Calculus

Anton Velinov

Takeaways

- Differentiation: rules, implicit differentiation
- Taylor and power series
- Maximization
- Definite and indefinite integrals
- Differential equations
- Multivariate function limits and continuity
- Partial derivatives, gradient and directional derivatives
- Constrained maximization
- Complex numbers
- Implicit function theorem (optional)

Differentiation I

Definition

Let $f : \mathbb{R} \to \mathbb{R}$. Its *derivative* with respect to a point x is defined by the following limit

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

Class Exercise

Use the above definition to calculate the derivative of

$$f(x) = a + bx$$

and

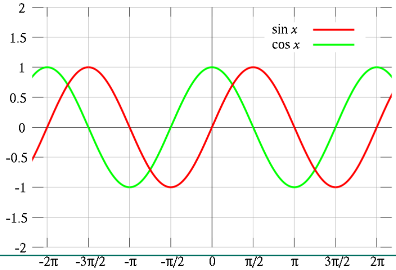
$$f(x) = \frac{x}{x+1}$$

with respect to x.

Differentiation II

Useful Rules

- $dx^k/dx = kx^{k-1}$
- $d(sin(x))/dx = cos(x)$ and $d(cos(x))/dx = -sin(x)$
- $d(e^x)/dx = e^x$
- $d(log(x))/dx = 1/x$



Differentiation III

Theorem

If a function $f(x)$ is differentiable at some point x, then it is continuous at x.

$$\text{differentiability} \Rightarrow \text{continuity}.$$

The converse is not true

$$\text{continuity} \nRightarrow \text{differentiability}.$$

Class Exercise

Show that $f(x) = |x - 2| + 1$ is continuous but not differentiable at $x = 2$.

Mean Value Theorem I

Theorem

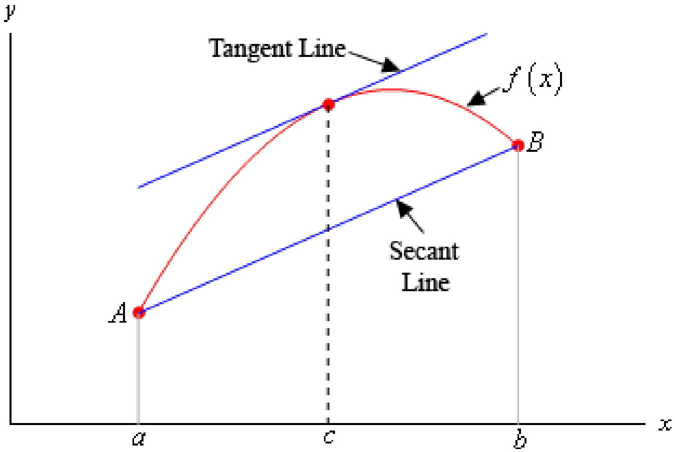
Let $f : [a, b] \to \mathbb{R}$ be a continuous function on the closed interval $[a, b]$ and differentiable on the open interval (a, b), where $a < b$. Then there exists some $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Class Exercise

1. Find c for the function $f(x) = x^2$ on the interval [1,3].
2. Find c for the function $f(x) = x^3 - 6x^2 + 12x$ on the interval [0,3].

Mean Value Theorem II



Differentiation Rules I

Product Rule

If functions $f(x)$ and $g(x)$ are differentiable then their product is differentiable and given as follows:

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

Quotient Rule

If functions $f(x)$ and $g(x)$ are differentiable then their quotient is differentiable and given as follows:

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

Differentiation Rules II

Example

Prove the product rule. **Hint**: Add and subtract $f(x+\Delta)g(x)$.

Chain Rule

For the *composite* function $g(x) = f(u(x))$ or $g(x) = (f \circ u)(x)$ the derivative is given as follows:

$$\frac{dg(x)}{dx} = \frac{df(\cdot)}{du}\frac{du(x)}{dx}.$$

Class Exercise

Use the chain rule to find the derivative of $f(x) = 2^x$ with respect to x.

Implicit Differentiation I

It is not always possible to differentiate x explicitly from $f(x)$. As a direct consequence of the chain rule we can therefore make use of *implicit* differentiation

Example

Find the derivative of

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

with respect to x and compare this to the explicit solution.

Class Exercise

Find the derivative of

$$(x^2 + y^2)^3 = 5x^2y^2$$

with respect to x.

Implicit Differentiation II

Class Exercise (optional)

Find the derivative of

$$y = x^x$$

with respect to x.

Higher Order Derivatives

Definition

The *second* derivative of $f(x)$ is given by

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

Definition

In general the kth-order derivative is the derivative of the $(k-1)$th-order derivative.

Taylor Series Approximation

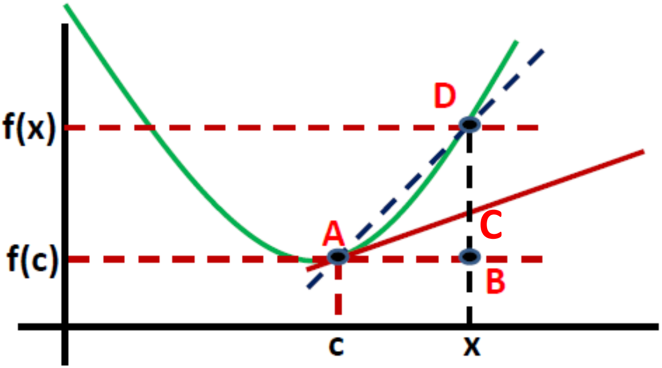
Definition

An *r-th-order Taylor series approximation* of a function $f(x)$ at $x = c + \Delta$ is as follows:

$$\begin{aligned} f(x) &= f(c) + \left.\frac{df}{dx}\right|_{x=c} \cdot \Delta + \frac{1}{2!}\left.\frac{d^2 f}{dx^2}\right|_{x=c} \cdot \Delta^2 + \\ &+ \frac{1}{3!}\left.\frac{d^3 f}{dx^3}\right|_{x=c} \cdot \Delta^3 + \cdots + \frac{1}{r!}\left.\frac{d^r f}{dx^r}\right|_{x=c} \cdot \Delta^r + R(c, x) \end{aligned}$$

provided that all r derivatives of $f(x)$ exist and are continuous in a neighbourhood of c.

First Order Taylor Series Approximation



Power Series I

Definition

A *Power series* assumes that the remainder term, $R(c, x)$ of the Taylor series goes to 0 as $r \to \infty$. It also centres the expansion at $c = 0$.

Class Exercise

Find the Power series of

$$f(x) = e^x$$

evaluated at $x = 1$.

Power Series II

Class Exercise (Optional)

Use the previous result to evaluate

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}.$$

Maxima and Minima

Definition

- The function $f : X \to \mathbb{R}$ has a *global maximum* at $x = c$ if $f(x) \leq f(c)$ for all $x \in X$
- The function $f : X \to \mathbb{R}$ has a *local maximum* at $x = c$ if $f(x) \leq f(c)$ for all x in some open interval around $x = c$.
- Analogous definitions hold for *global* and *local minima*.

Class Exercise

What does the Extreme Value Theorem have to do with the above definition?

Maximization I

Fermat's Theorem

If a function $f(x)$ has an extreme point at $x = c$ and $f'(c)$ exists, then $f'(c) = 0$.

This is a necessary condition for an interior maximum.

Class Exercise

Find an example where $f'(x) = 0$ does not lead to a maximum or a minimum.

Maximization II

Second Derivative Test

If the second derivative of $f(x)$ exists and if $x = c$ is a critical point of $f(x)$ then:

- If $f''(x) < 0, f(x)$ has a local maximum at x
- If $f''(x) > 0, f(x)$ has a local minimum at x
- If $f''(x) = 0$, the test is inconclusive.

Class Exercise

Use the second derivative test for the previous exercise.

Indefinite Integral

Definition

Informally speaking, in case of a continuous function, *integration*, denoted by $\int \cdot dx$, is the inverse of differentiation:

$$\int f'(x)dx = f(x).$$

However, we assume that we only know $f'(x)$ and hence, when integrating a function we always add a constant, C to it. This is called an *indefinite integral*. It is over $\pm\infty$.

Note

Integration is not quite the opposite of differentiation as functions that are non-differentiable can still be integrated.

Indefinite Integral Rules

Integration Rules

- $\int x^k dx = \frac{x^{k+1}}{k+1} + C$ for $k \neq -1$
- $\int x^{-1} dx = \log(|x|) + C$
- $\int cos(x)dx = sin(x) + C$
- $\int sin(x)dx = -cos(x) + C$
- $\int e^{ax} dx = (1/a)e^{ax} + C$
- $\int af(x)dx = a \int f(x)dx$ for some constant a not depending on x
- $\int [af(x) + bg(x)]dx = a \int f(x)dx + b \int g(x)dx$ for some constants a, b not depending on x.

Definite Integral

Definition

From the fundamental theorem of calculus a *definite integral* is defined as:

$$\int_a^b f(x)dx = \left[\int f(x)dx\right]\bigg|_{x=b} - \left[\int f(x)dx\right]\bigg|_{x=a} = F(b) - F(a).$$

Intuitively, it is the area under a function over the domain $[a, b]$.

Class Exercise

Find the integral of

$$\int_a^b x\frac{1}{b-a}dx.$$

Integration by Substitution

Definition

For the function that can be composed as $f(g(x))g'(x)$, its integral over the domain a, b is given as

$$\int_a^b f(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du.$$

This is known as *integration by substitution*. This definition applies to indefinite integration analogously.

Class Exercise

1. What is $\int x/(x^2+1)dx$?
2. What is $\int (x+1)^3 dx$?

Integration by Parts

Definition

Through the product rule we can derive the following formula for *integration by parts*

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx.$$

Class Exercise

Find the integral of $\ln x$.

Differential Equations I

Definition

A *differential equation* is an equation that involves an unknown function as its solution. It usually involves a general and a particular solution given some initial conditions.

Intuitively, we are given the *n*th order derivatives of a function in an equation and we need to solve this in order to find the original function.

In economics differential equations are mainly used in growth theory

Example

Find the solution to

$$\frac{dy}{dx} = \sqrt{x/y}$$

given the initial condition $y = 4, x = 1$.

Differential Equations II

Class Exercise

Find the solution to

$$\frac{dy}{dx} = x/y$$

given the initial condition $y = 1, x = 1$.

Definition

The previous two exercises deal with *separable differential equations*, i.e. when the xs and ys can be separated on both sides of the equation.

Homogeneous Differential Equations (Optional)

Definition

An equation of the form

$$\frac{dy}{dx} = f(x, y)$$

is referred to as a *homogeneous differential equation*. If the xs and ys are not separable, $v = y/x$ is used to solve the equation.

Class Exercise

Find the solution to

1. $\frac{dy}{dx} = \frac{x+y}{x}$
2. $\frac{dy}{dx} = \frac{x^2+3y^2}{2xy}$ (optional).

Multivariate Function Limits I

Definition

A *multivariate* or *multivariable* function, f has $n > 1$ arguments mapping onto $\mathbb{R}$, i.e. $f : \mathbb{R}^n \to \mathbb{R}$. Written as $f(x_1, x_2, \ldots, x_n)$ or $f(\mathbf{x})$.

Multivariate function limit

In Euclidean space f has a *limit* L as $\mathbf{x}$ approaches $\mathbf{x^0}$ if

$$\forall\ \varepsilon > 0\ \ \exists\ \delta(\varepsilon) > 0; \ \ |f(\mathbf{x}) - L| < \varepsilon \text{ whenever } 0 < \|\mathbf{x} - \mathbf{x^0}\|_2 < \delta(\varepsilon).$$

We write $f(\mathbf{x}) \to L$ as $\mathbf{x} \to \mathbf{x^0}$ or $\lim_{\mathbf{x} \to \mathbf{x^0}} f(\mathbf{x}) = L$.

Note

Unlike the univariate case, there are an **infinite** number of directions for $\mathbf{x}$ to approach $\mathbf{x^0}$.

Multivariate Function Limits II

Class Exercise

Determine whether the following limits exist

1. $\lim_{(x,y)\to(0,0)} \frac{x^2}{x^2+y^2}$
2. $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$.

Figure: $f(x, y) = xy/(x^2 + y^2)$

Multivariate Function Limits III

Class Exercise

Show that the following limit exists

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^2+y^2}.$$

Figure: $f(x,y) = x^2 y/(x^2+y^2)$

Continuity of Multivariate Functions

Epsilon-Delta Definition

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *continuous* at $\mathbf{x^0} \in \mathbb{R}^n$ if

$$\begin{aligned} \forall\, \varepsilon > 0 \;\; \exists\, \delta(\varepsilon, \mathbf{x^0}) > 0 \quad & ; |f(\mathbf{x}) - f(\mathbf{x^0})| < \varepsilon \\ \text{whenever} \quad & \|\mathbf{x} - \mathbf{x^0}\|_2 < \delta(\varepsilon, \mathbf{x^0}) \end{aligned}$$

Limit Definition

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *continuous* at $\mathbf{x^0} \in \mathbb{R}^n$ if

$$\lim_{\mathbf{x} \to \mathbf{x^0}} f(\mathbf{x}) = f(\mathbf{x^0}).$$

Theorem

Polynomial functions are continuous everywhere. Rational functions are continuous everywhere they are defined.

Partial Derivatives I

Definition

For the function $f : \mathbb{R}^n \to \mathbb{R}$, its *partial derivative* w.r.t. x_i is defined as

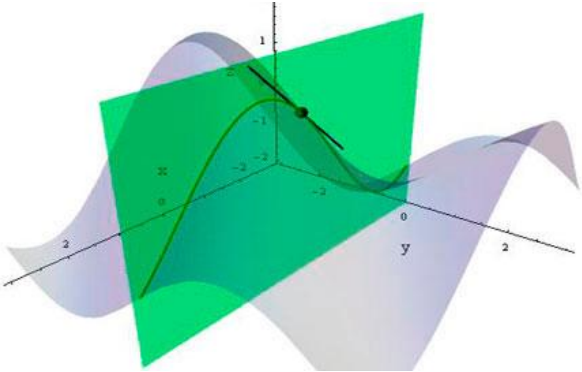
$$\begin{aligned} f_{x_i} = \frac{\partial f}{\partial x_i} &= \lim_{\Delta \to 0} [f(x_1, x_2, \ldots, x_{i-1}, x_i + \Delta, x_{i+1}, \ldots, x_n) \\ &- f(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n)]/\Delta. \end{aligned}$$

Class Exercise

Find the partial derivatives with respect to x and y for the following function

$$f(x, y) = x^4 + 6\sqrt{y} + 10.$$

Partial Derivatives II



Gradient I

Definition

The column vector of partial derivatives of the n dimensional function $f(\mathbf{x})$ is known as the *gradient* (denoted ∇) of f:

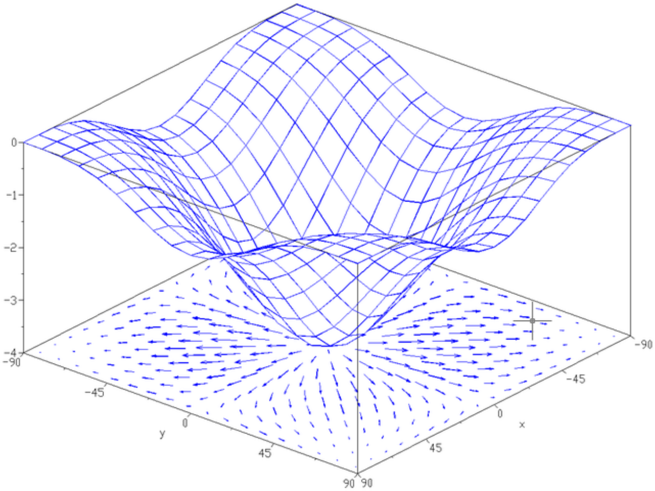
$$\partial f(\mathbf{x})/\partial \mathbf{x} = \nabla = \begin{bmatrix} \partial f(\mathbf{x})/\partial x_1 \\ \partial f(\mathbf{x})/\partial x_2 \\ \vdots \\ \partial f(\mathbf{x})/\partial x_n \end{bmatrix} = \frac{\partial f(\mathbf{x})}{\partial x_1}\mathbf{e}_1 + \frac{\partial f(\mathbf{x})}{\partial x_2}\mathbf{e}_2 + \cdots + \frac{\partial f(\mathbf{x})}{\partial x_n}\mathbf{e}_n,$$

where the $\mathbf{e}_i$s are orthogonal unit vectors.

Geometric Interpretation

The gradient is always orthogonal to the level set $f(\mathbf{x}) = k$. It gives the direction of the steepest ascent (in $\mathbb{R}^n$).

Gradient II



Directional Derivatives I

Definition

Partial derivatives examine the rate of change when allowing one variable to change while the rest are kept constant. *Directional derivatives* examine the rate of change when allowing all of a function's variables to vary. For the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ the directional derivative with the direction of the rate of change denoted by the n dimensional vector $\mathbf{u} = [u_1 \quad u_2 \quad \dots \quad u_n]'$ is given by

$$\begin{aligned} D_u f(x_1, x_2, \dots, x_n) &= \lim_{\Delta \to 0} [f(x_1 + u_1\Delta, x_2 + u_2\Delta, \dots, x_n + u_n\Delta) \\ &- f(x_1, x_2, \dots, x_n)]/\Delta \end{aligned}$$

Directional Derivatives II

The limit in the above definition is sometimes difficult to calculate so an equivalent formula based on the partial derivative is often used

$$\begin{aligned} D_u f(x_1, x_2, \ldots, x_n) &= (\partial f/\partial x_1)u_1 + (\partial f/\partial x_2)u_2 + \cdots + (\partial f/\partial x_n)u_n \\ &= \sum_{i=1}^{n} (\partial f/\partial x_i)u_i = \sum_{i=1}^{n} f_{x_i} u_i. \end{aligned}$$

Notice that this can be expressed in terms of the gradient vector as

$$D_u f(\mathbf{x}) = \boldsymbol{\nabla} \cdot \mathbf{u}$$

where $\cdot$ denotes the *dot product*.

Example (Optional)

Show that the above formula is equivalent to the one on the previous slide. **Hint**: Define the univariate function $g(z) = f(x_1^0 + u_1 z, x_2^0 + u_2 z, \ldots, x_n^0 + u_n z)$ and take its derivative at $z = 0$.

Directional Derivatives III

The **u** Vector

Note that the vector **u** can describe a rate of change in any infinite amount of directions. It is normalized to a unit length as

$$\|\mathbf{u}_N\|_2 = \sqrt{\sum_{i=1}^{n} u_{iN}^2} = 1,$$

where

$$\mathbf{u}_N = \frac{\mathbf{u}}{\|\mathbf{u}\|_2},$$

and $\| \cdot \|_2$ is the $L-2$ vector norm.

The Gradient and the Directional Derivative

Note

Unlike a directional derivative, the gradient is not a scalar but a vector.

Class Exercise

For the function

$$f(x, y, z) = x^2 z + y^3 z^2 - xyz$$

1. Calculate the gradient
2. Determine the directional derivative in the direction of $\mathbf{u} = [-1 \quad 0 \quad 3]'$
3. Would this be the direction of the steepest slope? If not, which direction would we need to go in?

Hessian I

Definition

The second derivative of the *n* dimensional function $f(\mathbf{x})$ is given by:

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} \quad \forall i, j = 1, \ldots, n.$$

These derivatives can be grouped in a matrix known as the *Hessian* matrix, **H**:

$$\mathbf{H} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}'} = \begin{bmatrix} \partial^2 f/\partial x_1 \partial x_1 & \partial^2 f/\partial x_1 \partial x_2 & \cdots & \partial^2 f/\partial x_1 \partial x_n \\ \partial^2 f/\partial x_2 \partial x_1 & \partial^2 f/\partial x_2 \partial x_2 & \cdots & \partial^2 f/\partial x_2 \partial x_n \\ \vdots & \vdots & \ldots & \vdots \\ \partial^2 f/\partial x_n \partial x_1 & \partial^2 f/\partial x_n \partial x_2 & \cdots & \partial^2 f/\partial x_n \partial x_n \end{bmatrix}.$$

Since the order of differentiation does not matter, **H** is a symmetric matrix.

Hessian II

Second Derivative Test

If $\mathbf{x} = \mathbf{c} \in \mathbb{R}^n$ is a critical point of $f(\mathbf{x})$ and if the matrix $\mathbf{H}$ exists, then it can be used as a second derivative test in determining local optima as follows:

- If $\mathbf{H}$ is positive definite, then $f(\cdot)$ attains a *local minimum* at $\mathbf{c}$
- If $\mathbf{H}$ is negative definite, then $f(\cdot)$ attains a *local maximum* at $\mathbf{c}$
- If the Eigenvalues of $\mathbf{H}$ alternate in sign, then $f(\cdot)$ has a *saddle point* at $\mathbf{c}$.

Jacobian and the Vector-Valued Function

Definition

Suppose we have the following *vector-valued function* $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, that is we have $f_i(x_1, x_2, \ldots, x_n) = f_i(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$ for $i = 1, \ldots, m$. Collect all the m functions in an $(m \times 1)$ column vector

$$f(\mathbf{x}) = [f_1(\mathbf{x}) \quad f_2(\mathbf{x}) \quad \cdots \quad f_m(\mathbf{x})]',$$

then the derivative of $f(\mathbf{x})$ w.r.t. $\mathbf{x}$ is called the *Jacobian* matrix:

$$\mathbf{J} = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}'} = \begin{bmatrix} \partial f_1/\partial x_1 & \partial f_1/\partial x_2 & \cdots & \partial f_1/\partial x_n \\ \partial f_2/\partial x_1 & \partial f_2/\partial x_2 & \cdots & \partial f_2/\partial x_n \\ \vdots & \vdots & \ldots & \vdots \\ \partial f_m/\partial x_1 & \partial f_m/\partial x_2 & \cdots & \partial f_m/\partial x_n \end{bmatrix}.$$

Note

Both **H** and **J** are sometimes defined as their respective determinants.

Matrix Differentiation Rules

Matrix Differentiation Rules

- For the $(n \times 1)$ vector $\mathbf{a}$ we have $\partial \mathbf{a}'\mathbf{x}/\partial \mathbf{x} = \partial \mathbf{x}'\mathbf{a}/\partial \mathbf{x} = \mathbf{a}$
- For the $(m \times n)$ matrix $\mathbf{A}$ we have $\partial \mathbf{A}\mathbf{x}/\partial \mathbf{x}' = \mathbf{A}$
- For the $(n \times n)$ matrix $\mathbf{C}$, the quadratic form $\mathbf{x}'\mathbf{C}\mathbf{x}$ has the derivative $\partial \mathbf{x}'\mathbf{C}\mathbf{x}/\partial \mathbf{x} = (\mathbf{C}' + \mathbf{C})\mathbf{x}$, which is $2\mathbf{C}\mathbf{x}$ if $\mathbf{C}$ is symmetric
- For the $(n \times n)$ matrix $\mathbf{C}$, $\partial \mathbf{C}^{-1}/\partial \mathbf{x} = -\mathbf{C}^{-1}(\partial \mathbf{C}/\partial \mathbf{x})\mathbf{C}^{-1}$

Class Exercise

Show that the first two points hold.

Multiple Integrals

Definition

Multiple integrals are of the form

$$\int\int\cdots\int f(x_1, x_2, \ldots, x_n)dx_1 dx_2 \cdots dx_n.$$

They can be solved by integrating x_1 out, then x_2 and so on. Note however that the order of integration does not matter.

Class Exercise

Solve

$$\int_0^2\int_1^2 e^{-(x+y)}dxdy$$

by integrating out x and y and then y and x. Compare your answers.

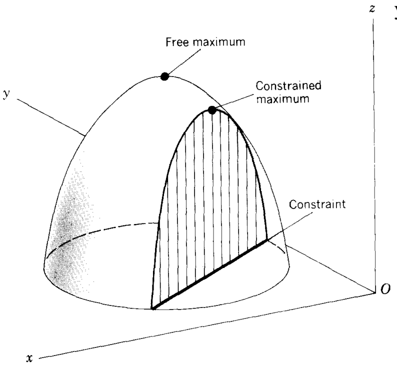
Constrained Maximization

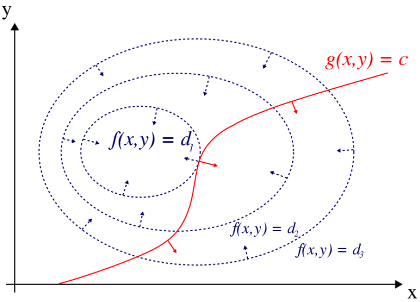
- In economics we often maximize functions under constraints
- The first order necessary conditions differ for constrained maximization
- **Lagrangian** method for equality constraints
- **Karush-Kuhn-Tucker (KKT) conditions** for equality and inequality constraints
- **Hamiltonian** function for dynamic optimization problems.

Lagrange Multipliers: Introduction

Consider two variables and one constraint, the optimization problem is:

$$\max f(x, y) \quad \text{s.t.} \quad g(x, y) = 0.$$





Lagrange Multipliers: Concept

- Going along $g(x, y) = 0$ a maximum point would be when f stops increasing. This could mean:
  1. The contour lines of f and g are parallel; we are walking along the contour line of f
  2. f does not change in any direction; we are at a level part of f, i.e. a critical point of f.
- If the contour lines of f and g are parallel, then:

$$\nabla f = \lambda \nabla g$$

  for the values of x and y at that point
- In the second case $\nabla f = 0$ and thus setting $\lambda = 0$ is a solution.

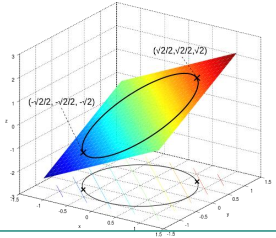
Lagrangian Function

The above two conditions can be incorporated in the *Lagrangian function*:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

Example

Maximize $f(x, y) = x + y$ subject to $x^2 + y^2 = 1$.



General Framework

In the general case the problem is stated as

$$\max f(x_1, x_2, \ldots, x_n) \ \text{ s.t.}$$
$$g_i(x_1, x_2, \ldots, x_n) = 0, \quad i = 1, \ldots, M.$$

The Lagrangian function then is

$$\begin{aligned}\mathcal{L}(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_M) &= f(x_1, x_2, \ldots, x_n) \\ &- \sum_{i=1}^{M} \lambda_i g_i(x_1, x_2, \ldots, x_n)\end{aligned}$$

Bordered Hessian

Sufficient conditions for an optimum are found through the signs of the principle minors of the bordered Hessian

$$\mathbf{H}^B = \begin{bmatrix} 0 & 0 & \cdots & 0 & g_1^1 & g_1^2 & \cdots & g_1^n \\ 0 & 0 & \cdots & 0 & g_2^1 & g_2^2 & \cdots & g_2^n \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 & g_M^1 & g_M^2 & \cdots & g_M^n \\ g_1^1 & g_2^1 & \cdots & g_M^1 & \mathcal{L}_{11} & \mathcal{L}_{12} & \cdots & \mathcal{L}_{1n} \\ g_1^2 & g_2^2 & \cdots & g_M^2 & \mathcal{L}_{21} & \mathcal{L}_{22} & \cdots & \mathcal{L}_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ g_1^n & g_2^n & \cdots & g_M^n & \mathcal{L}_{n1} & \mathcal{L}_{n2} & \cdots & \mathcal{L}_{nn} \end{bmatrix}$$

evaluated at the critical values, where

$$g_i^j = \frac{\partial g_i}{\partial x_j}, i = 1, \ldots, M, \; j = 1, \ldots, n \text{ and } \mathcal{L}_{lk} = \frac{\partial^2 \mathcal{L}}{\partial x_l \partial x_k}, \; l, k = 1, \ldots, n.$$

Sufficient Conditions

The second order sufficient conditions are:

- For a maximum the principle minors of $\mathbf{H}^B$ need to alternate in sign, with $|\mathbf{H}^B_{M+i}| = (-1)^{M+i}, i = 1, \ldots, M-n$
- For a minimum the principle minors of $\mathbf{H}^B$ need to have the same sign, namely $(-1)^M$.

Note, the first principle minor is $|\mathbf{H}^B_{M+1}|$, the determinant of the submatrix with $\mathcal{L}_{M+1,M+1}$ as the last element of the principle diagonal. The last principle minor is the determinant of $\mathbf{H}^B$ itself.

Example

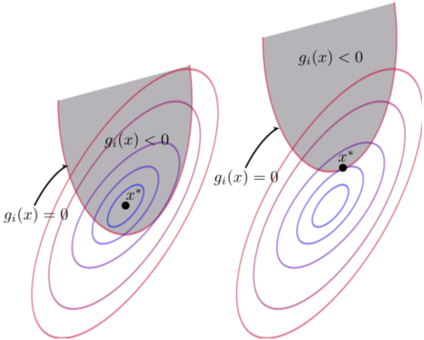
Evaluate the optima from the previous example using the second order conditions.

Inequality Constraints

Suppose now we have equality and inequality constraints:

$$\max f(\mathbf{x}) \text{ s.t.}$$

$$\begin{aligned} h_i(\mathbf{x}) &= 0, \quad i = 1, \ldots, M \\ g_j(\mathbf{x}) &\leq 0, \quad j = 1, \ldots, K \end{aligned}$$



The Karush-Kuhn-Tucker (KKT) Conditions

Given the Lagrangian

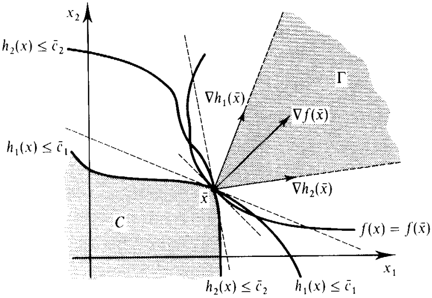
$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) - \sum_{i=1}^{M} \lambda_i h_i(\mathbf{x}) - \sum_{j=1}^{K} \mu_j g_j(\mathbf{x})$$

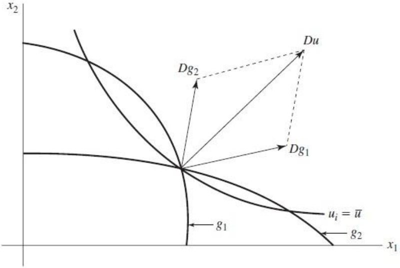
The necessary conditions for a local maximizer are

$$\begin{aligned} \nabla f &= \sum_{i=1}^{M} \lambda_i \nabla h_i + \sum_{j=1}^{K} \mu_j \nabla g_j \\ h_i(\mathbf{x}) &= 0 \quad i = 1, \ldots, M \\ g_j(\mathbf{x}) &\leq 0 \quad j = 1, \ldots, K \\ \mu_j &\geq 0 \quad j = 1, \ldots, K \\ \mu_j g_j(\mathbf{x}) &= 0 \quad j = 1, \ldots, K \ \text{ complementary slackness.} \end{aligned}$$

These are known as the Karush-Kuhn-Tucker (KKT) Conditions.

Sufficient Conditions





Sufficient conditions for a global maximum are that $f(\mathbf{x})$ is strictly quasiconcave and the constraints are convex sets.

Applications

Example

Maximize $f(x, y) = xy$ subject to $x + y^2 \leq 2, \quad x, y \geq 0$.

Class Exercise

1. Maximize $f(x, y) = 4x + 3y$ subject to $2x + y \leq 10, \quad x, y \geq 0$.
2. Suppose an individual has a quasiconcave utility function $u(\mathbf{x})$ for a vector of goods $\mathbf{x}$ with corresponding prices $\mathbf{p}$, and earns a wage of w. What would be the conditions to obtain an optimum utility level and what result do you derive?

In economics the multipliers are interpreted as the marginal or shadow value of relaxing the constraint.

Optimization in Continuous Time

Suppose we wish to maximize utility in continuous time subject to law of motion constraints:

$$\max_{c(t)} \int_0^\infty e^{-\rho t} u(c(t))dt \ \text{ s.t.}$$

$$\dot{k}(t) = f(k(t)) - c(t) - \delta k(t)$$

where

$$\dot{k}(t) = \frac{dk(t)}{dt}.$$

Here $c(t)$ is chosen by the agent and referred to as a *control* variable $k(t)$ describes the state of the economy and is referred to as a *state* variable.
Note, the constraint will thus involve a derivative of the state variable!

Hamiltonian

The problem is solved by setting up a *Hamiltonian* function and using Pontryagin's maximum principle

$$\mathcal{H}(c(t), k(t), \lambda(t)) = e^{-\rho t}u(c(t)) + \lambda(t)[f(k(t)) - c(t) - \delta k(t)]$$

with the necessary conditions

$$\frac{\partial \mathcal{H}}{\partial c(t)} = 0$$

$$\frac{\partial \mathcal{H}}{\partial k(t)} = -\dot{\lambda}(t)$$

$$\lim_{t\to\infty} [\lambda(t)k(t)] = 0 \text{ transversality condition.}$$

These conditions are also sufficient for an optimum if $u(\cdot)$ and $g(\cdot) = \dot{k}$ are concave in $c(t)$ and $k(t)$ respectively.

Example (Optional)

Derive the Euler equation in the above model.

Present Vs Current-Value Hamiltonian

The above Hamiltonian is known as the *present-value* Hamiltonian, valued at time 0. The *current-value*, time t Hamiltonian is

$$\mathcal{H}^*(c(t), k(t), \lambda(t)^*) = u(c(t)) + \lambda^*(t)[f(k(t)) - c(t) - \delta k(t)]$$

with the modified necessary conditions

$$\frac{\partial \mathcal{H}^*}{\partial c(t)} = 0$$

$$\frac{\partial \mathcal{H}^*}{\partial k(t)} = \rho\lambda^*(t) - \dot{\lambda}^*(t)$$

$$\lim_{t\to\infty} [\lambda^*(t)k(t)e^{-\rho t}] = 0 \text{ transversality condition.}$$

Multiple Variables

In case of a vector of control variables $\mathbf{c}(t) = [c_1(t) \ \cdots \ c_n(t)]'$ and state variables $\mathbf{k}(t) = [k_1(t) \ \cdots \ k_n(t)]'$ subject to constraints

$$\dot{k}_i = g_i(\mathbf{k}(t), \mathbf{c}(t)), \quad i = 1, \ldots, M$$

initial conditions $k_i(0) = k_{i0} > 0$ and no Ponzi games condition $\lim_{t\to\infty} k_i(t)e^{-\bar{r}(t)t} \geq 0$ the Hamiltonian is

$$\mathcal{H}(\mathbf{c}(t), \mathbf{k}(t), \boldsymbol{\lambda}(t)) = e^{-\rho t}u(\mathbf{c}(t)) + \sum_{i=1}^{M} \lambda_i(t)\dot{k}_i$$

with the necessary conditions

$$\frac{\partial \mathcal{H}}{\partial c_j(t)} = 0 \ \ j = 1, \ldots, n$$

$$\frac{\partial \mathcal{H}}{\partial k_i(t)} = -\dot{\lambda}_i(t) \ \ i = 1, \ldots, M$$

$$\lim_{t\to\infty} [\lambda_i(t)k(t)_i] = 0 \ \ \text{transversality condition.}$$

Logarithms in Economics

Definition

The *natural logarithm* is the inverse of e

$$\ln(e^x) = x.$$

For $x > 0$ it also follows that

$$x = e^{\ln(x)}.$$

In economics texts this is usually denoted as *log* instead of *ln*.

Class Exercise

From the above definition show that

$$\ln(x^a) = a\ln(x).$$

Logarithms as Percentages

Logarithms as Percentages

In economics the logarithm is often used to describe a percentage. For some small Δ

$$\ln(1+\Delta) \approx \Delta.$$

Class Exercise

Show that this holds by taking a first-order Taylor expansion of the function $f(x) = \ln(x)$ around $c = 1$ evaluated at $x = 1$.

Complex Numbers $\mathbb{C}$ I

Definition

A *complex number*, c is written in the form of

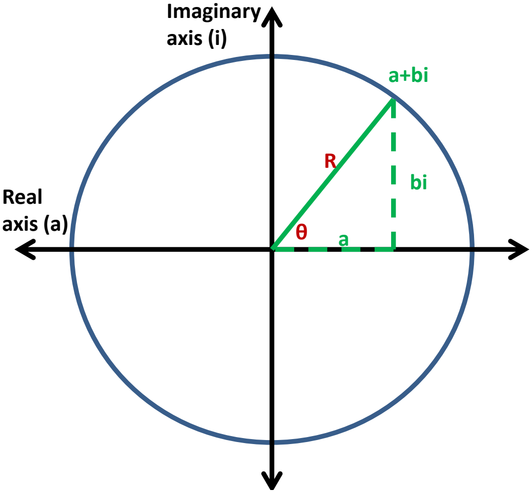
$$c = a + bi$$

where a and b are real numbers and i is an *imaginary* number such that $i = \sqrt{-1}$.

The Set $\mathbb{C}$

- The real number set is a subset of the complex number set, $\mathbb{R} \subset \mathbb{C}$
- $\mathbb{C}$ is a field obeying the field axioms (since $\mathbb{R}$ is also a field)
- $\mathbb{C}$ however does not satisfy the order axioms, unlike $\mathbb{R}$.

Complex Numbers $\mathbb{C}$ II: Complex Circle



Complex Numbers $\mathbb{C}$ III

Definition

A *complex conjugate*, of a complex number $a + bi$ is given by

$$a - bi.$$

Both numbers are known as a *complex conjugate pair*. Addition and multiplication of complex conjugate pairs produce real numbers.

Definition

The *modulus* or absolute value of a complex number, is its distance from the origin as measured by the $L - 2$ norm

$$R \equiv \|a + bi\|_2 = \sqrt{a^2 + b^2}.$$

Complex Numbers $\mathbb{C}$ IV

Polar Coordinates

A complex number can also be written in *Polar coordinate form*. Note that

$$cos(\theta) = a/R$$

and

$$sin(\theta) = b/R$$

hence,

$$a + bi = R[cos(\theta) + i \cdot sin(\theta)].$$

Complex Numbers $\mathbb{C}$ V

Exponent Form

Using the Power series it is also possible to write a complex number as

$$a + bi = Re^{i\theta}.$$

Example

Verify the above result by taking the Power series of $f(x) = \sin(x)$ and $f(x) = \cos(x)$. **Hint**: Also take the Power series of e^x evaluated at $x = i\theta$ and plug in all the results in the Polar coordinate form.

L' Hôpital's Rule

Definition

Let $f(\cdot)$ and $g(\cdot)$ be functions, which are differentiable on the open interval set X. Then for some point $x^0 \in X$ if

- $\lim_{x \to x^0} f(x) = 0$ and $\lim_{x \to x^0} g(x) = 0$ and $\lim_{x \to x^0} \frac{f'(x)}{g'(x)}$ exists or if
- $\lim_{x \to x^0} f(x) = \pm\infty$ and $\lim_{x \to x^0} g(x) = \pm\infty$ and $\lim_{x \to x^0} \frac{f'(x)}{g'(x)}$ exists

then $\lim_{x \to x^0} \frac{f(x)}{g(x)} = \lim_{x \to x^0} \frac{f'(x)}{g'(x)}$.

Class Exercise

Use L' Hôpital's Rule to evaluate

$$\lim_{x \to 0} \frac{f(x)}{g(x)} = \lim_{x \to 0} \frac{2\sin(x) - \sin(2x)}{x - \sin(x)}.$$

Homogenous Functions

Definition

The function $f(x_1, x_2, \ldots, x_n)$ is *homogenous of degree* $r \in \mathbb{R}$ if for any $t > 0$

$$f(tx_1, tx_2, \ldots, tx_n) = t^r f(x_1, x_2, \ldots, x_n).$$

Such functions are commonly used in economics. A function that is homogenous of degree 1 can be manipulated as

$$f(1, x_2/x_1, \ldots, x_n/x_1) = (1/x_1) f(x_1, x_2, \ldots, x_n).$$

Class Exercise

1. Give an example of a homogenous function in economics.
2. Show that $f(x_1, x_2) = \sqrt{x_1 x_2}$ is homogenous of degree 1.

Implicit Function Theorem I

Setting

Suppose we have a system of n functions $f_i : \mathbb{R}^{n+m} \to \mathbb{R}$, which depend on n 'parameters' $\mathbf{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]'$ and m 'endogenous' variables $\mathbf{y} = [y_1 \quad y_2 \quad \cdots \quad y_m]'$, i.e. $f_i(\mathbf{x}, \mathbf{y})$, where $\mathbf{x} \in X \subset \mathbb{R}^n$ and $\mathbf{y} \in Y \subset \mathbb{R}^m$. Further, suppose that

$$f_i(\mathbf{x}, \mathbf{y}) = 0 \quad i = 1, \ldots, n. \tag{1}$$

Assume the vectors $\mathbf{x}^0 \in X$ and $\mathbf{y}^0 \in Y$ satisfy (1).

Goal

We want to solve for $\mathbf{y}$ as a function of $\mathbf{x}$ *locally* around $\mathbf{x}^0$ and $\mathbf{y}^0$.

Implicit Function Theorem II

Definition

Equation (1) can be solved *locally* around $(\mathbf{x}^0, \mathbf{y}^0)$ so that $\mathbf{y}$ is a function of $\mathbf{x}$ if there are open neighbourhoods of $\mathbf{x}^0$ and $\mathbf{y}^0$ such that

$$f_i(\mathbf{x}, (g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_n(\mathbf{x}))) = 0, \quad i = 1, \ldots, n, \tag{2}$$

where

$$y_i = g_i(\mathbf{x}), \quad i = 1, \ldots, n,$$

provided that $\mathbf{x}$ is in its open neighbourhood. $g_i(\mathbf{x})$ are know as *unique implicit functions*.

Definition

An *open neighbourhood* of $\mathbf{x} \in \mathbb{R}^n$ is defined as

$$\{\mathbf{x}^0 \in \mathbb{R}^n : \|\mathbf{x}^0 - \mathbf{x}\|_2 < \varepsilon\}.$$

Implicit Function Theorem III

Implicit Function Theorem

Suppose all n functions are continuously partially differentiable w.r.t. their $n + m$ variables. A solution of the type of (2), involving unique implicit functions for the endogenous variables, is possible provided that

$$|J_{\mathbf{y}}|_{(\mathbf{x}=\mathbf{x}^0, \mathbf{y}=\mathbf{y}^0)} \neq 0.$$

Further, the effect of a change of $\mathbf{x}$ on $\mathbf{y}$ at $(\mathbf{x}^0, \mathbf{y}^0)$ can be expressed as

$$\frac{\partial g(\mathbf{x})}{\partial \mathbf{x}'} = -\left[\frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial \mathbf{y}}\right]^{-1} \frac{\partial f(\mathbf{x}, \mathbf{y})}{\partial \mathbf{x}},$$

evaluated at $(\mathbf{x}^0, \mathbf{y}^0)$.

Implicit Function Theorem IV

Intuitively, the IFT provides a *sufficient* condition for the existence of unique solutions (equilibria) and gives us the first-order comparative static effects of such a solution.

Example

For the unit circle, $x^2 + y^2 = 1 \Leftrightarrow f(x, y) = x^2 + y^2 - 1 = 0$ can we solve uniquely for x in terms of y at the point $(x^0, y^0) = (1, 0)$ s.t. $f(x, g(x)) = 0$?

Class Exercise

Can we do this at a point (x^0, y^0) where $y^0 > 0$ or where $y^0 < 0$?

Class Exercise

To illustrate that the IFT only gives a sufficient condition for the existence of unique solutions solve $f(x, y) = y^3 - x = 0$ for y in terms of x at the point $(x^0, y^0) = (0, 0)$.

Inverse Function Theorem

Inverse Function Theorem

When the number of parameters are the same as the number of endogenous variables, $m = n$ we have that

$$f_i(\mathbf{x}, \mathbf{y}) = g_i(\mathbf{x}) - y_i, \quad i = 1, \ldots, n.$$

which is known as the *Inverse Function Theorem*.

This is a special case of the IFT.

End of Theme 4

DIW Berlin – Deutsches Institut
für Wirtschaftsforschung e.V.
Mohrenstraße 58, 10117 Berlin
www.diw.de