Exercise 1 – Linear vs. non-linear duality. Consider the optimization problem (in $\mathbb{R}$)

$$\begin{cases} \max & \varphi(x) = x \\ & x \leq 0,\ x \in \{-2, 1\}. \end{cases}$$

a) Write the dual problem associated to relaxing the contraint $x \leq 0$. Show that the duality gap is 2.

b) Solve the convexified problem (with $x \in [-2, 1]$). Show that the convexified optimal value is equal to the optimal dual value.

c) Redo the two above questions with $\varphi(x) = -x^2$. Do we get the same final equality ?

Exercise 2 – Dualize other contraintes. With course notation, we consider

$$\begin{cases} \max & \varphi(x) \\ & x \in X \\ & c(x) \in B \end{cases}$$

where B is a subset of $\mathbb{R}^n$. We assume that we have an oracle solving $\theta(u) := \max_{x \in X} \varphi(x) - u^\top c(x)$.

a) Adding a slack variable, write the dual problem.

b) Apply the result to $B = \{0\}$, $B = \mathbb{R}^n_+$ and B the ℓ_2-ball of radius ε.

Exercise 3 – Relax or strengthen ?. Again with notation of the course, two independent questions.

a) Consider the problem

$$\begin{cases} \max & \varphi(x) \\ & c_1(x) = 0, c_2(x) = 0,\ x \in X. \end{cases}$$

Show that the duality gap for the relaxation of the two constraints $c_1(x) = 0$ and $c_2(x) = 0$ is larger (or equal) than the relaxation of $c_1(x) = 0$ only. What is the drawback of relaxing only $c_1(x) = 0$?

b) Assume that any $x \in X$, satisfying a constraint $c_1(x) = 0$, also satisfies an other constraint $c_2(x) = 0$. This assumption yields that the two following problems are equivalent :

$$(\mathrm{P}_1) \begin{cases} \max & \varphi(x) \\ & c_1(x) = 0,\ x \in X \end{cases} \qquad \text{et} \qquad (\mathrm{P}_{1,2}) \begin{cases} \max & \varphi(x) \\ & c_1(x) = 0,\ c_2(x) = 0,\ x \in X \end{cases}$$

Show that the Lagrangian relaxation of $(\mathrm{P}_{1,2})$ gives a gap always smaller (or equal) than the relaxation of (P_1).

Exercise 4 – Pricing for a mixed-integer problem. We consider the optimization problem in $\mathbb{R}^2$

$$F(d) := \begin{cases} \min & 5p_1 + 10p_2 \\ & p_1 + p_2 \geqslant d \\ & p \in \{0, 3\} \times [0, 1] \end{cases} \tag{P_d}$$

a) Find the optimal solution $p(d)$, depending on $d \in [0, 4]$. Draw the graph of F.

b) Write the optimization problem as a max and introducing the Lagrangian

$$L_0(p; u) := -5p_1 - 10p_2 - u(-p_1 - p_2)\,,$$

to dualize (P_0). Compute the optimal solution p^u of maximizing the Lagrangian, depending on $u \geqslant 0$. Draw the graph of the associated dual function $\theta_0(u)$.

c) Form the dual of (P_d), and express the dual function θ_d with the help of θ_0. What is the minimum of θ_d for $d = 2$?

d) Observe graphically that the dual optimal solution is the slope of the convex enveloppe of F.

Exercise 5 – Columns generation = dual cutting-plane. Consider $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and X a closed bounded subset of $\mathbb{R}^n$ (to fixed ideas, say that $X = \{0,1\}^n \cap P$, with P a simple polytope). We focus on the following problem

$$\begin{cases} \max \quad c^\top x\,, \\ \quad x \in X\,, \\ \quad Ax = b \in \mathbb{R}^m\,. \end{cases} \tag{1}$$

Assume that it is easy to maximize linear functions over X; denote formally by `optimX` the operator returning an optimal solution for a given linear objective function as an input.

a) We dualize the constaint $Ax = b$; write the Lagrangian $L(x; u)$, $u \in \mathbb{R}^m$ being the dual variable.

b) Using `optimX`, give an oracle for the dual function producing a linearization, at a given u.

We assume now that, for K dual points $u_1, \ldots, u_K$, we have associated primal points $x_1, \ldots, x_K$ (such that x_i maximize $L(\cdot, u_i)$ over X). We are interested in computing u_{K+1} and x_{K+1}.

c) Write the problem corresponding to an iteration of the cutting-plane algorithm to minimize the dual function. Show that it can be cast as

$$\begin{cases} \min \quad r\,, \\ \quad (u, r) \in \mathbb{R}^m \times \mathbb{R}\,, \\ \quad r \geqslant c^\top x_i - u^\top (Ax_i - b)\,, \quad i = 1, \ldots, K\,. \end{cases} \tag{2}$$

d) Let us dualize (2); rephrase the problem as a max (if it helps you) and write the Lagrangian $\widetilde{L}((u, r); \alpha)$ (with a vector $\alpha \in \mathbb{R}^K_+$ of dual variables).

e) What conditions should satisfy α to have

$$\max_{(u,r) \in \mathbb{R}^m \times \mathbb{R}} \widetilde{L}((u, r); \alpha) < +\infty \quad ?$$

Deduce the associated dual function.

f) Write the associated dual problem. Give a primal interpretation of it.

Exercise 6 – Augmented Lagrangian relaxation. We start this exercice with studying the simple optimization problem in $\mathbb{R}^2$

$$\begin{cases} \max \quad -x_1 - 2x_2 \\ \quad x_1 + x_2 = 3 \\ \quad x_1 \in [0, 2],\ x_2 \in \{0, 2\} \end{cases} \tag{P}$$

a) What is the optimal solution and the optimal value of (P)? Same question for the convexified problem where the constraint $x_2 \in \{0, 2\}$ is replaced by $x_2 \in [0, 2]$.

b) Write the Lagrangian and the dual function θ associated to the dualization of $x_1 + x_2 - 3 = 0$.

c) Draw the graph of θ. Deduce the dual optimal solution and the duality gap.

Let's now turn to the general framework of the course

$$\begin{cases} \max \quad \varphi(x) \\ \quad c(x) = 0,\ x \in X. \end{cases}$$

For a parameter $\rho > 0$, we define the augmented Lagrangian function by

$$L^\rho(x; u) := \varphi(x) - u^\top c(x) - \frac{\rho}{2} \|c(x)\|^2$$

and the associated augmented dual function by

$$\theta^\rho(u) := \max_{x \in X} L^\rho(x; u).$$

d) Show that $\theta^\rho \colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is convex. Show that weak duality holds : for any dual variable u and any primal feasible variable $x \in X$ such that $c(x) = 0$, we have $\theta^\rho(u) \geq \varphi(x)$.

e) Fix $\bar{u}$ and $x(\bar{u}) \in X$ such that $\theta^\rho(\bar{u}) = L^\rho(x(\bar{u}); \bar{u})$. Prove that, if $c(x(\bar{u})) = 0$, then $\bar{u}$ minimizes θ^ρ, $x(\bar{u})$ is a primal optimal solution, and that there is no duality gap.

Augmented Lagrangians have the following nice property. Contrary to *standard* Lagrangian relaxations, *augmented* Lagrangian relaxations always close the duality gap and recover primal solutions (when ρ is large enough). The aim of this exercise is to prove this property for (P) $\rho = 3$.

f) Write the augmented Lagrangian and the augmented dual function θ^3 (for $\rho = 3$) associated to the dualization of $x_1 + x_2 - 3 = 0$ in problem (P). Show that θ^ρ can be cast as

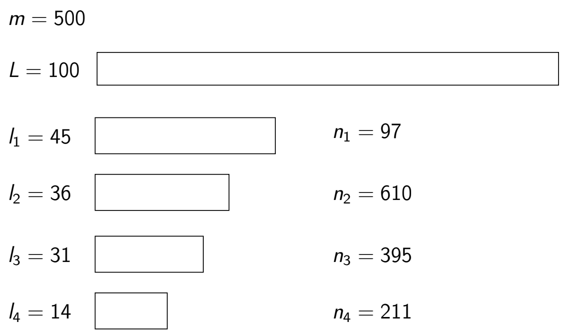
$$\theta^3(u) = \max\{\theta_0^3(u), \theta_2^3(u)\}$$

with two convex functions θ_0^3, θ_2^3.

g) Show that $\bar{u} = -1$ minimizes θ^3 and that there is no duality gap.

h) Thus solving the augmented Lagrangian relaxation allows us to solve the primal problem. But what is the big disadvantage of augmented Lagrangian (versus the usual Lagrangian) ?

Exercise 7 – Lagrangian decomposition for cutting-stock. The problem consists in minimizing the number of stock pieces of width L, used to meet demands $n_1, \ldots, n_I$, for items $i = 1, \ldots, I$, to be cut at their width $l_1, \ldots, l_I$. We assume that every l_j is smaller than L and that there are enough stock pieces, say m, available for a feasible cutting. We denote by $n \in \mathbb{R}^I$ (respectively $l \in \mathbb{R}^I$) the vector of entries n_i (resp. l_i) for all i. In the example drawn here : we have $m = 500$ pieces of width $L = 100$ where to cut $I = 4$ types of items ; the demand consists in different numbers of items n_i with different lengths $l_i \leq 100$ for the $I = 4$ types of items.



A possible formulation for the cutting-stock problem is the following integer linear optimization problem :

$$(\mathrm{P}) \quad \begin{cases} \min_{y,z} \quad \sum_{k=1}^m y^k \\ \sum_{k=1}^m z_i^k \geq n_i \qquad \text{for all } i = 1, \ldots, I \\ \sum_{i=1}^I z_i^k \, l_i \leq L \, y_k \qquad \text{for all } k = 1, \ldots, m \\ y^k \in \{0, 1\}, \; z_i^k \in \mathbb{N} \qquad \text{for all } i = 1, \ldots, I\,, \; k = 1, \ldots, m \end{cases}$$

a) Explain the modelling as (P) : what is the role of the variables ? and the meaning of the objective and the constraints ?

b) Let us dualize the I demand-covering constraints $\sum_{k=1}^m z_i^k \geq n_i$. Re-write the above problem as a max with the course's notation : introduce φ, c and X.

c) For a dual variable $u \in (\mathbb{R}_+)^I$, define the Lagrangian function and show that it is decomposable with respect to k.

d) Observe then that the associated dual function, denoted by θ, can be written as the juxtaposition of m identical max problems, that is,

$$\theta(u) = -n^\top u + \sum_{k=1}^{m} v(u) = -n^\top u + m\, v(u)$$

where $v(u)$ is the optimal solution of a max problem to be specified.

e) Show moreover that $v(u)$ can be explicitly written as :

$$v(u) = \begin{cases} 0 & \text{if } u^\top z(u) \le 1 \\ u^\top z(u) - 1 & \text{otherwise} \end{cases}$$

where $z(u)$ is the optimal solution of the following integer knapsack problem, parameterized by u

$$\begin{cases} \min \quad u^\top z \\ \quad l^\top z \le L, \ \ z \in \mathbb{N}^I. \end{cases}$$

f) Discuss the complexity and the practical difficulty of computing $\theta(u)$, compared to solving (P). Explain what would be an "oracle" for θ providing a linearization of the (convex) function θ.

g) Write the dual problem. What algorithm can we use to solve it ?

h) What does the dual optimal value correspond to, in the problem (P) ? How does it compare with the continuous relaxation consisting in relaxing all the integrity constraints of (P) (i.e. $y^k \in [0,1]$, $z_i^k \in [0, M]$ with M an upper bound).

Exercise 8 – Lagrangian relaxations for a warehouses location problem. We consider the following operations research problem. We have stores (indexed by $i \in \{1, \dots, I\}$) that must be regularly supplied by warehouses. These warehouses are not built yet, but we have a set of possible localizations (indexed by $j \in \{1, \dots, J\}$). For any store i, we know the demand d_i (supposed constant). For any possible localization $j \in J$, we know the cost o_j of opening a warehouse at j, the capacity k_j it would have, and the transportation cost a_{ij} of a unit of commodity from j to store i. Note that a store can be supplied by several warehouses.

The problem of opening warehouses and delivering stores at the minimal cost can be written as the following mixed-binary linear optimization problem :

$$(\mathrm{P}) \quad \begin{cases} \min_{x,y} \quad \displaystyle\sum_{i=1}^{I}\sum_{j=1}^{J} a_{ij}x_{ij} + \sum_{j=1}^{J} o_j y_j & \\ \displaystyle\sum_{j=1}^{J} x_{ij} = 1 & \text{for all } i \in \{1,\dots,I\} \qquad (1) \\ \displaystyle\sum_{i=1}^{I} d_i x_{ij} \le k_j y_j & \text{for all } j \in \{1,\dots,J\} \qquad (2) \\ x_{ij} \in [0,1], \ y_j \in \{0,1\} \quad \text{for all } i \in \{1,\dots,I\} \text{ and } j \in \{1,\dots,J\} & \end{cases}$$

a) Give the interpretation of the variables and the constraints of the problem (P).

b) We consider first the relaxation of the constraint (1). Rewrite the problem (up to a change of sign) as

$$\begin{cases} \max_{x,y} \quad \varphi(x,y) \\ \quad c_1(x,y) = 0, \\ \quad (x,y) \in \mathcal{X}_1 \end{cases}$$

to make appear the notation of the course. Give the definition of the Lagrangian function and its associated dual function, denoted θ_1.

c) Show that θ_1 decomposes along warehouses : computing $\theta_1(u)$ for $u \in \mathbb{R}^I$ reduces to solving J linear optimization problems. Thus, computing this bound is polynomial and numerically simple.

d) We now consider the relaxation of the constraint (2). Again write accordingly the problem, the Lagrangian and the dual function, denoted θ_2. Show that θ_2 is separable, first, between the variables (x, y) and, second, along the stores ($i \in \{1, \dots, I\}$).

e) Show that computing the value $\theta_2(v)$ for $v \in (\mathbb{R}_+)^J$ reduces to solving I trivial linear optimization problems which can be solved by sorting I vectors in $\mathbb{R}^J$. Thus, computing this bound is also polynomial and numerically very simple.

f) What algorithm could we use to compute (approximations of) the optimal dual values θ_1^* (the best of the dual bounds $\theta_1(u)$) and θ_2^* (the best of the dual bounds $\theta_2(v)$) ? Could you describe it *shortly*? How to update the inner subproblem of this algorithm from an iteration to the next one ?

g) Show that the optimal dual value θ_2^* coincide with the optimal value of the linear problem ($\bar{\text{P}}$) which corresponds to (P) where $y_j \in \{0, 1\}$ is replaced by $y_j \in [0, 1]$.

h) Explain why θ_1^* coincide with the optimal value of

$$\begin{cases} \max_{x,y} \quad \varphi(x, y) \\ \quad c_1(x, y) = 0, \\ \quad (x, y) \in \text{conv}\mathcal{X}_1. \end{cases}$$

Deduce that the duality gap for the relaxation of (1) is smaller (or equal) than the one for the relaxation (2).

i) Conclude about this approach by Lagrangian relaxation for this problem : what relaxation would you choose ? what would you do next to solve (P) to optimality ?

j) Bonus : in the special case $I = J = 1$ with $d < k$. Give the simplified expressions of θ_1 and θ_2 (we get rid of indexes of variables). Show that the dual optimal values are respectively $\theta_1^* = -a - o$ and $\theta_2^* = -a - o\, d/k$, so that the gap for (1) is strictly better.

Exercise 9 – Decomposition of network design problem. This exercise is about a classical problem in networks dealing with opening arcs in a graph to transfer commodities.

Consider a network formalized as a complete undirected graph (N, A) (with $n = |N|$ nodes). We have K commodities to transport over this network : for commodity k, we have a quantity W_k to transport from an origin node $o(k) \in N$ to a destination node $d(k) \in N$. We want to open arcs in the graph and to make the transfer of the commodities, at an optimal cost. The costs are the following :

- f_{ij} is the cost of opening the arc (ij),
- c_{ijk} is the nominal cost of transferring commodity k through (ij).

a) Show that the above-described problem can be cast as a mixed-binary linear problem with two types of variables : $y_{ij} \in \{0, 1\}$ and $x_{ijk} \in [0, W_k]$.

b) Show that the linear problem can be reformulated as

$$(\text{P}) \quad \min_{y \in \{0,1\}^{n \times n}} \quad \sum_{i,j=1}^{n} f_{ij} y_{ij} + \sum_{k=1}^{K} v_k(y)$$

where $v_k(y)$ are the optimal values of K independent sub-problems. What is the nature of these K sub-problems ? Are they easy problems ?

c) Justify why v_k is convex. Explain what is an oracle of it : explain what it would produce and what computation it would do.

d) Assume we are at iterate ℓ of a regularized Benders algorithm solving this problem. Write the optimization problem giving the next iterate $y_{\ell+1}$. What is the nature of this problem ? Is this an easy problem ? What are the modifications from this problem to the one at the next iteration ?

Exercise 10 – Benders decomposition for a localisation problem. This exercise is about a classical problem of localization in networks : we would like to position "hubs" that collect commodities from their origins, transporte them, and distribute them to their final destinations.

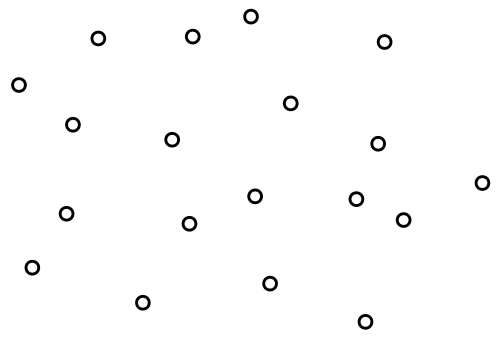
Consider a network formalized as a complete graph (N, A), where the distance between node i and node j is denoted d_{ij}. Here are the features of the problem, illustrated by side figures. We have $n = |N|$ nodes, and every node i could be a hub. We have K commodities to distribute over this network : for commodity k, we have a quantity W_k to transport from an origin node $o(k) \in N$ to a destination node $d(k) \in N$. Denote by χ, τ, δ the nominal costs of collect, transfer and distribution, respectively ; thus

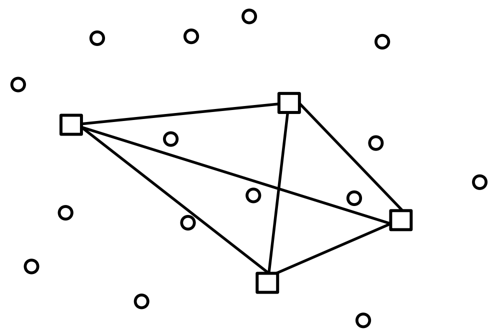
— $\chi d_{o(k)i}$ is the cost of collecting from the node $o(k)$ to a hub in node i,

— τd_{ij} is the cost of transferring from a hub in i et a hub in j,

— $\delta d_{jd(k)}$ is the cost of distributing from a hub in j toward the node $o(k)$.

Finally, denote by $f = (f_1, \ldots, f_n) \in \mathbb{R}^n$ where f_i is the cost of opening un hub at node i.

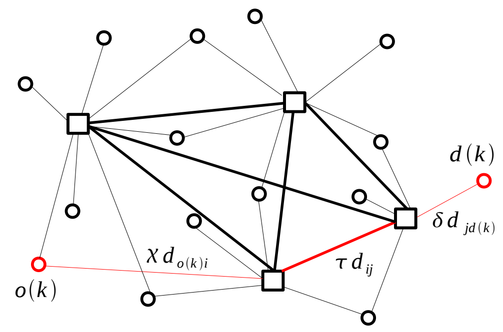




a) Show that the above-described problem can be cast as a mixed-integer linear problem with two types of variables :

— localization variables : $y_i \in \{0, 1\}$ which is equal to 1 if and only if a hub is localized at node i,

— path variables : $x_{ijk} \in [0, 1]$ which represents the fraction of the quantity of commodity k that go through the arc between i and j.



b) Show that the problem can be reformulated as

$$(\mathrm{P}) \quad \min_{y \in \{0,1\}^n} f^\top y + \sum_{k=1}^{K} v_k(y)$$

where $v_k(y)$ are the optimal values of K independent sub-problems. What is the nature of these K sous-problems ? Are they easy problems ?

c) Add to (P) a constraint expressing that there exists at least on hub in the network. Show that this constraint yields that we have no need of feasibility cuts.

Assume the framework of the course : the functions v_k are convex and we can compute a linearization at any y. We can then solve this problem by the regularized Benders algorithm.

d) Assume we are at iterate ℓ of the regularized Benders algorithm. Write the problem given the next iterate $y_{\ell+1}$. What is the nature of this problem. Is this an easy problem ? What are the modifications from this problem to the one at the next iteration ?