Optimization Techniques

Optimization joke

- At the end of his course on mathematical methods in optimization, the professor sternly looks at his students and says:
- "There is one final piece of advice I'm going to give you now: ...
- ...Whatever you have learned in my course – never ever try to apply it to your personal lives!"
- "Why?" the students ask.
- "Well, some years ago, I observed my wife preparing me breakfast, and I noticed that she wasted a lot of time walking back and forth in the kitchen. So, I went to work, optimized the whole procedure, and told my wife:
- "Dear, for some years, day by day, you are preparing me the breakfest each morning, but in a highly inefficient way.
- So, I kindly prepared for you (to show you how thankfull I am) the optimized breakfest-making procedure!"

Optimization joke (contd)

- "And what happened?!"

Optimization joke (contd)

- "And what happened?!"
- "Before I applied my expert knowledge in optimization, my wife needed about half an hour to prepare breakfast for the two of us . . .

Optimization joke (contd)

- "And what happened?!"
- "Before I applied my expert knowledge in optimization, my wife needed about half an hour to prepare breakfast for the two of us . . .
- And now, it takes **me** less than fifteen minutes . . . "

Introduction to Optimization

Definition

Optimization is any scenario in which you are trying to make certain decisions and reach the best possible outcome.

- Optimization is concerned with the study of **maximization and minimization of mathematical functions**. Oftenly, the arguments of these functions are subject to additional conditions or *constraints*.

Introduction to Optimization

Definition

Optimization is any scenario in which you are trying to make certain decisions and reach the best possible outcome.

- Optimization is concerned with the study of **maximization and minimization of mathematical functions**. Oftenly, the arguments of these functions are subject to additional conditions or *constraints*.
- Due to its great utility in such diverse areas as applied science, engineering, economics, finance, medicine, and statistics, optimization holds an important place in the practical world and the scientific world.

Introduction to Optimization

In the 18^{th} century, the swiss mathematician Leonhard Euler (1707-1783) proclaimed that . . .

Introduction to Optimization

In the 18^{th} century, the swiss mathematician Leonhard Euler (1707-1783) proclaimed that ...
nothing at all takes place in the Universe in which some rule of maximum or minimum does not appear.

Applications of Optimization

- **Economics:** Consumer theory / supplier theory
- **Finance:** Optimal hedging / pricing / risk management
- **Science / Engineering:** Aerospace, product design, data mining, traffic, communications and supply networks
- **Business decisions:** Scheduling, production, organizational decisions, logistics.
- **Government:** Military applications, fund allocation, etc
- **Personal decisions:** Sports, on-field decisions, player acquisition, marketing
- **WEB:** web minning, search engine optimization, web networks

Applications of Optimization

- The traditional portfolio optimization problem attempts to simultaneously minimize the risk and maximize the return
- The production manager will maximize profit and minimize cost
- A good sunroof design in a car could aim to minimize the noise the driver hears and maximize the ventilation
- In bridge construction, a good design is characterized by low total mass and high stiffness
- Aircraft design requires simultaneously optimization of fuel efficiency, payload and weight
- A good pipe network will allow the best flow at a minimal cost
- The logistic manager will maximize the transport load, schedule and routing and minimize the distance
- A good electronic circuit device will have a minimal power consumption with minimal size while satisfying manufacturing limits

Managing a Production Facility

Consider a production facility for a manufacturing company. The facility is capable of producing a variety of n products: $1, 2, \ldots, n$. These products are manufactured out of certain raw materials. Let us assume that there are m different raw materials: $1, 2, \ldots, m$.

- The facility has, for each raw material $i = 1, 2, \ldots, m$, an amount, b_i, on hand.

Managing a Production Facility

Consider a production facility for a manufacturing company. The facility is capable of producing a variety of n products: $1, 2, \ldots, n$. These products are manufactured out of certain raw materials. Let us assume that there are m different raw materials: $1, 2, \ldots, m$.

- The facility has, for each raw material $i = 1, 2, \ldots, m$, an amount, b_i, on hand.
- Each raw material i has a known unit market value ρ_i.

Managing a Production Facility

Consider a production facility for a manufacturing company. The facility is capable of producing a variety of n products: $1, 2, \ldots, n$. These products are manufactured out of certain raw materials. Let us assume that there are m different raw materials: $1, 2, \ldots, m$.

- The facility has, for each raw material $i = 1, 2, \ldots, m$, an amount, b_i, on hand.
- Each raw material i has a known unit market value ρ_i.
- Each product is made from known amounts of the various raw materials. That is, producing one unit of product j requires a certain known amount, say a_{ij} units, of raw material i. For example, producing product 7 will require

Managing a Production Facility

Consider a production facility for a manufacturing company. The facility is capable of producing a variety of n products: $1, 2, \ldots, n$. These products are manufactured out of certain raw materials. Let us assume that there are m different raw materials: $1, 2, \ldots, m$.

- The facility has, for each raw material $i = 1, 2, \ldots, m$, an amount, b_i, on hand.
- Each raw material i has a known unit market value ρ_i.
- Each product is made from known amounts of the various raw materials. That is, producing one unit of product j requires a certain known amount, say a_{ij} units, of raw material i. For example, producing product 7 will require
  - amount a_{17} of raw material 1
  - amount a_{27} of raw material 2
  - amount a_{37} of raw material 3
  - and so on, a_{m7} of raw material m.

Managing a Production Facility

Consider a production facility for a manufacturing company. The facility is capable of producing a variety of n products: $1, 2, \ldots, n$. These products are manufactured out of certain raw materials. Let us assume that there are m different raw materials: $1, 2, \ldots, m$.

- The facility has, for each raw material $i = 1, 2, \ldots, m$, an amount, b_i, on hand.
- Each raw material i has a known unit market value ρ_i.
- Each product is made from known amounts of the various raw materials. That is, producing one unit of product j requires a certain known amount, say a_{ij} units, of raw material i. For example, producing product 7 will require
  - amount a_{17} of raw material 1
  - amount a_{27} of raw material 2
  - amount a_{37} of raw material 3
  - and so on, a_{m7} of raw material m.
- Also, the j final product can be sold at the price of σ_j dollars per unit.

Managing a Production Facility

- We shall consider one optimization problems related to the efficient operation of this facility.
- Let us assume that production manager decides to produce x_j units of the j_{th} product, $j = 1, 2, \ldots, n$.
- The revenue associated with the production of one unit of product j is σ_j.
- But there is also a cost of raw materials that must be considered.
- The cost of producing one unit of product j is

$$\sum_{i=1}^{m} \rho_i a_{ij}$$

Managing a Production Facility

- Therefore, the **net revenue** associated with the production of one unit is the difference between the revenue and the cost.
- Let the net revenue obtained from one unit of product j be

$$c_j = \sigma_j - \sum_{i=1}^{n} \rho_i a_{ij}, \quad j = 1, 2, \ldots, n$$

- Now, the net revenue corresponding to the production of x_j units of product j is

$$\sum_{j=1}^{n} c_j x_j$$

Managing a Production Facility

- The manager's goal is to maximize this quantity.
- However, there are constraints on the production levels.
- For example, each production quantity x_j must be nonnegative, and so we have the constraint

$$x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

Managing a Production Facility

- The manager's goal is to maximize this quantity.
- However, there are constraints on the production levels.
- For example, each production quantity x_j must be nonnegative, and so we have the constraint

$$x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

- Secondly, we can't produce more products than we have raw material for.
- The amount of raw material i consumed by a given production schedule is

$$\sum_{j=1}^{n} a_{ij} x_j$$

Managing a Production Facility

- The manager's goal is to maximize this quantity.
- However, there are constraints on the production levels.
- For example, each production quantity x_j must be nonnegative, and so we have the constraint

$$x_j \geq 0, \quad j = 1, 2, \ldots, n.$$

- Secondly, we can't produce more products than we have raw material for.
- The amount of raw material i consumed by a given production schedule is

$$\sum_{j=1}^{n} a_{ij} x_j$$

- and so we must also consider the following constraints:

Managing a Production Facility

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, 2, \ldots, m$$

To summarize, the production manager's job is to determine production values x_j, $j = 1, 2, \ldots, n$, so as to maximize

$$\sum_{j=1}^{n} c_j x_j$$

subject to the constraints:

$$\begin{aligned} x_j &\geq 0, \quad j = 1, 2, \ldots, n \\ \sum_{j=1}^{n} a_{ij} x_j &\leq b_i, \quad i = 1, 2, \ldots, m \end{aligned}$$

This optimization problem is an example of a linear programming (LP) problem. This particular example is often called the resource allocation problem.

Managing a Production Facility reduced to LP problem

$$\text{Maximize } c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

$$\begin{aligned}
\text{subject to } a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &\leq b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &\leq b_2 \\
&\ldots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &\leq b_m \\
x_1, x_2, \ldots, x_n &\geq 0
\end{aligned}$$

For, example,

$$\text{Maximize } 5x_1 + 3x_2 - 7x_3$$

$$\begin{aligned}
\text{subject to } -x_1 + 4x_2 &\leq 5 \\
7x_1 - 11x_2 + x_3 &\leq -3 \\
x_1, x_2, x_3 &\geq 0
\end{aligned}$$

Power Distribution for a VLSI Circuit

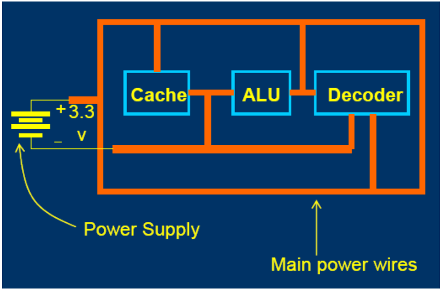
One application problem which generates large scale optimization problem is the problem of distributing power to the various parts of a Very Large Scale Integrated (VLSI) circuit processor. Usually a processor has nearly a million transistors, and millions of wires which transport signals and power. All one can really see by eye are the larger wires that carry power and patterns of wires that carry signals, boolean operations such as “and” and “or”.

A typical processor can be divided into a number of functional blocks:

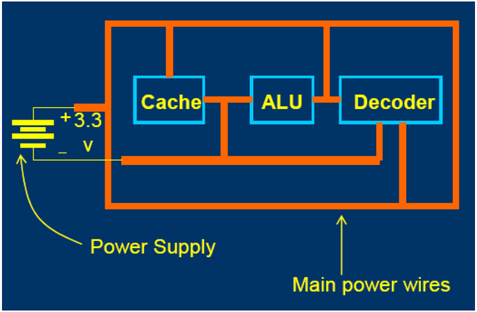
- **Caches**, which store copies of data and instructions from main memory for faster access.
- **Execution units** which perform boolean and numerical operations on data, such as “and”, “or”, addition and multiplication. These execution units are often grouped together and referred to as an Arithmetic Logic Unit (ALU).
- **Instruction decoder**, which translates instructions fetched from the cache into actions performed by the ALU.

Power Distribution for a VLSI Circuit

A simplified diagram of a typical processor with 3.3 volts power supply, the 3 main functional blocks, and the wires carrying power from the supply to the blocks is presented below. The wires are typically a micron thick, 10 microns wide and 1000 microns long (1 micron$= 10^{-4}$cm). The resistance of these wires is significant, and therefore even though the supply is 3.3 volts, these may not be 3.3 volts across each of the functional blocks. The main problem is whether or not each functional block has sufficient voltage to operate properly.



Design Objectives for the VLSI problem

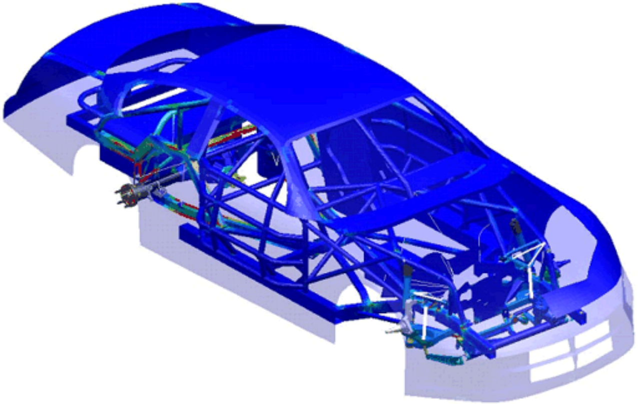


- Select topology and metal widths and lengths so that:
  - voltage across every function block > 3 volts;
  - and minimize the area used for the metal wires

Who uses VLSI Tools?

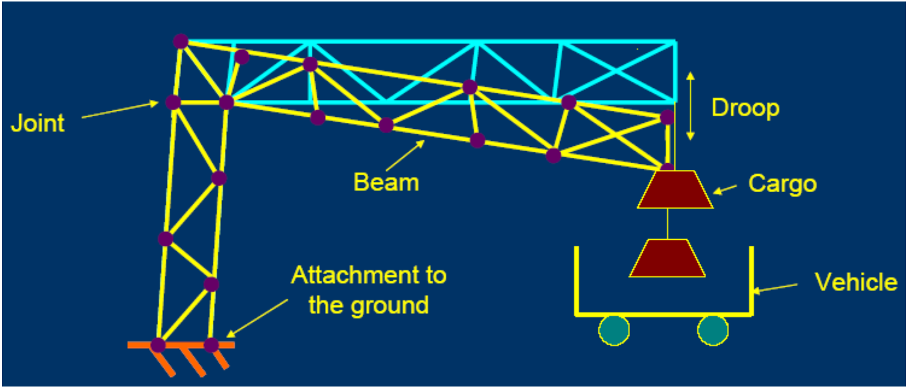
- IBM, Motorola, TI, Intel, Compaq, Sony, Hitachi and thousands of small and medium size companies.
- Once a VLSI circuit is designed, it is fabricated using a sequence of sophisticated deposition and etching processes which convert a wafer of Silicon into millions of transistors and wires.
- This processing can take more than a month.
- If the circuit does not function, the design flaw must be found and the fabrication process restarted from the beginning.
- Design errors can delay a product for months. In a competitive market, this delay can cost millions in lost revenue in addition to the cost of redesigning the circuit.
- In order to avoid fabricating designs with flaws, companies make extensive use of simulation tools to verify design functionality and performance.

Sructural design optimization



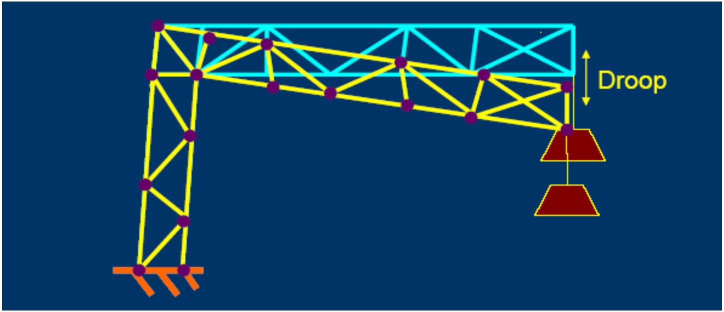
Structural design optimization

Consider designing a load bearing space frame:



Does the space frame droop too much under load?

Design objectives for a space frame

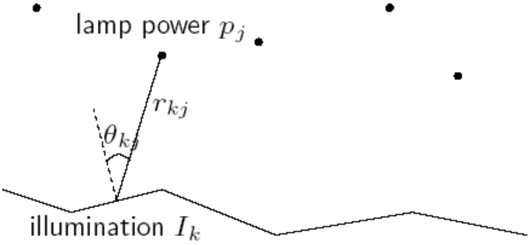


Select topology and strut widths and lengths such that

- Droop is small enough;
- Minimize the metal used.

Illumination problem

There are given m lamps illuminating n small flat pathches:



Intensity I_k at patch k depends linearly on lamp powers p_j:

$$I_k = \sum_{j=1}^{m} a_{kj} p_j$$

where

$$a_{kj} = \frac{\max\{\theta_{kj}, 0\}}{r_{kj}^2}$$

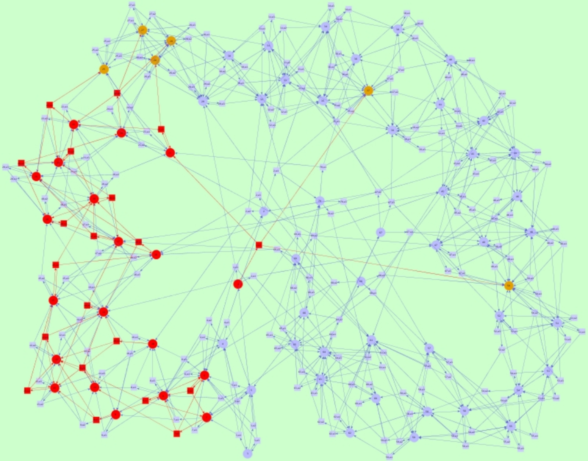
Illumination problem (contd)

Problem: Achieve the desired illumination intensity I_{des} with limited (bounded) lamp powers:

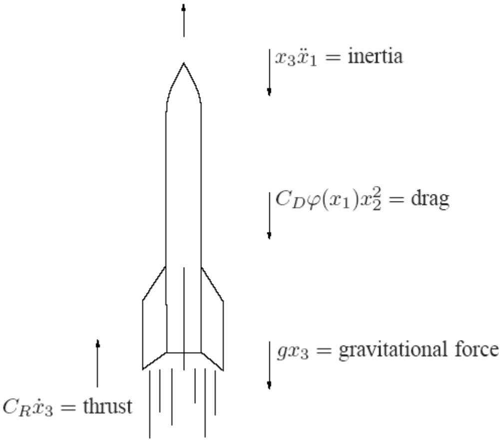
$$\text{minimize} \max_{k=1,2,\ldots,n} \left|\ln I_k - \ln I_{des}\right|$$

$$\text{subject to } 0 \leq p_j \leq p_{\max},\ j = 1, 2, \ldots m$$

Traffic and communication networks



Rocket science



Rocket science

Suppose that a rocket is to be launched vertically. The trajectory of a vertical launched rocket is controlled by adjusting the rate of fuel injection which generates the thrust force. Thus, the decision variable is the rate of fuel injection, which is function of time: $u = u(t)$.
Note that this is a dynamic problem (variables and some parameters depends on time! For example, weight of the rocket changes in time since it consummes the fuel.)

Introduce the variables:

Rocket science

Suppose that a rocket is to be launched vertically. The trajectory of a vertical launched rocket is controlled by adjusting the rate of fuel injection which generates the thrust force. Thus, the decision variable is the rate of fuel injection, which is function of time: $u = u(t)$.
Note that this is a dynamic problem (variables and some parameters depends on time! For example, weight of the rocket changes in time since it consummes the fuel.)

Introduce the variables:
$x_1(t)$ is the height of the rocket from the ground at time t;
$x_2(t)$ is the (vertical) speed at time t;
$x_3(t)$ is the rocket's weight ($=$ weight of remaining fuel) at time t.

Rocket science

Suppose that a rocket is to be launched vertically. The trajectory of a vertical launched rocket is controlled by adjusting the rate of fuel injection which generates the thrust force. Thus, the decision variable is the rate of fuel injection, which is function of time: $u = u(t)$.
Note that this is a dynamic problem (variables and some parameters depends on time! For example, weight of the rocket changes in time since it consummes the fuel.)

Introduce the variables:
$x_1(t)$ is the height of the rocket from the ground at time t;
$x_2(t)$ is the (vertical) speed at time t;
$x_3(t)$ is the rocket's weight ($=$ weight of remaining fuel) at time t.
We can consider that

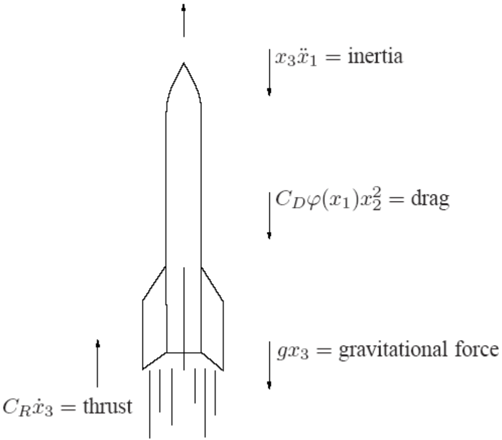
$$\frac{d}{dt}(x_3) = u(t)$$

Rocket science

The mathematical model and corresponding equations can be derived from the force equations under the assumption that there are four forces acting on the rocket, namely:

- **Inertia** $= x_3 \frac{d^2}{dt^2}(x_1) = x_3 \frac{d}{dt}(x_2)$;
- **Drag force** $= C_D \rho(x_1) x_2^2$, where C_D is the drag coefficient (a constant), $\rho(x_1)$ is a friction coefficient depending on atmospheric density which is a function of altitude, i.e. x_1;
- **Gravitational force** $= g x_3$ with g gravitational acceleration (assumed constant);
- **Thrust force** $= C_T \frac{d}{dt}(x_3)$, assumed proportional to the rate of fuel ejection.

Rocket science



Rocket science

The equations are:

$$\begin{aligned}\frac{d}{dt}(x_1(t)) &= x_2(t)\\ \frac{d}{dt}(x_2(t)) &= -\frac{C_D}{x_3(t)}\rho(x_1(t))x_2^2(t) - g + \frac{C_T}{x_3(t)}u(t)\\ \frac{d}{dt}(x_3(t)) &= -u(t)\end{aligned}$$

At time 0 we assume that $(x_1(0), x_2(0), x_3(0)) = (0, 0, M)$; that is, the rocket is on the ground, at rest, with initial fuel of weight M. At a prescribed final time t_f, it is desired that the rocket be at a position as high above the ground as possible.

Rocket science

Maximize $x_1(t_f)$ subject to

$$\begin{aligned}
\frac{d}{dt}(x_1(t)) &= x_2(t) \\
\frac{d}{dt}(x_2(t)) &= -\frac{C_D}{x_3(t)}\rho(x_1(t))x_2^2(t) - g + \frac{C_T}{x_3(t)}u(t) \\
\frac{d}{dt}(x_3(t)) &= -u(t) \\
(x_1(0), x_2(0), x_3(0)) &= (0, 0, M) \\
u(t) &\geq 0 \\
x_3(t) &\geq 0 \\
0 &\leq t \leq t_f
\end{aligned}$$

The decision variable (the unknown) is the function $u(t)$ defined on time interval $[0, t_f]$.

General optimization problems

A general optimization problem can be expressed mathematically as

$$\text{Maximize } f(x)$$

$$\text{subject to } x \in S$$

Function f is called the objective function, variable (or variables) x is called a decision variable and S is called the set of constraints.

General optimization problems

In this class we consider a particular class of optimization problems (very large though), i.e.

$$\text{Maximize } f(x)$$

$$\text{subject to } x \in S \subset \mathbb{R}^n$$

where the set of constraints S is specified in the form of

- equalities $h_i(x) = 0$, $i = 1, 2, \ldots, n$ or
- inequalities $g_i(x) \leq 0$, $j = 1, 2, \ldots, m$

Size and complexity

- number of decision variables
- number of constraints
- bit number to store the problem input data
- problem difficulty or complexity
- algorithm complexity

Solving optimization problems

General optimization problem are:

- usually no analytical solution
- very difficult to solve
- methods involve some compromise, e.g., very long computation time, or not always finding the solution

Exceptions: certain problem classes can be solved efficiently and reliably

- least-squares problems
- linear programming problems
- convex optimization problems

Model classification

Optimization problems are usually divided into various classifications

- **finite** dimensional vs **infinite** dimensional
  - depends on the dimension of the space that contains the constraints set S (and henceforth the decision variable).
- **unconstrained** vs **constrained**
  - unconstrained means that $S = \mathbb{R}^n$ (the entire space) or a box
  - constrained means that $S \subsetneq \mathbb{R}^n$
- **linear** vs **non-linear**
  - linear problem means that both objective function and constraints are linear functions
  - linearly constrained problems, constraints are linear fucntions
  - nonlinear problem

Model classification (contd)

- **differentiable** vs **non-differentiable**
  - functions involved are differentiable (more methods are available and the theory is well developed)
  - non-diferentiable: $\min |x|$
- **convex** vs **non-convex**
  - objective function and constraints are convex functions
  - non-convex optimization problems are much more difficult to solve
- **quadratic optimization**
  - objective function is quadratic (not linear, but convex)
  - used in finance and economics
  - Nobel prize in economics
- **integer optimization**
  - decision variables should be integers
  - mixed integer programming

Unconstrained optimization problem

$$\text{Maximize } f(x_1, x_2, \ldots, x_n)$$

$$\text{subject to } (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$$

$$\text{or } a_j \leq x_j \leq b_j, \; j = 1, 2, \ldots, m$$

Linear optimization problem

$$\text{Maximize } c_1x_1 + c_2x_2 + \ldots + c_nx_n$$

$$\text{subject to } a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \begin{Bmatrix} \leq \\ = \\ \geq \end{Bmatrix} b_i$$

$$i = 1, 2, \ldots, m$$

Definitions

Let $S \subset \mathbb{R}^n$ be a set and $f : S \to \mathbb{R}$ be a function in n variables.

Definitions

Let $S \subset \mathbb{R}^n$ be a set and $f : S \longrightarrow \mathbb{R}$ be a function in n variables.

Definition

The value x^* of the variable x solves the problem

$$\max_{x \in S} f(x)$$

if

$$f(x) \leq f(x^*) \text{ for all } x \in S$$

Definitions

Let $S \subset \mathbb{R}^n$ be a set and $f : S \rightarrow \mathbb{R}$ be a function in n variables.

Definition

The value x^* of the variable x solves the problem

$$\max_{x \in S} f(x)$$

if

$$f(x) \leq f(x^*) \text{ for all } x \in S$$

In this case we say that x^* is a maximizer of the function f subject to the constraint $x \in S$, and that $f(x^*)$ is the maximum (or maximum value) of the function f subject to the constraint $x \in S$.

Definitions

Let $S \subset \mathbb{R}^n$ be a set and $f : S \to \mathbb{R}$ be a function in n variables.

Definition

The value x^* of the variable x solves the problem

$$\max_{x \in S} f(x)$$

if

$$f(x) \leq f(x^*) \text{ for all } x \in S$$

In this case we say that x^* is a maximizer of the function f subject to the constraint $x \in S$, and that $f(x^*)$ is the maximum (or maximum value) of the function f subject to the constraint $x \in S$.
A minimizer is defined in an analogous manner.

Definitions

Definition

The variable x^* is a local maximizer of the function f subject to the constraint $x \in S$ if there is a number $\varepsilon > 0$ such that

$$f(x) \leq f(x^*)$$

for all $x \in S$ for which the distance between x and x^* is at most ε.

Definitions

Definition

The variable x^* is a local maximizer of the function f subject to the constraint $x \in S$ if there is a number $\varepsilon > 0$ such that

$$f(x) \leq f(x^*)$$

for all $x \in S$ for which the distance between x and x^* is at most ε.

A local minimizer is defined analogously. Sometimes we refer to a maximizer as a global maximizer to emphasize that it is not only a local maximizer. Every global maximizer is, in particular, a local maximizer (ε can take any value), and every minimizer is a local minimizer.

Existence of optimum

Let

$$f : S \subset \mathbb{R}^n \to \mathbb{R}$$

and consider the problem

$$\max_{x \in S} f(x)$$

Theorem (Weirstrass)

A continuous function on a compact (bounded and closed in $\mathbb{R}^n$) set attains both a maximum and a minimum on the set.

Existence of optimum

Let

$$f : S \subset \mathbb{R}^n \to \mathbb{R}$$

and consider the problem

$$\max_{x \in S} f(x)$$

Theorem (Weirstrass)

A continuous function on a compact (bounded and closed in $\mathbb{R}^n$) set attains both a maximum and a minimum on the set.

Note that this condition is sufficient without being necessary. In other words, a function can have a minimum or maximum without being continuous or defined on a compact set.

Necessary conditions for an interior optimum

Theorem (First-order conditions)

Let f be a differentiable function of n variables defined on the set S. If the point $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ in the interior of S is a local or global maximizer or minimizer of f then

$$\begin{aligned} \bigtriangledown f(x^*) &= 0 \\ \frac{\partial f}{\partial x_i}(x_1^*, x_2^*, \ldots, x_n^*) &= 0, \text{ for } i = 1, 2, \ldots, n \\ f_{x_i}(x^*) &= 0, \text{ for } i = 1, 2, \ldots, n \end{aligned}$$

Necessary conditions for an interior optimum

Theorem (First-order conditions)

Let f be a differentiable function of n variables defined on the set S. If the point $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ in the interior of S is a local or global maximizer or minimizer of f then

$$\begin{aligned} \bigtriangledown f(x^*) &= 0 \\ \frac{\partial f}{\partial x_i}(x_1^*, x_2^*, \ldots, x_n^*) &= 0, \ \textit{for } i = 1, 2, \ldots, n \\ f_{x_i}(x^*) &= 0, \textit{for } i = 1, 2, \ldots, n \end{aligned}$$

This result gives a **necessary** condition for a maximum (or minimum): if a point is a maximizer then it satisfies the condition. As in one dimensional case, the condition is called a *first-order condition*. Any point at which all the partial derivatives of f are zero is called a critical or stationary point of f.

Necessary conditions for an interior optimum

Note also, that the function f is supposed differentiable. In the case of ommiting the differentiability condition to the definition of critical points we should add also the points where gradient is not defined. Non-differntiability goes beyond the goals of this course.

Necessary conditions for an interior optimum

Note also, that the function f is supposed differentiable. In the case of ommiting the differentiability condition to the definition of critical points we should add also the points where gradient is not defined.
Non-differntiability goes beyond the goals of this course.

Procedure for solving a many-variable maximization problem
Let f be a differentiable function of n variables and let S be a set of n-vectors. If the problem $\max_{x \in S} f(x)$ has a solution, it may be found as follows:

- Find all the critical points of f in the constraint set S and calculate the value of f at each point.
- Find the largest and smallest values of f on the boundary of S
- The points x you have found at which the value of f is largest are the maximizers.