Suppose that you have an input-output analysis problem when the production of one good requires inputs of many intermediate goods. This relationship between the good or industries is represented by the technology matrix A with elements a_{ij}. Then the total demand for each good x_i would be the sum of intermediate demand $\sum_j a_{ij}x_j$ and final demand b_i. If X is the vector of all goods x_i, then this problem can be written as a system $X = AX + B$. To solve it, rewrite it as $(I - A)X = B$. Then if you define $\tilde{A} = I - A$, you would get the usual case of a linear system with $\tilde{A}X = B$ with the solution $X = \tilde{A}^{-1}B$.

Suppose $A = \begin{bmatrix} .25 & .24 & .08 \\ .15 & .05 & .08 \\ .10 & .18 & .04 \end{bmatrix}$ *and it represents the technology matrix for 3 industries, Farming, Construction, and Clothing. Suppose* $B = \begin{bmatrix} 50 \\ 79.9 \\ 85.4 \end{bmatrix}$ *is the final demand. Then* $X = \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} .25 & .24 & .08 \\ .15 & .05 & .08 \\ .10 & .18 & .04 \end{bmatrix} \right)^{-1} \begin{bmatrix} 50 \\ 79.9 \\ 85.4 \end{bmatrix} = \begin{bmatrix} 100 \\ 110 \\ 120 \end{bmatrix}$.

5. Cramer's rule for matrix solutions. Cramer's rule provides a simplified method of solving a system of linear equations through the use of determinants. The rule states that $x_i = \frac{|A_i|}{|A|}$, where A_i is the matrix formed from the original coefficients matrix A by replacing ith column of coefficients with the vector B.

If $x_3 = \frac{1}{|A|} \begin{vmatrix} 4 & 1 & 8 \\ -2 & 3 & 12 \\ 3 & -1 & 5 \end{vmatrix} = \frac{1}{98} 98 = 1$. *This is incomparably easier than finding* A^{-1}.

6. If you have calculated $\det A$ and it appeared that $\det A = 0$, this means that one of the rows of A is a linear combination of other rows of A (i.e. A is singular), meaning that one of the restrictions that the system imposes on the x-vector is trivial and in some sense identical to one of the other restriction. This could mean that either the system has no solutions or that the system has infinite number of solutions, but it definitely does not have a unique solution.

The logic is simple and best understood from an example. Suppose the system is $\begin{cases} x_1 + 2x_2 = 5 \\ 2x_1 + 4x_2 = 11 \end{cases}$. *Here,* $\det A = 1(4) - 2(2) = 0$ *and you can also see that the first row of A is equal to the second row of A, multiplied by* $\frac{1}{2}$. *If we simplify the 2nd row by 2, we will get* $x_1 + 2x_2 = 5.5$. *Thus, the system states that the same linear combination of* x_1 *and* x_2 *should simultaneously be equal to 5 and 5.5, which cannot be. Thus, the system does not have any solutions. If instead, in the original system we had 10 instead of 11, we would again simplify the 2nd row by 2 and it would become exactly the same as the 1st row, thus not imposing any additional restrictions on* x_1 *and* x_2. *So the only restriction would be that* $x_1 + 2x_2 = 5$ *and there will be infinite number of pairs* x_1 *and* x_2 *that would satisfy it.*

DIFFERENTIATION

THE CONCEPT OF LIMIT AND CONTINUITY

1. L is called the limit of $f(x)$ as x approaches x_0, denoted as $\lim_{x\to x_0} f(x) = L$, if for any infinitely small positive number ϵ there exists a small positive number δ such as for any $x \neq x_0$ satisfying $|x - x_0| < \delta$, $f(x)$ would satisfy $|f(x) - L| < \epsilon$. This is the usual, two-sided limit. To give the definition of the limit, we have used epsilon-neighborhood theory.

2. Limits when x approaches infinity are defined in a similar way. For example, $\lim_{x\to +\infty} f(x) = L$ means that for any infinitely small number ϵ there exists some A such as for any $x > A$, $|f(x) - L| < \epsilon$.

3. A function is infinitely small if its limit is 0; it's infinitely large, if its limit is infinity. If for two infinitely small $\alpha(x)$ and $\beta(x)$, $\lim_{x\to x_0} \frac{\alpha(x)}{\beta(x)} = L$ and $L = 0$, then $\alpha(x)$ is of the higher order of infinitesimals (smallness) than $\beta(x)$; if $L = \infty$, then $\beta(x)$ is of the higher order of smallness than $\alpha(x)$; if L is a finite number, they are of the same order of smallness.

> *$2x^2$ is of the higher order of smallness than $7x$, when $x \to 0$.*

4. A function $f(x)$ is called continuous at $x = x_0$ if these conditions hold:
   a. $f(x)$ is defined, i.e. exists, at $x = x_0$, implying there are no holes on the graph
   b. $\lim_{x\to x_0} f(x) = f(x_0)$, which also of course implies that it exists

   In plain English, this means that the function needs to have no breaks in its curve (but it still can have kinks, i.e. can be non-differentiable). This also implies that its left and right limits exist and are equal.

> *$f(x) = \frac{x-3}{x^2-9}$ at $x = 3$ does have a limit equal $\frac{1}{6}$, but is not defined at that point and thus is not continuous.*

5. Assuming the $\lim_{x\to x_0} f(x)$ and $\lim_{x\to x_0} g(x)$ both exists, the following rules hold:
   a. $\lim_{x\to x_0} k = k$
   b. $\lim_{x\to x_0} k\, f(x) = k \lim_{x\to x_0} f(x)$
   c. $\lim_{x\to x_0}[f(x) \pm g(x)] = \lim_{x\to x_0} f(x) \pm \lim_{x\to x_0} g(x)$
   d. $\lim_{x\to x_0}[f(x) \cdot g(x)] = \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x)$
   e. $\lim_{x\to x_0} \left[\frac{f(x)}{g(x)}\right] = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}$

> *$lim_{x\to 4}[(x+8)(x-5)] = lim_{x\to 4}(x+8) \cdot lim_{x\to 4}(x-5) = -12$. Another example: $\lim_{x\to 0} \frac{3x+8}{x-2} = -4$. And one more: $\lim_{x\to\infty} \frac{3x+8}{x-2} = 3$.*

Always look at the order of magnitude.

6. There could be defined left and right limits, denoted as $\lim_{x \to x_0-} f(x) = L_1$ and $\lim_{x \to x_0+} f(x) = L_2$, respectively. The left limit means that x approached x_0 from the left, and the right limit – from the right. When the limit exists, both left and right limits are equal. However, it's possible for both the left and right limits to exists and be different. Then the usual two-sided limit does not exist.

7. The slope of a curvilinear function. First, the slope is generally not constant, except for linear functions. The slope of $f(x)$ at some point x_0 is measured by the slope of a tangent line at that point. To understand better how it works, imagine a secant line going through $(x_0, f(x_0))$ and $(x_0 + \Delta x, f(x_0 + \Delta x))$. Its slope would be $\frac{f(x_0+\Delta x)-f(x_0)}{(x_0+\Delta x)-x_0} = \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$. As Δx approaches zero, the secant linear become tangent, defining the slope as $= \lim\limits_{\Delta x \to 0} \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$. In textbooks, you can often see h in place of Δx.

8. The rules of limits allow you calculating slopes for any functions.

> *Let's try to find the slope of $f(x) = 2x^2$. Of course, it will be a function of x:*
> $S = \lim\limits_{\Delta x \to 0} \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x} = \lim\limits_{\Delta x \to 0} \frac{2(x+\Delta x)^2-2x^2}{\Delta x} = \lim\limits_{\Delta x \to 0} \frac{2x^2+4x\Delta x+2\Delta x^2-2x^2}{\Delta x} =$
> $\lim\limits_{\Delta x \to 0} (4x + 2\Delta x) = 4x$. *Thus, at $x = 3$, the slope would be 12, and at $x = -1$, the slope would be -3.*

THE DERIVATIVE

1. Given a function $f(x)$, the derivative of the function, written as $f'(x)$ or $\frac{dy}{dx}$, is defined as $f'(x) = \lim\limits_{\Delta x \to 0} \frac{f(x+\Delta x)-f(x)}{\Delta x}$ if this limit exists. Other ways to write a derivative are y'_x, y_x, and $\frac{d}{dx} f(x)$. The derivative with respect to time t, $\frac{dy}{dt}$, is commonly denoted as $\dot{y}$.

2. A function is differentiable at a point if it's is both continuous at that point and smooth (there are no kinks). The derivative measures both the slope and the rate of change of the original function $f(x)$ at a given point.

3. Higher order derivatives. In practice, you would need to calculate the second order derivatives and more. Derivatives of orders 1 through 3 are usually denoted as y', y'', y''', while derivatives of higher order n are usually labeled $y^{(n)}$. Finding higher order derivatives is not more complicated than finding the first order derivative. For example, the second derivative is just the first derivative of the first derivative.

> *If you need to find y'', when $y = x^7$, you first find $y' = 7x^6$, and then differentiate it one more time, resulting in $y'' = 42x^5$.*

RULES OF DIFFERENTIATION OF A FUNCTION OF ONE VARIABLE

1. Differentiation is a process of finding a derivative of a function by applying several basic rules to a given function. The main rules are listed below. You should also keep in mind that each of those rules could be derived by finding the corresponding limit:

Function $y(x)$	Derivative $y'(x)$
k	0
$k \cdot f(x)$	$k \cdot f'(x)$
x^n	nx^{n-1}
e^x	e^x
$\ln x$	$\frac{1}{x}$
$\frac{1}{x}$	$-\frac{1}{x^2}$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\cot x$	$-\frac{1}{\sin^2 x}$

2. The rules for sums and differences. If $f(x) = g(x) \pm h(x)$, then $f'(x) = g'(x) \pm h'(x)$.

$$\left(x^5 - \sqrt{x}\right)' = 5x^4 + \frac{1}{2\sqrt{x}}.$$

3. The product rule. If $f(x) = g(x) \cdot h(x)$, then $f'(x) = g'(x)h(x) + g(x)h'(x)$.

$$[x^4(x-5)]' = 4x^3(x-5) + x^4 = 5x^4 - 20x^3.$$

4. The quotient rule. If $f(x) = \frac{g(x)}{h(x)}$, then $f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{[h(x)]^2}$.

$$\left(\frac{5x^3}{4x+3}\right)' = \frac{15x^2(4x+3) - 5x^3(4)}{(4x+3)^2} = \frac{40x^3 + 45x^2}{(4x+3)^2}.$$

RULES OF DIFFERENTIATION OF TWO OR MORE FUNCTIONS OF ONE VARIABLE

1. The chain rule. This is the most useful rule you need to know for differentiation. If you have a composite function, also called a function of a function, in which $y = f(u)$ is a function of u, and $u = g(x)$ in turn is a function of x, then $\frac{dy(u(x))}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$ or simply $y'_x = y'_u u'_x$.

> *If $y = (5x^2 + 3)^4$, you can think of it as being $y = u^4$ and $u = 5x^2 + 3$. Then $y'_x = 4u^3(10x) = 4(5x^2+3)^3(10x) = 40x(5x^2+3)^3$. Here, $5x^2 + 3$ is sometimes called the inner function, while $(\cdots)^4$ is called the outer function. Thus, when looking for y'_x, you first take the derivative of the outer function, and then multiply it by the derivative of the inner function.*

> *Here is a more complicated example: $y = \frac{3x(2x-1)}{5x-2}$. When calculating $y'(x)$, treat this function as $\frac{h(x)}{g(x)}$ and you already know how to deal with them:*
> $y' = \frac{(3x(2x-1))'(5x-2)-(5x-2)'(3x(2x-1))}{(5x-2)^2}$*. The next step is finding derivatives of $h(x)$ and $g(x)$:* $y' = \frac{12x(5x-2)-5(6x^2-3x)}{(5x-2)^2} = \frac{30x^2-24x+6}{(5x-2)^2}$

2. Combination of rules. There are many ways to solve each problem using different rules. You should pick the one that's easier for you. As long as you apply the rules correctly, they should all lead to the same answer. As an example, let's look again at the quotient rule. Even though this looks like a separate rule, it's just a special case of the product rule, when $f(x) = g(x) \cdot h^{-1}(x)$, from which it follows that $f'(x) = g'(x)h^{-1}(x) - g(x)h^{-2}(x) = \frac{g'(x)h(x)-g(x)h'(x)}{[h(x)]^2}$. So, if it's easier for you to use the product formula – you can do that, too.

> $y(x) = \left(\frac{3x+4}{2x+5}\right)^2$ *then* $y'(x) = \frac{42x+56}{(2x+5)^3}$

3. Differentiating matrices is somewhat tricky, but you could use the following formulas to simplify things. Assume that v and x are $(n \times 1)$ vector, A is an $(n \times n)$ matrix. Then, for a linear function, $\frac{\partial v'x}{\partial x} = \frac{\partial x'v}{\partial x} = v'$. Similarly, $\frac{\partial Ax}{\partial x} = \frac{\partial x'A}{\partial x'} = A$. For quadratic forms, $\frac{\partial x'Ax}{\partial x} = x'(A' + A) = x'A' + x'A$.

PARTIAL DERIVATIVES AND FUNCTIONS OF SEVERAL VARIABLES

1. Before, we looked at the examples where y was a function of only 1 variable, x. In economics, it's often the case that y can be a function of several variables, for example, x and z, which will still be called independent variables. To measure the effect of a single independent variable at y, holding other independent variables constant, we need to calculate partial derivatives. They are defined similarly: $\frac{\partial y}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x,z)-f(x,z)}{\Delta x}$ (note that we are using ∂ instead of d). And

$\frac{\partial y}{\partial z} = \lim_{\Delta z \to 0} \frac{f(x,z+\Delta z)-f(x,z)}{\Delta z}$. Partial derivatives are often denoted as y'_x and y'_z.

> *Given $y = 2x^2z + z^3$, find y'_x. If you treat z as being a parameter, $y'_x = 4xz$. What if instead of y'_x you are being asked to find y'_z? Then z is the variable you are interested in, and x should be treated as a parameter. So, here, $y'_z = 2x^2 + 3z^2$.*

2. All the above rules would still work for partial derivatives.

> *The product rule: $y = (3x+5)(2x+6z)$. Then $y'_x = 3(2x+6z) + (3x+5)2$ and $y'_z = (3x+5)6$.*

> *The quotient rule: $y = \frac{6x+7z}{5x+3z}$. Then $y'_x = \frac{6(5x+3z)-(6x+7z)5}{(5x+3z)^2}$ and $y'_z = \frac{7(5x+3z)-(6z+7z)3}{(5x+3z)^2}$.*

3. Second-order partial derivatives. For the above functions, it's possible to find second order derivatives with respect to either x or z or even find the cross (or mixed) partial derivative. $f_{xx} = (f_x)_x = \frac{\partial}{\partial x}\left(\frac{\partial y}{dx}\right) = \frac{\partial^2 y}{\partial x^2}$ and $f_{xx} = (f_z)_z = \frac{\partial}{\partial z}\left(\frac{\partial y}{dz}\right) = \frac{\partial^2 y}{\partial z^2}$. Similarly, $f_{xz} = (f_z)_x = \frac{\partial}{\partial x}\left(\frac{\partial y}{dz}\right) = \frac{\partial^2 y}{\partial x \partial z} = f_{zx} = (f_x)_z = \frac{\partial}{\partial x}\left(\frac{\partial y}{dz}\right) = \frac{\partial^2 y}{\partial z \partial x}$. Note that it doesn't matter whether you first differentiate with respect to x and then z, or vice versa.

> *$y = 7x^3 + 9xz + 2z^5$. For this function, $y_x = 21x^2 + 9z$, $y_z = 9x + 10z^4$, $y_{xx} = 42x$, $y_{zz} = 40z^3$, and $y_{xz} = y_{zx} = 9$.*

DIFFERENTIALS AND TOTAL DERIVATIVES

1. For a function of 2 or more variables, the total differential measures the change in the dependent variable brought by a small change in each of the dependent variables. If $y = f(x,z)$, then the total differential is $dy = y'_x dx + y'_z dz$. Each of the terms is called the partial differential and that's the change in the function from a small change in one of the variables when the other is held constant.

> *A firm's costs are related to its production of goods x and z as $y = x^4 + 8xz + 3z^3$. Then $y_x = 4x^3 + 8z$, $y_z = 8x + 9z^2$, and $dy = (4x^3 + 8z)dx + (8x + 9z^2)dz$. This means that if you plan to increase output by dx (say, 5 units), costs would go up by $(4x^3 + 8z)dx$ (say, $5(4x^3 + 8z)$) for your levels of x and z. This works well for only $dx \ll x$.*

2. Total derivative. Suppose as before, $y = f(x,z)$, but this time $z = g(x)$, i.e. x and z are not independent. The total derivative measures both the direct effect of x on y, measured by its partial derivative, $\frac{\partial y}{\partial x}$, plus the indirect effect of x on y through $z(x)$, $\frac{\partial y}{\partial z}\frac{\partial z}{\partial x}$. In brief, $\frac{dy}{dx} = y'_x + y'_z \frac{dz}{dx}$. In other

words, when x changes by the small amount, it directly affects y, and also it affects z by the amount z'_x, and thus change in z affects y by the amount y'_z, hence the formula. Note that to find total derivative, you can first find the total differential, and then divide it by dx.

> *If $y = f(x,z) = 6x^3 + 7z^2$ and $z = g(x) = 4x^2 + 3x + 8$. Then $\frac{dy}{dx} = 18x^2 + 14z(8x+3) = 18x^2 + 14(4x^2+3x+8)(8x+3)$.*

> *A more complicated example: $y = 8x^2 + 3z^2$ and $x = 4t, z = 5t$. Then $\frac{dy}{dt} = 16(4t)4 + 6(5t)5 = 406t$. Note that in both cases, the answer is written only in your variable(s) of interest, and does not contain any intermediate steps variables.*

IMPLICIT AND INVERSE FUNCTION RULES

1. Implicit functions are functions where $y = f(x)$ explicitly express y in terms of x and thus are explicit functions. Functions $f(x,y) = 0$ do not express explicitly the dependency between y and x, and thus are called implicit functions. If we find the total differential of the above equation, we would get $f'_x dx + f'_y dy = 0$, and assuming $y'_x \neq 0$, $\frac{dy}{dx} = -\frac{f'_x}{f'_y} = -\frac{1}{f'_y/f'_x}$.

> *$f(x,y) = 3x^4 - 7y^5 - 86 = 0$. $(12x^3)dx - (35y^4)dy = 0$, thus $\frac{dy}{dx} = \frac{12x^3}{35y^4}$.Another way of doing, which would come in handy while doing comparative statics, is the following. As you are being asked to find y'_x, it means that you first need to acknowledge explicitly that $y = y(x)$. Then, as you are looking for a derivative wrt. x, you have to differentiate the whole equation wrt. x, which would give you $12x^3 - 35y^4 y'_x = 0$ and then solve for y'_x.*
>
> *Another example: $f(x,y) = 3x^2 + 2xy + 4y^3$. For this function, $y'_x = -\frac{6x+2y}{12y^2+2x}$.*

2. Given a function $y = f(x)$, an inverse function $x = f^{-1}(y) = g(y)$ exists if each value of y yields one and only one value of x. Assuming this holds and $dy/dx \neq 0$, $\frac{dx}{dy} = \frac{1}{dy/dx}$. The proof is obvious.

> *$Q(p) = 20 - 2p^3$. $\frac{dp}{dQ} = \frac{1}{dQ/dp} = \frac{1}{-6p^2}$. Another example: $Q = P^3 + 2P^2 + 7P$, then $P'_Q = \frac{1}{3P^2+4P+7}$.*

COMPARATIVE STATICS

1. Comparative static analysis or simply comparative statics allows studying the effects of changes in various exogenous parameters of the model on the equilibrium levels of endogenous variables. Essentially, comparative statics comes down to finding an appropriate derivative.

2. Comparative statics with specific functions. When you are given a specific function and analytically finding the equilibrium is feasible, doing comparative statics is trivial – you solve for the equilibrium level of variables, and differentiate them with respect of the parameter of interest.

> *For example, if you are looking at the goods market and quantity demanded is $Q^D = a_1 - a_2P + a_3w$, where P is price and w is wage, and quantity supplied is $Q^S = b_1 + b_2P$, with $a_i, b_i > 0$, then finding the effect of w (exogenous parameter) on the equilibrium levels of Q^* and P^* is easy. First, you solve the model by equating $Q^D = Q^S = Q$, and find the equilibrium $P^* = \frac{a_1 - b_1 + a_3 w}{b_2 + a_2}$ and $Q^* = b_1 + b_2\frac{a_1 - b_1 + a_3 w}{b_2 + a_2}$. From here, to find the effect of w on P^*, you should simply find ${P^*}'_w = \frac{a_3}{b_2 + a_2} > 0$, as when income goes up, the equilibrium price level also goes up.*

3. Comparative statics with general functions. The advantage of comparative statics is that you can actually analyze the effect of exogenous parameters on equilibrium even when you cannot analytically solve the model (which is often the case). This is easiest to illustrate by an example.

> *Suppose as before, $Q^D = F(P, w)$, with $F'_P < 0$ and $F'_W > 0$, and suppose $Q^S = G(P)$, with $G'_P > 0$. First, let's find the equilibrium. It's determined by $F(P, w) = G(P)$, where P is a variable and w is a parameter. Suppose we are interested in the effect of w on P (P'_w, just as before). First, we need to rewrite the equilibrium explicitly assuming that $P = P(w)$: $F(P(w), w) = G(P(w))$. Then we need to differentiate the equilibrium equation with respect to the parameter of interest, : $F'_P P'_W + F'_W = G'_P P'_W$. Now let's solve it for our derivative of interest: $P'_W = -\frac{F'_W}{G'_P - F'_P} = -\frac{(+)}{(-)-(+)} > 0$. Thus, we were able to achieve the same conclusion as before even without specifying the specific demand and supply function. So, this conclusion is much more general.*

> *Another example comes from the IS-LM model, in particular, the LM curve. The model assume that supply of money equals $\frac{M}{P}$, while the demand of money is represented by a general function, L, which depends on the interest rate r and income w. In equilibrium, they are equal, thus $\frac{M}{P} = L(r, w)$ and $L'_r < 0, L'_W > 0$. What is the effect of w on the equilibrium interest rate r? To answer, assume $r = r(w)$, treat the rest of the variable as fixed, and differentiate the equation with respect to w: $0 = L'_r r'_W + L'_W$. Thus, $r'_W = -\frac{L'_W}{L'_r} = -\frac{(+)}{(-)} > 0$.*

> *Yet another example: Assume a two-sector income determination model expressed in general function as $C = C(Y)$ and $I = I_0$ with equilibrium $Y = C + I$. Estimate the effect on the equilibrium level of output, Y^*, of a change in investment I_0. $Y(I_0) - C(Y(I_0)) - I_0 = 0$. Then $Y'_I - C'_Y Y'_I - 1 = 0$ and therefore ${Y^*}'_I = \frac{1}{1 - C'_Y}$.*

4. Comparative statics with systems of equations. In the above examples, we have only 1 endogenous variable in the equilibrium, P^*. Comparative statics could be used even if we have more than 1 endogenous variable and more than 1 equation. In that case we would have to assume that both endogenous variables change as we change the parameter. Again, it's best understood by going over an example.

Suppose we are looking at the IS-LM model, this time – at both the IS and LM curves as a system. The LM curve is $\frac{M}{P} = L(r, w)$ *with* $L'_r < 0, L'_w > 0$. *The IS curve is* $w = E(w, r, G, T)$, *with* $E'_w > 0, E'_r < 0, E'_G > 0, E'_T < 0$. *The variables here are* r *and* w*; the rest are the parameters. How does the equilibrium* w *changes as* P *changes? Again, start by assuming* $r = r(P)$ *and* $w = w(P)$. *Now let's differentiate both equations:* $\begin{cases} -\frac{M}{P^2} = L'_r r'_P + L'_w w'_P \\ w'_P = E'_w w'_P + E'_r r'_P \end{cases}$. *Here, 2 derivatives are unknown:* w'_P *and* r'_P *with the latter being our primary derivative of interest. As we have 2 equations and 2 unknowns, we can easily find it:* $w'_P = -\frac{M}{P^2}\frac{E'_r}{E'_r L'_w + L'_r(1 - E'_w)}$.

OPTIMIZATION

INCREASING AND DECREASING FUNCTIONS, CONCAVITY AND CONVEXITY

1. A function $f(x)$ is said to be increasing at $x = x_0$, if in the neighborhood of x_0, its graph is rising. For smooth functions, a positive first derivative, evaluated at $x_0, f'(x_0) > 0$, indicates that the function is increasing. Similarly, if $f'(x_0) < 0$, the function is decreasing.

> *If $y = x^2 + 2x - 3$, $y' = 2x + 2$, so it's increasing for $\forall x > -2$ and decreasing otherwise.*

2. If $\forall x$: $f'(x) > 0$ (i.e. it's positive over its entire domain), the function is said to be monotonically increasing. If instead $\forall x$: $f'(x) \geq 0$, it's called monotonically non-decreasing or just non-decreasing. Similar terminology is used in case of negative first derivative.

> *If $y = e^{2x}$, $y' = 2e^{2x} > 0$, so this function is monotonically increasing.*

3. A function is called convex at $x = x_0$, if the tangent line is above the graph. This also corresponds to $f''(x_0) > 0$. For concave functions, $f''(x_0) < 0$. Note that the sign of the first derivate is irrelevant to concavity-convexity: both increasing and decreasing functions can be either convex or concave.

> *If $y = x^2 + 2x - 3$, $y' = 2x + 2$ and $y'' = 2 > 0$, so it's convex for $\forall x$.*

RELATIVE AND ABSOLUTE (GLOBAL) EXTREMA, UNCONSTRAINT OPTIMIZATION

1. A relative extremum is a point at which a function is at a relative maximum or minimum, i.e. it's flat and neither increasing nor decreasing at x_0. Therefore, it's necessary that at that point, $f'(x_0)$ should either be zero or be undefined, and this is called the first derivative test.

2. Points x_0 at which $f'(x_0)$ is zero or undefined are called critical points. One function can have multiple relative extrema and multiple critical points. To find all critical points, you have to find $\forall x_0 : f'(x_0) = 0 \cup f'(x_0) = \emptyset$.

3. If the function is "smooth" (i.e. differentiable), then you would only need to consider cases when $f'(x) = 0$ when looking for critical points.

4. Suppose $f'(x_0) = 0$, so this is a critical point. For this relative extremum to be a relative maximum, it's sufficient that $f''(x_0) < 0$, i.e. the function is concave at x_0. A sufficient condition for a relative minimum would be $f''(x_0) > 0$, i.e. the function is convex. This is the second derivative test.

5. To sum up, for $f(x_0)$ to be a relative maximum or minimum, these 2 conditions must hold:
   a. Either $f'(x_0) = 0$ or $f'(x_0) = \emptyset$
   b. $f''(x_0) < 0$ (for a maximum) or $f''(x_0) > 0$ (for a minimum)

Suppose revenue is $R(Q) = 1200Q - 2Q2$, *while cost is* $C(Q) = Q^3 - 61.25Q^2 + 1528.5Q + 2000$. *How much the producer needs to produce to maximize his profits? Profits* $= -Q^3 + 59.25Q^2 - 328.5Q - 2000$. $\pi'_Q = -3Q^2 + 118.5Q - 328.5 = 0$. *There are two solutions:* $Q_1^* = 3$ *and* $Q_2^* = 36.5$. *To differentiate between a minimum and maximum, we need to find the 2nd derivative:* $\pi''_{QQ} = -6Q + 118.5$. *If we plug in the critical points,* $\pi''_{QQ}(3) > 0$ *and* $\pi''_{QQ}(36.5) < 0$. *Thus,* Q_1^* *is a local minimum (which we are not interested in), and* Q_2^* *maximizes profits, and thus is the solution we are looking for.* $\pi(Q_2^*) = \$16318.44$.

Suppose $R = 4000Q - 33Q^2$ *and* $C = 2Q^3 - 3Q^2 + 400Q + 5000$. *Then profits* $\pi = -2Q^3 - 30Q^2 + 3600Q - 5000$. *Thus,* $\pi' = -6(Q^2 + 10Q - 600)$ *with roots* $Q_1 = -30$ *and* $Q_2 = 20$. *Check* $\pi''(20) = -300 < 0$, *thus it's a relative maximum.*

6. Suppose $f'(x_0) = 0$, so this is a critical point. Also suppose that $f''(x_0) = 0$. In this case, the second derivative test is inconclusive. You can either try building the graph, or proceed with finding derivatives of a higher-order. If $f^{(2n)}(x_0) > 0$, this was a minimum, if $f^{(2n)}(x_0) < 0$ – a maximum.

7. For any value of $f^{(2n+1)}(x_0)$ (i.e. not necessary a critical point) but $f''(x_0) = 0$, then this is an "inflection" point – a point at which concavity changes (i.e. it's neither a maximum nor a minimum). At the inflation point, the tangent line crosses the function.

8. If the function is strictly concave or strictly convex, this relative extremum would actually be absolute or global extremum (either maximum or minimum). In all other cases (even if you have several extrema) when you are not optimizing on a closed interval of x, you are not able to claim that the relative max (or min) is indeed the global one. In fact, in most cases the global extremum will not even exist.

9. Finding global maxima and minima is the goal of optimization. If a function is continuous on a closed interval, then by the extreme value theorem global maxima and minima exist. Furthermore, a global maximum (or minimum) either must be a local maximum (or minimum) in the interior of the domain, or must lie on the boundary of the domain. So a method of finding a global maximum (or minimum) is to look at all the local maxima (or minima) in the interior, and also look at the maxima (or minima) of the points on the boundary; and take the biggest (or smallest) one.

Find global max and min over $x \in [-2; +2]$ *for the following function:* $f(x) = x^2 + 2x - 11$. $f' = 2x + 2 = 0$. *Thus, the critical point is* $x = -1$. *This is a local minimum, as* $f''(-1) = 2 > 0$. *The value of function* f *evaluated at the minimum is* $f(-1) = -12$. *There are no more critical points. Now we need to find the value of the function on the borders of the interval.* $f(-2) = -11$, $f(+2) = -3$. *Thus, as our global minimum and maximum should be a part of this set:* $\{-12, -11, -3\}$. *Obviously,* -12 *is the minimum, while* -3 *is the maximum, which correspond to* $x = -1$ *and* $x = +2$, *respectively.*

$f(x) = -x^3 + 6x^2 + 15x - 32$. For this function, $f' = -3(x+1)(x-5)$ with the roots $x = -1$ and $x = 5$ which lead to $f(-1) = -40$ and $f(5) = 68$. $f'' = -6x + 12$ and $f''(-1) = 18 > 0$ thus, it's a relative minimum and $f''(5) = -18$ and so it's a relative maximum. If you goal is to, for example, find the global minimum on $x = [0; 10]$, you would have to compare all relative minima (it's only one here, at $x = 5$, as $x = -1$ is outside the allowed interval of xs: $f(5) = 68$ to $f(0) = -32$ and $f(10) = -282$. Thus, the global minimum is achieved at $x = 10$ and equals -282.

TAYLOR AND MACLAURIN SERIES

1. Taylor's series allows to express an arbitrary differentiable up to the order n function $f(x)$ as a polynomial of order n. Taylor's theorem states that the function $f(x)$ can be expanded around the point x_0 as $f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + R_n$, where R_n is the denotes the remainder, which is of higher order of smallness than the previous term, i.e. $\frac{f^{(n)}}{n!}(x - x_0)^n \gg R_n \approx 0$.

   $f(x) = \frac{1}{1+x}$ around $x = 1$ behaves like the following linear polynomial: $f(x) \approx \frac{1}{1+1} + \frac{-1}{(1+1)^2}(x - 1) = \frac{1}{2} - \frac{1}{4}(x - 1) = \frac{3}{4} + \frac{1}{4}x$.

2. Maclaurin's series is the Taylor's decomposition, but around zero, thus $x_0 = 0$. The decomposition states that for $x \to 0,\ f(x) = f(0) + f'(0)(x) + \frac{f''(0)}{2!}(x)^2 + \cdots + \frac{f^{(n)}(0)}{n!}(x)^n$

   $f(x) = \frac{1}{1+x}$ for small $x \approx 0$ behaves like the following linear polynomial: $f(x) = 1 - x$.

EXTREME VALUES OF A FUNCTION OF TWO VARIABLES

1. For a 2-variable function $y = f(x, z)$ to be a relative maximum or minimum, these 3 (versus 2 for a 1-variable case) conditions must hold:
   a. $f'_x = f'_z = 0$. Note that the critical values will actually be critical pairs of values, (x_0, z_0).
   b. $f''_{xx}, f''_{zz} < 0$ (for a maximum) or $f''_{xx}, f''_{zz} > 0$ (for a minimum)
   c. $f''_{xx} \cdot f''_{zz} - (f''_{xz})^2 > 0$ (for either a maximum or a minimum). Note that this expression equals to the determinant of a matrix that consists of 2nd derivatives of f.

2. If f''_{xx} and f''_{zz} are of different signs, we have a saddle-path stable point, which is both the minimum on one axes and maximum on the other. If $f''_{xx} \cdot f''_{zz} - (f''_{xz})^2 < 0$, we have an inflection point. If the function is strictly concave (convex) in both x and y, there will be only one global maximum (minimum).

Maximize profits $z = 60x + 34y - 4xy - 6x^2 - 3y^2 + 5$. The first order conditions result in the following system of equation: $\begin{cases} 60 - 4y - 12x = 0 \\ 34 - 6y - 4x = 0 \end{cases}$, which has the solution $x = 4$, $y = 3$, or simply $(4,3)$. The second derivatives, $z_{xx} = -12$ and $z_{yy} = -6$ are both negative, and thus, $(4,3)$ could be a maximum. To know this for sure, we need to check the last condition. As $z_{yx} = z_{yx} = -4$ and since $(-12) \cdot (-6) - (-4)^2 > 0$, it's indeed a maximum.

The cost of producing goods x and z is given by $y = 3x^2 - xz + 2z^2 - 4x - 7z + 12$. Optimize it. Start by finding the first derivatives and equating them to zero. $y_x = 6x - z - 4 = 0$ and $y_z = -x + 4z - 7 = 0$. This is a linear system with only 1 solution: $x = 1$ and $z = 2$, or $(1,2)$. This is the critical point. Now we need to find the 2nd derivatives and evaluate them at $(1,2)$. $y_{xx} = 6$ and $y_{zz} = 4$, both positive. Note that formally, we need to plug $(1,2)$, but they are constants, so we are not doing this. If they are positive, it means that we might have found a minimum. Now let's find the cross-derivative: $y_{xz} = -1$. Because $6(4) - (-1)^2 > 0$, the costs are minimized as $(1,2)$ and this is the global minimum.

EXTREME VALUES OF A FUNCTION OF N VARIABLES

1. Before we start, we need to introduce some terminology. As before, the critical points are those for which $f_{x1} = f_{x2} = \cdots = f_{xn} = 0$. For a function $y = f(x_1, x_2, \dots, x_n)$, we can build a matrix consisting of its 2nd cross-derivates, evaluated at the critical points, $H = \begin{bmatrix} f_{x1x1} & \cdots & f_{x1xn} \\ \vdots & \ddots & \vdots \\ f_{xnx1} & \cdots & f_{xnxn} \end{bmatrix}$, called the Hessian matrix.

2. Leading minors (determinants) for this matrix are $|H_1| = f_{x1x1}$, $|H_2| = \begin{vmatrix} f_{x1x1} & f_{x1x2} \\ f_{x2x1} & f_{x2x2} \end{vmatrix}$, …, $|H_n| = \begin{vmatrix} f_{x1x1} & \cdots & f_{x1xn} \\ \vdots & \ddots & \vdots \\ f_{xnx1} & \cdots & f_{xnxn} \end{vmatrix}$. The Hessian matrix is positive definite, if all its leading minors are positive, and negative definite, if all odd leading minors are negative and all even minors are positive.

3. For an $n > 2$ case, it's not possible to graph the function. However, the conditions for it maximum or minimum are similar to the two-variable case. Suppose our function is $y = f(x_1, x_2, \dots, x_n)$. Then, for this function to achieve a relative maximum or minimum, these conditions need to hold:
   a. $f_{x1} = f_{x2} = \cdots = f_{xn} = 0$ (Necessary first order conditions)
   b. $|H_1|, |H_3|, \dots, |H_{2m+1}| < 0 \ \cap |H_2|, |H_4|, |H_{2m}| > 0$ (for a maximum)
      $|H_1|, |H_2|, \dots, |H_n| > 0$ (for a minimum)

4. Note that this also holds for a 2-variable case. For example, for a minimum $|H_1| > 0$ means that $f_{xx} > 0$, and $|H_2| = f_{xx}f_{zz} - (f_{xz})^2 > 0$ means that the second and third conditions hold, too. As for the 3rd condition, it's trivial that it would hold here. As for the second condition that $f_{zz} > 0$, note that is required for the above expression to be positive (given that $f_{xx} > 0$).

Find the extreme values of $y = 2x_1^2 + x_1x_2 + 4x_2^2 + x_1x_3 + x_3^2 + 2$. *The first-order conditions for extremum solve the following system* $\begin{cases} 4x_1 + x_2 + x_3 = 0 \\ x_1 + 8x_2 = 0 \\ x_1 + 2x_3 = 0 \end{cases}$ *with the only solution* $x_1 = x_2 = x_3 = 0$. *The Hessian* $H = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 8 & 0 \\ 1 & 0 & 2 \end{bmatrix}$, *whose leading principles are* $|H_1| = 4, |H_2| = 31, |H_3| = 54$. *Thus,* $(0,0,0)$ *is the minimum which results in* $y^* = 2$.

A more complicated example: $z = -x_1^3 + 3x_1x_3 + 2x_2 - x_2^2 - 3x_3^2$. *Taking the first derivatives leads to the 3 equations, where* $f'_{x_2} = 0$ *immediately results in* $x_2 = 1$. *As the first equation is a quadratic equation, the solution to the system is a pair of the following triplets:* $(0,1,0)$ *and* $(\frac{1}{2}, 1, \frac{1}{4})$, *which imply* $z_1^* = 1$ *and* $z_2^* = \frac{17}{16}$, *respectively. The Hessian matrix is* $H = \begin{bmatrix} -6x_1 & 0 & 3 \\ 0 & -2 & 0 \\ 3 & 0 & -6 \end{bmatrix}$. *For the first solution, the* $(1,1)$ *element is zero, so the sufficient condition is not satisfied as* $|H_1| = 0$. *After more careful examination, it's possible to show that the first solution is an inflection point. As for the second solution,* $|H_1| = -3, |H_2| = 6$, *and* $|H_3| = -18$, *so it's a maximum.*

OPTIMIZATION WITH BOUNDING CONSTRAINTS

1. Suppose you need to maximize $f(x_1, x_2)$ subject to a constraint $g(x_1, x_2) = k$. This is called constraint maximization with bounding constraints (versus unconstraint maximization you have performed before and maximization with inequality constraints that we will study later). There are 2 ways to approach this problem.

2. First, if the constraint is simple in the sense that you can explicitly express $x_1(x_2)$ or $x_2(x_1)$, you can plug it then back into $f(x_1, x_2)$, and proceed with a simple one-variable unconstraint maximization you have already know to perform. This method allows decreasing dimensionality of the problem by 1, making it much easier. Also, if instead you had 3 variables and 2 constraints, you could have done the same.

$f(x_1, x_2) = (x_1 + 2)^2 + x_2$, subject to $g(x_1, x_2) = 2x_1^2 + x_2 = 5$. Then from the constraint, $x_2 = 5 - 2x_1^2$, and $f(x_1) = (x_1 + 2)^2 + 5 - 2x_1^2 = -x_1^2 + 4x_1 + 9$. $f' = -2x_1 + 4 = 0$, so $x_1^* = 2$. $f''(2) = -2$, thus this is indeed a maximum. $x_2^* = 5 - 2x_1^{2*} = 5 - 2(2^2) = -3$, and $f^* = f(2, -3) = 13$.

3. In a more general case, when it's impossible to express one variable in terms of another variable, the problem can be solved by building the Lagrangian L that consists of the function f and the constraint, pre-multiplied by an artificial variable λ, called the lagrange multiplier: $L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda(g(x_1, x_2) - k)$. Then L is being treated as an general function that we need to maximize with respect to

3 variables, $x_1, x_2, \ \lambda$, without any additional constraints. Thus, we need to find the FOCs, solve the system of equations, and check the 2nd order conditions using the bordered Hessian (which I will not cover in this course).

> Maximize $f(x,y) = 4x^2 + 3xy + 6y^2$, *subject to* $g(x,y) = x + y = 56$. $L = (x_1 + 2)^2 4x^2 + 3xy + 6y^2 - \lambda(x + y - 56)$. *FOCs will result in the following system:* $\begin{cases} 8x + 3y - \lambda = 0 \\ 3x + 12y - \lambda = 0 \\ x + y = 56 \end{cases}$. *Note that FOC with respect to* λ *will always give you the constraint itself. The next step should be using the first 2 equations (that are the FOC with respect to real variables* x *and* y*, rather than the artificial variable* λ*) and eliminate* λ *from them, as we are not interested in it. Usually, this is done by expressing* λ *from one equation and plugging it into the second equation.* $\lambda = 8x + 3y = 3x + 12y$. *Thus, we end up with the system* $\begin{cases} 5x - 9y = 0 \\ x + y = 56 \end{cases}$, *which you can solve to find* $x^* = 36$ *and* $y^* = 20$.

> *Maximize* $U = x^{0.6}y^{0.25}$ *given that* $P_x = 8$, $P_y = 5$, *and a consumer can spend* $B = 680$. *This lead to the following Lagrangian:* $L = x^{0.6}y^{0.25} - \lambda(8x + 5y - 680)$, *which will eventually result in* $x^* = 60$ *and* $y^* = 40$.

4. λ is often referred to as the *shadow value of the constraint* (in economic applications it is called the marginal utility of income). It tells us by how much the maximum value of the objective function f changes when we relax the constraint, k, by one unit: $\frac{df^*}{dk} = \lambda$.

5. Note that in case you are asked to minimize a function (say, costs) instead of maximizing a function, you can multiply it by (-1), and then proceed with maximization. Another issue: some textbooks make a big deal of the sign before λ in the Lagrangian. It's actually not that important – in case you have a bounding constraint, the solution will stay the same whichever sign you use.

6. The case of 3 variables and 1 constraint. It's very similar to what we did before. Again, you would have to find the FOCs (this time – 4 of them), then use the first 3 to eliminate λ (this will leave you with 2 equations only in your main variables), which you will combine with the last FOC (the constraint) to get the system of 3 equations with 3 unknowns.

> *Maximize* $x^2 + 3xy + z^2$, *subject to* $x + y + z = 10$. $L = x^2 + 3xy + z^2 - \lambda(x + y + z - 10)$, *thus FOCs are* $\begin{cases} 2x + 3y - \lambda = 0 \\ 3x - \lambda = 0 \\ 2z - \lambda = 0 \\ x + y + z = 10 \end{cases}$. *The first 3 equations could be reduced to 2 equations if you for example, express* λ *from the first equation, and plug it into the second the third equations:* $\lambda = 2x + 3y$, *thus the system transforms into*

$\begin{cases} 3x - (2x + 3y) = 0 \\ 2z - (2x + 3y) = 0 \\ x + y + z = 0 \end{cases}$, *which is the system of 3 equations with 3 unknowns, with the solution* $(3, 1, 6)$.

7. The case of multiple constraints. Often, you need to maximize or minimize $f(x_1, x_2, \dots, x_n)$ subject to a set of (at most $n-1$) constraints $g_1(x_1, x_2, \dots, x_n) = k_1, g_{2(x_1, x_2, \dots, x_n)} = k_2, \dots, g_{n-1}(x_1, x_2, \dots, x_n) = k_{n-1}$. If you have several constraints, the solution is similar, but you have to have a separate λ_i for each of the constraints (so you have to introduce several artificial variables). Thus, $= f(x_1, x_2, \dots, x_n) - \lambda_1(g_1(x_1, x_2, \dots, x_n) - k_1) - \lambda_2(g_2(x_1, x_2, \dots, x_n) - k_2) - \dots - \lambda_{n-1}(g_{n-1}(x_1, x_2, \dots, x_n) = k_{n-1})$, which will give you n FOCs with respect to each x_i and at most $n-1$ additional restrictions from the constraints (each is the derivative $L'_{\lambda_i} = 0$).

Maximize $F = xyz$, subject to $x^2 + y^2 = 1$ and $x + z = 1$. $L = xyz - \lambda_1(x^2 + y^2 - 1) - \lambda_2(x + z - 1)$, which is a function of 5 variables. The FOCs will consist of 2 blocks of equations, FOCs with respect to x, y, z and to λ_1, λ_2 (which are just the constraints themselves). The first 3 equations are: $\begin{cases} yz - 2\lambda_1 x - \lambda_2 = 0 \\ xz - 2\lambda_1 y = 0 \\ xy - \lambda_2 = 0 \end{cases}$ *. As before, the best way to proceed it to concentrate on these equations first, disregarding the constraints. You goal is to eliminate λs. For example, you can express λ_2 from the last equation, λ_1 from the second equation, and plug them back into the first equation. This way the system will collapse into only 1 equation, which, however, will contain only x, y, z as variables. If you do this, and recall that you have 2 constraints, you will get the following system of 3 equations and 3 unknowns:* $\begin{cases} yz - 2x\frac{xz}{2y} - xy = 0 \\ x^2 + y^2 = 1 \\ x + z = 1 \end{cases}$*, which should be possible to solve. In fact, it will give you 4 triplets of solutions, but when you plug them back into $F = xyz$, you would see that only one of them $(-0.7676, -0.6409, 1.7676)$ leads to the maximum of F.*

OPTIMIZATION WITH INEQUALITY CONSTRAINTS

1. Kuhn-Tucker conditions. A typical problem with inequality constrains would be maximizing $f(x, y)$ subject to $g(x, y) \geq 0$ and $x, y \geq 0$. So solve this problem, you would have to use the Kuhn-Tucker conditions, which I will not cover in this class, but you can read about them in any of the recommended textbooks.

2. In many cases, this problem can be reduced to the one you already know how to solve. In the essence, this kind of strategy should lead you to the correct solution:
   a. Disregard the constraints and maximize $f(x, y)$.
   b. If the globally best x^*, y^* do not satisfy the inequality constraints solve several separate problems with equality constraints by using all possible combinations of constraints.
   c. Among all local extrema in (a) and (b), choose that duplet that leads to the highest value of $f(x, y)$.

Maximize utility $U = c^2 h$ from consuming Cola (c) and hotdogs (h) given that they cost $1 and $2, respectively, and you have only $12 to spend, meaning that you budget constraint is $c + 2h \leq 12$. Obviously, $c, h \geq 0$. Let's first disregard the constraint and maximize $U(c, h) = c^2 h$. It's an increasing function of both arguments, and it goes to infinity as $c, h = \infty$, which is outside the budget constraint. Thus, let's now turn to constraint maximization restricting $c + 2h = 10$. Then $L = c^2 h - \lambda(c + 2h - 12)$, which results in the following system: $\begin{cases} 2ch - \lambda = 0 \\ c^2 - 2\lambda = 0. \\ c + 2h = 12 \end{cases}$ Using the first 2 equations to eliminate λ leads to $\begin{cases} c^2 = 4ch \\ c + 2h = 12 \end{cases}$, which you can solve to find $h^* = 2$ and $c^* = 8$. This solution automatically satisfies $c, h \geq 0$, so we can stop here.

INTEGRATION

INDEFINITE INTEGRALS

1. If the derivative (or the rate of change) $f'(x)$ is known and the goal is to find the original function $f(x)$, we can achieve that by integration or anti-differentiation.

2. The original function $f(x)$ is called the integral or antiderivative of $f'(x)$.

Letting $y(x) = F'(x)$ for simplicity, the integral of $y(x)$ is expressed as $\int y(x)dx = F(x) + c$, where c is the constant of integration. The s-shaped symbol is called an integral sign, and $y(x)$ is the integrand function. Here, the left hand side of the equation is read "the indefinite integral of $y(x)$ with respect to x."

3. The indefinite integrals will still be functions of x and thus they could possibly take multiple values; hence the name.

RULES OF INTEGRATION

1. If you can differentiate well, integration shouldn't cause you any trouble. In fact, you should already know many important integrals as you already know which functions need to be differentiated to obtain them. The main rules are listed below:

Function $y(x)$	Its integral
$ky(x)$	$k\int y(x)dx$
$y(x) + g(x)$	$\int y(x)dx + \int g(x)dx$
k	$kx + c$

$\boldsymbol{x^n}$	$\frac{1}{n+1}x^{n+1} + c, n \neq 1$
$\boldsymbol{x^{-1}}$	$\ln\|x\| + c$
$\boldsymbol{e^{kx}}$	$\frac{1}{k}e^{kx} + c$
$\boldsymbol{\sin x}$	$-\cos x + c$
$\boldsymbol{\cos x}$	$\sin x + c$

2. Note that whenever an integral is computed, you have to add an arbitrary constant of integration, c, to the answer. Generally, you cannot find the specific value for this constant. The reason is that the derivative of a constant is zero. Thus, whenever you differentiate the right hand side expression in the above table, the constant will drop out and won't affect the answer.

3. Also note that the left hand side of the above table is just the derivates of the right hand side expressions. You can use this to check yourself whenever you are calculating an integral.

> *$\int 9e^{-3x}dx = 9\int e^{-3x}dx = 9\frac{1}{-3}e^{-3x} + c = -3e^{-3x} + c$. Let's check it by differentiation that we are correct: $(-3e^{-3x} + c)' = 9e^{-3x} + 0 = 9e^{-3x}$.*

> *Another example: $\int (3x^3 - x + 1)dx = \frac{3}{4}x^4 - \frac{1}{2}x^2 + x + c$.*

> *And two more examples: $\int \left(1 - \frac{2}{x}\right)dx = \int dx - \frac{2\int dx}{x} = x - 2\ln x + c$. And: $\int (\cos 3x + x^3)\, dx = \int \cos 3x dx + \int x^3 dx = \frac{1}{3}\sin 3x + \frac{1}{4}x^4 + c$.*

4. In many problems, in additional to the integral itself, you are given an initial $y(x = 0) = y_0$ or a boundary $y(x = x_0) = y_0$ condition. In that case you can pin down the exact value of the constant of integration, c, by plugging in the condition into the final equation and solving for c.

> *Given the boundary condition $y = 11$ when $x = 4$, find the integral $y(x) = \int xdx$. In general, $\int xdx = \frac{1}{2}x^2 + c$. Also, we know that when we plug $x = 4$, we should get $y = 11$. Thus, $11 = \frac{1}{2}4^2 + c$, leading to $c = 3$. So, our answer is $y = \frac{1}{2}x^2 + 3$, as out of the group of function with the derivative x, only this one satisfied the boundary condition we have.*

5. Note that even though you have identified the value of c, the answer is still a function of x and thus it still would be called an indefinite integral. The definite integrals would be those that have an exact value and don't depend on x.

INTEGRATION BY SUBSTITUTION

1. Often, the integrals look complicated and cannot be computed simply by looking at the table of integrals; then you would have to simplify the integral. One of the most popular methods is integration by substitution when you replace a complicated expression with another variable.

2. If the integrand can be expressed as a constant multiple of another function, u, and its derivative du/dx, integration by substitution is possible.

3. This could seen from this expression: $\int f(x)dx = \int \left(u\frac{du}{dx}\right) dx \rightarrow \int udu = F(u) + c.$

4. When integrating by substitution and replacing one variable with another, make sure you not only rewrite (a part of) the function $y(x)$ in terms of another variable, u, but also rewrite dx in terms of du:

> *For this integral, $\int 12x^2(x^3+2)dx$, let's define $(x^3+2) = u$. Then, if you calculate the differential, you would get $3x^2dx = du$. Now you can immediately rewrite the integral as $4\int udu = 2u^2 + c = 2(x^3+2)^2 + c$. So, your goal is to define a variable in such a way that when you would have to calculate du, it simplifies the integral even further.*

> *Some basic example: $\int (x+1)^2dx = \frac{1}{3}(x+1)^2 + c$. $\int \frac{1}{x+1}dx = ln(x+1) + c$.*

> *Another example would be $\int xe^{x^2}dx$. Here, you could define $u = x^2$, leading to $du = 2xdx$, which simplifies the original integral to $\int xe^{x^2}dx = \frac{1}{2}\int e^u du = \frac{1}{2}e^{x^2}$. And one more example: $\int (ln\,x^2)\frac{dx}{x} = [u = ln\,x^2, du = \frac{2x}{x^2}dx = \frac{2}{x}dx] = \frac{1}{2}\int u\,du = \frac{1}{4}u^2 + c = \frac{1}{4}(ln\,x^2)^2 + c$.*

INTEGRATION BY PARTS

1. Some integrals are impossible to calculate by substitution. For those, you could use another method of integration – by parts. To do this, you would have to split the integrand function into two parts, one of which you would have to drag under the "d" sign.

2. The formula works in the following way. Suppose your goal is to compute $\int y(x)g(x)dx$ and you could easily calculate $\int g(x)dx = G(x) + c$. Then $\int y(x)g(x)dx = \int y(x)d[G(x)] = y(x)G(x) - \int G(x)d[y(x)]$. Or alternatively, $\int y(x)g(x)dx = y(x)G(x) - \int G(x)y'(x)dx$.

3. This might look absolutely inefficient, until you realize that by using this complicated procedure you can decrease the order of $y(x)$ by 1, which is extremely efficient if $y(x)$ is a polynomial.

Suppose you need to find $\int x(x+1)^3 dx$. This is not an easy integral to calculate as it consists of 2 different functions, x and $(x+1)^3$. But this is a perfect example how integration by parts can help by getting rid of one of the two functions, x. In general, if you see a polynomial function multiplied by some other function, you have to use integration by parts and define $y(x)$ as that polynomial. Here, $y(x) = x$, $g(x) = (x+1)^3$, and $G(x) = \int (x+1)^3 dx = \frac{1}{4}(x+1)^4$. Thus, $\int x(x+1)^3 dx = \int xd\left[\frac{1}{4}(x+1)^4\right] = x\frac{1}{4}(x+1)^4 - \frac{1}{4}\int (x+1)^4 d[x] = \frac{1}{4}x(x+1)^4 - \frac{1}{4}\int (x+1)^4 dx$. Now this new integral is much easier to calculate than the old integral. For example, you can do this by substitution assuming $(x+1) = u$ and $dx = du$. Then $\int x(x+1)^3 dx = \frac{1}{4}x(x+1)^4 - \frac{1}{20}(x+1)^5 + c$.

$\int \frac{3x}{(x+1)^2} dx = \int (-1)3x\, d\left(\frac{1}{x+1}\right) = -\frac{3x}{x+1} + \int \frac{1}{x+1} d(3x)$ with the second integral being trivial.

In a similar way, $\int xe^x dx = \int xd[e^x] = xe^x - \int e^x dx = xe^x - e^x = e^x(x-1) + c$. To check this is true, simply differentiate this expression.

DEFINITE INTEGRALS AND AREA UNDER THE CURVE

1. Suppose you need to calculate the area under an irregularly shaped curve, $y = f(x)$, from $x = a$ to $x = b$. One way to do this would be dividing the x-interval into many small increments with coordinates $x_1, x_2, x_3, \dots, x_n$. with the distance between them being as small as possible. In that case area under the curve would be the following sum: $\sum_{i=1}^{n} f(x_i)\Delta x_i$, where Δx_i is the distance between the x's. This sum is called a Riemann's sum.

2. As Δx_i approaches zero, or alternatively, as $n \to \infty$, this sum closer and closer approximates area under the curve, and it becomes possible to express the Riemann's sum as a definite integral: $\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x_i$. The left hand side is read "the integral from a to b of f of x dx." Here, a is the lower limit and b is the upper limit.

3. The fundamental theorem of calculus states that the numerical value of the definite integral of a continuous function $f(x)$ over the interval from a to b is given by the indefinite integral $F(x) + c$ evaluated at the upper limit b, minus the same indefinite integral $F(x) + c$ evaluated at the lower limit of integration a. Since c is common to both, it could be dropped. Thus, $\int_a^b f(x)dx = F(x)|_a^b = F(b) - F(a)$.

$\int_1^4 10xdx = 5x^2|_1^4 = 5(4)^2 - 5(1)^2 = 75$. Another example: $\int_1^3 2x^3dx = \frac{1}{2}x^4|_1^3 = \frac{1}{3}3^4 - \frac{1}{2}1^4 = 40$.

4. Properties of definite integrals are quite similar to those of indefinite integrals. Some of them that still worth noting include $\int_a^b f(x)dx = -\int_b^a f(x)dx$ (which is obvious from the formula above) and $\int_a^a f(x)dx = 0$ (which is again trivial).

5. Another advantage of definite integrals is that they allow you calculating areas between the curves (not only below a curve). To do that, you would generally have to first draw a sketch of the graph of the functions to determine which of the two curves lies above the other. Then the area between the curves would be the difference between the areas under each of the curve.

For example, if you goal is to calculate the area between $y_1 = 3x^2 - 6x + 8$ and $y_2 = -2x^2 + 4x + 1$ for $x = [0; 2]$, you first should note that if you try plotting the curves, y_1 would be above y_2, and thus the area between the curved would be equal to $\int_0^2(3x^2 - 6x + 8)dx - \int_0^2(-2x + 4x + 1)dx = 7$.

6. The ability to calculate the area between the curves is extremely useful when you need to calculate consumer and producer surpluses. For example, suppose that $P_d = 25 - Q^2$ is the "demand" and $P_s = 2Q + 1$ is the "supply". In equilibrium, demand equals supply, and thus $P_d = P_s$. This results in $P^* = 9$ and $Q^* = 4$. If you now plot supply and demand, then the area between the demand curve and $P = 9$ would be the consumer surplus, and the area between $P = 9$ and the supply curve would be the producer surplus, for $Q = [0; 4]$.

IMPROPER INTEGRALS

1. A definite integral with infinity (positive or negative) for either an upper or lower limit of integration is called an improper integral: $\int_a^\infty f(x)dx$ or $\int_{-\infty}^b f(x)dx$.

2. Formally, you cannot substitute an infinity into $F(x)$, thus improper integral are defined as limits: $\int_a^\infty f(x)dx = \lim_{b\to\infty}\int_a^b f(x)dx = \lim_{b\to\infty} F(b) - F(a)$.

$\int_1^\infty \frac{3}{x^2}dx = lim_{b\to\infty}\int_1^b \frac{3}{x^2}dx = lim_{b\to\infty}\left(-\frac{3}{b}\right) - \left(-\frac{3}{1}\right) = -lim_{b\to\infty}\left(\frac{3}{b}\right) + 3 = 3$.

$\int_1^\infty e^{-x}dx = -e^{-x}|_1^\infty = -e^\infty + e^{-1} = \frac{1}{e}$.

L'HOPITAL'S RULE

1. Sometime, when calculating a limit (not necessary for the case of improper integrals), you would get an indeterminacy when both numerator and denominator approach either 0 or infinity (the so called $\frac{0}{0}$ or $\frac{\infty}{\infty}$ indeterminacy). For these cases, you cannot calculate the limit using traditional methods.

> $lim_{x\to 4}\frac{x-4}{16-x^2}=\frac{0}{0}$

2. You can cope with cases like that using the following rule: $\lim_{x\to a}\frac{g(x)}{h(x)}=\lim_{x\to a}\frac{g'(x)}{f'(x)}$, which is the L'Hopital's rule. Again, if one of the functions is a polynomial function, this simplifies things a lot as differentiation decreases the degree of that function by 1.

> $lim_{x\to 4}\frac{x-4}{16-x^2}=lim_{x\to 4}\frac{1}{(-2x)}=-\frac{1}{8}$. *Another example:* $\lim_{x\to 0}\left(\frac{1-e^{2x}}{x}\right)=\lim_{x\to 0}\left(\frac{-2e^{2x}}{1}\right)=$ -2.

3. Sometimes you have to apply the rule several times before you get rid of $\frac{0}{0}$ or $\frac{\infty}{\infty}$ indeterminacy. Note, however, that one the indeterminacy has disappeared, it's incorrect to apply L'Hopital's rule; moreover, it could lead to an incorrect answer.

> $\lim_{x\to\infty}\left(\frac{2x^3+x^2-x}{x^3-88}\right)=\lim_{x\to\infty}\left(\frac{6x^2+2x-1}{3x^2}\right)=\lim_{x\to\infty}\left(\frac{12x+2}{6x}\right)=\lim_{x\to\infty}\left(\frac{12}{6}\right)=2$. *Note that if you had applied the rule to the last ratio, you would have gotten a new indeterminacy of 0/0 instead of the correct answer "2".*

DIFFERENTIAL AND DIFFERENCE EQUATIONS

BASICS OF DIFFERENTIAL EQUATIONS

1. A differential equation is an equation which expresses as explicit or implicit relationship between a function $y=f(x)$ and one or more of its derivatives or differentials:

> *Some examples are* $\frac{dy}{dx}=5x+9$ *and* $y''-2y'+19=0$ *and* $y'=2y$. *For this last case, you can easily imagine a function whose derivative would be equal to the function itself, multiplied by "2": it's* $y=e^{2x}$. *Moreover, this would be true for the whole family of functions,* $y=ce^{2x}$.

2. The solution to a differential equation is any equation (function) without derivates that satisfies the differential equation for all the values of the independent variable(s) in the interval. There are 2 main ways to solve an equation: to "guess" the general form of the solution and then pin down the exact parameters, or to solve the equation by "brute force" by integration.

3. Equations with only one single independent variable are called ordinary differential equations. Differential equations which describe a relationship between a function of several variables and its partial derivatives are called partial differential equations (we will not cover them here). An ordinary differential equation

can include only a function of y on the RHS as in $y' = F(y)$; this will be called an autonomous (y depends only on a function of itself) or time independent (when x is time) differential equation.

4. A differential equation can also include only a function of x as in $y' = F(x) = x^2$ (in this case the equation can be solved by simple integration) or both the dependent and independent variables simultaneously as $y' = F(y, x) = y + x$. These are examples of non-autonomous or time-dependent equations.

5. The order of a differential equation is the order of the highest derivative in the equation. The degree of a differential equation is the highest power to which the derivative of highest order is raised.

> *This would be the first-order fourth-degree equation:* $\left(\frac{dy}{dx}\right)^4 - 5t^2 = 0$

The order of the equation determines the number of constants in the general solution.

LINEAR FIRST ORDER DIFFERENTIAL EQUATIONS

1. A simple example of a linear first order equation: $y' = ay$. The general solution to this equation is known and is $y(t) = ce^{at}$. You can "prove" this by simply plugging it back into the equation. To actually solve this equation by "brute force," you can use the separation of variables method described below.

2. A non-homogenous version of the above equation would be $y' = ay + b$, and it's called this way as we have an additional term b. Here, the solution would be $y(t) = -\frac{b}{a} + ce^{at}$. Note that $-\frac{b}{a}$ is also one of the solutions itself (called a "particular" y_p solution, rather than a "general" family of solutions). Moreover, this solution is constant in time – if $y = y_p$, the y variable stays the same indefinitely as t changes. This is called the "steady state solution" (sometimes, also called the stationary point, rest point, or equilibrium).

3. In general, when you have a non-homogenous equation, the general solution could be expressed as the sum of a particular solution of this non-homogenous equation ($y_p(t) = -\frac{b}{a}$) and the general solution to a homogenous version of this equation ($y_g(t) = ce^{at}$). This rule also holds for other more complicated differential equations.

4. Here is an example of a non-autonomous version of the above equation (the equation with time-varying coefficient): $y' = a(t)y$. Its general solution is also known and could be found as $y(t) = ce^{\int a(t)dt}$.

5. Finally, if you have a non-autonomous equation with both time-varying coefficients as in $y' = a(t)y + b(t)$. For this case, there exists a general solution $y(t) = e^{\int a(t)dt}\left(c + \int b(t)\, e^{-\int a(t)dt} dt\right)$ where c is an arbitrary constant. You can get this solution by first, multiplying the original equation by $e^{-\int a(t)dt}$ and then noting that $y'e^{-\int a(t)dt} - a(t)ye^{-\int a(t)dt} = \frac{d}{dt}\left(ye^{-\int a(t)dt}\right)$, hence the solution.

SEPARATION OF VARIABLES AND OTHER SPECIAL CASES

1. First-order nonlinear differential equation and separation of variables. In general, nonlinear equations are hard to solve. However, often, you have cases when you can rewrite the equation so that one of the variables appears on the left side of the equation and the other – on the right. This would be true for equations that look like $y' = g(y) \cdot h(x)$ and this procedure is called "separation of variables." The equation can then be solved by simple integration.

> *$\frac{dy}{dx} - y^2x = 0$. You can rewrite this as $\frac{dy}{dx} = y^2x$ and then as $\frac{dy}{y^2} = xdx$. Now integrate both sides, giving you $\int \frac{dy}{y^2} = \int xdx$, which results in $-\frac{1}{y} + c_1 = \frac{1}{2}x^2 + c_2$ or after you express $y = y(x)$ and group c_1 and c_2, the solution would become $y = -\frac{2}{x^2+c}$.*

Often, you can convert even more complicated differential equations to a one with separable variables and solve it exactly in the same way as the equation above.

> *$\frac{y'(x)}{x+1} = 2y(x) + 5$. Let's rearrange the terms as $\frac{dy}{dx} \cdot \frac{1}{2y+5} = x + 1$. Thus, $\frac{dy}{2x+5} = (x+1)dx$ and $\frac{1}{2}ln(2y+5) = \frac{1}{2}x^2 + x + c$. Therefore, $\sqrt{2y+5} = c \cdot e^{\frac{1}{2}x^2+x}$ and $y = \frac{1}{2}(c^2 \cdot e^{x^2+2x} - 5)$.*

2. Sometimes, non-linear equations for which separation of variables doesn't work could be converted into one of "easy" types by making a "smart" variable substitution. For example, if you have a general homogenous equation of the following form $y' = f\left(\frac{y}{x}\right)$, you can convert it to an equation with separable variables by first defining a new variable u such as $u = \frac{y}{x}$ (and thus $y' = u'x + u$). So the equation will become $u'x + u = f(u)$ or $\frac{du}{f(u)-u} = \frac{dx}{x}$.

> *$2y' = \frac{x+y}{x}$. Note that here you cannot simply separate the variables. However, you can take into account that this equation could be rewritten in terms of $\frac{y}{x}$ as $2y' = 1 + \frac{y}{x}$. Let's change the variables as described above: $2(u'x + u) = 1 + u$ which leads to $\frac{2du}{1-u} = \frac{dx}{x}$. Thus, $-2\,ln(1-u) = ln(x) + c$.*

3. The Bernoulli's equation. If you have an equation $y' = P(x)y + Q(x)y^n$, you can transform it into a linear equation after the following substitution: $y^{1-n} = u$. Indeed, you can rewrite the original equation as $\frac{y'}{y^n} = P(x)y^{1-n} + Q(x)$. If you differentiate the substitution, you would get $(1-n)y^{-n}dy = du$, which leads to $\frac{u'}{1-n} = P(x)u + Q(x)$ that you already know how to solve.

> *$y' = x^2y + xy^2$. This is another example where you cannot separate the variables. However, you can apply the above method to solve it by substituting $u = y^{1-n} =$*

$y^{1-2} = \frac{1}{y}$. Thus, $u' = -\frac{1}{y^2}y' = -u^2y'$, leading to $-\frac{u'}{u^2} = \frac{x^2}{u} + \frac{x}{u^2}$ or $-u' = x^2u + x$, which can be solved by using the first-order linear differential equation with varying coefficients method.

PHASE DIAGRAMS AND FIRST ORDER AUTONOMOUS DIFFERENTIAL EQUATIONS

1. Often in economics you are dealing with equations that do not have a closed form solution (for example, because the resultant integral could only be computed numerically. This is often true for non-linear equations). In those cases you can still say a lot about the equation and its solution using the "graphical" method, the so called "phase diagram" (or phase portrait).

2. Phase diagrams allow finding "steady states" – a very important in Economics type of solutions in which the system could stay indefinitely if not being "disturbed." For an autonomous equation $\dot{y} = F(y)$, this would mean that $\dot{y} = F(y) = 0$ as y should be constant in time while in a steady state.

3. In general, there could be 2 types of steady states: stable and unstable. If the system is in a stable steady state and is being hit by a small external shock (which takes it away from the steady state) the system will converge back to it. If, instead, that were an unstable steady state, the system would diverge from it.

4. The phase diagram is a plot of $\dot{y} = F(y)$ with $\dot{y}$ being the vertical axis and y being the horizontal axis. In some sense, you can think you are plotting $y = F(x)$ where $y = \dot{y}$ and $x = y$. Ideally, you would want to draw $F(y)$, but a sketch would suffice also. Your goal is to identify all the points when $\dot{y} = 0$ and then look for which intervals of y it is decreasing and for which – increasing.

For example, let the equation be $\dot{y} = y(2 - y)$. Our goal is to identify the steady states and their stability. The steady states are those ys for which $\dot{y} = 0$. Thus, we have 2 of them: $y = 0$ and $y = 2$. The function $F(y) = 2y - y^2$ is a flipped-over parabola. You can see that to the left of $y = 0$ and to the right of $y = 2$, it is negative, while between $y = 0$ and $y = 2$, it is positive. A negative $\dot{y} = F(y)$ (for example, to the left of $y = 0$) means that if y drops below 0, it will keep decreasing even further. Similarly, if it goes above 0, it will keep going up. You can draw arrows on the horizontal axis to help you visualize that. Thus, $y = 0$ is an unstable steady state. Using the same logic, you can show that $y = 2$ would be a stable steady state.

Another example: Suppose $\dot{y} = e^y(y - y^3)$. This function clearly doesn't have a nice looking closed form solution. So we have to study it graphically. The RHS of this equation has 3 roots: 0, +1, -1, and thus there are 3 steady states. Formally, now we need to draw $F(y) = e^y(y - y^3)$, however, all what really matters is where $F(y)$ is positive and where it's negative. Since e^y is always positive, $F(y)$ has the same sign as $y - y^3$ for all y. This implies that this equation has the same phase diagram as $\dot{y} = y - y^3$; the only difference is the speed of the motion. You can show that -1 and +1 are stable steady states, while 0 is an unstable steady state.

5. Another way to determine whether a steady state is stable or unstable is by finding the first derivate of the function and evaluating it at the steady state y_0. If $F'(y_0) < 0$ then $y_0$is stable and if $F'(y_0) > 0$ then y_0 is unstable. You can also see that from the previous example when you plot the function: When the function slopes downwards at a steady state, that steady state is a stable one. And vice versa.

> *For example, if $\dot{y} = \frac{y^2-1}{y^2+1}$, then there are 2 equilibria: $y = +1$ and $y = -1$. The derivative of the RHS is $F'(x) = \frac{4y}{(y^2+1)^2}$ and it's positive when $y = +1$ and negative when $y = -1$. Thus, the former is an unstable and the latter is a stable equilibrium.*

6. Note that very infrequently you can get a case when the steady state is neither stable nor unstable. This happens when both arrows point in the same direction. Those steady states are called either unstable from the right or unstable from the left, depending on the direction of the arrows.

HIGHER ORDER DIFFERENTIAL EQUATIONS

1. Second order linear differential equations. This equation has the following form $y'' + a(t)y' + b(t)y = f(t)$. If $f(t) \equiv 0$, the equation is called homogenous; otherwise it's called heterogeneous. If $a(t)$ and $b(x)$ are constants, the equation is called an equation with constant coefficients; otherwise, it's called an equation with varying coefficients.

2. To solve a linear homogenous equation with constant coefficients $y'' + ay' + by = 0$, you have to look for the solution of the form $y = e^{kt}$, which would lead to the following quadratic equation $k^2 + ak + b = 0$. This equation is called the "characteristic equation," and it has two roots, k_1 and k_2. The solution to the original equation will depend on nature of k_1 and k_2.

3. If the roots are real and different $k_1 \neq k_2$, then the solution to the equation is $y(t) = c_1e^{k_1t} + c_2e^{k_2t}$. You can easily check that by plugging this solution in the original equation. What you'll get are two quadratic equations in terms of k_1 and k_2, each of which would be equal to zero as both ks are the solutions.

> *Suppose $y'' + 9y' + 14y = 0$, then $k_1 = -2$ and $k_2 = -7$. Thus, the solution is $y(t) = c_1e^{-2t} + c_2e^{-7t}$. Another example: Suppose $y' - y' - 2y = 0$. The roots are 2 and -1. Thus, $y(t) = c_1e^{2t} + c_2e^{-t}$.*

4. If the roots are real and equal $k_1 = k_2 = k$, then the solution is $y(t) = c_1e^{kt} + c_2te^{kt}$. Note that there is an extra "t" after the second constant, c_2. Similarly, if they are the same but complex, the solution would be $y(t) = c_1e^{\alpha t}\cos(\beta t) + c_1te^{\alpha t}\cos(\beta t)$.

Suppose $y'' - 12y' + 36y = 0$. Then the roots to the characteristic equation are $k_1 = k_2 = 6$. Thus, the solution would be $y(t) = c_1 e^{6t} + c_2 t e^{6t}$. Another example: $y'' - 4y' + 4y = 0$. The roots are 2. Thus, $y(t) = 2e^{2t} + te^{2t}$.

5. If the roots are complex and different $k_1 \neq k_2$ and equal $\alpha \pm \beta i$, then the solution to the equation is still $y(t) = c_1 e^{(\alpha+\beta i)t} + c_2 e^{(\alpha-\beta i)t}$. However, you can go further and rewrite it as $y(t) = c_1 e^{\alpha t} e^{i\beta t} + c_2 e^{\alpha t} e^{i\beta t} = c_1 e^{\alpha t}\cos(\beta t) + c_2 e^{\alpha t}\sin(\beta t) = e^{\alpha t}(c_1 \cos\beta t + c_2 \sin\beta t)$ using the Euler formula that $e^{ix} = \cos x + i\sin x$ and some tricks with choosing the constants c_i. Note that this solution is real, there is no imaginary part i. Also note that this solution would experience periodicity.

Suppose $y'' + 2y' + 10y = 0$, then the roots of the characteristic equation are $k_1 = -1 + 3i$ and $k_2 = -1 - 3i$. Therefore, the solution would be $y(t) = e^{-t}(c_1 \cos 3t + c_2 \sin 3t)$. Another example: $2y'' + y = 0$. Here, both roots are imaginary without the real part: $k = \pm\sqrt{1/2}$. Thus, $y(t) = c_1 \cos\frac{1}{2}t + c_2 \sin\frac{1}{2}t$.

6. In general, higher order equations are solved in a similar way. If you have $y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = 0$, you first need to solve the characteristics equation in ks and find all n solutions. Then the general solution is $y(t) = \sum c_i e^{k_i t}$. For any real root of multiplicity p there will correspond p linearly-independent particular solutions $e^{kt}, te^{kt}, \ldots, t^{p-1}e^{kt}$. While for any pair of complex conjugate roots $\alpha \pm \beta i$ of multiplicity q there corresponds $2q$ linearly-independent particular solutions $e^{\alpha t}\cos\beta t, e^{\alpha t}\sin\beta t,\ te^{\alpha t}\cos\beta t, te^{\alpha t}\sin\beta t, \ldots, t^{q-1}e^{\alpha t}\cos\beta t, t^{q-1}e^{\alpha t}$.

7. If in the equation the right hand side is not a zero but some constant, c, so that we have a non-homogenous equation $y'' + ay' + by = c$, the solution to the equation is very similar and would be $y(t) = c_1 e^{k_1 t} + c_2 e^{k_2 t} + \frac{c}{b}$.

Suppose $y''(t) + 9y'(t) + 14y(t) = 7$. Then the characteristic equation would be $k^2 + 9k + 14 = 0$, with the roots $k_1 = -2$ and $k_2 = -7$. Thus, the general solution is $y(x) = c_1 e^{-2t} + c_2 e^{-7t} + \frac{1}{2}$. Here, again, $\frac{1}{2}$ is a particular solution.

8. Now suppose you have a non-homogenous equation with the RHS being not just a constant term, but a more general function of t: $y'' + ay' + by = g(t)$. Then a general solution is the sum of any particular solution y_p and a general solution of the corresponding homogenous equation $y'' + ay' + by = 0$.

9. You already know how to solve homogenous equations. The problem right now is to be able to find a particular solution for the heterogeneous equation. This is usually done using the "methods of undetermined coefficients." In this method, you look for a particular solution that has the same "form" as the function $g(x)$.

Let $y'' - 2y' - 3y = 9t^2$. The general solution to the homogenous equation is $y_g(t) = c_1 e^{3t} + c_2 e^{-t}$. To find a particular solution, let's assume it to be quadratic function of a general from: $y_p(t) = At^2 + Bt + C$. Let's plug it into the equation: $9t^2 = (2A) - 2(2At + B) - 3(At^2 + Bt + C) = (-3A)t^2 + (-4A - 3B)t +$

$(2A-2B-3C)$. Now we have to compare the similar terms and equate their coefficients. This gives us the following system: $\begin{cases} 9=-3A \\ 0=-4A-3B \\ 0=2A-2B-3C \end{cases}$ *whose solution is $A=-3, B=4, C=-\frac{14}{3}$. Thus, the general solution to the complete heterogeneous equation would be $y(t)=c_1e^{3t}+c_2e^{-t}-3t^2+4t-\frac{14}{3}$.*

Another example: $y''-2y'-3y=8e^{-t}$. To find a particular solution, you could start with $y_p=Ae^{-t}$. When plugged, it would give you $8e^{-t}=-4Ae^{-t}$. Thus, $A=-2$.

10. Note that as before, if, for example, you are searching for a solution in the form $y_p=Ae^{Bt}$ and you are not able to identify A, you would have to start searching for the solution in the form $y_p=Ate^{Bt}$ by including an extra t. Cases like this would appear if one of the solutions to the homogenous equation would be the same as the RHS part of the nonhomogenous equation.

For example, let $y''-y'-2y=4e^{-t}$. If you search for the solution in the form $y_p=Ae^{-t}$and plug it into the equation, you would get $A+A-2A=4e^{-t}$ thus not being able to identify A. This is the case because one of the characteristic roots is "-1" which lead to c_1e^{-t} which is similar to the RHS. Then your next step should be searching for a solution in the form $y_p=Ate^{t}$ and you could show that $A=-\frac{4}{3}$.

DIFFERENCE EQUATIONS IN DISCRETE TIME

1. A difference equation expresses a relationship between a dependent variable and a lagged independent variable which changes at discrete (rather than continuous) intervals of time. For example, $y_t=f(y)$. The order of the equation is determined by the greatest number of periods lagged. The change in y as t changes from t to $t+1$ is called the first difference of y and written as $\frac{\Delta y}{\Delta t}=\Delta y_t=y_{t+1}-y_t$. The solution defined y for every value of t and does not contain a difference expression or lagged values of y such as y_{t-k}.

2. Simple difference equations could be solved iteratively, by substitution. For example, if $y_{t+1}=by_t$ and the initial value of y is $y(t=0)=y_0$, then you can infer that $y_1=by_0, y_2=by_1=b^2y_0$ etc., from which you could guess the solution to the equation $y_t=b^ty_0$.

3. First-order linear difference equations. Given that you have a linear equation $y_t=b\,y_{t-1}+a$, the general formula for a definite solution is $y_t=\left(y_0-\frac{a}{1-b}\right)b^t+\frac{a}{1-b}$ when $b\neq 1$ and $y_t=y_0+at$ when $b=1$.

$y_t=-7\,y_{t-1}+16$ and $y_0=5$. Since $b\neq 1$, the solution is $y_t=3(-7)^t+2$.The time path oscillated and explosive.

Another example: $x_t + 3x_{t-1} + 8 = 0$ and $x_0 = 16$. First you need to rearrange the terms as in $x_t = -3x_{t-1} - 8$, thus, $b = -3$ and $a = -8$. So the solution is $x_t = \left(16 + \frac{8}{1+3}\right)(-3)^t - \frac{8}{1+3} = 18(-3)^t - 2$. With $b = -3$ which is negative and higher than 1, the time path will be oscillating and explosive.

4. The above solution could be expressed in a general form as $y_t = Ab^t + c$, from which you could infer A and c by simply substituting the solution in the equation. As you could see from the general form, if $|b| > 1$ the solution would be explosive, i.e. the value of y would be indefinitely increasing in magnitude (either gradually, for positive b, or switching signs, for negative b). If instead $|b| < 1$, the time path will be damped and move towards the equilibrium (converging path). For negative b, it will again be oscillating. If $|b| = 1$ we will say that the path is divergent (though not explosive).

If $y_t = 6\left(-\frac{1}{4}\right)^t + 6$, then since $b < 0$, the time path oscillates, and since $|b| < 1$, the time path converges.

5. For some equations $y_t = f(y_{t-1})$, it's impossible to find an algebraic solution. However, you can get a nice idea of how the solution looks like (whether it's stable, etc) by constructing a phase diagram. That's done is several steps: finding the steady state at which $y_t = y_{t-1}$, finding the first and second derivatives of $y_t = f(y_{t-1})$ with respect of y_{t-1}, and finally, drawing the diagram with y_{t-1} on the horizontal axis and y_t on the vertical axis.

Let $y_t = \sqrt{y_{t-1}}$. There are 2 steady state solutions as $y_t = \sqrt{y_t}$ has 2 solutions, $y = 0$ and $y = 1$. The first derivative f is $f' = \frac{0.5}{\sqrt{y_{t-1}}} > 0$ and $f'' = -\frac{0.25}{\sqrt{y_{t-1}^3}} < 0$. Thus, you can draw a 45 degree line which will represent all possible steady state solutions, and the phase diagram itself, $y_t = \sqrt{y_{t-1}}$. You see that they intersect twice, and $y = 0$ is an unstable steady state, while $y = 1$ is a stable steady state.

Another example: $k_t = k_{t-1}^3$ where k is capital, and thus should be positive. As before, start with finding the steady state solutions. At steady state, $k_t = k_{t-1} = k$. Thus, the solutions are 0, 1. The function is a convex function that crosses the 45 degree line at those steady states. 0 would be a stable steady state, and 1 would be an unstable steady state. You could have used the derivatives, too. $\frac{dy_t}{dy_{t-1}} = 3y_{t-1}^2 > 0$ so the slope is positive and $\frac{d^2y_t}{dy_{t-1}^2} = 6y_{t-1} > 0$ so the function is convex. The first derivative evaluated at 0 is 0 and thus smaller than 1 in absolute value and thus is locally stable. When it's evaluated at 1, it's 3 and it is higher than 1 in absolute value and thus that equilibrium is unstable. As both derivatives are positive, there is no oscillation.

OPTIMAL CONTROL THEORY AND THE HAMILTONIAN

1. Before, you have learned how to optimize a function subject to constraints in discrete time. Now when you know how to work with integrals and differential equation, we can take the optimization problem to another level – optimization in continuous time. It's best illustrated by an example.
2. Suppose a consumer maximizes lifetime utility V which is now continuous instead of discrete, $V = \int_0^\infty e^{-\rho t} U(c(t))dt$, subject to a constraint called the evolution equation $a(t)' = ra(t) + w - c$, where a is the level of assets that increases due to interest r earned on current assets holdings and wages w, and decreases due to consumption c. ρ is the discount factor.
3. Note that when you solve the differential equation in a, you would get a more familiarly-looking budget constraint $a_0 + \int_0^\infty e^{-rt} w dt = \int_0^\infty e^{-rt} c dt$
4. This optimization problem could be solved by using a Hamiltonian (versus the Lagrangian that you use in discrete problems). To proceed, we need to define a couple of terms. Variables that an agent can choose are called control; here, $c(t)$ is a control variable, variable of choice. Variables that an agent can affect but cannot choose are called state variables; here, $a(t)$ is a state variable. The utility maximization problem is carried out in 5 steps.
5. First, you construct the Hamiltonian by adding the instantaneous (the integrant) part of V to the right hand side of the transition equation, pre-multiplied by a Lagrange multiplier, λ: $H = e^{-\rho t} U\big(c(t)\big) + \lambda(t)\big(ra(t) + w - c(t)\big)$.
6. Second, you take the derivative of H with respect to all control variables (here: c) and set them equal to zero: $H'_c = e^{-\rho t} U'\big(c(t)\big) - \lambda(t) = 0$
7. Third, you take the derivative of H with respect to all state variables (here: a) and set them equal to the negative derivative of λ: $H'_a = \lambda(t) r = -\lambda'_t$, from which it follows that $\frac{\lambda'_t}{\lambda} = -r$.
8. Fourth, you check the transversality condition which rules out infinite debt schemes: For finite-horizon problems, you restrict $\lambda(t)a(a) = 0$, for infinite horizon with discounting, you make sure that $\lim_{t\to\infty}[\lambda(t)a(t)] = 0$ and for infinite horizons without discounting: $\lim_{t\to\infty}[H(t)] = 0$.
9. Fifth, you differentiate H with respect to λ and equate the results to zero to recover the transition equation. So, you last condition is basically the transition equation.

Suppose you need to maximize $V = \int_0^\infty e^{-\rho t} U(c(t))dt$ for the case when $U(c(t)) = \frac{c(t)^{\sigma-1}-1}{\sigma-1}$, which is the CRRA utility function, subject to $a(t)' = ra(t) + w - c$. First, you go over the 5 steps above. Next, the best strategy is getting rid of λ. You can achieve that by first differentiating H'_c equation (from the second step) with respect to time, and then forming $\frac{\lambda'_t}{\lambda}$ to then equate it to the equation from the third step: $(H'_c)'_t = -\rho e^{-\rho t} U'\big(c(t)\big) + e^{-\rho t} U''\big(c(t)\big) c'_t = \lambda'_t$. If you now divide this equation to the original equation from the second step, you would get $\frac{\lambda'_t}{\lambda} = \frac{-\rho e^{-\rho t} U'(c(t)) + e^{-\rho t} U''(c(t)) c'_t}{e^{-\rho t} U'(c(t))} = -\rho + \frac{U''(c(t))}{U'(c(t))} c'_t = -r$ (from the third step). Thus, the Euler equation becomes $\frac{c'_t}{c(t)} = -\frac{U'(c(t))}{c(t)U''(c(t))}(r-\rho)$. For our specific CRRA utility, this transforms into $c' = c\left[\frac{1}{\sigma}(r-\rho)\right]$, which has the solution $c_t = c_0 e^{\frac{1}{\sigma}(r-\rho)t}$, where

c_0 could be recovered from the budget constraint $c_0 = \frac{1}{\sigma}(\rho - (1-\sigma)r)(a_0 + \int_0^\infty e^{-rt} w dt)$.