Chapter 3: Analysis I

Multivariate Calculus

Julian Klix | E600 Mathematics Fall Semester 2026

UNIVERSITÄT MANNHEIM

Outline

In this chapter, we discuss

- A formal introduction to multi-dimensional functions
- Key function properties: invertability, convexity (and concavity)
- Multivariate differentiation (main focus)
    - Formal definition and derivation
    - Application
- Multivariate integration: concept and key theorems

Key Concepts

- Function $f: X \mapsto Y$ with domain X, codomain Y and image $\text{im}(f) = f[X]$
  - A function assigns exactly one value in the codomain to every value in its domain
  - $X \subseteq \mathbb{R}$: univariate function
  - $X \subseteq \mathbb{R}^n$: multivariate function
  - $Y \subseteq \mathbb{R}$: real-valued function
  - $Y \subseteq \mathbb{R}^m$: vector-valued function
- Examples:
  - Multivariate, real-valued function: $x \mapsto \|x\|$, $x \mapsto x'Ax$, $(x, y) \mapsto x \cdot y$
  - Multivariate, vector-valued function: $x \mapsto Ax$
- Graph:

$$G(f) = \{(x, y) \in X \times Y : y = f(x)\} = \{(x, f(x)) : x \in X\}$$

Invertability of Functions

- Mapping properties of a function $f : X \mapsto Y$

  Surjectivity: for every $y \in Y$ at least one x maps to y

  $\forall y \in Y \exists x \in X: \; f(x) = y$

  Injectivity: for every $y \in Y$ at most one x maps to y

  $\forall x_1, x_2 \in X, \; x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$

  Bijective: Injective + Surjective

- **Inverse function** f^{-1} of f: $f(f^{-1}(y)) = y$ and $f^{-1}(f(x)) = x$
  - More formally: $f^{-1} \circ f = Id_X$, $f \circ f^{-1} = Id_Y$ [$Id_Z : Z \mapsto Z, z \mapsto z$ is the identity function]
  - Consistent with our usual notion of inversion "$x \cdot x^{-1} = 1$"

Invertibility of a Function

A funtion $f : X \mapsto Y$ is invertible if and only if f is bijective

Intermediate Value Theorem

Intermediate Value Theorem

Consider an interval $I = [a, b]$ of real numbers and a continuous function $f : I \mapsto \mathbb{R}$. Then, if u is a number and $u \in (f(a), f(b))$, then there exists a $c \in (a, b) : f(c) = u$

$\Rightarrow$ Useful for showing surjectivity of continuous functions

Bolzano's Theorem

Consider an interval $I = [a, b]$ of real numbers and a continuous function $f : I \mapsto \mathbb{R}$. Then, if a and b are of opposite signs, there exists at least one point x_0 such that $f(x_0) = 0$

Convexity (and Concavity)

Convex and Concave Real Valued Function

Let $X \subseteq \mathbb{R}^n$ be a convex set. A function $f : X \to \mathbb{R}$ is **convex** if for any $x, y \in X$ and $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Moreover, if for any $x, y \in X$ such that $y \neq x$ and $\lambda \in (0, 1)$,

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

we say that f is **strictly** convex.
Moreover, we say that f is (strictly) concave if $-f$ is (strictly) convex.

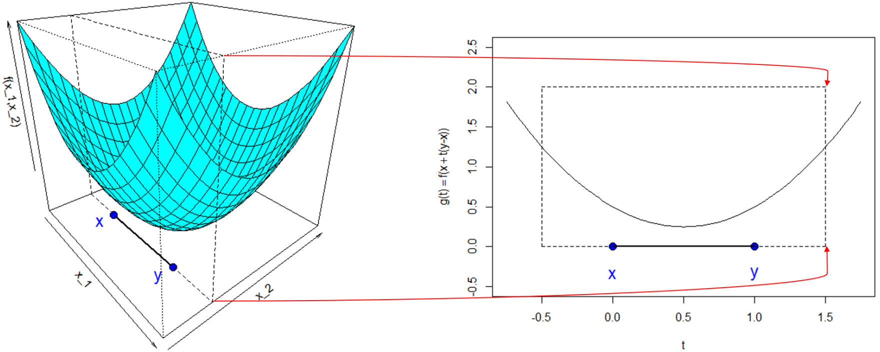
$\Rightarrow$ For concavity (line 1) and strict concavity (line 2) reverse respective inequalities

Convexity – Univariate Intuition

- In what follows: focus on convexity
- Recall: $\lambda x + (1-\lambda)y$ ($\lambda \in [0,1]$) is a convex combination of x and y
  - Convexity of functions = statement about convex combinations across domain and codomain of the function!
  - $f(\lambda x + (1-\lambda)y)$ must always be well-defined $\rightarrow$ convex domain
- $G(f) \subseteq \mathbb{R}^2$ (i.e. f univariate, real-valued function):
  - $(1-\lambda)x + \lambda y$, $\lambda \in [0,1]$ defines an interval between x and y
    - $\lambda = 0$: start from x
    - increasing λ moves away from x towards y
  - $(1-\lambda)f(x) + \lambda f(y)$ is the line piece connecting $f(x)$ and $f(y)$

  $\Rightarrow$ Line piece must lie above function graph

Convexity of Multivariate Functions



- Multivariate function $f : X \mapsto \mathbb{R}$, $X \subseteq \mathbb{R}^n$
- For any fixed $x, y \in X$, $(1-\lambda)x + \lambda y = x + \lambda(y-x)$ expands in a single direction
- Gives univariate function $t \mapsto f(x + t(y-x)) \rightarrow$ univariate definition applies

Disproving Convexity

Disproving Convexity

Let $X \subseteq \mathbb{R}^n$ be a convex set and $f: X \mapsto \mathbb{R}$. Then, if there exist $x_0 \in X$ and $i \in \{1, \ldots, n\}$ such that $g : \mathbb{R} \mapsto \mathbb{R}, t \mapsto f(x_0 + t \cdot e_i)$ is not (strictly) convex, then f is not (strictly) convex.

- Recall: e_i canonical basis vector of $\mathbb{R}^n$
- Necessary condition of convexity: convex in every *fundamental direction* of $\mathbb{R}^n$
- Example: $f: \mathbb{R}^n_+ \mapsto \mathbb{R}^n$ with $f(x) = h(x_1, \ldots, x_{n-1}) \cdot \sqrt{x_n}$
  where $h(x_1, \ldots, x_{n-1}) \geq 0$ is an arbitrary, unspecified function

Level Sets – Definition

- Optimization: convexity immensely helpful, but restrictive concept
- Most desirable properties will be preserved under quasi-convexity.
- To define quasi-convexity, we need level sets in the domain of f.

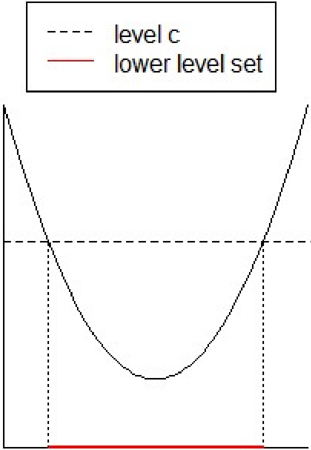
Lower and Upper Level Set of a Function

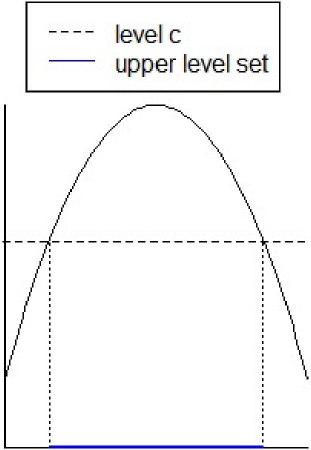
Let $X \subseteq \mathbb{R}^n$ be a convex set and $f : X \to \mathbb{R}$ be a real-valued function. Then, for $c \in \mathbb{R}$, the sets

$$L_c^- := \{x \in X : f(x) \leq c\} \text{ and } L_c^+ := \{x \in X : f(x) \geq c\}$$

are called the lower level and upper level set of f at c, respectively.

Level Sets – Illustration





Quasi-Convexity – Definition

Quasiconvexity of real-valued Functions

Let $X \subseteq \mathbb{R}^n$ be a convex set. A real-valued function $f: X \to \mathbb{R}$ is quasiconvex if and only if

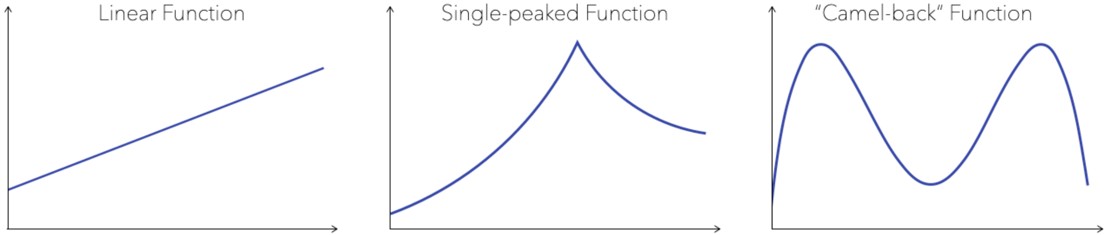
$$\forall x, y \in X \ \forall \lambda \in [0,1] : f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\}$$

Conversely, f is quasiconcave if and only if

$$\forall x, y \in X \ \forall \lambda \in [0,1] : f(\lambda x + (1-\lambda)y) \geq \min\{f(x), f(y)\}$$

$\Rightarrow$ For strict versions, replace respective inequalities with strict inequalities

Quasi-Convexity – Examples and Alternative Definition



Quasiconvexity and Level Sets

Let $X \subseteq \mathbb{R}^n$ be a convex set. A real-valued function $f : X \to \mathbb{R}$ is quasiconvex if and only if $\forall c \in \mathbb{R}$ the lower level set L_c^- is a convex set.

$\Rightarrow$ For quasiconcavity, replace lower with upper level sets

Differentiation – Univariate Introduction

- As before: start from univariate real-valued functions and generalize.
- If $X \subseteq \mathbb{R}$, what is "the slope" of $f : X \mapsto \mathbb{R}$?
- **Relative** rate of change of $f(x)$ given variation in x at $x_0 \in X$:

$$\frac{\Delta f(x)}{\Delta x} := \frac{f(x) - f(x_0)}{x - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}, \quad h := x - x_0 \in \mathbb{R}$$

$\Rightarrow$ Difference quotient

- **Marginal** rate of change: make difference arbitrarily small

$$\frac{df(x)}{dx} := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

$\Rightarrow$ Differential quotient

Differentiation – Univariate Defintion

Univariate Real-Valued Function: Differentiability and Derivative

Let $X \subseteq \mathbb{R}$ and consider the function $f: X \mapsto \mathbb{R}$. Let $x_0 \in X$. If

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists, f is said to be differentiable at x_0, and we call this limit the **derivative of f at x_0**, denoted by $f'(x_0)$.
If for all $x_0 \in X$, f is differentiable at x_0, f is said to be differentiable (over X). Then, the function $f' : X \mapsto \mathbb{R}, x \mapsto f'(x)$ is called the **derivative** of f.

- Differentiability: point-specific vs. domain

  Examples: absolute value function, square root function

Differentiation – The Differential Operator

Univariate Real-Valued Function: Differential Operator

Let $X \subseteq \mathbb{R}$, define $D^1(X, \mathbb{R}) = \{f : X \mapsto \mathbb{R} : f$ is differentiable over X$\}$, and let $F_X := \{f : X \mapsto \mathbb{R}\}$. Then, the differential operator is defined as the function

$$\frac{d}{dx} : D^1(X, \mathbb{R}) \mapsto F_X, f \mapsto f'$$

where f' denotes the derivative of $f \in D^1(X, \mathbb{R})$.

- (Differential) Operator: function between function spaces
- $f' = \frac{d}{dx}(f)$ is a specific value in the codomain of $\frac{d}{dx}$
- Levels of objects in differentiation: operator, function, value
- Please **don't** write $\frac{df(x)}{dx}$ [either write $\frac{df}{dx}(x)$ or $\frac{d}{dx}f(x)$]

Operations in the Function Space

- Set of Functions F_X can form a (vector) space as well
  $\Rightarrow$ allows to define (algebraic) operations between functions

- **Function sum**: $f+g$ is such that $\forall x \in X : (f+g)(x) = f(x) + g(x)$
- **Scalar product**: λf is such that $\forall x \in X : (\lambda f)(x) = \lambda \cdot f(x)$
- **Function product**: $f \cdot g$ is such that $\forall x \in X : (f \cdot g)(x) = f(x) \cdot g(x)$
- **Function chain**: $f \circ g$ is such that $\forall x \in X : (f \circ g)(x) = f(g(x))$

- **Function quotient**: $f/g = f \cdot g^{-1}$ [requires invertability]

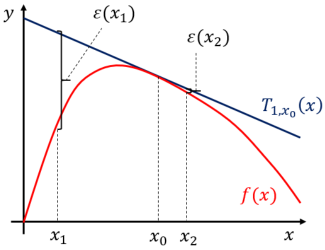
Differentiation Rules for Univariate Derivatives

Rules for Univariate Derivatives

Let $X \subseteq \mathbb{R}$, $f, g \in D^1(X, \mathbb{R})$ and $\lambda, \mu \in \mathbb{R}$. Then,

1. (Linearity) $\lambda f + \mu g$ is differentiable and
   $\frac{d}{dx}(\lambda f + \mu g) = \lambda \frac{df}{dx} + \mu \frac{dg}{dx}$,
2. (Product Rule) The product fg is differentiable and
   $\frac{d}{dx}(fg) = \frac{df}{dx} \cdot g + f \cdot \frac{dg}{dx}$
3. (Quotient Rule) If $\forall x \in X$, $g(x) \neq 0$, the quotient f/g is differentiable and
   $\frac{d}{dx}(f/g) = \frac{\frac{df}{dx} \cdot g - f \cdot \frac{dg}{dx}}{g \cdot g}$
4. (Chain Rule) if $g \circ f$ exists, the function is differentiable and
   $\frac{d}{dx}(g \circ f) = \left(\frac{dg}{dx} \circ f\right) \cdot \frac{df}{dx}$.

Local Approximation using Derivatives – Taylor's Theorem



- Crucial theorem in many fields (incl. Economics)
- First order **Taylor approximation** to f at x_0:

$$T_{1,x_0}(x) = f(x_0) + f'(x_0)(x - x_0)$$

- Error: $\varepsilon_{1,x_0}(x) := f(x) - T_{1,x_0}(x)$ (formula: next slide)

- "Good" approximation:
  - $\lim_{x \to x_0} \frac{\varepsilon_1(x)}{x - x_0} = 0$ "minimize" (relative) errors locally
  - $T^{(n)}_{N,x_0}(x_0) = f^{(n)}(x_0)\ \forall n < N$ (locally) match as many "moments" of f as possible
- Taylor expansion of first order: decomposition of f into linear and error term

$$f(x) = T_{1,x_0}(x) + \varepsilon_{1,x_0}(x)$$

Taylor Expansion – Formal Definition

Taylor Expansion for Univariate Functions

Let $X \subseteq \mathbb{R}$ and $f \in D^d(X, \mathbb{R})$ where $d \in \mathbb{N} \cup \{\infty\}$. For $N \in \mathbb{N} \cup \{\infty\}$, $N \leq d$, the Taylor expansion of order N for f at $x_0 \in X$ is given by

$$f(x) = T_{N,x_0}(x) + \varepsilon_{N,x_0}(x) = f(x_0) + \sum_{n=1}^{N} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \varepsilon_{N,x_0}(x),$$

where $\varepsilon_{N,x_0}(x)$ is the approximation error of T_{N,x_0} for f at $x \in X$.

Then, the approximation quality satisfies $\lim_{h\to 0} \varepsilon_{N,x_0}(x_0 + h)/h^N = 0$ where $h = x - x_0$. Further, if f is $N+1$ times differentiable, $\exists \lambda \in (0,1)$ such that

$$\varepsilon_{N,x_0}(x_0 + h) = \frac{f^{(N+1)}(x_0 + \lambda h)}{(N+1)!} h^{N+1}.$$

Differentiation – Helpful Theorems

Mean Value Theorem

Let $f : X \mapsto \mathbb{R}$ be differentiable on $[x_1, x_2] \subset X$. Then, there exists $x^* \in (x_1, x_2)$ with

$$f'(x^*) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Inverse Function Theorem

Let $f : X \mapsto \mathbb{R}$ be continuously differentiable on $[x_1, x_2] \subset X$ and $f'(x) \neq 0$ for all $x \in [x_1, x_2]$. Then, the inverse $g = f^{-1}$ is differentiable at $y \in f\big([x_1, x_2]\big)$ and

$$g'(y) = \frac{1}{f'(g(y))}.$$

Differentiation – Multivariate Introduction

- In univariate case (when $n = 1$), d^* is the derivative of f at x_0 if

$$d^* = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

$\Rightarrow$ Problem: if $n > 1$, the denominator has a *vector*; not defined

- But: in $\mathbb{R}$, expression is *equivalent* to

$$\lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - d^* \cdot h|}{||h||} = 0$$

where $||\cdot||$ is a norm on $\mathbb{R}$; and norms extend to $\mathbb{R}^n$!

$\Rightarrow$ Generalize definition (together with norm) to $\mathbb{R}^n$ [still **real-valued** case]

Differentiation – Multivariate Definition

Multivariate Derivative of Real-valued Functions

Let $X \subseteq \mathbb{R}^n$ and $f : X \mapsto \mathbb{R}$. Further, let $x_0 \in int(X)$ (interior point). Then, f is differentiable at x_0 if there exists $d^* \in \mathbb{R}^{1\times n}$ such that

$$\lim_{||h|| \to 0} \frac{|f(x_0 + h) - f(x_0) - d^* h|}{||h||} = 0.$$

Then, we call d^* the derivative of f at x_0, denoted $\frac{df}{dx}(x_0)$ or $D_f(x_0)$.
If f is differentiable at any $x_0 \in X$, we say that f is differentiable, and we call $\frac{df}{dx} : X \mapsto \mathbb{R}, x \mapsto \frac{df}{dx}(x)$ the derivative of f.

- Interior point: able to consider balls around x_0 on which f is defined
- Most textbooks use $D_f(x_0)$ rather than $\frac{df}{dx}(x_0)$, is the same thing!
- Derivative operator as before: mapping between function spaces

Differentiation – Partial Derivatives and Gradients

- **Partial derivative** of f at x_0 with respect to x_j:

$$\frac{\partial f}{\partial x_j}(x_0) = \frac{df_{e_j,x_0}}{dt}(0) = \frac{d}{dt} f(x_0 + te_j)|_{t=0}$$

  - Variation along j-th axis around x_0 ("holding x_i, $i \neq j$ constant")
  - Also: j-th partial derivative (of f at x_0); sometimes denoted $f_j(x_0)$

- **Gradient**: *ordered* collection of partial derivatives (**row** vector!)

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \frac{\partial f}{\partial x_2}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right)$$

Differentiation – Overview of Concepts

- Partial differentiability
  - $f : X \mapsto \mathbb{R}$ partially differentiable (p.d.) at x_0:
    all partial derivatives $\frac{\partial f}{\partial x_j}(x_0)$ and therefore the gradient at $x_0 \in X$, $\nabla f(x_0)$, exists
  - "point-specific to general": $f : X \mapsto \mathbb{R}$ p.d.: f p.d. at any $x_0 \in X$
  - Set of p.d. functions from X to $\mathbb{R}$: $D_p^1(X, \mathbb{R}) = \{f : X \mapsto \mathbb{R} : f \text{ is p.d.}\}$
- Recall: univariate derivative is a real-valued function
  - $\frac{\partial f}{\partial x_j} : X \mapsto \mathbb{R}, x_0 \mapsto \frac{\partial f}{\partial x_j}(x_0)$ is a real-valued function
  - $\nabla f : X \mapsto \mathbb{R}^{1\times n}, x_0 \mapsto \nabla f(x_0)$ is a (real row-)vector-valued function
- associated operators: mappings between function spaces
  - $\frac{\partial}{\partial x_j} : D_p^1(X, \mathbb{R}) \mapsto F_X, f \mapsto f_j = \frac{\partial f}{\partial x_j}$
  - $\nabla : D_p^1(X, \mathbb{R}) \mapsto F_X^{1\times n}, f \mapsto \nabla f$

Differentiation – Gradient and Derivative

The Gradient and the Derivative

Let $X \subseteq \mathbb{R}^n$ and $f : X \mapsto \mathbb{R}$ such that f is differentiable at $x_0 \in int(X)$. Then, all partial derivatives of f at x_0 exist, and $\frac{df}{dx}(x_0) = \nabla f(x_0)$.

Partial Differentiablility and Differentiability

Let $X \subseteq \mathbb{R}^n$, $f : X \mapsto \mathbb{R}$ and $x_0 \in int(X)$. If all partial derivatives of f at x_0 exist and are continuous, then f is differentiable.

- Set of **continuously differentiable functions**:

$$C^1(X, \mathbb{R}) := \left\{ f : X \mapsto \mathbb{R} : \left(\forall j \in \{1, \ldots, n\} : \frac{\partial f}{\partial x_j} \text{ is continuous} \right) \right\}$$

Differentiation – Conclusion

- Partial differentiability and differentiability
  - Generally, if f is differentiable, the derivative is equal to the gradient
    $\Rightarrow$ if the gradient does not exist, f is not differentiable
  - Theoretically: may encounter weird D^1 but not C^1 functions
    $\Rightarrow$ issue not too relevant in (economic) practice

- In applications: taking the derivative of $f : X \mapsto \mathbb{R}$, $X \subseteq \mathbb{R}^n$
  1. Compute all partial derivatives $\frac{\partial f}{\partial x_j}$
  2. Are all partial derivatives continuous? If so: ∇f is the derivative!
  3. Stack the partial derivatives into the gradient ∇f

$\Rightarrow$ How about multivariate **vector-valued** functions?

Differentiation – Vector-valued Introduction

- Consider $X \subseteq \mathbb{R}^n$, $f: X \mapsto \mathbb{R}^m$
- f is ordered collection of real-valued functions which we know how to handle:

$$f = \begin{pmatrix} f^1 \\ \vdots \\ f^m \end{pmatrix} \text{ so that } \forall x \in X: f(x) = \begin{pmatrix} f^1(x) \\ \vdots \\ f^m(x) \end{pmatrix} \text{ where } \forall i \in \{1, \dots, m\}, f^i : X \mapsto \mathbb{R}$$

- Idea: ordered collection of derivatives, i.e.

$$\frac{df}{dx} = \begin{pmatrix} \nabla f^1 \\ \vdots \\ \nabla f^m \end{pmatrix} = \begin{pmatrix} f_1^1 & \dots & f_n^1 \\ \vdots & \ddots & \vdots \\ f_1^m & \dots & f_n^m \end{pmatrix}$$

Differentiation – Vector-valued Definition

Multivariate Derivative of Vector-valued Functions (Jacobian)

Let $X \subseteq \mathbb{R}^n$, $f : X \mapsto \mathbb{R}^m$ and $x_0 \in X$. For $i \in \{1, \ldots, m\}$, let $f^i : \mathbb{R}^n \mapsto \mathbb{R}$ such that $f = (f^1, \ldots, f^m)'$. Then, if at x_0, $\forall i \in \{1, \ldots, m\}$, f^i is partially differentiable with respect to any x_j, $j \in \{1, \ldots, n\}$, we call

$$J_f(x_0) = \begin{pmatrix} \nabla f^1(x_0) \\ \nabla f^2(x_0) \\ \vdots \\ \nabla f^m(x_0) \end{pmatrix} = \begin{pmatrix} f^1_1(x_0) & f^1_2(x_0) & \ldots & f^1_n(x_0) \\ f^2_1(x_0) & f^2_2(x_0) & \ldots & f^2_n(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f^m_1(x_0) & f^m_2(x_0) & \ldots & f^m_n(x_0) \end{pmatrix}$$

the **Jacobian** of f at x_0. If the above holds at any $x_0 \in X$, we call the mapping $J_f : \mathbb{R}^n \mapsto \mathbb{R}^{n \times m}, x_0 \mapsto J_f(x_0)$ the Jacobian of f.

Differentiation – Derivative and Jacobian

The Jacobian and the Derivative

Let $X \subseteq \mathbb{R}^n$, $f : X \mapsto \mathbb{R}^m$ and $f^1, \dots, f^m : X \mapsto \mathbb{R}$ such that $f = (f^1, \dots, f^m)'$. Further, let $x_0 \in int(X)$ (interior point), and suppose that f is differentiable at x_0. Then, for any f^i, $i \in \{1, \dots, m\}$, all partial derivatives of f^i at x_0 exist, and $\frac{df}{dx}(x_0) = J_f(x_0)$.

- Analogue to gradient: differentiability $\Rightarrow$ existence partial derivatives
  $\Rightarrow$ reverse also true: all partial derivatives continuous $\Rightarrow$ Jacobian is derivative
- Jacobian collects expansion in all fundamental directions of all sub-functions f^i
  $\Rightarrow$ generalizes derivative (and gradient) to vectorized functions

Vector-valued Differentiation Rules

Rules for Multivariate Derivatives

Let $X \subseteq \mathbb{R}^n$, $f, g : X \mapsto \mathbb{R}^m$ and $h : \mathbb{R}^m \mapsto \mathbb{R}^k$. Suppose that f, g and h are differentiable functions. Then,

- Linearity: For all $\lambda, \mu \in \mathbb{R}$, $\lambda f + \mu g$ is differentiable and $\frac{d(\lambda f + \mu g)}{dx} = \lambda \frac{df}{dx} + \mu \frac{dg}{dx}$.
- Product Rule: $f^T \cdot g$ is differentiable and $\frac{d(f^T g)}{dx} = f^T \cdot \frac{dg}{dx} + g^T \cdot \frac{df}{dx}$.
- Chain Rule: $h \circ f$ is differentiable and $\frac{d(h \circ f)}{dx} = (\frac{dh}{dx} \circ f) \cdot \frac{df}{dx}$.

- Chain rule special case for $f(g(x)) = f(y(x), x)$:

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

Second Derivative – Introduction

- Thus far: first derivative operator $(\cdot)'$ generalized to ∇ and J
- In univariate, real-valued case: $f'' = (f')'$, we can generalize this logic
- Recall: derivative increases order in codomain
  - Derivative of $f : \mathbb{R}^n \mapsto \mathbb{R}$ is vector-valued: $\nabla f : \mathbb{R}^n \mapsto \mathbb{R}^{1\times n}$
  - Derivative of $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is matrix-valued: $J_f : \mathbb{R}^n \mapsto \mathbb{R}^{m\times n}$

  $\Rightarrow$ Derivative of $\nabla f^T : \mathbb{R}^n \mapsto \mathbb{R}^n$ should be matrix-valued : $J_{\nabla_f} : \mathbb{R}^n \mapsto \mathbb{R}^{n\times n}$

  $\Rightarrow$ Derivative of $J_f : \mathbb{R}^n \mapsto \mathbb{R}^{m\times n}$ should be **tensor**-valued [we will disregard these]

- Expectation: first derivative is vector $\to$ second is matrix
  - First derivative = gradient: $\nabla f : \mathbb{R}^n \mapsto \mathbb{R}^{1\times n}$
  - Second derivative = derivative of transposed gradient: $\frac{d}{dx}(\nabla f)'$

Second Derivative – The Hessian

- If $\frac{\partial f}{\partial x_i}$ is differentiable at x_0, the (i, j)-second order partial derivative at x_0 is

$$f_{i,j}(x_0) = \frac{\partial f_i}{\partial x_j}(x_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$$

Hessian or Hessian Matrix

Let $X \subseteq \mathbb{R}^n$ and $f: X \mapsto \mathbb{R}$. Further, let $x_0 \in int(X)$, and suppose that f is differentiable at x_0 and that all second order partial derivatives of f at x_0 exist. Then, the **Hessian** of f at x_0 is the matrix

$$H_f(x_0) = \begin{pmatrix} \nabla f_1(x_0) \\ \nabla f_2(x_0) \\ \vdots \\ \nabla f_n(x_0) \end{pmatrix} = \begin{pmatrix} f_{1,1}(x_0) & f_{1,2}(x_0) & \cdots & f_{1,n}(x_0) \\ f_{2,1}(x_0) & f_{2,2}(x_0) & \cdots & f_{2,n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ f_{n,1}(x_0) & f_{n,2}(x_0) & \cdots & f_{n,n}(x_0) \end{pmatrix}$$

Second Derivative – Helpful Theorems

- Generalize continuous differentiability sets:

$$C^k(X) = C^k(X, \mathbb{R}) = \{f : X \mapsto \mathbb{R} : \text{All } k\text{-th order part. deriv's are continuous}\}$$

Schwarz's Theorem/Young's Theorem

Let $X \subseteq \mathbb{R}^n$ be an open set and $f : \mathbb{R}^n \mapsto \mathbb{R}$. If $f \in C^k(X)$, then the order in which derivatives up to order k are taken can be permuted.

Hessian and Gradient

Let $X \subseteq \mathbb{R}^n$ and $f \in C^2(X)$. Then, the Hessian is symmetric and corresponds to the Jacobian of the transposed gradient: $H_f = J_{(\nabla f)^T}$.

Multivariate Taylor Approximation

Second Order Multivariate Taylor Approximation

Let $X \subseteq \mathbb{R}^n$ be an open set and consider $f \in C^2(X)$. Let $x_0 \in X$. Then, the second order Taylor approximation to f at $x_0 \in X$ is

$$T_{2,x_0}(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)' \cdot H_f(x_0) \cdot (x - x_0).$$

The error $\epsilon_{2,x_0}(x) = f(x) - T_{2,x_0}(x)$ approaches 0 at a faster rate than $\|x - x_0\|^2$, i.e. $\lim_{\|h\| \to 0} \frac{\epsilon_{2,x_0}(x+h)}{\|h\|^2} = 0$.

- Analogous to the univariate Taylor approximation
- Error formula for first order: there exists $\lambda \in (0, 1)$ so that

$$\epsilon_{1,x_0}(x_0 + h) = \frac{1}{2} h' \cdot H_f(x_0 + \lambda h) \cdot h$$

Total Derivative

- Directional derivative of f at x_0 in direction $z \neq \mathbf{0}$ (Chain Rule):
  $\Rightarrow$ instead of canonical direction of partial derivative

$$\frac{d}{dt} f(x_0 + tz)\Big|_{t=0} = \nabla f(x_0) \cdot z = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) \cdot z_i$$

- Notation: $z = (dx_1, \ldots, dx_n)$ as vector of *relative variation* in the arguments
  $\rightarrow$ df as resulting *relative* induced marginal change

$$df = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} dx_i \quad \text{or alternatively} \quad df(x_0) = \sum_{i=1}^{n} f_i(x_0) dx_i$$

- In economics:
  - variation in fixed ratios/specific directions $\rightarrow$ comparative statics
  - Consideration is *relative*: fix one reference variable j with $dx_j = 1$

Second Derivative and Convexity

- For univariate functions ($X \subseteq \mathbb{R}$, $f \in C^2(X)$):
  - f is convex if and only if $\forall x \in X : f''(x) \geq 0$ (equivalent condition)
  - If $\forall x \in X : f''(x) > 0$, then f is strictly convex (sufficient condition)
- For multivariate f: we can study $g : \mathbb{R} \mapsto \mathbb{R}, t \mapsto f(x + tz)$ for $x, z \in \mathbb{R}^n$, $z \neq \mathbf{0}$
  - If $f \in C^2(X)$ then especially $g \in C^2(\mathbb{R})$ for fixed x, z
  - Second derivative (applying chain rule):

$$g''(t) = z^T H_f(x + tz)z$$

  - This implies:

    $\Rightarrow \forall y \in X : (H_f(y)$ pos. semi-definite$) \Leftrightarrow f$ convex (proof in script)

    $\Rightarrow \forall y \in X : (H_f(y)$ pos. definite$) \Rightarrow f$ strictly convex
  - Intuition: definiteness of the symmetric Hessian $\widehat{=}$ sign

Multivariate Differentiation – Final Remarks

- A lot of notation and definitions...
- Key take-aways:
  1. Gradients and Jacobians are the derivatives of multivariate functions
     - ...if the components (partial derivatives) are continuous; i.e. almost always
     - Intuition: summary of variation in fundamental directions
  2. Univariate differentiation rules translate to multivariate equivalents
  3. Taylor expansions give "good" polynomial approximations "close to" the approximation point
  4. Second derivatives of real-valued multivariate functions ("Hessian") can be obtained from differentiating the (transposed) gradient
  5. The definiteness of the Hessian determines convexity/concavity

Univariate Integration

- f is the instantaneous change of its accumulation

  $\Rightarrow$ If the integral measures accumulation, the function itself should be the integral's derivative!

- Idea: obtain integral operator by inverting the derivative operator

$$\frac{d}{dx} : D^1(X) \mapsto F_X, f \mapsto \frac{df}{dx}$$

- Issue: recall that inversion requires injectivity ("one-to-one")
  - $f(x) = 2x + 3$ vs. $g(x) = 2x \quad \rightarrow f'(x) = g'(x) = 2$
  - Problem: constants cancel out when taking the derivative
  - Derivative is unique up to the constant!

Indefinite Integrals

- Restrict attention to univariate, real-valued $f : X \mapsto \mathbb{R}$
- We can't invert $\frac{d}{dx}$, let's do the next best thing:

$$\int : F_X \mapsto \mathcal{P}(D^1(X)), f \mapsto \{\tilde{F} : X \mapsto \mathbb{R} : \frac{d\tilde{F}}{dx} = f\}$$

where $\mathcal{P}$ denotes the power set (set of all subsets)

$\Rightarrow$ Correspondence: set-valued mapping, **not** a function!

- We write $\int f = \{\tilde{F} : X \mapsto \mathbb{R} : \frac{d\tilde{F}}{dx} = f\}$ (called pre-image of f under $\frac{d}{dx}$)
- Any $\tilde{F} \in \int f$ has the form $\tilde{F}(x) = F(x) + C$ for a $C \in \mathbb{R}$
  - F has no constant, i.e. $F(\min X) = 0$ [or technically $\lim_{x \to \inf X} F(x) = 0$]
  - Notation: $\tilde{F}(x) = \int f(x)dx = F(x) + C$ for an unspecified C

Indefinite Integrals – Helpful Rules

Rules for Indefinite Integrals

Let f, g be two integrable functions and let $a, b \in \mathbb{R}$ be constants, $n \in \mathbb{N}$. Then

- $\int (af(x) + g(x))dx = a \int f(x)dx + \int g(x)dx$,
- $\int x^n dx = \frac{x^{n+1}}{n+1} + C$ *if* $n \neq -1$ *and* $\int \frac{1}{x} dx = \ln(x) + C$,
- $\int e^x dx = e^x + C$ *and* $\int e^{f(x)} f'(x)dx = e^{f(x)} + C$,
- $\int (f(x))^n f'(x)dx = \frac{1}{n+1}(f(x))^{n+1} + C$ *if* $n \neq -1$ *and* $\int \frac{f(x)}{f'(x)} dx = \ln(f(x)) + C$.

Integration by parts (Reverse Product Rule)

Let u, v be two differentiable functions. Then,

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx.$$

Definite Intergral

- Accumulation is unique up to initial level C :
  For any $\tilde{F} = F + C \in \int f$ and any $x, y \in X$: $\tilde{F}(y) - \tilde{F}(x) = F(y) - F(x)$
- Uniquely defined **Definite Integral**:

$$\int_x^y f(t)dt = \tilde{F}(y) - \tilde{F}(x), \quad \text{where} \quad \tilde{F}(x) \in \frac{d}{dx}^{-1} f$$

Fundamental Theorem of Calculus

Let X be an interval in $\mathbb{R}$ with $a = \inf(X)$ and $f : X \mapsto \mathbb{R}$. Suppose that f is integrable, and define $F := X \mapsto \mathbb{R}, x \mapsto \int_a^x f(t)dt$. Then, F is differentiable, and

$$\forall x \in X : F'(x) = \frac{dF}{dx}(x) = f(x).$$

Multivariate Integration

Fubini's Theorem

Let X and Y be two intervals in $\mathbb{R}$, let $f: X \times Y \to \mathbb{R}$ and suppose that f is continuous. Then, for any $I = I_x \times I_y \subseteq X \times Y$ with intervals $I_x \subseteq X$ and $I_y \subseteq Y$,

$$\int_I f(x, y) d(x, y) = \int_{I_x} \left(\int_{I_y} f(x, y) dy \right) dx$$

and all the integrals on the right-hand side are well-defined.

$\Rightarrow$ extendable to n-dimensional integration

$$\int_I f(x_1, \ldots, x_n) d(x_1, \ldots, x_n) = \int_{I_1} \left(\cdots \left(\int_{I_n} f(x_1, \ldots, x_n) dx_n \right) \cdots \right) dx_1.$$

Multivariate Integration – Helpful Theorem

Integration of Multiplicatively Separable Functions

Let $X_f \in \mathbb{R}^n, X_b \in \mathbb{R}^m$, $f : X_f \to \mathbb{R}$, $g : X_b \to \mathbb{R}$ continuous functions. Then, for any intervals $A \subseteq X_f$ and $B \subseteq X_g$

$$\int_{A \times B} f(x)g(y)d(x,y) = \left(\int_A f(x)dx\right)\left(\int_B g(y)dy\right).$$

- As with derivatives: multivariate integrals can be simplified
  1. Check for multiplicative separability of the integrand (and separate if applicable)
  2. Transform individual integrals into series of univariate integrals
  3. Compute “simple” integrals from inside out
- Order of sequential integration can be interchanged

Recap Chapter 3

Multivariate Functions

- Invertibility of one-to-one functions
- Convexity and quasi-convexity

Differentiation

- Logic extends to multivariate cases
- Vectors (Gradient) and matrices (Jacobian) as derivatives

Integration

- Again, logic extends to multivariate cases
- Transform multivariate integral into series of univariates

That's all Folks!

Please take a look at the problem set that we will discuss tomorrow morning.

julian.klix@uni-mannheim.de

L7 3-5, Room 3.43

UNIVERSITÄT MANNHEIM