Chapter 4: Analysis II

Optimization

Julian Klix | E600 Mathematics | Fall Semester 2026

UNIVERSITÄT MANNHEIM

Outline

In this chapter, we discuss

- The formal basics of mathematical optimization
- Unconstrained optimization and its justification
- Optimization with one equality constraint and its justification
- Generalization to more complex problems
- Solution techniques (especially: simplification)

Constrained Optimization – Introduction

- Unconstrained optimization problem

$$\underset{x \in \operatorname{dom}(f)}{\text{maximize}} \; f(x)$$

- Optimization problem with equality constraints

$x_1 + x_2 =$ const.

$$\max_{\operatorname{dom}(f)} \; f(x) \qquad \text{subject to} \quad g_i(x) = 0,\; i = 1, \ldots, m$$

→ Lagrangian

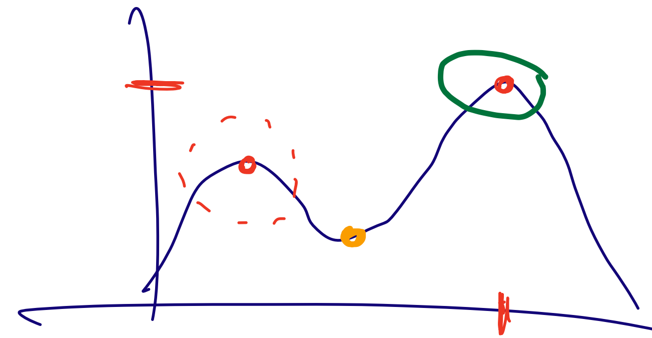
- General constrained optimization problem

$$\max_{\operatorname{dom}(f)} \; f(x) \qquad \text{s.t.} \quad g_i(x) = 0,\; i = 1, \ldots, m$$
$$h_i(x) \leq 0,\; i = 1, \ldots, k$$

→ Karush-Kuhn-Tucker

$\Rightarrow$ We focus on maximization problems $\rightarrow$ fully analogous for minimization!

Language for Optimization Problems



- **(Global) maximum** for f if $\forall x \in X : f(x_0) \geq f(x)$

  $\Rightarrow$ maximum refers to function value $f(x_0) = \max f(x)$

  $\Rightarrow$ maximizer refers to value $x_0 = \arg\min f(x)$

- **Local maximum** if $\exists \varepsilon > 0$ such that $\forall x \in X \cap B_\varepsilon(x_0) : f(x_0) \geq f(x)$

  $\Rightarrow$ Strict versions: inequalities strict for all $x \neq x_0$

- Global maximization implies local maximization

  $\Rightarrow$ Definitions analogously for minima and minimizers

- **Extremum** (global or local): Either maximum or minimum

- **Critical point**: potential candidate for an extremum (e.g. FOC)

Definitions for Optimization Problems

- **Maximizer** (alternatively maximum argument):

$$\arg\max_{C(\mathcal{P})} := \{x^* \in C(\mathcal{P}) : f(x^*) = \max_{C(\mathcal{P})} f(x)\}$$

$x^* \in \arg\max x$

$(x^* = \arg\max x)$

uniqueness

$\Rightarrow$ Generally a set $\rightarrow$ treated as the element for singletons ($\Leftrightarrow$ uniqueness)!

- **Constraint set** of a problem $\mathcal{P}$: subset in domain of f

$$C(\mathcal{P}) := \{x \in X : \forall i \in \{1, \dots, m\} \wedge \forall j \in \{1, \dots, k\} : g_i(x) = 0 \wedge h_j(x) \leq 0\}$$

- **Restricted function** for a subset $A \subseteq \text{dom}(f)$: $f|_A : A \mapsto \mathbb{R}, x \mapsto f(x)$

- **Constrained maximizer** of f in the problem $\mathcal{P}$: maximizer of $f|_{C(\mathcal{P})}$

Extremum Existence

Weierstrass Extreme Value Theorem

Suppose that $X \subseteq \mathbb{R}^n$ is compact, and that $f : X \mapsto \mathbb{R}$ is continuous. Then, f assumes its maximum and minimum on X, such that

$$\arg\max_{x \in X} f(x) \neq \varnothing \quad \text{and} \quad \arg\min_{x \in X} f(x) \neq \varnothing$$

- Includes multivariate (real-valued) functions
- Recall: In Euclidean vector spaces $\mathbb{X} = (\mathbb{R}^n, +, \cdot)$ compact $\Leftrightarrow$ closed & bounded
  - Closedness: Extremum as corner solution is included
  - Boundedness: Extremum is reached (e.g. monotone functions)
  - Continuity: Extremum isnt part of a jump

Roadmap for Unconstrained Optimization

- General maximization approach
  $\Rightarrow$ Reduce domain to set of candidate points
  $\Rightarrow$ Find all local maximizers and see which are global (comparing values)

- Outline
  - "Good" necessary conditions give a relatively narrow set of candidates
    $\Rightarrow$ Generally, first order conditions (FOC) are used
  - Further candidates: boundary points, points of non-differentiability
  - Sufficient conditions further restrict the set of candidates
  - Existence is useful to guarantee that at least one candidate is a solution

Necessary Conditions for Local Maximizers: FOC

$f(x^2)$ FOC $f'(x) = 2x \overset{!}{=} 0$ $f'' \gtrless 0$

Critical Point or Stationary Point

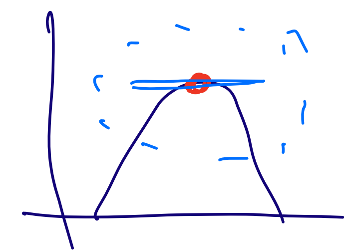
Let $X \subseteq \mathbb{R}^n$, $f: X \mapsto \mathbb{R}$ and $x^* \in X$. Then, if f is differentiable at x^* and $\nabla f(x^*) = \mathbf{0}$, we call x^* a critical point of f or a stationary point of f.

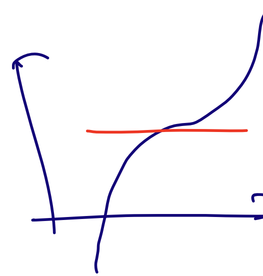
$\Rightarrow$ FOC

- Necessary FOC: all *interior* local maxima are critical points
- Not sufficient: more points feature $\nabla f(x) = 0$

  $\Rightarrow$ Local minima and "saddle points"
- More insight from second derivative?
  - $f \in C^2(\mathbb{R})$: f' positive before and negative after local maximizer x^*
    $\Rightarrow$ f' decreasing around x^*: $f''(x^*) < 0$
  - Recall: definiteness $\approx$ "sign" of symmetric matrix

Hf



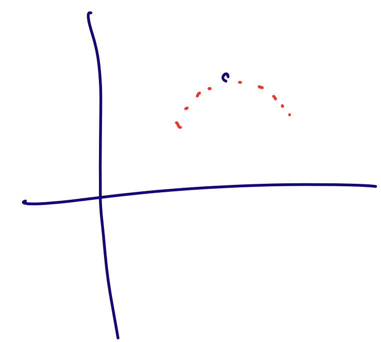


Sufficient Conditions for Local Maximizers: SOC

Unconstrained Local Maximum – Sufficient Condition

Let $X \subseteq \mathbb{R}^n$, $f \in C^2(X)$ and $x^* \in int(X)$. Suppose that x^* is a critical point of f, and that $H_f(x^*)$ is negative definite. Then, x^* is a strict local maximizer of f.

- For univariate functions: $H_f(x^*)$ negative definite $\Leftrightarrow$ $f''(x) < 0$
- Minimum: FOC + Hessian *positive definiteness*
- Careful: what about x^* where
  1. the FOC holds: $\nabla f(x^*) = 0$,
  2. $H_f(x^*)$ is negative semi-definite, but
  3. $H_f(x^*)$ is not negative definite?



$\Rightarrow$ Cannot rule out as a solution, compare values to other candidates!

A Shortcut for Definiteness

- Define leading principal minor D_k

$$D_k = det(A_k) \text{ where } A_k = (a_{ij})_{i \in \{1,\dots,k\}, j \in \{1,\dots,k\}}$$

$\Rightarrow$ Submatrix consisting of first k rows and columns

det det det

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & \ddots & \cdots \\ \vdots & & \ddots \end{pmatrix}$$

Definiteness and Leading Principal Minors

Let $A \in \mathbb{R}^{n \times n}$ be a symmetrical matrix. Then the following is true

1. A is positive definite if and only if $D_k > 0$ for all k
2. A is negative definite if and only if $(-1)^k D_k > 0$ for all k

$$A = \begin{pmatrix} 5 & 1 \\ 1 & 6 \end{pmatrix}$$

det $D_1 < 0$ det $D_2 > 0$ det $D_3 < 0$

- More general version for semi-definiteness exists (but less useful in practice)

Global Maximization – Helpful Theorems

max → min $-f$ $\qquad$ min f = max $-f$

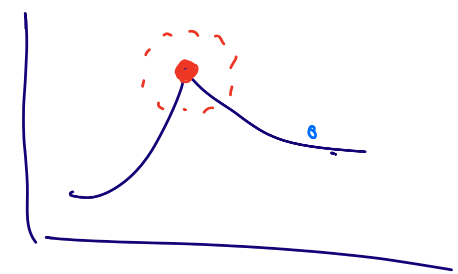
Sufficiency for the Global Unconstrained Maximum

convex

Let $X \subseteq \mathbb{R}^n$ be a convex set, and $f \in C^2(X)$. Then, if f is **concave** and for $x^* \in int(X)$, it holds that $\nabla f(x^*) = \mathbf{0}$, then x^* is a global maximizer of f.

minimizer

- Does not require compactness
- Quasi-concavity not sufficient $\quad \rightarrow$ saddle points
- For global minimizers, f has to be convex instead



Maximum and Quasi-Convexity

quasi-convex

Let $X \subseteq \mathbb{R}^n$ be a convex set, and $f \in C^2(X)$. Then if f is quasi-concave and x_0 a local maximizer, then x_0 is also a global maximizer of f.

minimizer $\qquad$ minimizer

A Cookbook Recipe for the Global Maximum

→ Compact
→ continuity

-1. Determine whether a solution exists at all (optional)
0. Collect border candidates: boundary points, non-differentiability, limits
1. Interior solutions: Finding and eliminating candidates
   - If existence guaranteed and at any step, only one candidate (including border) remains, stop, you found the maximum!
   - Necessary FOC: Initial set of candidates = critical values
   - Necessary SOC: rule out those that violate it
   - If only one candidate (including border) remains
     $\Rightarrow$ Existence guaranteed? Or: sufficient condition holds?
2. Multiple candidates remaining: compare values, check existence if not already done

Optimization Problems with Equality Constraints

A Simple Special Case

max f s.t. $g(x) = 0$

- Example with single constraint
- Re-write the constrained problem to use our approach to unconstrained problems

$p_1 x_1 + p_2 x_2 - y = 0$

$$\max_{x \in \mathbb{R}^2} u(x_1, x_2) \text{ s.t. } y = p_1 x_1 + p_2 x_2 \Leftrightarrow \max_{x_1 \in \mathbb{R}} u\left(x_1, \frac{y - p_1 x_1}{p_2}\right)$$

$x^2 + y^2 = 1$

...having solved the constraint for $x_2 = \frac{y - p_1 x_1}{p_2}$

- Works if we can find an expression of the constraint for x_1 that is...
  - explicit: we can write down the equation for x_1 in terms of x_{-1}
    ...vs. implicit: we know that there is some function $x_1 = h(x_{-1})$
  - global: the function applies to the whole domain
    ...vs. local: applies around a local maximizer
- Particularly promising for $m < n$ linear constraints (see example)

The Lagrangian – A General Approach

Lagrange's Necessary First Order Condition [Single Constraint]

Consider the constrained problem $\max f|_{C(g)}$ where $X \subseteq \mathbb{R}^n$, $f, g \in C^1(X)$ and x^* such that $g(x^*) = 0$. Suppose that $\nabla g(x^*) \neq \mathbf{0}$. Then, x^* is a local maximizer of the constrained problem only if there exists $\lambda \in \mathbb{R} : \nabla f(x^*) = \lambda \nabla g(x^*)$.
If such $\lambda \in \mathbb{R}$ exists, we call it the Lagrange multiplier associated with x^*.

$\nabla f = 0$

$= \lambda g(x^*)$

$g(x) = 0$

- Here: only constraint $g : \mathbb{R}^n \mapsto \mathbb{R}$ instead of multiple constraints
- **Lagrangian function** (or: "Lagrangian"): $\mathcal{L}(\lambda, x) = f(x) - \lambda g(x)$ $\to$ FOC

  $\nabla f - \lambda \nabla g = 0 \iff \nabla f = \lambda \nabla g$
- Note, $x^* \in X$ satisfies the FOC iff for a $\lambda \in \mathbb{R}$, (λ, x) is a critical value of $\mathcal{L}(\lambda, x)$!

  $\Rightarrow$ FOC for $x : \nabla f(x^*) - \lambda \nabla g(x^*) = \mathbf{0} \iff \nabla f(x^*) = \lambda \nabla g(x^*)$ [Lagrangian FOC]

  $\Rightarrow$ FOC for $\lambda : g(x^*) = 0$ [Constraint is satisfied]

The Lagrangian – An Example

$\max f(x) \iff \max t(f(x))$ str. incr. fct. $t(x)$

$\max e^x \iff \max x$ $t = e^x$

Example 1 Find the vector with minimum Euclidean length $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ in the $\mathbb{R}^2$ that satisfies $x_1 + x_2 = 1$, i.e. solve

$$\underbrace{\min_{x \in \mathbb{R}^2} \|x\|_2}_{f(x)} \text{ subject to } x_1 + x_2 = 1 \iff \max_{\mathbb{R}^2} -\|x\|_2 \text{ s.t. } x_1 + x_2 = 1$$

$\Rightarrow \underbrace{x_1 + x_2 - 1}_{g(x)} = 0$

$x_1 + x_2 = 1$

$x_2 = 1 - x_1$

Alternative (Simple) Solution Rewrite constraint as $x_2 = 1 - x_1$ and solve

$$\min_{x_1 \in \mathbb{R}} x_1^2 + (1 - x_1)^2$$

$\min \sqrt{x_1^2 + x_2^2} \iff \min x_1^2 + x_2^2$

$= x_1^2 + (1 - x_1)^2$

1. Finding critical values via FOC: $4x_1 - 2 = 0 \iff x_1 = \frac{1}{2} \Rightarrow x_2 = \frac{1}{2}$
2. Checking SOC: $4 > 0 \Rightarrow x^* = \left(\frac{1}{2}, \frac{1}{2}\right)$ (global) minimizer

$$\min \quad x_1^2 + x_2^2 \qquad g(x) = x_1 + x_2 - 1 = 0$$

$$\mathcal{L}(x_1, x_2, \lambda) = \overbrace{x_1^2 + x_2^2}^{f} \overbrace{- \lambda}^{-\lambda} \overbrace{(x_1 + x_2 - 1)}^{g}$$

$$\mathrm{FOC}_{x_1}: \quad 2x_1 - \lambda = 0$$

$$\mathrm{FOC}_{x_2}: \quad 2x_2 - \lambda = 0$$

$$\rightarrow 2x_1 - \lambda = 2x_2 - \lambda$$

$$\Leftrightarrow x_1 = x_2$$

$$\mathrm{FOC}_{\lambda}: \quad x_1 + x_2 - 1 = 0$$

$$x_1 + x_1 - 1 = 0 \quad \Rightarrow x_1 = \frac{1}{2}, \; x_2 = \frac{1}{2} \qquad (\lambda = 1)$$

Shadow cost of the constraint

The Lagrangian FOC for multiple constraints

Lagrange's Necessary First Order Condition [Multiple Constraints]

Consider the constrained problem $\max f|_{C(g)}$ where $X \subseteq \mathbb{R}^n$ and $f \in C^1(X)$, $\nabla g = 0$
$g \in C^1(X, \mathbb{R}^m)$. Let x^* be such that $g(x^*) = \mathbf{0}$ and suppose that $\operatorname{rk}(J_g(x^*)) = m$.
Then, x^* is a local maximizer of the constrained problem only if there exists
$\Lambda = (\lambda_1, \ldots, \lambda_m)' \in \mathbb{R}^m : \boxed{\nabla f(x^*) = \Lambda' J_g(x^*)}$.
If such $\Lambda \in \mathbb{R}^m$ exists, we call λ_i the Lagrange multiplier for the i-th constraint.

- Direct generalization of the single-constraint case
- **Multi-constraints Lagrangian function**: $\mathcal{L}(\Lambda, x) = f(x) - \underline{\Lambda' g(x)}$

$\Rightarrow$ Again, find critical values by checking FOC of $\mathcal{L}$

$g_1(x) = 0$
$g_2(x) = 0$

FOC: $\mathcal{L}(\lambda_1, \lambda_2, x) = f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x)$

Optimization Problems with Inequality Constraints

The General Constrained Optimization Problem

$-\mu h = 0$

- Optimization problem with equality and inequality constraints

$$\max_{\text{dom}(f)} f(x) \qquad \text{s.t.} \quad g(x) = \mathbf{0}, \quad g : \mathbb{R}^n \mapsto \mathbb{R}^m$$
$$h(x) \leq \mathbf{0}, \quad h : \mathbb{R}^n \mapsto \mathbb{R}^k$$

$h(x) \leq 0$

$h(x) < 0$ $\mu_j = 0$ $h(x) = 0$ $\mu_j \neq 0$

slack binding

- Lagrange problem except for additional inequality constraints
- Inequality constraints can be either binding or slack

  $\Rightarrow$ Binding constraint $\approx$ equality constraint (with $\mu_j \approx$ Lagrange multiplier)

  $\Rightarrow$ Slack constraint: No restriction on optimization (no value cost $\mu_j = 0$)
- Complementary slackness: either $h_j(x) = 0, \mu_j \neq 0$ **or** $h_j(x) < 0, \mu_j = 0$
- **Solve individual Lagrangian case-by-case for each slackness-combination!**

Karush-Kuhn-Tucker – A General Approach

Karush-Kuhn-Tucker Theorem

Consider the constrained problem $\max f|_{C(g,h)}$ where $X \subseteq \mathbb{R}^n$ and $f \in C^1(X)$, $g \in C^1(X, \mathbb{R}^m)$, $h \in C^1(X, \mathbb{R}^k)$. Then, if x^* is a local maximum of the constrained problem for which the set $\{\nabla h_j(x^*) : h_j(x^*) = 0\} \cup \{\nabla g_i(x^*) : i \in \{1, \dots, m\}\}$ is linearly independent, there exist $\Lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^k$ such that (x^*, Λ^*, μ^*) satisfy the optimality conditions:

$\nabla g = 0$

$x_i \geq 0$

1. (Feasibility) $\forall j \in \{1, \dots, k\} : h_j(x) \leq 0$ and $\forall i \in \{1, \dots, m\} : g_i(x) = 0$,
2. (FOC for x) $\nabla f(x) = \sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{j=1}^k \mu_j h_j(x)$,
3. (Complementary Slackness) $\forall j \in \{1, \dots, k\} : \mu_j h_j(x) = 0$.

$-x_i \leq 0$

- Linear independence condition fulfils similar role as rank condition before
- **Karush-Kuhn-Tucker Lagrangian function**: $\mathcal{L}(\lambda, \mu, x) = f(x) - \Lambda' g(x) - \mu' h(x)$

Karush-Kuhn-Tucker – An Example

Example 2 Find the vector with maximum Euclidean length $\|x\|_2 = \sqrt{x_1^2 + x_2^2}$ in the $\mathbb{R}^2_+$ that satisfies $x_1 + x_2 = 1$, i.e. solve

$$\max_{x \in \mathbb{R}^2} \|x\|_2 \quad \text{s.t.} \quad x_1 + x_2 = 1$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$

$\sqrt{x_1^2 + x_2^2}$

$\rightarrow \quad g(x) = 0 \qquad x_1 + x_2 - 1 = 0$

$-x_1 \leq 0$

$h(x) \leq 0 \qquad -x_2 \leq 0$

or, put differently (noting that $\sqrt{\cdot}$ is a monotone function)

$$\max_{x \in \mathbb{R}^2} x_1^2 + x_2^2 \quad \text{s.t.} \quad 1 - x_1 - x_2 = 0 \qquad g$$
$$\begin{pmatrix} -x_1 \\ -x_2 \end{pmatrix} \leq \mathbf{0} \qquad h$$

$$\mathcal{L}(x_1, x_2, \lambda, \mu_1, \mu_2) = x_1^2 + x_2^2 - \lambda(1 - x_1 - x_2) - \mu_1(-x_1) - \mu_2(-x_2)$$

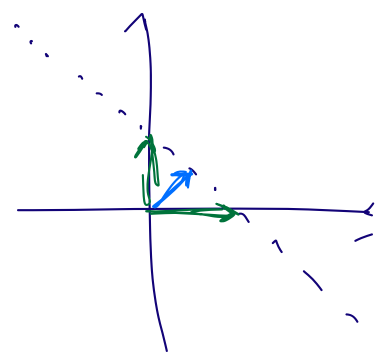
1) h_1, h_2 binding $\quad x_1 = 0 \quad x_2 = 0$ $\qquad$ No solution

$\Rightarrow \quad x_1 = 0 \quad x_2 = 0 \quad x_1 + x_2 = 1 \quad \perp$

2) h_1, h_2 non-binding $\quad x_1 > 0, \; x_2 > 0 \quad \Rightarrow \mu_1, \mu_2 = 0$

$$\mathcal{L} = x_1^2 + x_2^2 - \lambda(1 - x_1 - x_2)$$ Example 1 ← Minimum

$x_1 = \frac{1}{2} = x_2 \qquad \binom{1/2}{1/2}$ $\qquad$ No solution $\quad \left\| \binom{1/2}{1/2} \right\| = \frac{1}{4}$



3) h_1 binding $\quad h_2$ slack $\quad \mu_2 = 0$

$$\mathcal{L} = x_1^2 + x_2^2 - \lambda(1 - x_1 - x_2) + \mu_1 x_1$$

$FOC_{x_1} \quad 2x_1 + \lambda + \mu_1 = 0 \qquad \mu = 2$

$FOC_{x_2} \quad 2x_2 + \lambda = 0 \qquad \lambda = -2$

$FOC_{\lambda} \quad x_1 + x_2 = 1 \qquad x_2 = 1$

$FOC_{\mu} \quad x_1 = 0 \qquad \Rightarrow \binom{0}{1}$

4) h_1 slack, h_2 binding $\Rightarrow \quad x_1 > 0 \quad \boxed{x_2 = 0}$

$x_1 + x_2 = 1$

$x_1 = 1 \quad \Rightarrow \binom{1}{0}$

$$\Rightarrow \left\| \binom{1}{0} \right\| = \left\| \binom{0}{1} \right\| = 1$$

$$\arg\max = \left\{ \binom{1}{0}, \binom{0}{1} \right\}$$

Recap Chapter 4

Unconstrained Optimization

- Locally solveable by FOCs & SOCs
- Global solution requires further single-point comparisons

Equality-constrained Optimization → Lagrange approach

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

- Additional constraints that are *always* binding
- Use FOCs of Laplace-function (internalizing constraints)

Inequality-constrained Optimization → Karush-Kuhn-Tucker approach

$\mathcal{L} = f(x) - \lambda g(x) - \mu h(x)$

- Additional constraints that are *sometimes* binding
- Use FOCs of Laplace-function of all binding constraints

-> bordered Hessian

That's all Folks!

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

julian.klix@uni-mannheim.de

L7 3-5, Room 3.43

UNIVERSITÄT MANNHEIM