# Local and Global Optimization

## Understanding Optima in Complex Landscapes

Dr. Helga Ingimundardóttir ⟨helgaingim@hi.is⟩

**VÉL113F**: Design and Optimization

INDUSTRIAL ENGINEERING

# Introduction

UNIVERSITY OF ICELAND

Delve into the dynamics of optimization landscapes, differentiating between local optima and the overarching global optimum, and the techniques to approach each.

# Mathematical Programming

**Calculus Methods**: Using mathematical tools to solve problems involving change and motion.

- Determining the fastest moment a car was going during a trip.

**Calculus of Variations**: A field of mathematical analysis that deals with maximizing or minimizing functionals.

- Finding the shortest path a beam of light can take to reflect off a surface.

**Nonlinear Programming**: Optimizing functions that are not straight lines.

- Maximizing profit when the production cost changes as you produce more.

**Geometric Programming**: Optimization technique based on polynomial equations.

- Designing a soda can with the least material but a specific volume.

**Quadratic Programming**: Optimizing quadratic functions, which are polynomials of degree 2.

- Minimizing the cost of production given certain constraints.

# Mathematical Programming II

**Linear Programming**: A method to achieve the best outcome in a mathematical model whose requirements are represented by linear relationships.

- Finding the best combination of products to manufacture to maximize profit.

**Dynamic Programming**: Breaking down a problem into simpler parts and solving each part only once.

- Figuring out the most efficient way to store data to minimize retrieval time.

**Integer Programming**: Optimization where some of the variables are restricted to integer values.

- Deciding the number of buses a school should deploy, as you can't have a fraction of a bus.

**Stochastic Programming**: Making decisions in the face of uncertainty.

- Planning for the future stock of a store when future demand is uncertain.

# Mathematical Programming III

UNIVERSITY OF ICELAND

**Separable Programming**: A nonlinear program where the objective and constraint functions are separable.

- Maximizing crop yield by optimizing the amount of water and fertilizer used when the effects of each are independent.

**Multi-objective Programming**: Making decisions while considering multiple goals simultaneously.

- Designing a product that's both low-cost and high-quality.

**Network Methods, CPM and PERT**: Tools for project planning and control.

- Organizing tasks when building a house to ensure it's done efficiently and on time.

**Game Theory**: Studying mathematical models of strategic interactions among rational decision-makers.

- Two businesses deciding on the price of a product, considering what the other might charge.

# Stochastic and Statistical Methods

**Stochastic Processing Techniques**: Techniques to handle processes that involve uncertainty.

- Predicting stock prices based on past fluctuations.

**Statistical Decision Theory**: Making decisions using data analysis.

- Choosing to launch a product based on customer survey data.

**Markov Processes**: Processes where the next state depends only on the current state.

- Predicting tomorrow's weather based on today's.

**Queuing Theory**: Studying the behavior of waiting lines.

- Optimizing supermarket cashiers to reduce wait times.

**Renewal Theory**: Statistics of the time to events in processes.

- Predicting machine failure based on past breakdowns.

**Simulation Methods**: Imitating a real-world process using a model.

- Predicting drug spread using a computer model.

# Stochastic and Statistical Methods II

**Reliability Theory**: Predicting and enhancing system durability.

- Determining average lifespan of a car part.

**Regression Analysis**: Examining relationships between variables.

- Determining sales relation to advertising.

**Cluster Analysis & Pattern Recognition**: Grouping based on similarities.

- Grouping customers by buying habits.

**Design of Experiments**: Planning experiments to get valid data.

- Testing if new fertilizer improves growth.

**Discriminant Analysis/Factor Analysis**: Breaking down data into core influences.

- Understanding factors influencing grades.

# Modern Optimization Techniques

**Genetic Algorithms**: Optimization using natural selection principles.

- Finding best airplane wing design by "evolving" designs.

**Simulated Annealing**: Probabilistic optimization mimicking the annealing process.

- Optimizing delivery routes.

**Ant Colony Optimization**: Optimization using ant behavior.

- Optimizing city traffic flow.

**Particle Swarm Optimization**: Optimization based on flock behavior.

- Adjusting wind turbine design for efficiency.

**Neural Networks**: Algorithms designed to recognize patterns.

- Recognizing faces in photos.

**Fuzzy Optimization**: Optimization using fuzzy logic rather than binary.

- Adjusting car heat based on "warm" or "cold".

# Classification of Optimization Problems

**Equations Involved:** Based on the nature of equations.

- Linear vs. Nonlinear equations.

**Design Variables Values:** Based on permissible values.

- Discrete variables in selecting warehouse locations.

**Deterministic Nature:** Based on variables' determinacy.

- Predictable machine outputs vs. unpredictable stock prices.

**Existence of Constraints:** Whether constraints are present.

- Maximizing revenue with a limited budget.

# Classification of Optimization Problems II

**Design Variables Nature:** Nature of the variables involved.

- Binary choices in a network design.

**Physical Structure:** Based on the problem's inherent structure.

- Structural engineering optimizations.

**Separability of Functions:** If functions can be separated.

- Independent departmental budgets in a company.

**Number of Objectives:** Based on the number of goals.

- Balancing cost, quality, and time in project management.

# Objective Function: Maximize or Minimize

The choice between maximizing and minimizing a function can often be translated by considering the function's opposite or a scaled version of the function.

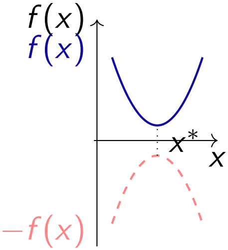



Figure: $\min_x f(x) \iff \max_x -f(x)$

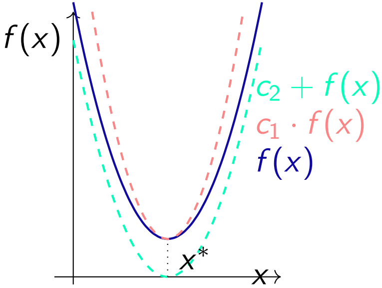
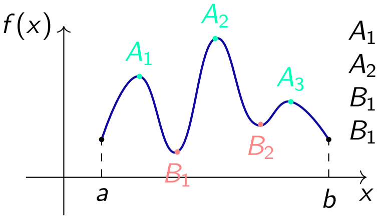



Figure: Scaled $c_1 \cdot f(x)$ and translated $c_2 + f(x)$ both retain the original shape.

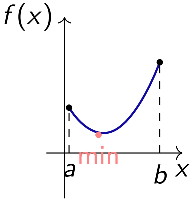
## Single variable optimization with no constraints

Determine the value of $x = x^*$ within the interval $[a, b]$ that minimizes the function $f(x)$.

Necessary Condition If a function $f(x)$ is defined in the interval $a \leq x \leq b$ and has a relative minimum at $x = x^*$, where $a < x^* < b$, and if the derivative $\frac{df(x)}{dx} = f'(x)$ exists as a finite number at $x = x^*$, then $f'(x^*) = 0$.




$A_1, A_2, A_3$ are local maxima
$A_2$ is the global maximum
$B_1, B_2$ are local minima
$B_1$ is the global minimum




local minimum is also global minimum

## Considerations

- The proof holds even if $x^*$ is a local maximum.
- The derivative's existence at $x^*$ is not guaranteed for every minimum or maximum.
- Extrema at the endpoints of the function's definition interval are not covered.
- A zero derivative doesn't guarantee the presence of an extremum; it could be an inflection (or saddle) point.

UNIVERSITY OF ICELAND

## Sufficient Condition for Extrema

Let's assume the first $n-1$ derivatives at point $x^*$ are zero, but the $n$th derivative is not zero.

- If the $n$th derivative at $x^*$ is positive and $n$ is even, then $f(x^*)$ is a minimum.
- If the $n$th derivative at $x^*$ is negative and $n$ is even, then $f(x^*)$ is a maximum.
- If $n$ is odd, $f(x^*)$ is neither a maximum nor a minimum.

Think of the even $n$ as giving the function "another chance" to decide if it's curving up or down. Odd $n$ means the function hasn't settled into a curve direction.

## Multivariable Optimization: Necessity vs. Sufficiency

In single-variable optimization, a zero derivative suggests potential extrema. Similarly, in the multivariable case, all first partial derivatives should be zero at a stationary point, constituting the **necessary condition**. To further discern if it's genuinely a maximum or minimum (and not a saddle point), we turn to the Hessian matrix for the **sufficient condition**.

Simply put, for a point to be a high or low point in multiple dimensions, the function shouldn't be rising or falling in any of those directions.

# Multivariable Optimization II

Necessary Condition For a function $f(\mathbf{x})$ to have an extreme point at $\mathbf{x} = \mathbf{x}^*$:

$$\frac{\partial f}{\partial x_1}(\mathbf{x}^*) = \frac{\partial f}{\partial x_2}(\mathbf{x}^*) = \cdots = \frac{\partial f}{\partial x_n}(\mathbf{x}^*) = 0$$

Sufficient Condition To determine the nature of a stationary point $\mathbf{x}^*$ of the function $f(\mathbf{x})$:

- If the Hessian matrix at $\mathbf{x}^*$ is positive definite, $\mathbf{x}^*$ is a local minimum.
- If it's negative definite, $\mathbf{x}^*$ is a local maximum.

This helps us be certain about the nature of the stationary point.

## Refresher: The Hessian Matrix

The Hessian matrix $H$ of a function $f(\mathbf{x})$ is the matrix of its second-order partial derivatives. It provides insight into the curvature of the function:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

# Multivariable Optimization IV

A matrix $A$ is considered:

- **Positive definite** if all its eigenvalues are positive. This suggests the function is curving upwards, indicating a minimum. That means all values of $\lambda$ that satisfy the determinantal equation $|A - \lambda I| = 0$ should be positive.
- **Negative definite** if all its eigenvalues are negative. This suggests the function is curving downwards, indicating a maximum.

## Determining Definiteness Using Determinants

To assess the definiteness of a matrix $A$ of order $n$, evaluate the determinants of all the leading principal minors.

Example: For a $n \times n$ matrix $A$:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

# Multivariable Optimization VI

The determinants of the leading principal minors are:

$$A_1 = |a_{11}|, \quad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \cdots \quad A_n = \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix}$$

The matrix $A$ will be:

- **Positive definite** if all $A_1, A_2, \ldots, A_n$ are positive.
- **Negative definite** if the sign of $A_j$ alternates starting with negative, i.e., the sign of $A_j$ is $(-1)^{j+1}$ for $j = 1, 2, \ldots, n$.

Laplace expansion is a technique to compute matrix determinants using cofactors. For more, see Wikipedia: Laplace Expansion.

UNIVERSITY OF ICELAND

### Multivariable Optimization with Equality Constraints

**Goal**: Minimize a function $f(\vec{x})$ where $\vec{x} = [x_1, x_2, \ldots, x_n]$.
**Constraints**: Equations given by $g_j(\vec{x}) = 0$, dictating the rules our solution must follow.

Note If the number of constraints $(m)$ is greater than the elements in our solution list $(n)$, the problem is **overdefined**. Typically $m > n$ means no solution exists.

UNIVERSITY OF ICELAND

## Example

Find the dimensions of a box of largest volume that can be inscribed in a sphere of unit radius.

### Setup

- Let the origin of the Cartesian coordinate system $x_1, x_2, x_3$ be at the center of the sphere.
- Let the sides of the box be $2x_1, 2x_2$, and $2x_3$.
- The volume of the box is:

$$f(x_1, x_2, x_3) = 8x_1x_2x_3$$

# Optimization with Equality Constraints

- Since the corners of the box lie on the sphere:

$$x_1^2 + x_2^2 + x_3^2 = 1$$

## Constraints and Reformulation

- This problem has three design variables and one equality constraint.
- Use the equality constraint to eliminate one variable:

$$x_3 = \sqrt{1 - x_1^2 - x_2^2}$$

- The objective becomes:

$$f(x_1, x_2, x_3) = 8x_1x_2\sqrt{1 - x_1^2 - x_2^2}$$

# Optimization with Equality Constraints IV

Necessary Conditions for the maximum of $f$ provide a system of equations.

$$\frac{\partial f}{\partial x_1} = 8x_2 \left[ \sqrt{1 - x_1^2 - x_2^2} - \frac{x_1^2}{\sqrt{1 - x_1^2 - x_2^2}} \right] = 0$$

$$\frac{\partial f}{\partial x_2} = 8x_1 \left[ \sqrt{1 - x_1^2 - x_2^2} - \frac{x_2^2}{\sqrt{1 - x_1^2 - x_2^2}} \right] = 0$$

After simplification, the equations become:

$$1 - 2x_1^2 - x_2^2 = 0,$$
$$1 - x_1^2 - 2x_2^2 = 0.$$

Solving this system of equations yields the solution $x_1^* = x_2^* = \dfrac{1}{\sqrt{3}}$ and hence $x_3^* = \dfrac{1}{\sqrt{3}}$. Therefore, the volume of the box is $f(\vec{x}^*) = \dfrac{8}{3\sqrt{3}}$.
To verify that this is a maximum, we need to check the sufficient conditions to $f(x_1, x_2)$. The second-order partial derivatives of $f$ at $\vec{x}^*$ are given by

$$\frac{\partial^2 f}{\partial x_1^2}(\vec{x}^*) = -\frac{32}{\sqrt{3}}, \quad \frac{\partial^2 f}{\partial x_2^2}(\vec{x}^*) = -\frac{32}{\sqrt{3}}, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(\vec{x}^*) = -\frac{16}{\sqrt{3}},$$

and since $\dfrac{\partial^2 f}{\partial x_1^2} < 0$ and $\dfrac{\partial^2 f}{\partial x_1^2}\dfrac{\partial^2 f}{\partial x_2^2} - \left(\dfrac{\partial^2 f}{\partial x_1 \partial x_2}\right)^2 > 0$ then the Hessian matrix of $f$ at $\vec{x}^*$ is negative definite and hence $\vec{x}^*$ is a maximum point of $f$.