E600 Mathematics

Chapter 4: Optimization

Martin Reinhard

August 31/September 1, 2022

- This chapter discusses
  - 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)

- Economics (broadly): study of optimal ("efficient") allocations given limited resources → Sub-discipline of constrained optimization?!
  - Optimization: no better allocation ...
  - Constrained: taking as given constraints to the resources
- Examples
  - Utility maximization subject to budget constraint
  - Cost minimization given output target
  - Labor/leisure choice given time budget
  - Public good provision subject to government budget constraint
  - ...

1. Introduction

Mathematical Constrained Optimization Problem

$$(\mathcal{P}_{min}) \qquad \begin{aligned} \underset{x \in \text{dom}(f)}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) = 0, \; i = 1, \dots, m. \\ & h_i(x) \leq 0, \; i = 1, \dots, k. \end{aligned}$$

$$(\mathcal{P}_{max}) \qquad \begin{aligned} \underset{x \in \text{dom}(f)}{\text{maximize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) = 0, \; i = 1, \dots, m. \\ & h_i(x) \leq 0, \; i = 1, \dots, k. \end{aligned}$$

We focus on $\mathcal{P}_{max}$ (easy to show equivalence of solutions)!

- Step 1: Unconstrained Problem – formal idea in simplest scenario
- Step 2: Problem with one equality constraint: justify Lagrangian formally
- Step 3: General equality-constrained problems: multivariate generalization of step 2
- Step 4: Take the intuition of 3 to solve general problems with inequality constraints

- Maximum/minimum of a set X: (Extremum: min or max!)
  - $x = \max(X) \Leftrightarrow (x = \sup(X) \wedge x \in X)$
  - $x = \min(X) \Leftrightarrow (x = \inf(X) \wedge x \in X)$
- Local and global maximizers: $X \subseteq \mathbb{R}^n$, $f : X \mapsto \mathbb{R}$. Then, $x_0 \in X$ is a
  - Global maximizer for f if $\forall x \in X : f(x_0) \geq f(x)$
  - Local maximizer if $\exists \varepsilon > 0$ such that $\forall x \in X \cap B_\varepsilon(x_0) : f(x_0) \geq f(x)$
  - Strict versions: inequalities strict for all $x \neq x_0$
  - Global implies local
  - Graphically?

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

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

- **Restricted function**: $A \subseteq \text{dom}(f)$

$$f|_A : A \mapsto \mathbb{R}, x \mapsto f(x)$$

- **Constrained maximizer** of f in the problem $\mathcal{P}$: maximizer of $f|_{C(\mathcal{P})}$ (global maximizer = "**solution**")
  - $\arg\max f$: Set of global maximizers of f ($\arg\max f \subseteq \text{dom}(f)$)
  - $\max f$: Value at the maximum, $\max f = f(x^m)$, $x^m \in \arg\max f$
  - Solutions: $\arg\max f|_{C(\mathcal{P})}$, alternatively $\arg\max_{x \in C(\mathcal{P})} f(x)$

- Arg max and maximum summarized again:

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

  - If $\arg\max_{x \in C(\mathcal{P})} f(x) = \emptyset$, then $\max_{x \in C(\mathcal{P})} f(x)$ does not exist!
  - If there is only a single arg max x^*, we write

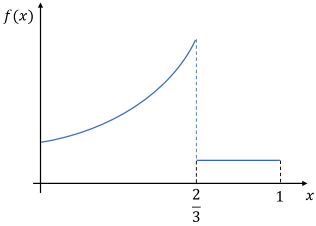
$$x^* = \arg\max_{x \in C(\mathcal{P})} f(x)$$

  - Constrained maximization problem: finding $\arg\max_{x \in C(\mathcal{P})} f(x)$!
- Actually: standard maximization of restricted function, but:
- restriction may not transfer appealing properties from f to $f|_{C(\mathcal{P})}$ (e.g. continuity) $\rightarrow$ isolated investigation

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

- Why $\text{dom}(f)$ compact = closed + bounded? And why continuous?:



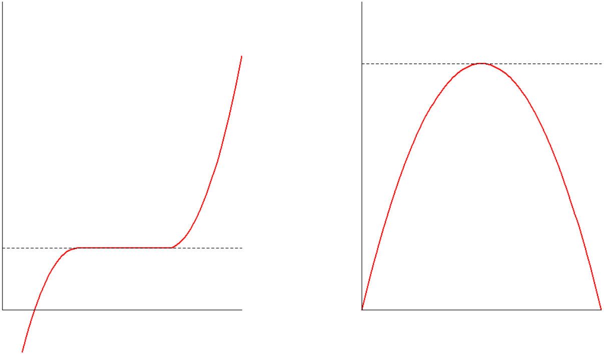
- Recall: intervals $[a, b] \subseteq \mathbb{R}$ are compact
- Example: $f : [0, 1] \mapsto \mathbb{R}$,

$$f(x) = \begin{cases} x^2 + 2 & x < 2/3 \\ 1 & x \geq 2/3 \end{cases}$$

- General maximization approach
  - Find all local maximizers and see which are global (comparing values)
  - Ideally: easy-to-check equivalent conditions
  - Alternatively: combination of necessary and sufficient conditions
- Outline
  - Equivalent conditions are generally hard to come by
  - "Good" necessary conditions ("If x_0 is a local maximizer, then...") give a relatively narrow set of candidates
  - 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

2. Unconstrained Optimization

Necessary Conditions for Local Maximizers



- Local maximizer of continuous function: flat or hill $\rightarrow$ zero slope!
- First order (first derivative) necessary condition: $f'(x^*) = 0$
- $\mathbb{R}^n$ ("no slope in any direction"): $\nabla f(x^*) = \mathbf{0}$ (FOC)

Definition (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.

- Necessary FOC: all *interior* local maxima are critical points
  - Proof: example for contrapositive method
- Not sufficient: more points feature $\nabla f(x) = 0$
  - Same logic applies to local minimum
  - "Saddle points": minimum in one and maximum in other direction
- 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^*: $f''(x^*) \leq 0$
  - Recall: definiteness ≈ "sign" of symmetric matrix

- Second Order Necessary Condition (SOC): If $f \in C^2(X)$

$$(x^* \text{ is loc. maximizer}) \Rightarrow (H_f(x^*) \text{ is neg. semi-definite})$$

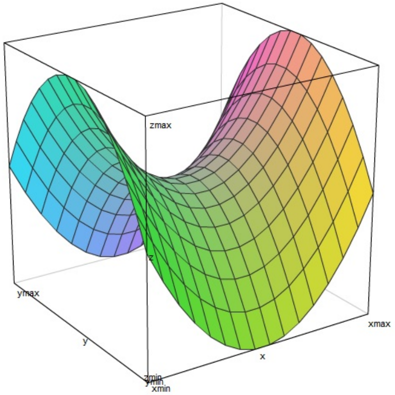
  - For minimum: H_f pos. semi-definite
- Example: $f(x_1, x_2) = x_1^2 - x_2^2$
  - Gradient: $\nabla f(x_1, x_2) = (2x_1, -2x_2)$, Hessian:

$$H_f(x) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

  - Definiteness:

$$z' H_f(x) z = (z_1, z_2) \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = (z_1, z_2) \begin{pmatrix} 2z_1 \\ -2z_2 \end{pmatrix} = 2(z_1^2 - z_2^2)$$

  - Indefinite everywhere $\Rightarrow$ no maximum or minimum!



- Left: graph of $f(x_1, x_2) = x_1^2 - x_2^2$
- Indefinite Hessian: critical values may be *saddle points*
- Why only necessary? Consider $n = 1$, $f(x) = x^3$
  - FOC: $f'(x^*) = 2(x^*)^2 = 0 \Leftrightarrow x^* = 0$
  - SOC: $f''(x^*) = 6x^*$, $f''(0) = 0$
  - $\forall v \in \mathbb{R} : v' f''(0) v = v^2 \cdot 0 = 0$

  $\rightarrow$ SOC holds, but not local maximum!

Theorem (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.

- Minimum: FOC + Hessian *positive definiteness*
- Careful: what about x^* where
  - (i) the FOC holds: $\nabla f(x^*) = 0$,
  - (ii) $H_f(x^*)$ is negative semi-definite, but
  - (iii) $H_f(x^*)$ is not negative definite?

  $\Rightarrow$ Cannot rule out as a solution, compare values to other candidates!
- Necess./suff. SOC proof: 1st order Taylor Expansion (see script)

Corollary (Sufficiency for the Global Unconstrained Maximum)

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.*

- Limit Behavior (Non-compact optimization):
  - For univariate functions: $\lim_{x \to \pm\infty} f(x) = c$ breaks the largest interior maximum $f(x^*)$ as the global maximum if and only if $f(x^*) < c$
  - For multivariate functions:
    - Our usual objectives f vanish asymptotically ($\lim_{\|x\| \to \infty} f(x) = -\infty$)
    - Else (not needed often): compare largest interior maximum $f(x^*)$ to $\lim_{\lambda \to \infty} f(\lambda v^*(\lambda))$ where $v^*(\lambda)$ is the direction v that maximizes $f(\lambda v)$ for fixed v with $\|v\| = 1$
    - Example: online exercises

-1. Determine whether a solution exists at all (optional)
0. Collect border candidates: boundary points, non-diff'ability, 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
     - Existence guaranteed? Or: sufficient condition holds? Done ✓
2. Multiple candidates remaining: compare values, check existence if not already done

- Last formal part, everything beyond will generalize rather easily
- We consider a problem of the form

$$(\mathcal{P}) \qquad \max_{x \in C(\mathcal{P})}\text{imize } f(x) \quad \text{where} \quad C(\mathcal{P}) = \{x \in \text{dom}(f) : g(x) = 0\}$$

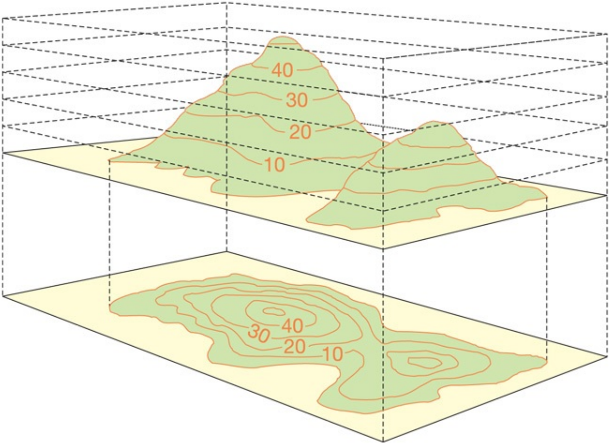
Definition (Level Set)

Let $X \subseteq \mathbb{R}^n$, $g : X \mapsto \mathbb{R}$, and $c \in \mathbb{R}$. Then, we call $L_c(g) = \{x \in X : g(x) = c\}$ the c-level set of g.

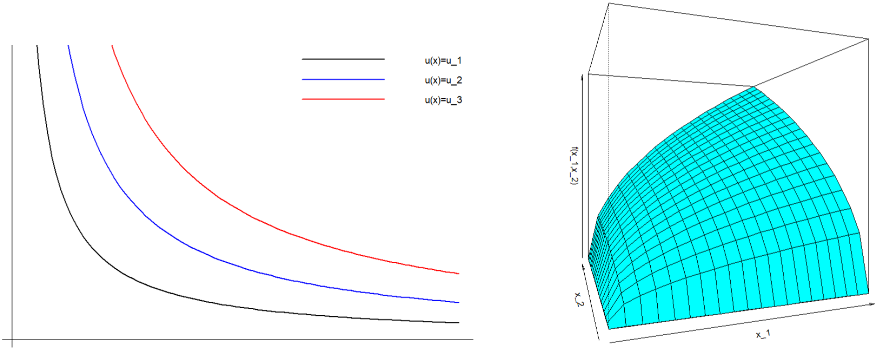
- Constraint set is zero-level set of g: $C(\mathcal{P}) = L_0(g)$
- Constrained maximization problem: find $\arg\max f|_{L_0(g)}$

3. Constrained Optimization: One Equality Constraint

Level Sets 1/2



Level sets of "coordinates $\mapsto$ height" (taken from http://canebrake13.com/fieldcraft/map_compass.php)



Indifference curves of utility function $u(x_1, x_2) = \sqrt{x_1 x_2}$

3. Constrained Optimization: One Equality Constraint

Intuition 1/3

- Can't we re-write the constrained problem to use our approach to unconstrained problems? e.g.

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

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

- 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

$\Rightarrow$ We want to use the more general concepts (implicit, local), but the idea is exactly this!

3. Constrained Optimization: One Equality Constraint

Intuition 2/3

- Recall: constrained local maximizer (on the level set):

$$x^* \in L_0(g) : (\exists \varepsilon > 0 : (\forall x \in B_\varepsilon(x^*) \cap L_0(g) : f(x^*) \geq f(x)))$$

- Deriving conditions like before: start from local maximizer x^* and consider neighborhoods $B_\varepsilon(x^*)$
- Only new issue: how to stay on the level set?
  - x^* lies on the level set: $g(x^*) = g(x_1^*, x_2^*, \dots, x_n^*) = 0$
  - Suppose we marginally vary $x_{-1}^* := (x_2^*, x_3^*, \dots, x_n^*)'$, but not x_1^*
  - If g is differentiable, g should still be "really close" to zero
  - If $\frac{\partial g}{\partial x_1}(x^*) \neq 0$, we may pull g back to zero using x_1:

$$g(h(x_2, x_3, \dots, x_n), x_2, x_3, \dots, x_n) = 0 \quad \text{for all} \quad x_{-1} \in B_\varepsilon(x_{-1}^*)$$

    - h: local "pull-back" function around x^*

- h: implicit function
  - Because $x^* = (h(x^*_{-1}), x^*_{-1})$ is a constrained local maximizer:
    - Points around x^* on $L_0(g)$ have the form $x = (h(x_{-1}), x_{-1})$, and thus

$$f(x^*) \geq f(h(x_{-1}), x_{-1}) \quad \text{for all} \quad x_{-1} \in B_\varepsilon(x^*_{-1})$$

  $\Rightarrow$ Thus, x^*_{-1} is a (unconstrained!) local maximizer of $f(h(x_{-1}), x_{-1})$

  $\Rightarrow$ Have reduced issue to unconstrained problem that we can handle!

  - Label "implicit": no explicit formula derived/known
  - The procedure of course works with any j where $\frac{\partial g}{\partial x_j}(x^*) \neq 0$
  - Complication: h unknown, does it have nice properties?

Theorem (Univariate Implicit Function Theorem)

Let $X_1 \subseteq \mathbb{R}$, $X_2 \subseteq \mathbb{R}^{n-1}$ and $X := X_1 \times X_2$, and $g : X \mapsto \mathbb{R}$. Suppose that $g \in C^1(X)$, and that for a $(y^*, z^*) \in X_1 \times X_2$, $g(y^*, z^*) = 0$. Then, if $\frac{\partial g}{\partial y}(y^*, z^*) \neq 0$, there exists an open set $U \subseteq \mathbb{R}^{n-1}$ such that $z^* \in U$ and $h : U \mapsto \mathbb{R}$ for which $y^* = h(z^*)$ and $\forall z \in U : g(h(z), z) = 0$. Moreover, it holds that $h \in C^1(U)$ with derivative

$$\nabla h(z) = -\left(\frac{\partial g}{\partial y}(h(z), z)\right)^{-1} \frac{\partial g}{\partial z}(h(z), z) \qquad \forall z \in U.$$

- Verbally:
  - Given our intuitive conditions (differentiability, "g moves with y")...
  - there exists a "pull-back" function h locally around x^*...
  - that is continuously differentiable with a **known gradient**
- Notes
  - $y \mathrel{\widehat{=}} x_1$: variable to be replaced, $z \mathrel{\widehat{=}} x_{-1}$: remaining variables
  - $\frac{\partial g}{\partial z}$ is derivative w.r.t. $n-1$ variables (gradient without $\frac{\partial g}{\partial y}$)

- Necessary first order condition in the constrained problem
  - if $x^* = (y^*, z^*) \in L_0(g)$ is a constrained local maximizer and $\nabla g(x^*) \neq \mathbf{0}$, then...
    - By the implicit function theorem, there exists $h(\cdot)$ around x^*, where...
    - $y^* = h(z^*)$, and z^* is a local maximizer of $f(h(z), z)$
    - Unconstrained theorem: $\frac{d}{dz} f(h(z), z) = \mathbf{0}$ for $z = z^*$; Chain rule gives

$$\exists \lambda \in \mathbb{R} : \nabla f(x^*) = \lambda \nabla g(x^*)$$

    - Note that this implies $\lambda = \left(\frac{\partial g}{\partial x_i}(x^*) \right)^{-1} \frac{\partial f}{\partial x_i}(x^*)$ for any $i \in \{1, \dots, n\}$!
- Additional border candidates (singularities): x^s so that $\nabla g(x^s) = 0$ (why?)

Theorem (Lagrange's Necessary First Order Condition)

Consider the constrained problem $\max_{x \in L_0(g)} f(x)$ *where* $X \subseteq \mathbb{R}^n$ *and* $f, g \in C^1(X)$. *Let* $x^* \in L_0(g)$ *and 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^*.

- Equivalently: $\exists \lambda \in \mathbb{R} : \nabla f(x^*) - \lambda \nabla g(x^*) = \mathbf{0}$
- Lagrangian function (or: "Lagrangian")

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

- The FOC for λ is $g(x) = 0$, i.e. $x \in L_0(g)$
- Thus, $x^* \in X$ satisfies the necessary FOC if and only if for a $\lambda \in \mathbb{R}$, (λ, x) is a critical value of the Lagrangian function!

First Order Conditions - Summary of Method

- Unconstrained problem: Taylor expansions give necessary FOC based on gradient of f
- Constrained problem: we transfer the unconstrained approach by...
  - locally "re-writing" the constraint and plugging it into the objective,
  - ...which gives an unconstrained problem in one less variable
  - We may not know the "plug-in function", but we know its derivative

  $\rightarrow$ enough to derive gradient-based FOC (Chain rule)!

- Constrained maximization = Lagrangian maximization?
  - Identical first order necessary condition
  - But: Lagrangian has only saddle points (see script)
  - Can still derive *sufficiency* criterion for constrained maxima/minima from second derivative = Hessian of the Lagrangian
    - Constrained optimization: no *necessary* SOC!
  - Issue: rather ugly conditions (see slides appendix)
  - ...but we might also use our intuition for the problem and the sign of multipliers λ

- Lagrangian Multiplier: value cost of the constraint
  - Example: shadow cost of the budget constraint
  - Here: $\lambda =$ change in objective due to marginally "relaxing" constraint
  - $\lambda < 0 \rightarrow$ constraint limits our ability to *lower* objective; not maximizer!

- Why (and when) does the multiplier trick work?
  - Requires equivalence to inequality-constrained problem ($g(x) \leq 0$!)
  - Directional derivative in direction $z \neq \mathbf{0}$:

$$\left[\frac{d}{dt}f(x^* + tz)\right]\bigg|_{t=0} = \nabla f(x^*)z = \lambda \nabla g(x^*)z = \lambda \left[\frac{d}{dt}g(x^* + tz)\right]\bigg|_{t=0}$$

    - If z points to the interior of the constraint set: $\nabla g(x^*)z < 0$
    - $x^* \rightarrow z$ marginally decreases (increases) f if $\lambda > 0$ ($\lambda < 0$)

  $\rightarrow$ $\lambda^* > 0$ ($\lambda^* < 0$) rules out local minimizers (maximizers)!
- Notes of caution: avoid sign errors!
  - Use a minus in the Lagrangian: $\mathcal{L}(x, \lambda) = f(x) - \lambda g(x)$
  - Consider an equivalent "smaller-or-equal" problem (if $g(x) \geq 0$, multiply both sides by -1 first)
  - if $g(x) = c - \tilde{g}(x) \leq 0$, use the gradient of g, not the one of $\tilde{g}$
- Application in exercise sessions

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

$$\max_{x \in \mathbb{R}^2} -\|x\|_2 \quad \text{subject to} \quad x_1 + x_2 = 1$$

or equivalently,

$$\max_{x \in \mathbb{R}^2} -(x_1^2 + x_2^2) \quad \text{subject to} \quad x_1 + x_2 - 1 = 0$$

- Multiple equality constraints $m \in \mathbb{N}$: stacked in

$$g = \begin{pmatrix} g^1 \\ \vdots \\ g^m \end{pmatrix}$$

so that the constraint becomes $g(x) = \mathbf{0}$

  - Looks alot like single constraint (recall: vector space)
  - *Only* adjustment: varying x_i generally moves all m directions of $g(x)$

  $\rightarrow$ need to adjust m arguments in the implicit function:

$$g(h(x_{-m}), x_{-m}) = 0) \quad \text{for} \quad x_{-m} = (x_{n-m+1}, x_{n-m+2}, \ldots, x_n)$$

  - Adjustment must reach all m directions $x_1, \ldots, x_m$: $\text{rk}(J_g(x^*)) = m$

4. Constrained Optimization: Multiple Equality Constraints

Theorem (Multivariate Implicit Function Theorem)

Let $X_1 \subseteq \mathbb{R}^m$, $X_2 \subseteq \mathbb{R}^{n-m}$ and $X := X_1 \times X_2$, and $g : X \mapsto \mathbb{R}^m$. Suppose that $g \in C^1(X, \mathbb{R}^m)$, and that for a $(y^*, z^*) \in X_1 \times X_2$, $g(y^*, z^*) = \mathbf{0}$. Then, if $rk\left(\frac{\partial g}{\partial y}(y^*, z^*)\right) = m$, there exists an open set $U \subseteq \mathbb{R}^{n-1}$ such that $z^* \in U$ and $h : U \mapsto \mathbb{R}^m$ for which $y^* = h(z^*)$ and $\forall z \in U : g(h(z), z) = \mathbf{0}$. Moreover, it holds that $h \in C^1(U, \mathbb{R}^m)$ with derivative

$$J_h(z) = -\left(\frac{\partial g}{\partial y}(h(z), z)\right)^{-1} \frac{\partial g}{\partial z}(h(z), z) \qquad \forall z \in U.$$

- Intuition identical to univariate theorem, just "vectorized" version
- Rank condition for the "partial Jacobian" $\frac{\partial g}{\partial y}(y^*, z^*)$ can be met if rk $J_g(x^*) = m$

4. Constrained Optimization: Multiple Equality Constraints

Theorem (Lagrange's Multiple First Order Necessary Condition)

Consider the constrained problem $\max_{x \in L_0(g)} f(x)$ *where* $X \subseteq \mathbb{R}^n$ *and* $f \in C^1(X), g \in C^1(X, \mathbb{R}^m)$. *Let* $x^* \in L_0(g)$ *and suppose that* $rk(J_g(x^*)) = m$. *Then,* x^* *is a local maximizer of the constrained problem only if there exists* $\Lambda = (\lambda_1, \ldots, \lambda_m)' \in \mathbb{R}^m : \nabla f(x^*) = \Lambda' J_g(x^*)$. *If such* $\Lambda \in \mathbb{R}$ *exists, we call* λ_i *the Lagrange multiplier associated with* x^* *for the i-th constraint.*

- FOC for x^*: $\nabla f(x^*) = \Lambda' J_g(x^*)$ with $\Lambda = (\lambda_1, \ldots, \lambda_m)$, multiplying out gives

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*)$$

$\rightarrow$ Straightforward to see the generalization of the univariate case

-1. Determine whether a solution exists at all (optional)
0. Collect border candidates: ~~boundary points,~~ non-diff'ability, singularities of the level set
1. Interior solutions: Finding and eliminating candidates
   - Necessary FOC: Initial set of candidates = Lagrangian critical values (No necessary SOC!)

optional: Sufficient condition for Lagrangian multipliers?

- Single candidate (including border) + existence guaranteed? Done ✓

optional: Sufficient SOC: rule out those identified as a strict local minimum

- If only one candidate (including border) remains
  - Existence guaranteed? Or: sufficient condition holds? Done ✓

2. Multiple candidates remaining: compare values, check existence if not already done; limits when constraint set is not bounded

- Recall: maximization problem $\mathcal{P}$ with general constraint set

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

- Examples
  - (Government) Budget constraints: $p \cdot x \leq y$
  - Non-negativity constraints: $\forall i \in \{1, \ldots, n\} : x_i \geq 0$
  - Production possibility frontier
  - etc.
- Our approach:
  1. Formal theorem to deal with inequality constraints
  2. Simpler: replace inequality with equality or eliminate the constraint

- Recall Lagrangian FOC: ("FOC + feasibility")

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) \quad \wedge \quad \forall i \in \{1, \ldots, n\} : g_i(x) = 0$$

- Binding constraint: $h_j(x^*) = 0$ "like an equality constraint" at x^*
- Slack constraint $h_j(x^*) < 0$ has no value cost: zero multiplier μ_j
  - "Complementary slackness" condition: $\mu_j h_j(x^*) = 0$
  - Set of binding inequality constraints varies across candidates, but...
  - $\mu_j h_j(x^*)$ does not! Lagrangian FOC is always

$$\nabla f(x^*) = \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{j=1}^{k} \mu_j h_j(x) \quad \wedge \quad \forall i \in \{1, \ldots, n\} : g_i(x) = 0$$

  because $\mu_j = 0$ for all non-binding (irrelevant) constraints
- Add feasibility: $h_j(x^*) \leq 0$ for non-binding constraints and we're done

5. Constrained Optimization: Inequality Constraints

Theorem (Karush-Kuhn-Tucker Theorem)

For $\Lambda = (\lambda_1, \ldots, \lambda_m)' \in \mathbb{R}^m$ *and* $\mu = (\mu_1, \ldots, \mu_k)' \in \mathbb{R}^k$, *consider the* **optimality conditions**

(i) *(Feasibility)* $\forall j \in \{1, \ldots, k\} : h_i(x) \leq 0$ *and* $\forall i \in \{1, \ldots, m\} : g_i(x) = 0$,

(ii) *(FOC for x)* $\nabla f(x) = \sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{j=1}^k \mu_j h_j(x)$,

(iii) *(Complementary Slackness)* $\forall j \in \{1, \ldots, k\} : \mu_j h_j(x) = 0$.

Then, if $x^* \in dom(f)$ *is a local maximum of the constrained problem for which the set* $\{\nabla h_j(x^*) : h_j(x^*) = 0\} \cup \{\nabla g_i(x^*) : i \in \{1, \ldots, m\}\}$ *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.*

- In words: the optimality conditions are (first-order) necessary
- Rank condition: set of binding constraints varies across locations

- Necessary FOC may be sufficient if solution exists and a unique candidate remains
- Sufficient conditions: convex optimization (using concave objectives and/or quasi-convex inequality constraints)
  - More details in script
- We may be able to simplify the issue to an equality-constrained one!
- Intuition: set of solutions unchanged by imposing equality or removing constraint

- Imposing equality: "not really an inequality constraint at all"
  - Formally: x^* local maximizer $\Rightarrow$ constraint binding
  - Contrapositive: constraint not binding $\Rightarrow$ x^* not local maximizer
  - Show either and you may impose equality
- Dropping the constraint: "irrelevant"
  - Formally: constraint binding $\Rightarrow$ x^* not local maximizer

$\Rightarrow$ Problem modification preserves the set of solutions!

- Example: constrained utility maximization with "regular" utility function ($\frac{\partial u}{\partial x_j}$ strictly monotonically increasing, $\lim_{x_j \to 0} \frac{\partial u}{\partial x_j}(x) = \infty$)

SLIDES APPENDIX

Definition (Leading Principal Minor)

Consider a symmetric matrix $A = (a_{ij})_{i,j\in\{1,\dots,n\}} \in \mathbb{R}^{n\times n}$*. Then, for* $k \leq n$*, the* k*-th leading principal minor of* A*, or the leading principal minor of* A *of order* k *is the matrix obtained from eliminating all rows and columns with index above* k *from* A*, i.e. the matrix* $M_k^A = (a_{ij})_{i,j\in\{1,\dots,k\}} \in \mathbb{R}^{n\times n}$*.*

Example:

$$A = \begin{pmatrix} 1 & 4 & 3 & 2 \\ 2 & 0 & 0 & 0 \\ 3 & 4 & -1 & -2 \\ 0 & 1 & e & \pi \end{pmatrix}.$$

$\rightarrow$ leading principal minors ("top-left squares") of A:

$$M_1^A = (1), \quad M_2^A = \begin{pmatrix} 1 & 4 \\ 2 & 0 \end{pmatrix}, \quad M_3^A = \begin{pmatrix} 1 & 4 & 3 \\ 2 & 0 & 0 \\ 3 & 4 & -1 \end{pmatrix}, \quad M_4^A = A.$$

Theorem (Lagrange's Sufficient Conditions)

Consider the constrained problem $\max_{x \in L_0(g)} f(x)$ *where* $X \subseteq \mathbb{R}^n$ *and* $f, g \in C^2(X)$*. Let* $x^* \in L_0(g)$ *and* $\lambda^* \in \mathbb{R}$ *such that* $\nabla f(x^*) = \lambda^* \nabla g(x^*)$ *and* $g(x^*) = 0$*. If* $m = 1$ *is the number of equality constraints, denote by* $M_{n-m+1}^{H_{\mathcal{L}}}(\lambda^*, x^*), \dots, M_n^{H_{\mathcal{L}}}(\lambda^*, x^*)$ *the last* $n - m$ *principal minors of* $H_{\mathcal{L}}(\lambda^*, x^*)$*. If*

- $\forall j \in \{n-m+1, \dots, n\} : sgn(\det(M_j^{H_{\mathcal{L}}})) = (-1)^m$*, then* x *is a local minimizer of the constrained problem.*
- $\forall j \in \{n-m+1, \dots, n\} : sgn(\det(M_j^{H_{\mathcal{L}}})) = (-1)^j$*, then* x *is a local maximizer of the constrained problem.*

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

- $H_{\mathcal{L}}$: Bordered Hessian, structure:

$$H_{\mathcal{L}}(\lambda, x) = \begin{pmatrix} \frac{\partial^2 \mathcal{L}}{\partial \lambda^2}(\lambda, x) & \frac{\partial^2 \mathcal{L}}{\partial \lambda \partial x}(\lambda, x) \\ \frac{\partial^2 \mathcal{L}}{\partial x \partial \lambda}(\lambda, x) & \frac{\partial^2 \mathcal{L}}{\partial x^2}(\lambda, x) \end{pmatrix} = \begin{pmatrix} 0 & -\nabla g(x) \\ -(\nabla g(x))' & H_f(x) - \lambda H_g(x) \end{pmatrix}$$

- Can use sign of submatrices to determine extremum type
- No necessary SOC: typically, not too many candidates anyways

Theorem (Lagrange's Multiple Sufficient Conditions)

Suppose additionally that $f \in C^2(X), g \in C^2(X, \mathbb{R}^m)$, *and that* (Λ^*, x^*) *is a critical point of the Lagrangian function, i.e.* $\nabla f(x^*) = (\Lambda^*)' J_g(x^*)$ *and* $g(x^*) = \mathbf{0}$. *Denote by* $M^{H_\mathcal{L}}_{n-m+1}(\lambda^*, x^*), \ldots, M^{H_\mathcal{L}}_n(\lambda^*, x^*)$ *the last* $n-m$ *principal minors of* $H_\mathcal{L}(\lambda^*, x^*)$. *If*

- $\forall j \in \{n-m+1, \ldots, n\} : sgn(\det(M^{H_\mathcal{L}}_j)) = (-1)^m$, *then* x *is a local minimizer of the constrained problem.*
- $\forall j \in \{n-m+1, \ldots, n\} : sgn(\det(M^{H_\mathcal{L}}_j)) = (-1)^j$, *then* x *is a local maximizer of the constrained problem.*

- Why $n - m$ last principal minors?
  - Generally: first derivative condition rk $J_g(x^*) = m$ ($m = 1$?)
  - $\rightarrow$ constraints *linearly independent* at x^*
  - m constraints restrict x in m dimensions, $n - m$ "free variables"
  - One condition for every free direction!