✗

2.1 Multi-period energy production planning

Consider the problem of supplying energy for the Italian market over a ten year ($T = 10$) time horizon. Let d_t be the expected energy consumption in year $t \in T$. The energy demand in 2009 ($t = 1$) is known and equal to 36.4 GW, the expected growth is 1.81% per year. The *in-house* energy production cost of energy, averaged out on different sources, such as water, thermal, geothermal, wind and solar, is estimated equal to c_1 euro per MW for a production between 0 and l_1 MW; it decreases to $c_2 < c_1$ euro per MW if the amount of energy produced is between l_1 and l_2 MW, and it increases, due to the reduced efficiency, up to $c_3 > c_1$ euro per MW, for an amount greater than l_2 MW. Even if the country is able to meet completely its own energy need, the energy can be partially bought from abroad (mostly from France and Swiss) at an average cost c_4 euro per MW. The amount of imported energy cannot be higher than 13.3% of the overall energy used in 2009.

Advise: we can start with nonlinear programming and then we linearize it

Give an integer linear programming formulation of the problem of minimizing the overall energy supplying cost. *Hint:* Use binary variables to model the piecewise linear cost function.

= use an indicator function

2.2 Home care service planning

A home care service provider must organize the every day activity of the nurse staff. The nurse staff is indexed by the set I. For each nurse $i \in I$ the maximum daily working time v_i is given. The home care service is in charge of a set of patients P. Each patient requires a daily visit and he/she must be visited by one nurse: the length of the visit for patient $k \in P$ is τ_k. Besides, the travel time t_{kl} from the home of patient k to that of patient l is known, for all pairs k, l in P. The home care service provider must assign patients to nurses and decide the order according to which patients are visited by each nurse. Nurses start their daily tour from the home care service head office, and must return to it at the end of the tour. The sum of the travel and visit time of each nurse i cannot exceed his/her available working time v_i. The aim of the home care management is to keep the travel time as small as possible, so as to reduce the amount of nurses' working time wasted for traveling. Give an integer linear programming formulation for the problem.
Hint: Note that the tour carried out by each nurse is a circuit visiting all the patients assigned to him/her (not necessarily Hamiltonian).

Variant: Suppose that patients require different skills and that nurses have different skills: for each patient k let $I_k \subseteq I$ be the subset of nurses who have the skills required by k. How should the formulation be modified to account for the different skills?

2.3 Comparison of directed and undirected formulations for the minimum cost Steiner tree

Given an undirected graph $G = (V, E)$, a cost function $c : E \to \mathbb{R}^+$ and a subset of nodes $T \subset V$, the so-called terminal nodes, the *minimum Steiner Tree* problem asks for finding a minimum cost acyclic subgraph of G (a tree) that spans all the nodes in T, possibly spanning some nodes in $V \setminus T$. Variants of this problem have applications in areas such as circuit layout and network design, for instance in the design of networks for *multicast traffic*.

a) Give an ILP formulation for the problem.

b) Let $G' = (V, A)$ be a directed graph corresponding to G, where each edge $e = \{i, j\}$ is replaced by two arcs $(i, j), (j, i)$, whose cost is $c_{ij} = c_{ji} = c_e$. Give an ILP formulation for the problem of finding the minimum cost Steiner arborescence (i.e., directed subtree that contains a directed path from a given root node r to every other terminal in T) in G'. Show that any solution s_1 of this directed problem corresponds to a solution s_2 of the above undirected problem with equal cost, describing how to derive s_2 from s_1.

c) Show that the feasible region of the linear programming relaxation of the directed formulation is included in the feasible region of the linear programming relaxtion of the undirected formulation.

d) Verify that the inclusion is strict.

2.4 Staff rostering and totally unimodular matrices

a) Consider the following hospital staff rostering problem. Let d_i denote the minimum number of nurses who must be on shift in each day $i = 1, \ldots, 7$ of the week. Each nurse works for 5 consecutive days and then has two days off. The problem consists in assigning shifts to nurses for the next weeks so as to minimize the number of nurses involved, while meeting the daily personnel requirements.

   Give an integer linear programming formulation in matrix form. Is the constraint matrix totally unimodular? What does it imply?

b) Consider the variant of the problem where, instead of cyclically planning a week, only the first 7 days are considered, assuming that the ward was closed in the previous weeks. How do the ILP formulation and the constraint matrix change? Use the necessary and sufficient conditions presented in class to establish whether the constraint matrix is totally unimodular.

2.5 Ideal formulations

Consider the uncapacitated facility location problem. The candidate site set N where a facility can be installed is given, as well as the set of clients M, the fixed cost f_j for using the depot located in candidate site $j \in N$ and the transportation cost c_{ij} of serving the whole demand of client i from depot in candidate site j. The problem consists in deciding where to locate the depots so as to minimize the total installation and transportation costs, while meeting the whole client demands. In the ILP formulation proposed in class, consider the subproblem involving only the constraints linking the location variables y_j and the transportation variables x_{ij}, namely

$$\begin{aligned} x_{ij} &\le y_j && \text{for } i \in M, j \in N \\ x_{ij} &\in \{0,1\} && \text{for } i \in M, j \in N \\ y_j &\in \{0,1\} && \text{for } j \in N, \end{aligned}$$

and the polyhedron corresponding to the feasible region of its linear relaxation.

Are all the vertices of this polyhedron integer (all the coordinates are integer)? Does a similar property hold for the feasible region of the linear relaxation of the overall uncapacitated facility location problem?

2.6 Comparison of formulations for the minimum cost spanning tree

Given an undirected graph $G = (V, E)$, representing a telecommunication network, and a cost function $c : E \to \mathbb{R}^+$, representing the cost of activating the edge e, we want to select the minimum cost subset of edges to be used for a broadcast service which sends messages from a node to any other node of the network. The resulting *Minimum Spanning Tree (MST)* problem asks for finding a minimum cost acyclic subgraph of G that spans all the nodes in V.

a) Give an integer linear programming (ILP) formulation for the problem using cut-set (CUT) constraints suitably modified. Show that the formulation is correct.

b) Replace the cut-set (CUT) constraints with sub-tour elimination constraints (SEC) and verify that such formulation is correct.

c) Let P^0_{SEC} and P^0_{CUT} be the feasible regions of the linear programming relaxations of the above formulations. Show that $P^0_{SEC} \subseteq P^0_{CUT}$.

d) Show that the inclusion is strict by exhibiting a (fractional) solution of the linear relaxation that belongs to P^0_{CUT} but not to P^0_{SEC}.

#1

$T = \{1, 2, \dots, 10\}$ time periods

$S = \{1,2,3\}$ indeces associated with ranges : $1 \to [0, l_1]$, $2 \to [l_1, l_2]$, $3 \to [l_2, \infty)$

$c_{1,2,3,4}$ = cost of MW

d_t = demand in period $t \in T$: $d_1 = 36{,}400$ MW $\quad d_t = 1.0181\, d_{t-1}$

Variables : x_t = amount of energy locally produced in $t \in T$

y_t = " " purchased from aboroard during $t \in T$

$z_t^s \in \{0,1\} = \begin{cases} 1 & \text{if in period } t \text{ the locally produced energy is in } s \in S \\ 0 & \text{if not} \end{cases}$

Model :

$$\min \sum_{t\in T} c_4 y_t + \sum_{t \in T} \sum_{s \in S} c_s x_t z_t^s$$

it's like an indicator function

s.t.

$x_t + y_t \geq d_t \quad \forall t \in T$ (demand)

$y_t \leq 0.133\,(x_1 + y_1) \quad \forall t \in T$ (imported energy constraint)

$$\boxed{\begin{array}{lll} x_t \geq 0 & & \\ x_t \geq z_t^2 l_1 & \forall t \in T & \text{lower bound} \\ x_t \geq z_t^3 l_2 & & \\ x_t \leq l_1 + (1 - z_t^1) M & \forall t \in T & \text{upper bound} \\ x_t \leq l_2 + (1 - z_t^2) M & & \end{array}}$$

! $\sum_{s \in S} z_t^s = 1 \quad \forall t \in T$ (we always have to produce)

$z_t^s \in \{0,1\}, \quad y_t \geq 0, \quad x_t \geq 0$

implies : $\begin{cases} z_t^1 = 1 \Rightarrow 0 \leq x_t \leq l_1 \\ z_t^2 = 1 \Rightarrow l_1 \leq x_t \leq l_2 \\ z_t^3 = 1 \Rightarrow x_t \geq l_2 \end{cases}$

But it's not linear.

We have to adjust the product $x_t z_t^s$:

$k_t^s := c_s x_t \mathbb{1}_{\{x_t \text{ in } s \in S\}}$; $\quad z_t^s = 1 \to k_t^s = c_s x_t$

$\longrightarrow \begin{cases} k_t^s \leq c_s x_t + M(1 - z_t^s) \\ k_t^s \geq c_s x_t - M(1 - z_t^s) \end{cases} \quad \forall s \in S \quad \forall t \in T$

$\longrightarrow$ Model : $\min \sum_{t\in T} c_4 y_t + \sum_{t \in T} \sum_{s \in S} k_t^s$

$(\dots)$

$k_t^s \leq c_s x_t + M(1 - z_t^s)$

$k_t^s \geq c_s x_t + M(1 - z_t^s)$

2

$I = \{\text{nurses}\} \ni i$

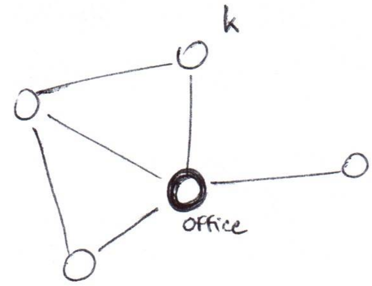
$V_i =$ max daily working for nurse i

$P = \{\text{patients}\} \ni k$

$\tau_k =$ length of the visit of patient k

$t_{kl} =$ time to travel from patient k to pat. l

$(\text{travel time})_i + (\text{visit time})_i \leq V_i \quad \forall i$



The problem can be represented with a graph $G = (N, A)$:

$N = P \cup \{0\} = \{$patients, office$\}$

$A =$ arcs between the nods $=$ travels

Aim : minimize the travel time (ILP formulation)

Variables :

$$x_{ik} = \begin{cases} 1 & \text{if patient } k \text{ is assigned to nurse } i \\ 0 & \text{if not} \end{cases}$$

$$y^i_{lk} = \begin{cases} 1 & \text{if nurse } i \text{ goes from } l \in N \text{ to } k \in N \\ 0 & \text{if not} \end{cases}$$

Objective function :

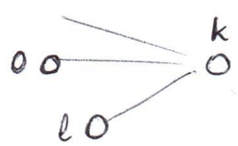
$$\min \sum_{i \in I} \Big[\underbrace{\sum_{k \in P} \sum_{l \in P} y^i_{kl} t_{kl}}_{\text{time between patients}} + \underbrace{\sum_{k \in P} y^i_{k0} t_{k0}}_{\text{time to go back to office}} + \underbrace{\sum_{k \in P} y^i_{0k} t_{0k}}_{\text{time to go to the first patient after office}} \Big]$$

Constraints ;

- $\forall k \in P$ (patient) must be visited by exactly 1 nurse : $\displaystyle\sum_{i \in I} x_{ik} = 1 \quad \forall k \in P$

- $\forall i \in I$ (nurse) starts and end in the office : $\displaystyle\sum_{k \in P} y^i_{k0} = 1 \ , \ \sum_{k \in P} y^i_{0k} = 1 \quad \forall i \in I$

- node-arc connection : if a nurse i goes to k then one arc must go to k : $\displaystyle\sum_{l \in P} y^i_{lk} + y^i_{0k} = x_{ik} \quad \forall i \in I \quad \forall k \in P$



- if a nurse enter a node then she must go out : $\displaystyle\sum_{\substack{l \in P \\ l \neq k}} y^i_{lk} + y^i_{0k} = \sum_{\substack{l \in P \\ l \neq k}} y^i_{kl} + y^i_{k0} \quad \forall i \in I \ \forall k \in P$

- subtour elimination : $\displaystyle\sum_{l,k \in S} y^i_{kl} \leq \sum_{k \in S} x_{ki} - 1 \quad \forall S \subset P, \ 2 \leq |S|$

- time constraints for every nurse :

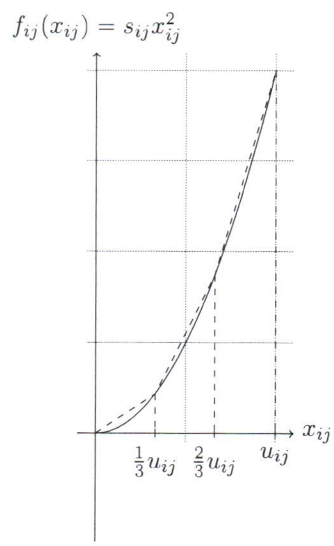
$$\underbrace{\sum_{k \in P} x_{ik} \tau_k}_{\text{visit time}} + \underbrace{\sum_{k \in P} \sum_{l \in P} y^i_{kl} t_{kl}}_{\text{time travel throug patients}} + \underbrace{\sum_{k \in P} t_{0k} y^i_{0k} + \sum_{k \in P} t_{k0} y^i_{k0}}_{\text{office-patients time travel}} \leq V_i \quad \forall i$$

3.1 Capacitated network design

A company must design a telecommunication network. Let V denote the set of nodes in the network and A the set of potential links. A set of demands K must be served: each demand $k \in K$ is described by an origin node s_k and a destination node t_k and is associated with a traffic amount $d_k \geq 0$. Arc capacity must be dimensioned by installing suitable devices on each arc. Each device provides one unit capacity and installing one device on arc $(i, j) \in A$ costs c_{ij}. Let u_{ij} denote the maximum amount of capacity that can be installed on arc (i, j). Let s_{ij} denote the cost of routing one unit of traffic along arc $(i, j) \in A$. To protect the network from failures, we require that there exist at least two link-disjoint paths, using only arcs with non zero capacity, from each origin s_k to each destination t_k, namely, the graph composed of arcs with non zero capacity must be biconnected.

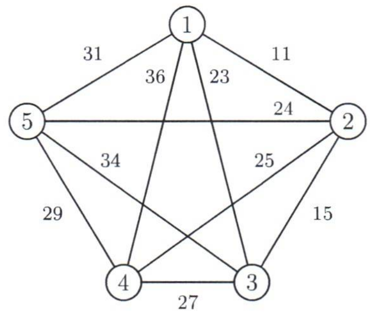
a) Give a mixed integer linear programming formulation for the problem of determining how much capacity to install on each arc so as to minimize the overall installation and routing cost, while satisfying the demand of each origin-destination pair and the survivability requirement. How many constraints are there in the formulation?

b) Suppose that the routing cost on each arc $(i, j) \in A$ is described by the quadratic function $f_{ij} : [0, u_{ij}] \to \mathbb{R}^+$, defined as $f_{ij}(x_{ij}) := s_{ij}x_{ij}^2$, where x_{ij} is the amount of data routed on (i, j). How can the formulation be updated to account for an approximation of the objective function with a piecewise linear function with three pieces, as shown in the figure?



3.2 Branch and Bound for TSP

a) Consider the following graph.



Solve the symmetric TSP problem applying the Branch and Bound method. Use the 2-tree relaxation and compute an initial heuristic solution with the nearest neighbor algorithm. Propose a branching strategy. Are there alternative branching strategies?

b) Consider the asymmetric TSP instance described by the following distance matrix:

$$c_{ij} = \begin{pmatrix} - & 4 & 6 & 3 \\ 3 & - & 6 & 9 \\ 2 & 3 & - & 1 \\ 7 & 9 & 4 & - \end{pmatrix}$$

Solve it by applying the Branch and Bound method. Compute the lower bound by solving a minimum cost assignment problem by inspection (instead of using an ad hoc algorithm or LP solver). Propose a branching strategy.

3.3 Branch and Bound for ILP

Consider the following Integer Linear Programming problem:

$$\begin{aligned} \max \quad & x_1 + 2x_2 \\ \text{s.t.} - & x_1 + x_2 \leq \frac{3}{2} \\ & 2x_1 + 4x_2 \leq 9 \\ & x_1, x_2 \in \mathbb{N}. \end{aligned}$$

Find the optimal solution applying the Branch and Bound method with a best-first strategy (at each iteration the node with the most promising bound is considered), solving the linear relaxations graphically.

4.1 Hub location

In air transportation, usually there is not a direct connection between every pair of airports. Indeed, passengers of intercontinental flights leaving from small airports usually travel on small aircrafts toward the so-called *hub* airports, where they stop over and are boarded on higher capacity aircrafts for the intercontinental flight. If their final destination is a small airport, they stop over in a second hub from which they reach their final destination. This allows the airline company to fully exploit the high capacity of intercontinental flights by aggregating several passengers with different origins and destinations.

An airline company serves a set of cities C. Suppose the number t_{ij} of passengers traveling from city $i \in C$ to $j \in C$ is known. The company must decide where to open its p hubs among the airports in the set C. The distance d_{ij} between airports i and j is given. The cost of one passenger traveling from i to j is proportional to the distance d_{ij}. However, due to the economy of scale, the cost per kilometer of an inter-hub flight is lower than the one of a flight between a non-hub and a hub airport: the cost of the travel between i and j is βd_{ij} if either i or j is not a hub, and it is equal to αd_{ij} if i and j are both hubs, where $\alpha < \beta$.

Give an Integer Linear Programming (ILP) formulation for the problem of deciding where to locate the p hubs and how to assign the non-hub airports to the hubs so as to minimize the overall traveling cost. Consider the following two cases:

a) *multiple assignment*: each non-hub airport can send (and receive) its passengers to (and from) more than one hub.

b) *single assignment*: each non-hub airport can send (and receive) its passengers to (and from) a single hub.

4.2 Chvàtal-Gomory cuts for the maximum clique problem

Given an undirected graph $G = (V, E)$, the maximum clique problem consists in finding a *clique*, i.e., a completely connected subgraph, of maximum cardinality. This well-known NP-hard problem arises in a number of real-world settings including, for instance, social networks, bioinformatics and computational chemistry. In social networks, where the graph's nodes represent people and the graph's edges represent mutual acquaintance, maximum cliques correspond to the largest groups of mutual friends.

If, for every node $i \in V$, the binary variable x_i indicates whether node i is selected in the clique, a natural ILP formulation is as follows:

$$\max \quad \sum_{i \in V} x_i \tag{1}$$

$$\text{s.t.} \quad x_i + x_j \leq 1 \qquad \forall \{i,j\} \notin E \tag{2}$$

$$x_i \in \{0,1\} \qquad \forall i \in V. \tag{3}$$

Let $\mathcal{S}$ denote the set of all the *stable sets* (or independent sets) in G, i.e., the subsets of nodes such that no two nodes are adjacent. Consider the stable set inequalities

$$\sum_{i \in S} x_i \leq 1 \qquad \forall S \in \mathcal{S}.$$

a) Verify that these inequalities are valid. Are they tighter than those in (2)?

b) Which stable set inequalities can be obtained by applying a single iteration of the Chvàtal-Gomory procedure?

c) Show that all the stable set inequalities can be obtained applying the Chvàtal-Gomory procedure. Which is the Chvàtal-Gomory rank of a stable set inequality of cardinality k?

4.3 Cover inequalities for the 0-1 knapsack problem

Consider the 0-1 knapsack problem with the feasible region

$X = \{\underline{x} \in \{0,1\}^6 : 12x_1 + 9x_2 + 7x_3 + 5x_4 + 5x_5 + 3x_6 \leq 14\}$.

a) List all the minimal cover inequalities for X.

b) Assuming $x_1 = x_2 = x_4 = 0$, consider the cover inequality $x_3 + x_5 + x_6 \leq 2$, which is valid for $X' = X \cap \{\underline{x} \in \{0,1\}^6 : x_1 = x_2 = x_4 = 0\}$. Show that this cover inequality defines a facet of $conv(X')$.

c) Apply the sequential lifting procedure to the cover inequality $x_3 + x_5 + x_6 \leq 2$ according to the order x_1, x_2, x_4 to obtain an inequality $\alpha_1 x_1 + \alpha_2 x_2 + x_3 + \alpha_4 x_4 + x_5 + x_6 \leq 2$ valid for X. Does this inequality define a facet of $conv(X)$?

d) Describe the separation problem for the cover inequalities and propose a simple heuristic to identify violated cover inequalities. Given the optimal solution of the current linear relaxation with the cover inequalities generated so far, how do we need to tackle the above separation problem to make sure that no other violated cover inequality exist?

4.4 Valid inequalities for the maximum stable set problem

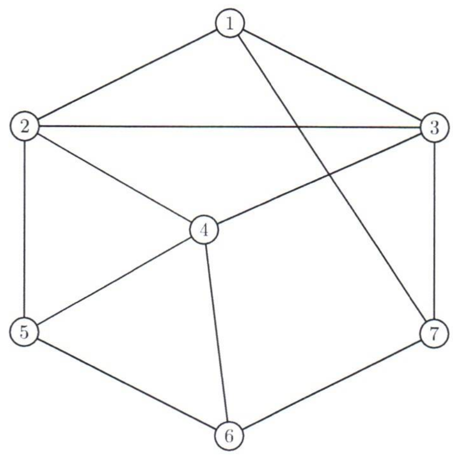
Consider the maximum stable set problem where, given an undirected graph $G = (V, E)$, we wish to find a stable set, i.e., a subset of nodes that are not pairwise adjacent, of maximum cardinality. This NP-hard problem arises in many fields of application where the edges represent incompatibilities between the entities associated to the corresponding nodes.

If, for every node $i \in V$, the binary variable x_i is equal to 1 if node i is selected in the stable set and 0 otherwise, a natural ILP formulation is as follows:

$$\begin{aligned} \max \quad & \sum_{i \in v} x_i \\ \text{s.t.} \quad & x_i + x_j \leq 1 && \forall \{i,j\} \in E \\ & x_i \in \{0,1\} && \forall i \in V. \end{aligned}$$

a) Which valid inequality can be associated to every clique $K \subseteq V$ of G? Such valid inequalities are referred to as *clique inequalities*.

b) Consider the instance of the maximum stable set problem defined by the following undirected graph:



Add to the above natural ILP formulation all the clique inequalities of cardinality 3, that is, containing 3 nodes.

c) Show, by exhibiting a fractional optimal solution x^*_{LP} of the linear relaxation of the ILP formulation obtained in (b), that the clique inequalities are not sufficient to describe the convex hull of all integer solutions (incidence vectors of the stable sets).

d) Determine the dimension of the stable set polytope P_{Stab}, i.e., the convex hull of the incidence vectors of all stable sets of the instance in (b). Show that the clique inequality associated to a clique K defines a facet of P_{Stab} if and only if K is a maximal clique (in terms of inclusion).

e) State the separation problem for the clique inequalities and mention how it can be solved?

f) A subset $H \subseteq V$ is a *hole* of G if the induced subgraph is a simple cycle, i.e., the nodes of H are connected via a cycle and for each pair of non adjacent nodes i and j we have $\{i, j\} \notin E$. Which valid inequality can be associated to any odd hole H of G, that is, a hole of G with an odd number of nodes? Explain why such odd hole inequalities are valid.

g) What about the odd hole inequalities and of cardinality 3 and the clique inequalities of cardinality 3? List all the odd hole inequalities of cardinality 5 for the instance in (b).

h) Consider the instance in (b) and the odd hole inequality corresponding to the nodes 2, 3, 7, 6 and 5, and lift the coefficient of the variable x_4 associated to node 4. Note that this odd hole of cardinality 5 with node 4 constitutes a "wheel".

5.1 Base station location and power management in cellular networks

Consider the problem of locating a set of *Base Stations* (BSs) in a UMTS network. Let $I = \{1, \ldots, m\}$ denote the set of *Test Points* (TPs) which represent the origins of the traffic. Let $S = \{1, \ldots, n\}$ denote the set of candidate sites in which a BS can be installed. Let c_j be the cost of installing one BS in site $j \in S$. Let $g_{ij} \in [0, 1]$ denote the transmission gain between TP $i \in I$ and candidate site $j \in S$: if TP $i \in I$ is transmitting with power p_i, the power received by the BS located in candidate site $j \in S$ is $g_{ij}p_i$. We consider the *uplink* part of the communications, namely the connections from cellular phones (TPs) to the base stations.

To guarantee suitable quality of service, the signal-to-interference ratio (SIR) between each TP and the BS to which it is assigned must be greater or equal to a threshold SIR_{min}. The interference between TP $i \in I$ and the BS located in $j \in S$ is the sum of the power received from all the other TPs $h \in I \setminus \{i\}$.

The location and power management problem consists in deciding where to locate the BSs and, for each TP, the transmission power (between 0 and p_{max}) and the BS to which it is assigned so as to minimize the total installation cost for the BSs.

a) Give a mixed integer nonlinear programming formulation of the problem, formulating the SIR constraints as described above.

b) How can the formulation be modified to obtain a mixed integer linear programming formulation? *Hint:* Note that equation $\frac{a}{b} \geq c$, with $b > 0$, can be written as $a \geq bc$, thus obtaining a formulation involving only a product of a binary and a continuous variable.

5.2 Lagrangian relaxation for the symmetric traveling salesman problem

Consider the undirected complete graph with 5 nodes K_5, whose edge costs are given in the following table:

	2	3	4	5
1	1	6	7	2
2	.	4	3	8
3	.	.	6	2
4	.	.	.	8

Apply 4 iterations of the subgradient method to the Held and Karp Lagrangian dual, using the following updating rules:

$$u_i^{k+1} = u_i^k + \alpha_k(2 - \sum_{e \in \delta(i)} x_e)$$

where $\alpha_k = \alpha_0 \rho^k$. The parameters are initialized as follows: $\underline{u}^0 = \underline{0}$, $\alpha_0 = 10$ and $\rho = \frac{1}{2}$.

5.3 Lagrangian relaxation for the minimum spanning tree problem with node degree constraints

A telecommunication company must design a network to connect a set of V nodes. The network is represented by a graph $G = (V, E)$, where E is the set of edges which can be activated. For each edge $e \in E$, let c_e denote the cost of installing the needed transmission devices on the edge. The company must select the subset of edges to be activated so as to connect each pair of nodes. The goal is to minimize the overall cost of the transmission devices. For technical reasons, each node can deal with at most d active connections.

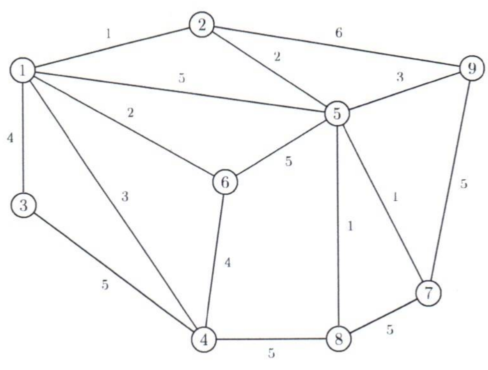
a) Give an ILP formulation for the problem.

b) Propose a Lagrangian relaxation for the problem, clearly describe the Lagragian subproblem and the Lagrangian dual problem. How can the Lagrangian subproblem be solved?

c) Apply the first 3 iterations of the subgradient method to tackle the Lagrangian dual of the instance reported below, with $d = 3$. Update the multipliers according to the rule:

$$u_i^{k+1} = \max\{u_i^k + \alpha\gamma_i^k, 0\},$$

where γ_i^k is the i-th component of the subgradient of the dual function at $\underline{u}^k$. Initialize the parameters as follows: $\underline{u}^0 = \underline{0}$ and $\alpha = 1$.



5.4 Choice of the Lagrangian dual

Consider the generalized assignment problem with a set I of processes and a set J of machines. Denote by

- c_{ij} the cost for executing process $i \in I$ on machine $j \in J$,
- w_{ij} the amount of resource required to execute process $i \in I$ on machine $j \in J$,

- b_j the total amount of resource available on machine $j \in J$.

The goal is to assign the processes to the machines so as to minimize the total cost, while respecting the resource constraints. Assume that once started the processes cannot be interrupted.

An ILP formulation for the problem is as follows:

$$
\begin{aligned}
\min\ z = & \sum_{i\in I}\sum_{j\in J} c_{ij}x_{ij} \\
\text{s.t.} \quad & \sum_{j\in J} x_{ij} = 1 && \forall i \in I \\
& \sum_{i\in I} w_{ij}x_{ij} \leq b_j && \forall j \in J \\
& x_{ij} \in \{0,1\} && \forall i \in I, j \in J,
\end{aligned}
$$

where $x_{ij} = 1$ if process $i \in I$ is assigned to machine $j \in J$, and $x_{ij} = 0$ otherwise.

a) Propose two Lagrangian relaxations in which different groups of constraints are relaxed.

b) For each one of the two Lagrangian relaxations, describe the Lagrangian subproblem and explain how it can be solved.

c) Write the Lagrangian duals of the two relaxations and establish whether the associated bounds can be stronger (tighter) than that provided by the LP relaxation. Recall the general result presented in class and apply it to both Lagrangian duals.

5.5 Lagrangian duality and Linear Programming duality

Consider the following generic Linear Programming problem in canonic form:

$$
\begin{aligned}
\min \quad & \underline{c}^T\underline{x} \\
\text{s.t.} \quad & A\underline{x} \geq \underline{b} \\
& \underline{x} \geq \underline{0},
\end{aligned}
$$

where $A \in \mathbb{R}^{m\times n}$, $\underline{c} \in \mathbb{R}^n$, $\underline{x} \in \mathbb{R}^n$ and $\underline{b} \in \mathbb{R}^m$ and formulate the Lagrangian dual by relaxing the constraints $A\underline{x} \geq \underline{b}$. Consider now the Linear Programming dual:

$$
\begin{aligned}
\max \quad & \underline{u}^T\underline{b} \\
\text{s.t.} \quad & \underline{u}^T A \leq \underline{c}^T \\
& \underline{u} \geq \underline{0}.
\end{aligned}
$$

What is the relationship between the two dual problems?

1.1 Crude oil mix

An oil refinery is composed of a basin and two blenders, as shown in Figure 1.

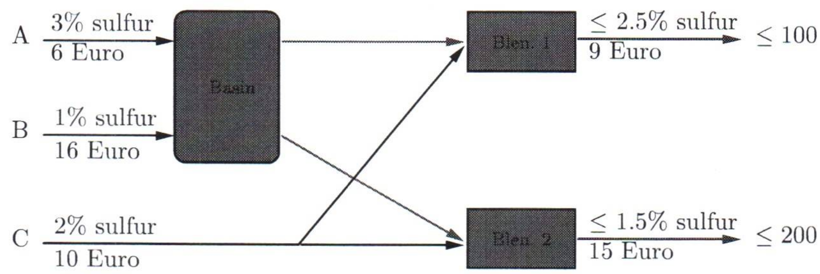


Figure 1: Crude oil mix

The basin has two entrances corresponding to two types of crude oil, A and B, whose costs per barrel are 6 and 16 euros, respectively. The percentage of sulfur are 3% and 1%, respectively. The oil goes from the basin to the blenders which have a third entrance C, whose cost per barrel is 10 euros and whose percentage of sulfur is 2%. The first blender must produce a crude oil with a maximum guaranteed percentage of sulfur equal to 2.5% and a cost per barrel equal to 9 euro. The second blender must produce a higher quality oil, whose cost is 15 euros per barrel and with a guaranteed percentage of sulfur equal to 1.5%. Maximum market request is 100 barrels for the first oil type, and 200 barrels of the second oil type.

a) Give a mathematical programming formulation for the problem of determining the amount of crude oil of type A, B and C needed in order to maximize the revenue. Is it a convex problem?

b) **Optional.** Consider a variant where the percentage of sulfur coming out of the basin can be reduced by filtering the oil. Filtering the oil costs 1.1 euro per each filtered barrel of sulfur. (Note that the amount of oil coming out of the basin is reduced by the amount of filtered sulfur).

1.2 Service center resource allocation

A web service provider, which owns a set N of servers, receives a set R of requests. Each request $r \in R$ consists in a software application (for instance, online banking or flight reservation) to be run on a server. Requests are served by virtual machines which are instantiated on the servers. Due to memory constraints on each server $i \in N$ at most M_i virtual machines can be instantiated. Each virtual machine on server $i \in N$ has an operational cost c_i (to be paid if requests are assigned to it) and a service rate μ_i (namely the number of requests that can be served in one unit of time). The service provider must decide how many virtual machines are instantiated on each server. Requests can be rejected: if a request $r \in R$ is rejected a penalty p_r must be paid. To guarantee a suitable level of service, the service provider must guarantee that the average delay on each server is below a threshold τ. Service provider engineers decided to approximate the delay on a server as the inverse of the difference between the total service rate for all the virtual machines instantiated on the server and the average number of requests assigned to each virtual machine. Give a mixed integer non linear programming formulation for the problem of allocating the virtual machines on the servers and deciding which requests are served and by which server so as to minimize the total costs.

Variant. Generalize the model to consider classes of requests. Requests are divided into classes, described by the set K. The number of requests in class k is a_k. All the requests in the same class k have the same penalty p_k. The delay constraint must be satisfied for each server and for each class: the threshold for class k is τ_k. The service rate of virtual machines on server i depends on the class: let μ_i^k be the service rate of virtual machines on server i for requests in class k.

1.3 Convexity and epigraph

Let $f : C \to \mathbb{R}$ be a function and $C \subseteq \mathbb{R}^n$ a convex set. Show that f is convex if and only if $\text{epi}(f) \subseteq \mathbb{R}^{n+1}$ is convex.

1.4 Convexity

(a) Discuss the convexity of $f(x_1, x_2) = x_1^2 + kx_2^2$, with k constant, using equivalent definitions of convexity.

(b) The sound intensity perceived by a listener is directly proportional to the power of the source and inversely proportional to the square of the distance between the listener and the source. In the one-dimensional case, the function is:

$$I_s(x) = \eta \frac{1}{(x - x_s)^2}$$

where x_s is the position of the source and η is a parameter that depends on the power of the source and on the transmission mean. Discuss the convexity of I_s.

(c) Consider two loudspeakers s and t. The sound intensity function is

$$J(x) = I_s(x) + I_t(x) = \eta_s \frac{1}{(x - x_s)^2} + \eta_t \frac{1}{(x - x_t)^2}$$

where η_s and η_t are the η parameters as described above for source s and t, respectively. Is the function J convex on the interval $(-\infty, \min\{x_s, x_t\})$?

1.5 Operators preserving convexity

(a) Consider a convex set $C \subseteq \mathbb{R}^n$ and convex functions $f_i : C \to \mathbb{R}$ with $i \in \{1, \ldots, m\}$. No assumption on differentiability is considered. Are the following functions convex?

- $g(\underline{x}) = \sum_{i=1}^{m} \lambda_i f_i(\underline{x})$ with $\lambda_i \geq 0$ for $i \in \{1, \dots, m\}$
- $g(\underline{x}) = \max_i f_i(\underline{x})$
- $g(\underline{x}) = \min_i f_i(\underline{x})$

(b) Let $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be two functions in class C^2. Under which conditions the function $h = g \circ f = g(f(\underline{x}))$ is convex? Does the convexity change if $f : \mathbb{R}^n \to \mathbb{R}$?

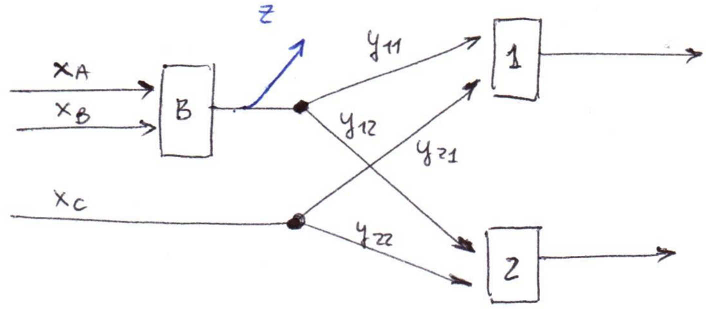
(c) Consider a convex set $C \subseteq \mathbb{R}^n$.

1) Is function $\frac{1}{f}$ convex if $f : C \to \mathbb{R}_+ \backslash \{0\} \in C^2$ is concave?

2) Is function $\frac{1}{f}$ convex if $f : C \to \mathbb{R}_+ \backslash \{0\} \in C^2$ is convex?

#1

a)



x_A, x_B, x_C = quantity of crude oil coming from A, B, C

y_{11} = quantity going from the barril to 1

y_{12} = from the barril to 2

y_{21} = from C to 1

y_{22} = from C to 2

p = percentage of the sulfur out of the barril

Model :

$$\max \left(9(y_{11}+y_{21}) + 15(y_{12}+y_{22})\right) - (6x_A + 16x_B + 10x_C)$$

such that :

- $x_A + x_B = y_{11} + y_{12}$
- $x_C = y_{21} + y_{22}$ } balance
- $y_{11} + y_{21} \leq 100$
- $y_{22} + y_{12} \leq 200$ } demand
- $0.03x_A + 0.01x_B = p(y_{11}+y_{12})$ } balance of % of sulfur
- $py_{11} + 0.02y_{21} \leq 0.025(y_{11}+y_{21})$
- $py_{21} + 0.02y_{22} \leq 0.015(y_{12}+y_{22})$ } % of sulfur accepted
- $x_A, x_B, x_C \geq 0$
- $y_{11}, y_{12}, y_{21}, y_{22} \geq 0$
- $p \geq 0$

The problem is not convex due to the presence of the bilinear terms in p and y.

b) Filtering of the oil (add $\xrightarrow{z}$)

z = quantity of the sulfur extracted

y_{11} and y_{12} are considered after the filtering

Model :

$$\max \left(9(y_{11}+y_{21}) + 15(y_{22}+y_{12})\right) - (6x_A + 16x_B + 10x_C) - 1.1z$$

such that :

- $x_A + x_B = y_{11} + y_{12} + z$ } new balance
- $0.03x_A + 0.01x_B = z + p(y_{11}+y_{12})$ } balance of % of sulfer
- $z \leq 0.03x_A + 0.01x_B$
- $z \geq 0$ } z conditions
- (previous model)

2

N = servers $i \in N$

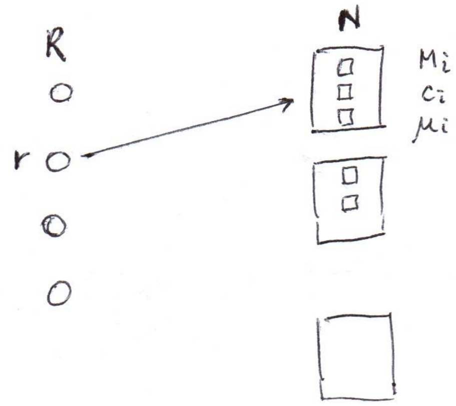
R = requests $r \in R$

c_i = cost of 1 virtual machine

M_i = # virtual machines in $i \in N$

p_r = penalty for rejection

μ_i = # requests served in 1 unit of time



request $r \in R$ —accepted→ $i \in N$ servers
M_i virtual machines
c_i cost for VM
μ_i #request per unit of time

—rejected→ p_r penalty

Variables:

$$x_{ir} = \begin{cases} 1 & \text{if } i \text{ accepts the request } r \in R \\ 0 & \text{if not} \end{cases}$$

y_i = number of VM on server $i \in N$

Model:

$$\min \sum_{i \in N} y_i c_i \qquad \sum_{r \in R} p_r \left(1 - \sum_{i \in N} x_{ir}\right)$$

such that:

$$\sum_{i \in N} x_{ir} \leq 1$$ a request is assigned at most once

$$y_i \leq M_i$$ max num of VM on i

$$\frac{1}{\mu_i y_i - \frac{\sum_{r \in R} x_{ir}}{y_i}} \leq \tau$$ delay

#3

f convex : $f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \leq \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2)$

epi(f) : $= \{ (\underline{x}, y) \in C \times \mathbb{R} : f(\underline{x}) \leq y \} \subseteq \mathbb{R}^{n+1}$

epi(f) convex : $\alpha(\underline{x}_1, y_1) + (1-\alpha)(\underline{x}_2, y_2) \in C \quad \forall (\underline{x}_1, y_1), (\underline{x}_2, y_2) \in C$

f convex $\Longleftrightarrow$ epi(f) convex

$(\Longrightarrow)$ $(\underline{x}_1, y_1), (\underline{x}_2, y_2) \in \text{epi}(f)$: $f(\underline{x}_1) \leq y_1, \quad f(\underline{x}_2) \leq y_2$

so: $\left.\begin{array}{l} \alpha f(\underline{x}_1) \leq \alpha y_1 \\ (1-\alpha) f(\underline{x}_2) \leq (1-\alpha) y_2 \end{array}\right\} \Rightarrow \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2) \leq \alpha y_1 + (1-\alpha) y_2$

since f is convex:

$$\begin{aligned} f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) &\leq \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2) \\ &\leq \alpha y_1 + (1-\alpha) y_2 \end{aligned}$$

$(\Longleftarrow)$ $(\underline{x}_1, y_1), (\underline{x}_2, y_2) \in \text{epi}(f) \longrightarrow \alpha(\underline{x}_1, y_1) + (1-\alpha)(\underline{x}_2, y_2) \in \text{epi}(f)$

$\Longrightarrow f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \leq \alpha y_1 + (1-\alpha) y_2 \qquad (*)$

and since $f(\underline{x}_1) \leq y_1$ and $f(\underline{x}_2) \leq y_2$, $(*)$ holds even for

$y_1 = f(\underline{x}_1)$

$y_2 = f(\underline{x}_2) \quad \Longrightarrow f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \leq \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2)$

#4

(a) $f(\underline{x}) = f(x, y) = x^2 + ky^2$

1. $\boxed{f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \leq \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2)}$:

$$\begin{aligned} f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) &= f\left(\alpha \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + (1-\alpha) \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \\ &= f\left(\begin{bmatrix} \alpha x_1 + (1-\alpha) x_2 \\ \alpha y_1 + (1-\alpha) y_2 \end{bmatrix}\right) \\ &= (\alpha x_1 + (1-\alpha) x_2)^2 + k(\alpha y_1 + (1-\alpha) y_2)^2 \\ &= [..] \\ &= \alpha^2(x_1^2 + ky_1^2) + (1-\alpha)^2 (x_2^2 + ky_2^2) + \alpha(1-\alpha)(x_1 x_2 + k y_1 y_2) \end{aligned}$$

$f(\alpha \underline{x}_1 + (1-\alpha)\underline{x}_2) \overset{?}{\leq} \alpha f(\underline{x}_1) + (1-\alpha) f(\underline{x}_2) = \alpha(x_1^2 + ky_1^2) + (1-\alpha)(x_2^2 + ky_2^2)$

$\Longrightarrow \alpha^2(x_1^2 + ky_1^2) + (1-\alpha)^2(x_2^2 + ky_2^2) + \alpha(1-\alpha)(x_1 x_2 + k y_1 y_2) \overset{?}{\leq} [..]$

$\Longrightarrow [..]$

$\Longrightarrow \alpha(1-\alpha)\left[(x_1 - x_2)^2 + k(y_1 - y_2)^2\right] \geq 0$

$\Longrightarrow$ for $k \geq 0$ the function is convex

(a) 2. $\boxed{f \text{ convex} \iff f(\underline{y}) \geq f(\underline{x}) + \nabla^T f(\underline{x})(\underline{y}-\underline{x}) \quad \forall \underline{x}, \underline{y} \in C}$

$\nabla f(\underline{x}) = [2x, 2ky]^T$

We want: $f(\underline{x}_1) \geq f(\underline{x}_2) + \nabla^T f(\underline{x}_2)(\underline{x}_1 - \underline{x}_2)$

$$\Longrightarrow x_1^2 + ky_1^2 \overset{?}{\geq} x_2^2 + ky_2^2 + [2x_2, 2ky_2]\begin{bmatrix} x_1 - x_2 \\ y_1 - y_2 \end{bmatrix}$$

$$\overset{?}{\geq} x_2^2 + ky_2^2 + 2x_2(x_1 - x_2) + 2ky_2(y_1 - y_2)$$

$$\overset{?}{\geq} -x_2^2 - ky_2^2 + 2x_1x_2 + 2ky_1y_2$$

$$\Longrightarrow (x_1 - x_2)^2 + k(y_1 - y_2)^2 \overset{?}{\geq} 0 \quad \text{always true if } k \geq 0$$

3. $\boxed{f \text{ convex} \iff \nabla^2 f(\underline{x}) = \left(\frac{\partial^2 f(\underline{x})}{\partial x_i \partial x_j}\right) \text{ is positive def.}}$

f must be two times continuously differentiable

$$\nabla^2 f(\underline{x}) = \begin{bmatrix} 2 & 0 \\ 0 & 2k \end{bmatrix} \Longrightarrow$$

- $\underline{x}^T A \underline{x} \geq 0$:
  $[x \ y]\begin{bmatrix} 2 & 0 \\ 0 & 2k \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = 2x^2 + 2ky^2 \geq 0$
- $\text{eigen}(A) = \{2, 2k\}$

In both cases f convex if $k \geq 0$

(b) $f(x) = \dfrac{\eta}{(x-x_S)^2}$, $f'(x) = \dfrac{-2\eta}{(x-x_S)^3}$, $f''(x) = \dfrac{6\eta}{(x-x_S)^4}$

There is a discontinuity in $x = x_S$.

In $(-\infty, x_S)$ $f''(x)$ is always $\geq 0 \Longrightarrow f(x)$ convex in $(-\infty, x_S)$

In (x_S, ∞) $f''(x)$ is always $\geq 0 \Longrightarrow f(x)$ convex in $(x_S, +\infty)$

It's not possible to determine if $f(x)$ is convex in $(-\infty, +\infty)$ since it's not defined in that interval.

(c) $f''(x) = \dfrac{6\eta_S}{(x-x_S)^4} + \dfrac{6\eta_t}{(x-x_t)^4} \Longrightarrow f(x)$ convex in $(-\infty, \min\{x_S, x_t\})$

5

(a) C convex $(C \subseteq \mathbb{R}^n)$, $f_i : C \to \mathbb{R}$ convex $\forall i \in \{1, \dots, m\}$

- $g(\underline{x}) = \sum_{i=1}^m \lambda_i f_i(\underline{x}) \quad \lambda_i \geq 0 \ \forall i$ is convex:

$$g(\alpha\underline{x}_1 + (1-\alpha)\underline{x}_2) = \sum_{i=1}^m \lambda_i f_i(\alpha\underline{x}_1 + (1-\alpha)\underline{x}_2) \leq \sum_{i=1}^m \lambda_i\left(\alpha f_i(\underline{x}_1) + (1-\alpha) f_i(\underline{x}_2)\right)$$

$$\leq \alpha\left(\sum_{i=1}^m \lambda_i f_i(\underline{x}_1)\right) + (1-\alpha)\left(\sum_{i=1}^m \lambda_i f_i(\underline{x}_2)\right)$$

$$= \alpha\, g(\underline{x}_1) + (1-\alpha)\, g(\underline{x}_2)$$

$\Longrightarrow$ The sum of two convex functions is still convex

- $g(\underline{x}) = \max_i f_i(\underline{x})$ is convex

$$\text{epi}(g) = \{(\underline{x}, y) : \max f_i(\underline{x}) \leq y\} = \{(\underline{x}, y) : f_i(\underline{x}) \leq y \ \forall i\}$$

$$= \cap_{i=1}^m \text{epi}(f_i)$$

$\text{epi}(f_i)$ is convex $\forall i \Longrightarrow$ the $\boxed{\text{intersection}}$ of convex functions is convex

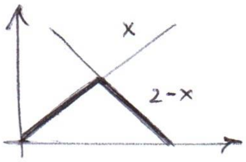
$\longrightarrow \text{epi}(g)$ is convex $\Longrightarrow g$ is convex

arbitrary intersection

#5 (2)

(a) • $g(\underline{x}) = \min_i f_i(\underline{x})$ is <u>not</u> convex

let $f_1(x) = x$, $f_2(x) = 2-x$;



The function $\min\{f_1(x), f_2(x)\}$ is concave

(b) $f: \mathbb{R} \to \mathbb{R}$, $g: \mathbb{R} \to \mathbb{R}$, $f, g \in C^2$ $\Big\}$ $h = g \circ f = g(f(x))$ is convex?

$$h'(x) = g'(f(x))\, f'(x)$$
$$h''(x) = g''(f(x))(f'(x))^2 + g'(f(x))\, f''(x)$$

h convex $\iff$ $h'' \geq 0$

$(f'(x))^2 \geq 0$,

g''	g'	f''	h''
+	+	+	+
+	+	−	?
+	−	+	?
+	−	−	+
−	+	+	?
−	+	−	−
−	−	+	−
−	−	−	?

Here it depends on the cases

h(x) is
- convex if:
  - g convex, $g \uparrow$, f convex
  - g convex, $g \downarrow$, f concave
- concave if:
  - g concave, $g \uparrow$, f concave
  - g concave, $g \downarrow$, f convex

Alternatively we can study h without the derivatives:

$$g(f(\alpha x_1 + (1-\alpha)x_2)) \leq \alpha\, g(f(x_1)) + (1-\alpha)\, g(f(x_2))$$

(c) We are studying $\frac{1}{x} \circ f$ with $f: C \to \mathbb{R}_+ \setminus \{0\}$.
By setting $g(x) = \frac{1}{x}$ we can use the point (b).

$g'(x) = -\frac{1}{x^2}$, $g''(x) = \frac{2}{x^3}$ $\qquad g'' > 0$, $g' < 0$

$\Longrightarrow$ f :
- • $f'' \geq 0 \Longrightarrow$ $1/f$ can be anything (*)
- • $f'' \leq 0 \Longrightarrow$ $1/f$ is convex

(*) : $f(x) = \frac{1}{x^2} \longrightarrow \frac{1}{f}(x) = x^2$ convex

$f(x) = \frac{1}{\sqrt{x}} \Longrightarrow \frac{1}{f}(x) = \sqrt{x}$ concave

6.1 Clustering

Data mining, which is the interdisciplinary subfield of computer science aiming at discovering patterns in large data sets, has a wide range of applications. An important problem arising in data mining is the so-called clustering problem which consists in subdividing a given set of data points into subsets and assigning a center to each subset. Clustering is applied for instance in health care, with the aim of dividing a set of patients into clusters, based on their available clinical data, and deciding a diagnosis for each patient.

Consider the problem of determining, given a set of points $A = \{\underline{a}_1 \dots \underline{a}_m\} \subseteq \mathbb{R}^n$, a partition of A in k subsets (the so-called *clusters*) and assigning to each cluster A_j a *center* $\underline{c}_j \in \mathbb{R}^n$, which may not coincide with an element of A_j, so as to minimize the sum, for all the points $\underline{a}_i \in A$, of the squared Euclidean distance between $\underline{a}_i$ and the center $\underline{c}_j$ of the corresponding cluster A_j.

(a) Give a mixed integer nonlinear programming formulation for the problem.

(b) Consider the two subproblems: the partition of the data points into subsets (clusters) and the determination of the centers of each cluster. Suppose that the partition of the data points, and therefore the elements of each cluster, is given: can the subproblem of determining the k centers be easily solved? Analogously, if the centers of each cluster are given, can the partition subproblem be easily solved? Propose a heuristic approach based on the iterative solution of the two subproblems.

6.2 Traffic assignment problem

Consider a traffic network $G = (V, A)$ whose arcs represent the roads of a city and whose nodes represent the crossroads. Citizens travel every day from some origins toward some destinations in the city: let d_k be the amount of cars which leave point $s_k \in V$ and move toward destination $t_k \in V$. The *free flow time* of arc (i, j) (FFT_{ij}) is the time needed to travel from i to j on arc (i, j) if the road (i, j) is empty. The traveling time c_{ij} increases with the increasing flow on the arc, and in general in the network. A congestion threshold u_{ij} is given for each arc (i, j) which represents the maximum flow on arc (i, j) before congestion.

a) Consider the case where the traveling time c_{ij} is equal to the FFT_{ij} increased by the fourth power of the ratio between the flow on the arc and its congestion threshold. Give a nonlinear programming formulation of the problem of assigning traffic to roads so as to minimize the total traveling time.

b) Consider the case where the amount of flow on arc (j, i) has an impact on the traveling time on (i, j), namely, beside the flow on (i, j) also the flow on (j, i) increases the congestion. How does the formulation change?

c) And how does it change if also the flow on the roads incident in j contribute to the (i, j) congestion?

6.3 Convergence order and rate

Analyze the convergence order and rate of the following sequences:

a) $x_k = 1 + \frac{1}{k} \to 1$

b) $x_k = 1 + \frac{1}{2^k} \to 1$

c) $x_k = 1 + \frac{1}{k^k} \to 1$

d) $x_k = 1 + \frac{1}{2^{2^k}} \to 1$

(*Hint:* study the following limit $\lim_{k\to\infty} \frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^p} = r$.)

6.4 Gradient method step

Consider the following function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} \frac{3}{4}(1-x)^2 - 2(1-x) & \text{per } x > 1 \\ \frac{3}{4}(1+x)^2 - 2(1+x) & \text{per } x < -1 \\ x^2 - 1 & \text{per } -1 \le x \le 1. \end{cases}$$

(a) Draw the graph of f. Analyze convexity and differentiability of f. Determine a minimizer of f.

(b) Show that if the starting point is x_0 with $|x_0| > 1$, and the step size is $\alpha_k = 1, \forall k$, the gradient method does not converge to the minimum, even if f is convex.

(c) Show that if x_k is such that $|x_k| > 1$ the step size is $\alpha_k = 1$ satisfies the descent property $f(x_{k+1}) < f(x_k)$.

6.5 Wolfe conditions for quadratic functions

Let $f(\underline{x}) = \underline{x}^T Q \underline{x}$, with Q positive definite. Consider the first Wolfe condition

$$f(\underline{x}_k + \alpha_k \underline{d}_k) \le f(\underline{x}_k) + c_1 \alpha_k (\nabla f(\underline{x}_k))^T \underline{d}_k. \tag{1}$$

Find the maximum value of c_1 such that the condition (1) is satisfied if the step size α_k is computed via the exact line search.

Hint. Compute the value of α_k as the minimum of $f(\underline{x}_k + \alpha_k \underline{d}_k)$. Replace the obtained value in the Wolfe condition.

6.6 Affine invariance of Newton method

Consider the following linear mapping $\underline{x} = D\underline{y}$, where D is an $n \times n$ nonsingular matrix. Apply the Newton method to the functions $f(\underline{x})$ and $f(D\underline{y})$, and let $\underline{d}_{\underline{x}} = -(\nabla^2_{\underline{x}} f(\underline{x}))^{-1} \nabla_{\underline{x}} f(\underline{x})$ and $\underline{d}_{\underline{y}} = -(\nabla^2_{\underline{y}} f(D\underline{y}))^{-1} \nabla_{\underline{y}} f(D\underline{y})$ the corresponding Newton directions.

Show that $\underline{x} + \underline{d}_{\underline{x}} = D(\underline{y} + \underline{d}_{\underline{y}})$.

Hint. Show that $\nabla_{\underline{y}} f(D\underline{y}) = D^t \nabla_{\underline{x}} f(\underline{x})$ and $\nabla^2_{\underline{y}} f(D\underline{y}) = D^t \nabla^2_{\underline{x}} f(\underline{x}) D$.

7.1 The largest polygon

Give the mathematical programming formulation for the following problem: given a positive integer n, find the polygon with n edges with diameter (the maximum distance between two vertices) less or equal to 1, of maximum area.

Hint. Use the polar coordinates. Note that, given the length a, b of two edges of a triangle and the angle α between them, the area of the triangle is $\frac{1}{2}ab\sin(\alpha)$. Note that, given two vectors, their lengths a, b and the angle α between them, their square Euclidean distance is $a^2 + b^2 - 2ab\cos(\alpha)$.

7.2 Optimal control

A robot must move an object, whose mass is M, from the initial position $\underline{x}^i \in \mathbb{R}^3$ to the final position $\underline{x}^f \in \mathbb{R}^3$, avoiding a parallelepiped obstacle Y: the position and size of the obstacle, mass and initial and final position of the object are known. The movement must be performed in T seconds, as shown in Figure 1. For the sake of simplicity, the time interval $[0 \ldots T)$ is divided into n time slots $I_k = \left[\frac{(k-1)T}{n}, \frac{kT}{n}\right)$ $(k = 1 \ldots n)$ of equal length. We assume that the force to be applied on the object in each direction is constant during each time slot. Determine two input control sequences for the robot, which optimize the "comfort" of the movement and the overall work performed. Optimizing the "comfort" consists in minimizing the force variations between each time-slot and the next one.

Hint. Represent the *state* of the system with two vectors, $\underline{x}$ and $\underline{v}$, representing the position and the speed at the beginning of each time slot. The constraint which avoids collisions between the object and the obstacle can be relaxed, by applying the constraint only to such instants.

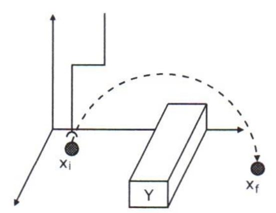


Figure 1: Optimal robot control.

7.3 Optimizing along conjugate directions

Apply two iterations of the conjugate directions method to the following problem

$$\min f(x_1, x_2) = -12x_2 + 4x_1^2 + 4x_2^2 - 4x_1x_2$$

starting from the initial point $\underline{x}_0 = (-\frac{1}{2}, 1)$ and from direction $\underline{d}_0 = (1, 0)$. Let $\underline{d}_1$ be a direction conjugated to $\underline{d}_0$ with respect to the Hessian matrix of f. What happens if we optimize first along $\underline{d}_1$ and then along $\underline{d}_0$? What happens if we start from a different initial point?

7.4 Conjugate directions method for quadratic functions

Consider the function

$$q(\underline{x}) = \frac{1}{2}\underline{x}^{\mathsf{T}} Q \underline{x} - \underline{b}^{\mathsf{T}} \underline{x}$$

where Q is symmetric and positive definite. Let $\underline{x}_0 \in \mathbb{R}^n$ be an initial point and $\{\underline{x}_k\}$ a sequence of points generated by the conjugate gradient method with respect to Q-conjugated directions $\underline{d}_0, \ldots, \underline{d}_{n-1}$. Let α_i be the step size computed by the exact line search at i-th iteration. Show that for each $k = 1, \ldots, n$ the point

$$\underline{x}_k = \underline{x}_0 + \sum_{i=0}^{k-1} \alpha_i \underline{d}_i$$

is the global minimum of $q(\underline{x})$ on the affine subspace $V_k = \{\underline{x} \in \mathbb{R}^n : \underline{x} = \underline{x}_0 + \underline{v}, \underline{v} \in \mathrm{span}\{\underline{d}_0, \ldots, \underline{d}_{k-1}\}\}$. In particular, $\underline{x}_n$ is the global minimum of $q(\underline{x})$ on $\mathbb{R}^n$.

Hint. Since $q(\underline{x})$ is convex, it is sufficient to verify that $\nabla q(\underline{x}_k)$ is orthogonal to V_k, due to the necessary and sufficient optimality conditions for convex problems.

8.1 Advertising campaign

An advertising agency must propose an advertising campaign using two media: radio and newspapers. Several (m) radio stations can be considered; the cost of having a one minute commercial on one of the radio stations depends on the total number of minutes purchased: the cost is 1000 euros per minute, but it is reduced by 20 euros for each minute purchased up to 25 minutes (e.g. if 3 minutes are purchased, the per minute cost is 940 euros). If more than 25 minutes of commercials are purchased, the per minute cost is constant and equal to 500 euros. Each radio station can broadcast at most 3000 minutes of commercials. Suppose that n newspapers are available. Each half page of newspaper advertising costs 600 euros. A fixed cost of 300 euros must be paid to a newspaper if it is used in the advertising campaign, independently of the number of pages purchased.

Based on statistics, a one minute commercial on a radio is expected to reach 15000 people, while a half page on a newspaper is expected to reach 18000 people. The whole campaign must reach at least 4 million people. Further, due to contractual agreements, at least $\frac{1}{3}$ of the budget must be spent for radio commercials.

Give a mathematical programming model of the problem of planning the minimum cost campaign.

8.2 Optimality conditions (constraint qualification and KKT)

Consider the following problem:

$$
\begin{aligned}
\min \quad & x_1^3 + x_2^2 && (1)\\
\text{s.t.} \quad & x_1^2 + x_2^2 = 9 && (2)\\
& x_2 \geq 0. && (3)
\end{aligned}
$$

a) Draw the feasible region.

b) Find all the feasible points where the constraint qualification condition holds.

c) Find all the solutions of the KKT system.

d) Are the KKT conditions necessary to guarantee local optimality? Are they sufficient?

e) Find an optimal solution.

8.3 Optimality conditions 2

Consider the following quadratic function

$$f(x_1, x_2) = -x_1^2 + x_2^2 + 2x_1$$

and the feasible region $X = \{\underline{x} \in \mathbb{R}^2 \,:\, x_1 + x_2 = 2,\ x_2^2 - x_1 \leq 0\}$.

a) Draw the feasible region.

b) Find all the feasible points where the constraint qualification condition holds. Justify the answer.

c) State the optimality conditions, and say if they are necessary and/or sufficient.

d) Find the maximizers and minimizers of f on X.

e) Write the Lagrangian dual of the problem of minimizing f on X, and show the connection between primal and dual problem. Motivate the answer.

f) Is it possible to simplify the problem by exploiting the equality constraint so as to remove one variable?

8.4 Optimality conditions 3

Consider the following problem:

$$
\begin{array}{lll}
\max & -8x_1^2 - 10x_2^2 + 12x_1x_2 - 50x_1 + 80x_2 & \\
\text{s.t.} & x_1 + x_2 \leq 1 & (1) \\
 & 8x_1^2 + x_2^2 \leq 2 & (2) \\
 & x_2 \geq 0 & (3) \\
 & x_1 \geq 0 & (4)
\end{array}
$$

a) Draw the feasible region.

b) Is the problem convex? Are the constraint qualification conditions satisfied at each point?

c) Write the KKT conditions and say if they are sufficient.

8.5 Optimality conditions 4

Consider the following problem:

$$
\begin{array}{l}
\min x_1^2 + 2x_2^2 + 2x_1 \\
x_1 + 2x_2 \geq 5 \\
x_1, x_2 \in \mathbb{R}.
\end{array}
$$

a) Draw the feasible region.

b) All the feasible points satisfy the constraint qualification condition. Why?

c) Are the KKT conditions sufficient for guaranteeing optimality for this problem? Find the points satisfying the KKT conditions and find the optimal solution.

9.1 Portfolio optimization

Given a set of n assets and a budget B, consider the problem of constructing a portfolio by selecting a subset of assets and the fraction of B to invest in each asset. For each i in $\{1,\dots,n\}$, the rate of return on investment i is modeled as a random variable r_i with known expected value $\bar{r}_i$ and covariance matrix $Q \in \mathbb{R}^{n\times n}$.

Two conflicting goals are usually considered in these problems: maximizing the expected return and minimizing the risk. Harry Markowitz, who won the Nobel Prize in Economics in 1990, in 1952 proposed one way to cope with the two goals, that is minimizing the risk and imposing a constraint on the expected return.

a) How can the risk and the expected return be measured from a statistical point of view?

b) Give a nonlinear programming formulation for the above portfolio optimization problem. Which kind of problem is it and which constrained optimization method can be used to solve it?

c) Write the Karush-Kuhn-Tucker (KKT) conditions. Are they sufficient?

d) Suppose that a minimum amount of money d_i is required for each asset i, with $i = 1,\dots,n$, and that at most k assets can be selected, with $k < n$, so as to limit fragmentation. Give the modified mathematical programming formulation. Which kind of problem is it?

9.2 Active set methods for quadratic programming

Solve the following quadratic programming problem

$$\begin{array}{ll} \min & 125 - 10x_1 - 20x_2 + x_1^2 - 2x_1x_2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \le 6 \\ & 0 \le x_1 \le 3 \\ & 0 \le x_2 \le 5 \end{array}$$

using the active set method and starting from the feasible solution $(2.5, 0)$.

9.3 Quadratic penalty and augmented Lagrangian methods

Consider the nonlinear optimization problem:

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 = 1. \end{array}$$

a) Write the KKT conditions and find an optimal solution.

b) Write the objective function of the quadratic penalty problem. Study its condition number as $\mu \to 0$, computing the Hessian matrix and its eigenvalues.

c) Apply three iterations of the quadratic penalty method starting with $\mu = 0.5$. Update the penalty parameter as follows: $\mu_{k+1} = 0.5\mu_k$.

d) Apply three iterations of the augmented Lagrangian method starting from $\mu_0 = 1$ and $v = 0$, where v is the multiplier associated with the only constraint of the problem. Update the penalty parameter as follows: $\mu_{k+1} = 0.5\mu_k$.

9.4 Quadratic penalty and augmented Lagrangian methods

Consider the nonlinear optimization problem:

$$\begin{aligned} \min \quad & x_1 x_2 \\ \text{s.t.} \quad & x_1 - 2x_2 = 3. \end{aligned}$$

a) Solve the problem analytically: exploit the equality constraint to obtain an unconstrained optimization problem. Derive an optimal solution.

b) Write the optimality conditions for the quadratic penalty problem depending on the penalty parameter μ. Are they sufficient?

c) Apply three iterations of the augmented Lagrangian method starting with $\mu_0 = 1$ and $v = 0$, where v is the multiplier associated with the only constraint of the problem. Update the penalty parameter as follows: $\mu_{k+1} = 0.5\mu_k$.

9.5 Sequential quadratic programming

Consider the following nonlinear programming problem:

$$\begin{aligned} \min \quad & x_1 - x_2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 \leq 1. \end{aligned}$$

Apply two iterations of the sequential quadratic programming method starting from the feasible solution $(-1, 0)$. The initial value of the Lagrange multiplier is $u = 1$.