MATH CAMP – SESSION 2

TOPOLOGY & OPTIMIZATION FOR ECONOMICS

Department of Economics

Duke University

Topology for Economics

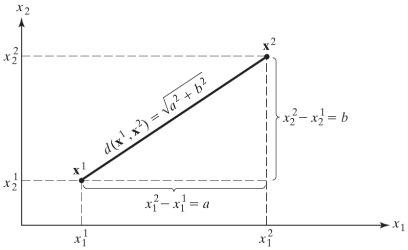
What is Topology?

- Topology is the study of fundamental properties of sets and mappings. In this section, we introduce a few basic topological ideas and use them to establish some important results about sets, and about continuous functions from one set to another.
  - We confine ourselves to considering sets in $\mathbb{R}^n$, i.e., sets that contain real numbers or vectors of real numbers.

Metric Space I

- A *metric* is simply a measure of distance. A **metric space** is just a set with a notion of distance defined among the points within the set. The real line, $\mathbb{R}$, together with an appropriate function measuring distance, is a metric space.
- One such distance function, or metric, is just the absolute-value function. For any two points x^1 and x^2 in $\mathbb{R}$, the distance between them, denoted $d(x^1, x^2)$, is given by
$$d(x^1, x^2) = |\, x^1 - x^2 \,|.$$
- The Cartesian plane, $\mathbb{R}^2$, is also a metric space. A natural notion of distance defined on the plane is inherited from Pythagoras. Choose any two points $\mathbf{x}^1 = (x_1^1, x_2^1)$ and $\mathbf{x}^2 = (x_1^2, x_2^2)$ in $\mathbb{R}^2$.

Metric Space II



- Construct the right triangle connecting the two points. If the horizontal leg is of length a and the vertical leg is length b, Pythagoras tells us the length of the hypotenuse – the distance between the points x^1 and x^2 – is equal to $\sqrt{a^2 + b^2}$.
- Now a^2 is just the square of the difference between the x_1 components of the two points, and b^2 is the square of the difference in their x_2 components. The length of the hypotenuse, or $d(\mathbf{x}^1, \mathbf{x}^2)$, is therefore

$$d(\mathbf{x}^1, \mathbf{x}^2) = \sqrt{a^2 + b^2} = \sqrt{(x_1^2 - x_1^1)^2 + (x_2^2 - x_2^1)^2}.$$

A Generalization in $\mathbb{R}^n$: Euclidean Metric

- Pythagoras tells us that we can again measure the distance between two points as the square root of a product of the difference between the two points, this time the vector product of the vector difference of two points in the plane.
- Analogous to the case of points on the line, we can therefore write $d(\mathbf{x}^1, \mathbf{x}^2) = \sqrt{(\mathbf{x}^1 - \mathbf{x}^2) \cdot (\mathbf{x}^1 - \mathbf{x}^2)}$ for $\mathbf{x}^1$ and $\mathbf{x}^2$ in $\mathbb{R}^2$.
- The distance between any two points in Rn is just a direct extension of these ideas. In general, for $\mathbf{x}^1$ and $\mathbf{x}^2$ in $\mathbb{R}^n$,

$$\begin{aligned} d(\mathbf{x}^1, \mathbf{x}^2) &\equiv \sqrt{(\mathbf{x}^1 - \mathbf{x}^2) \cdot (\mathbf{x}^1 - \mathbf{x}^2)} \\ &\equiv \sqrt{(x_1^1 - x_1^2)^2 + (x_2^1 - x_2^2)^2 + \cdots + (x_n^1 - x_n^2)^2}, \end{aligned}$$

  which we can summarize with the notation $d(\mathbf{x}^1, \mathbf{x}^2) = \|\mathbf{x}^1 - \mathbf{x}^2\|$.
- We call this formula the **Euclidean metric** or **Euclidean norm**. Naturally enough, the metric spaces $\mathbb{R}^n$ that use this as the measure of distance are called **Euclidean spaces**.

Open and Closed ε-Balls

- Once we have a metric, we can make precise what it means for points to be 'near' each other. If we take any point $\mathbf{x}^0 \in \mathbb{R}^n$, we define the set of points that are less than a distance $\varepsilon > 0$ from $\mathbf{x}^0$ as the *open ε-ball* with centre $\mathbf{x}^0$.
- The set of points that are a distance of ε or less from $\mathbf{x}^0$ is called the *closed ε-ball* with centre $\mathbf{x}^0$.
- Formally:

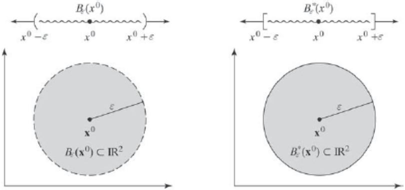
1. The open ε-ball with centre $\mathbf{x}^0$ and radius $\varepsilon > 0$ (a real number) is the subset of points in $\mathbb{R}^n$:

$$B_\varepsilon(\mathbf{x}^0) \equiv \{\mathbf{x} \in \mathbb{R}^n \mid d(\mathbf{x}^0, \mathbf{x}) < \varepsilon)\}$$

2. The closed ε-ball with centre $\mathbf{x}^0$ and radius $\varepsilon > 0$ is the subset of points in $\mathbb{R}^n$:

$$B^*_\varepsilon(\mathbf{x}^0) \equiv \{\mathbf{x} \in \mathbb{R}^n \mid d(\mathbf{x}^0, \mathbf{x}) \leq \varepsilon)\}$$

Examples



- On the real line, the open ball with centre x^0 and radius ε is just the open interval $B_\varepsilon(x^0) = (x^0 - \varepsilon, x^0 + \varepsilon)$. The corresponding closed ball is the closed interval $B_\varepsilon^*(x^0) = [x^0 - \varepsilon, x^0 + \varepsilon]$.
- In $\mathbb{R}^2$, an open ball $B_\varepsilon(\mathbf{x}^0)$ is a disc consisting of the set of points inside, or on the interior of, the circle of radius ε around the point $\mathbf{x}^0$.

Open and Closed Sets in $\mathbb{R}^n$

Definition (Open Sets)

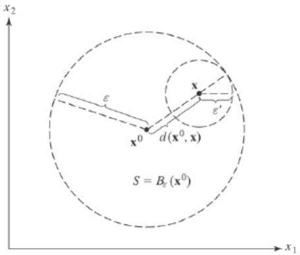
$S \subset \mathbb{R}^n$ is an open set if, for all $\mathbf{x} \in S$, there exists some $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{x}) \subset S$.

- The definition says that a set is open if around any point in it we can draw some open ball – no matter how small its radius may have to be – so that all the points in that ball will lie entirely in the set.

Definition

S is a closed set if its complement, S^c, is an open set.

Example: Open Ball in $\mathbb{R}^2$

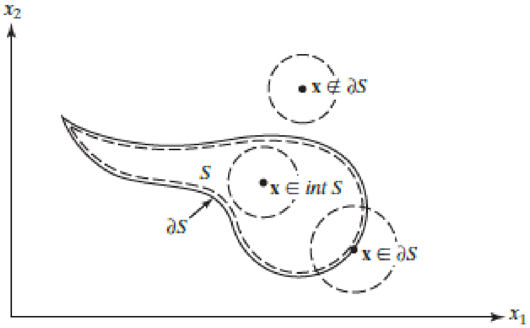


- Let S be the open ball with centre $\mathbf{x}^0$ and radius ε as in the figure above. If we take any other point $\mathbf{x}$ in S, it will always be possible to draw an open ball around $\mathbf{x}$ whose points all lie within S by choosing the radius of the ball around $\mathbf{x}$ carefully enough.
- Because $\mathbf{x}$ is in S, we know that $d(\mathbf{x}^0, \mathbf{x}) < \varepsilon$. Thus, $\varepsilon - d(\mathbf{x}^0, \mathbf{x}) > 0$. If we let $\varepsilon' = \varepsilon - d(\mathbf{x}^0, \mathbf{x}) > 0$, then it will always be the case that $B'_\varepsilon(x) \subset S$ no matter how close we take $\mathbf{x}$ to the edge of the circle, as required.

Boundary and Interior Points

- Loosely speaking, a set in $\mathbb{R}^n$ is open if it does not contain any of the points on its boundary, and is closed if it contains all of the points on its boundary.
- More precisely, a point $\mathbf{x}$ is called a **boundary point** of a set S in $\mathbb{R}^n$ if every ε-ball centred at $\mathbf{x}$ contains points in S as well as points not in S. The set of all boundary points of a set S is denoted ∂S.
- A set is open if it contains none of its boundary points; it is closed if it contains all of its boundary points.
- A point $\mathbf{x} \in S$ is called an **interior point** of S if there is some ε-ball centred at $\mathbf{x}$ that is entirely contained within S, or if there exist some $\varepsilon > 0$ such that $B_\varepsilon(x) \subset S$.
- The set of all interior points of a set S is called its interior and is denoted $int\ S$. Looking at things this way, we can see that a set is open if it contains nothing but interior points, or if $S = int\ S$.
- By contrast, a set S is closed if it contains all its interior points plus all its boundary points, or if $S = int\ S \cup \partial S$.

Graphical Example



Functions

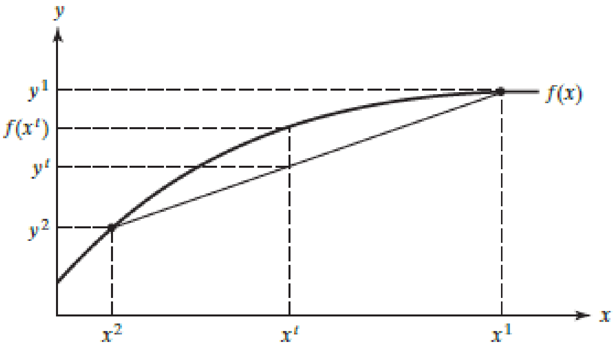
Concave Functions

- We will restrict our attention to real-valued functions whose domains are convex sets.

Definition (Concave Function)

A function $f : X \to \mathbb{R}$ is concave if $\forall \mathbf{x}^1, \mathbf{x}^2$ and $t \in [0, 1]$

$$f(t\mathbf{x}^1 + (1-t)\mathbf{x}^2) \geq tf(\mathbf{x}^1) + (1-t)f(\mathbf{x}^2).$$



Quasiconcave Functions

Definition (Quasiconcave Function)

A function $f : X \to \mathbb{R}$ is quasi-concave if $\forall \mathbf{x}^1, \mathbf{x}^2 \in X$ and $t \in [0, 1]$,

$$f(t\mathbf{x}^1 + (1-t)\mathbf{x}^2) \geq \min\{f(\mathbf{x}^1), f(\mathbf{x}^2)\},$$

or equivalently, if

$$f(\mathbf{x}^1) \geq f(\mathbf{x}^2) \implies f(t\mathbf{x}^1 + (1-t)\mathbf{x}^2) \geq f(\mathbf{x}^2).$$

Moreover, f is strictly quasi-concave if the inequality is strict for $\mathbf{x}^1 \neq \mathbf{x}^2$ and $t \in (0, 1)$.

Theorem

f is quasi-concave if and only if its better-than set $\{\mathbf{x} \in X \mid f(\mathbf{x}) \geq K\}$ is convex for all $K \in \mathbb{R}$, where the "better-than" set is $B(\mathbf{x}) = \{\mathbf{y} \in X \mid \mathbf{y} \succsim \mathbf{x}\}$.

Proof. Omitted for now – will be covered in ECON 601. □

Homogeneous Functions

Definition (Homogeneous Function)

A real-valued function $f(x)$ is called homogeneous of degree k if

$$f(t\mathbf{x}) \equiv t^k f(\mathbf{x}) \quad \forall t > 0.$$

- Two special cases are worthy of note:
  - $f(\mathbf{x})$ is homogeneous of degree 1, or linear homogeneous, if $f(t\mathbf{x}) \equiv tf(\mathbf{x}) \; \forall \; t > 0$;
  - it is homogeneous of degree zero if $f(t\mathbf{x}) \equiv f(\mathbf{x}) \; \forall \; t > 0$.
- Homogeneous functions display very regular behaviour as all variables are increased simultaneously and in the same proportion.
- When a function is homogeneous of degree 1, for example, doubling or tripling all variables doubles or triples the value of the function.
- When homogeneous of degree zero, equiproportionate changes in all variables leave the value of the function unchanged.

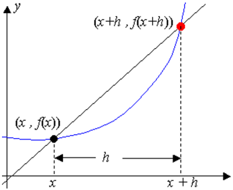
Calculus & Optimization

Differentiation

Definition (Derivative)

The derivative of $f(x)$ with respect to x is the function $f'(x)$ and is defined as,

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



- A function $f(x)$ is called *differentiable at a point a* if $f'(a)$ exists.
- A function $f(x)$ is called *differentiable on an interval* if the derivative exists for each point in that interval.
- A function $f(x)$ is called a *differentiable function* if the function is differentiable at every point in the domain.

Single Variable Functions: Rules of Differentiation I

A function $y = f(x)$ is differentiable if it is both continuous and 'smooth', with no breaks or kinks.

- Sums:

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

- Power Rule:

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

- Product Rule:

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

- Quotient Rule:

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

- Chain Rule:

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

Single Variable Functions: Rules of Differentiation II

Some common derivatives in economics are the following:

- Logarithmic Functions:

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

- Exponential Functions:

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

$$\frac{d}{dx} a^{x} = \ln a \cdot a^{x}$$

Examples

Differentiate $h(x)$ with respect to x, where $h(x) = \frac{4\sqrt{x}}{x^2-2}$.

$$
\begin{aligned}
h'(x) &= \frac{4\left(\frac{1}{2}\right)x^{-\frac{1}{2}}(x^2-2) - 4x^{\frac{1}{2}}(2x)}{(x^2-2)^2} \\
&= \frac{2x^{\frac{3}{2}} - 4x^{-\frac{1}{2}} - 8x^{\frac{3}{2}}}{(x^2-2)^2} \\
&= \frac{-6x^{\frac{3}{2}} - 4x^{-\frac{1}{2}}}{(x^2-2)^2}
\end{aligned}
$$

□

Find the derivative of $f(x)$, where $f(x) = \ln(x^{-4} + x^4)$

$$
\begin{aligned}
g'(x) &= \frac{1}{x^{-4}+x^4}(-4x^{-5} + 4x^3) \\
&= \frac{-4x^{-5} + 4x^3}{x^{-4}+x^4}
\end{aligned}
$$

□

Multivariate Functions: Partial Derivative

Definitions:

$$f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$
$$f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}$$

Note: Allow one variable to move and hold all other variables constant.

$$f(x,y) \implies f_x(x,y) = \frac{\partial f}{\partial x} \quad \& \quad f_y(x,y) = \frac{\partial f}{\partial y}$$

Example: Find the partial derivatives of $f(x,y) = 2x^2y^3$ with respect to x and y.

Solution:

$$f_x(x,y) = 4xy^3$$
$$f_y(x,y) = 6x^2y^2$$

Total Derivative and Multivariate Chain Rule

Given the function $f(x, y)$, the differential df is given by:

$$df = f_x dx + f_y dy$$

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

The chain rule for a multivariate function $f(y(x), z(x))$ is as follows:

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

Example

Find the derivative of $f(y, z)$ with respect to x:

$$f(y, z) = ye^{yz}; \quad y(x) = x^2; \quad z(x) = \frac{1}{x}$$

Solution:

$$\begin{aligned}
\frac{df}{dx} &= \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} \\
&= (e^{yz} + zye^{yz})(2x) + y^2 e^{yz}(-x^{-2}) \\
&= 2x(e^{yz} + zye^{yz}) - x^{-2} y^2 e^{yz}
\end{aligned}$$

Multivariate Homogeneous Functions

Example: Is $f(x, y) = x^2y^2 + y^4$ an HDk function? If yes, find the value of k.

Solution:

$$\begin{aligned} f(tx, ty) &= (tx)^2(ty)^2 + (ty)^4 \\ &= t^4x^2y^2 + t^4y^4 \\ &= t^4(x^2y^2 + y^4) \\ &= t^4f(x, y) \end{aligned}$$

So, $f(x, y)$ is homogeneous of degree 4, or HD4.

Euler Equation

Definition

A differentiable function $f(\mathbf{x})$ is HDk if and only if

$$\sum_{i=1}^{n} \frac{\partial f(\mathbf{x})}{\partial x_i} x_i = kf(\mathbf{x}) \quad \forall \mathbf{x}.$$

Moreover, if f is HDk, then $\frac{\partial f}{\partial x_i}$ is HD$(k-1)$.

Example: Is $f(x, y) = x^2y^2 + y^4$ an HDk function? If yes, find the value of k.

Solution:

$$\begin{aligned}\frac{\partial f}{\partial x}x + \frac{\partial f}{\partial y}y &= (2xy^2)x + (2yx^2 + 4y^3)y = 2x^2y^2 + 2y^2x^2 + 4y^4 \\ &= 4x^2y^2 + 4y^4 = 4f(x, y).\end{aligned}$$

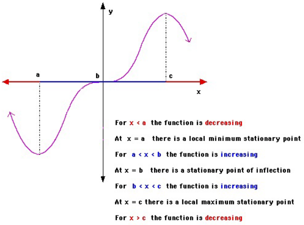
So, the result is the same as earlier using the Euler equation.

Unconstrained Optimization

You can use simple derivatives to find the maximum or minimum of a function, $f(x)$, when there are no constraints on the underlying variable, x. For instance, to solve,

$$\max_x \; f(x) \text{ where } x \in \mathbb{R}.$$

- **First-order Condition (FOC):** Find x where $f'(x) = 0$
- **Second-order Condition (SOC):** Check the value of $f''(x)$ at this x.
  - If this $f''(x) > 0$, then this function is at a minima;
  - If this $f''(x) < 0$, then this function is at a maxima.

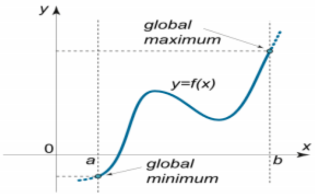


Constrained Optimization

- However, in most cases we encounter, x can't take any value. For instance, traveling to new places might give you a lot of joy but your budget may not allow you to visit all the places you want. So, you would try to pick the best few locations keeping in mind the money you can spend.
- In effect, you are trying to solve,

$$\max_x f(x) \text{ where } x \in [a, b].$$

- In addition to the FOC and SOC, you would add a **third-step**. We have to check the value of $f(x)$ at the boundary points, which might be the maxima/minima.



Lagrange Multiplier

- While that may work for functions of one variable, we need a more general technique to deal with **multivariate cases**.
- For instance, to maximize an objective function, $f(\mathbf{x})$, with constraints, $g(\mathbf{x}) \geq 0$, where $\mathbf{x} \in \mathbb{R}^n$ we can create a **Lagrangian** function, $\mathcal{L}(\mathbf{x}; \lambda)$, and use the **Kuhn-Tucker FOC** for each of the inputs x_i,

$$\max_{\mathbf{x}} f(\mathbf{x}) \text{ such that } g(\mathbf{x}) \geq 0 \text{ and } x_i \geq 0, \forall x_i \in \mathbf{x}$$

$$\max_{x} \min_{\lambda} \mathcal{L}(x, \lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x}) \text{ such that } x_i \geq 0 \ \forall x_i \in \mathbf{x}, \lambda \geq 0$$

Example

Find the maximum values of $f(x, y) = 4x^2 + 10y^2$ on the disk $x^2 + y^2 \leq 4$.

Solution: Find all the critical points that are in the disk, i.e., satisfy the constraint. Here are the two first-order partial derivatives:

$$\begin{aligned} f_x = 8x \implies 8x = 0 \implies x = 0 \\ f_y = 20y \implies 20y = 0 \implies y = 0 \end{aligned}$$

So, the only critical point is $(0, 0)$ and it does not satisfy the inequality.

At this point we proceed with Lagrange Multipliers and we treat the constraint as an equality instead of the inequality. We only need to deal with the inequality when finding the critical points. So, here is the system of equations that we need to solve.

$$\begin{aligned} 8x &= 2\lambda x \\ 20y &= 2\lambda y \\ x^2 + y^2 &= 4 \end{aligned}$$

Example (Contd.)

From the first equation, we get: $2x(4 - \lambda) = 0 \implies x = 0$ or $\lambda = 4$.

If we have $x = 0$, then the constraint gives us $y = \pm 2$.

If we have $\lambda = 4$, the second equation gives us: $20y = 8y \implies y = 0$.

The constraint then tells us that $x = \pm 2$. If we'd performed a similar analysis on the second equation we would arrive at the same points.

So, Lagrange Multipliers gives us four points to check : $(0, 2)$, $(0, -2)$, $(2, 0)$, and $(-2, 0)$.

To find the maximum and minimum we need to simply plug these four points along with the critical point in the function and the maximum is reached at

$$f(0, 2) = f(0, -2) = 40.$$

In this case, the minimum was interior to the disk and the maximum was on the boundary of the disk.

Finding Marhsallian Demand

Assume we are given a utility function, $u(\mathbf{x})$, and it is differentiable.

$$
\begin{aligned}
&\max_{\mathbf{x}\in\mathbb{R}^n_+} u(\mathbf{x}) \\
\text{subject to: } & \mathbf{p}\cdot\mathbf{x} \le I \iff I - \mathbf{p}\cdot\mathbf{x} \ge 0 \ (\lambda) \\
\text{and: } & x \ge 0 \qquad\qquad\qquad x \ge 0
\end{aligned}
$$

The above is a constrained optimization – common in Economics. So, we turn it into an unconstrained problem via the Lagrangian method.

$$\mathcal{L}(x,\lambda) = u(x) + \underset{\substack{\uparrow \\ \text{Lagrange Multiplier (new variable)}}}{\lambda} [I - p\cdot x]$$

$$\min_{\lambda}\max_{x} \mathcal{L}(x,\lambda) \quad \text{s. to } \mathbf{x} \ge 0 \text{ and } \lambda \ge 0.$$

Kuhn-Tucker First-Order (Necessary) Conditions

$$\frac{\partial \mathcal{L}}{\partial x_i} = \frac{\partial u}{\partial x_i} - \lambda p_i \leq 0, \qquad \underbrace{\frac{\partial \mathcal{L}}{\partial x_i} x_i^*}_{\text{Complementary slackness}} = 0,$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = I - \mathbf{p} \cdot \mathbf{x}^* \geq 0, \qquad \frac{\partial \mathcal{L}}{\partial \lambda} \lambda^* = 0,$$

$$\mathbf{x}^* \geq 0, \text{ and} \qquad \lambda^* \geq 0.$$