



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



INFORMATION ENGINEERING DEPARTMENT

COMPUTER ENGINEERING - AI & ROBOTICS

OPERATIONS RESEARCH 1 2022/2023

Francesco Crisci 2076739

Convex Programming

Definition 1.1.1

Given $x, y \in R^n$, the point $z := \lambda x + (1 - \lambda)y$ is said to be a convex combination of x, y for all $\lambda \in [0, 1]$. The combination is said to be strict if $0 < \lambda < 1$.

More generally, the convex combination of k points $x^1, \dots, x^k \in R^n$ is defined as $\sum_{i=1}^k \lambda_i x^i$, with $\lambda_i, \dots, \lambda_k \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$.

Definition 1.1.2

A set $X \subseteq R^n$ is said to be convex if $\forall x, y \in X$ we have that X contains all the convex combinations of x and y , i.e. $z := [\lambda x + (1 - \lambda)y], \forall \lambda \in [0, 1]$.

Proposition 1.1.1

The intersection of two convex set $A, B \in R^n$ is still a convex set.

Proof:

Given $x, y \in A \cap B$, for all $\lambda \in [0, 1]$ we have $z := \lambda x + (1 - \lambda)y \in A$ by convexity of A , $z \in B$ for convexity of B , hence $z \in A \cap B$, as requested.

Definition 1.1.3

A function $f : X \rightarrow R$ defined on a convex set $X \subseteq R^n$ is said to be convex if $\forall x, y \in X$ and $\forall \lambda \in [0, 1]$ we have that $f(z) \leq \lambda f(x) + (1 - \lambda)f(y)$, where $z = \lambda x + (1 - \lambda)y$.

Theorem 1.1.1

Let $X = \{x \in R^n : g_i(x) \leq 0, i = 1, \dots, m\}$. If for all $i \in \{1, \dots, m\}$ the functions $g_i : R^n \rightarrow R$ are convex, then the set X is convex.

Proof:

Clearly $X = \cap_{i=1}^m X_i$, where $X_i := \{x \in R^n : g_i(x) \leq 0\}$. By proposition 1.1.1, it is then sufficient to prove that each set X_i is convex. Indeed, given any two elements x and y of X_i and a generic point $z = \lambda x + (1 - \lambda)y, \lambda \in [0, 1]$, by the convexity hypothesis of the function g_i we can write $g_i(z) = g_i(\lambda x + (1 - \lambda)y) \leq \lambda g_i(x) + (1 - \lambda)g_i(y) \leq 0$, where the latter inequality is valid since $g_i(x) \leq 0, g_i(y) \leq 0$, and $0 \leq \lambda \leq 1$. It follows that $g_i(z) \leq 0$, hence $z \in X_i$. Given the arbitrariness of x, y and z , one thus has that X_i is convex, as requested.

Theorem 1.1.2

Consider a convex programming problem, i.e., a problem $\min\{f(x) : x \in X\}$ where $X \in R^n$ is a convex set and $f : X \rightarrow R$ is a convex function. Every locally optimal solution is also a globally optimal solution.

Proof:

Let \tilde{x} be any locally optimal solution. By the local optimum definition, there exists then $\epsilon > 0$ such that $f(\tilde{x}) \leq f(z)$

for all $z \in I_\epsilon(\tilde{x}) := \{x \in X : |x - \tilde{x}| \leq \epsilon\}$. We have to prove that $f(\tilde{x}) \leq f(y)$ for all $y \in X$. Given any $y \in X$, consider the point z belonging to the segment that connects \tilde{x} to y and define as $z := \lambda\tilde{x} + (1 - \lambda)y$, where $\lambda < 1$ is chose very close to the value 1 so that $z \in I_\epsilon(\tilde{x})$ and hence $f(\tilde{x}) \leq f(z)$. By the convexity hypothesis of f it follows that $f(\tilde{x}) \leq f(z) = f(\lambda\tilde{x} + (1 - \lambda)y) \leq \lambda f(\tilde{x}) + (1 - \lambda)f(y)$, from which, dividing by $(1 - \lambda) > 0$, we obtain $f(\tilde{x}) \leq f(y)$, as requested.

Simplex Algorithm

Definition 4.1.1

The sets $\{x \in R^n : \alpha^T x \leq \alpha_0\}$ and $\{x \in R^n : \alpha^T x = \alpha_0\}$ are called affine half-space and hyperplane, respectively, induced by (α, α_0) .

Definition 4.1.2

A (convex) polyhedron is defined as the intersection of a finite number of affine half-spaces and hyperplanes. The sets of feasible solutions of Linear Programming problems are hence polyhedra.

Definition 4.1.3

A bounded (i.e. there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in P$) polyhedron P is called polytope.

Definition 4.1.4

A point x of a polyhedron P is said to be an extreme point or a vertex of P if it cannot be expressed as a strict convex combination of other two points of the polyhedron, i.e, if there exist no $y, z \in P, y \neq z$ and $\lambda \in (0, 1)$ such that $x = \lambda y + (1 - \lambda)z$.

Theorem 4.1.1 Minkowski-Weyl theorem

Every point of a polytope can be obtained as the convex combination of its vertices.

Theorem 4.1.2

If the set P of the feasible solutions of the linear programming problem $\min\{c^T x : x \in P\}$ is bounded, then there exists at least one optimal vertex of P .

Proof:

Let x^1, \dots, x^k be the vertices of P and $z^* := \min\{c^T x^i : i = 1, \dots, k\}$. Given any $y \in P$, we need to prove that $c^T y \geq z^*$. Indeed, $y \in P$ implies the existence of multipliers $\lambda_1, \dots, \lambda_k \geq 0, \sum_{i=1}^k \lambda_i = 1$, such that $y = \sum_{i=1}^k \lambda_i x^i$. Hence we have $c^T y = c^T \sum_{i=1}^k \lambda_i x^i = \sum_{i=1}^k \lambda_i (c^T x^i) \geq \sum_{i=1}^k \lambda_i z^* = z^*$.

Definition 4.1.5

A collection of m linearly independent columns of A is said to be a basis of A . The x_j variables associated with the basic columns are called basic variables; the remaining variables are called non-basic variables.

Definition 4.1.6

The solution obtained imposing $x_F = 0$ and $x_B = B^{-1}b$ is said to be the basic solution associated with basis B. The basic solution (and by extension, basis B itself) is said to be feasible if $x_B = B^{-1}b \geq 0$.

Definition 4.1.7

A basis B is said to be degenerate if $B^{-1}b$ has one or more zero components.

Theorem 4.1.3

A point $x \in P$ is a vertex of the not empty polyhedron $P := \{x \geq 0 : Ax = b\}$ if and only if x is a basic feasible solution of the system $Ax = b$.

Proof:

Let us prove the implication x is a basic feasible solution then x is a vertex. Let $x = [x_1, \dots, x_k, 0, \dots, 0]^T$ be any basic feasible solution associated with basis B of A, where $k \geq 0$ is the number of non-zero components of x. It follows that columns A_1, \dots, A_k must be part of B, possibly together with other columns. Let us assume by contradiction that x is not a vertex. There exist thus $y = [y_1, \dots, y_k, 0, \dots, 0]^T \in P$ and $z = [z_1, \dots, z_k, 0, \dots, 0]^T \in P$ with $y \neq z$ such that $x = \lambda y + (1 - \lambda)z$ for any $\lambda \in (0, 1)$, which implies that $k \geq 1$. Note that both y and z must have the last components set to zero, otherwise their convex combination cannot give x. For the hypothesis, we then have $y \in P \Rightarrow Ay = b \Rightarrow A_1 y_1 + \dots + A_k y_k = b$ and $z \in P \Rightarrow Az = b \Rightarrow A_1 z_1 + \dots + A_k z_k = b$. By subtracting the second equation from the first we obtain $(y_1 - z_1)A_1 + \dots + (y_k - z_k)A_k = \alpha_1 A_1 + \dots + \alpha_k A_k = 0$, where $\alpha_i = y_i - z_i, i = 1, \dots, k$. Hence there exist $\alpha_1, \dots, \alpha_k$ scalars not all zero such that $\sum_{i=1}^k \alpha_i A_i = 0$, thus columns A_1, \dots, A_k are linearly dependent and cannot be part of the basis B (\Rightarrow contradiction).

We will now prove the implication x is a vertex then x is a basic feasible solution; the fact that the basic solution is also feasible obviously derives from the hypothesis that $x \in P$. Writing, as before, $x = [x_1, \dots, x_k, 0, \dots, 0]^T$ with $x_1, \dots, x_k > 0$ and $k \geq 0$, we have that $x \in P \Rightarrow Ax = b \Rightarrow A_1 x_1 + \dots + A_k x_k = b$. From here two cases can occur:

1. Columns A_1, \dots, A_k are linearly dependent (or $k = 0$): by arbitrarily selecting other $m-k$ linearly independent columns (which, as is well known, is always possible), we obtain basis $B = [A_1, \dots, A_k, \dots]$ whose basic associated solution is indeed x (which satisfies $Ax=b$ and has non-basic components all equal to zero), thus concluding the proof.
2. columns A_1, \dots, A_k are linearly dependent: we will prove that this case cannot actually happen. Indeed, if the columns were linearly dependent, then there would exist $\alpha_1, \dots, \alpha_k$ not all zero such that $\alpha_1 A_1 + \dots + \alpha_k A_k = 0$. With some math manipulation we can obtain, with $\epsilon > 0$, the following $(x_1 + \epsilon \alpha_1)A_1 + \dots + (x_k + \epsilon \alpha_k)A_k = b$ and $(x_1 - \epsilon \alpha_1)A_1 + \dots + (x_k - \epsilon \alpha_k)A_k = b$. By defining $y := [x_1 - \epsilon \alpha_1, \dots, x_k - \epsilon \alpha_k, 0, \dots, 0]^T$ and $z := [x_1 + \epsilon \alpha_1, \dots, x_k + \epsilon \alpha_k, 0, \dots, 0]^T$, we would have $Ay=b$ and $Az=b$, while choosing a sufficient small ϵ we would have $y, z \geq 0$ and thus $y, z \in P, y \neq z$. But since by construction we have $x = \frac{1}{2}y + \frac{1}{2}z$, thus would mean that vertex x can be expressed as the strict convex combination of two distinct points of P (\Rightarrow contradiction).

Corollary 4.1.1

Every problem $\min\{c^T x : Ax = b, x \geq 0\}$ defined on a polytope $P = \{x \geq 0 : Ax = b\} \neq \emptyset$ has at least one optimal solution coinciding with a basic feasible solution.

Proof:

According to theorem 4.1.2, there always exists an optimal solution coinciding with a vertex of P and thus, according to theorem 4.1.3, with a basic feasible solution.

Definition 4.2.1

The vector $\bar{c}^T := c^T - c_B^T B^{-1}A = [c_B^T - c_B^T B^{-1}B = 0, c_F^T - c_B^T B^{-1}F]$ is called the reduced cost vector with respect to basis B.

Theorem 4.2.1

Let B be a feasible basis. If $\bar{c}^T := c^T - c_B^T B^{-1} A \geq 0^T$, then the basic solution associated with B is optimal.

Proof:

Rewriting the objective function as $c^T x = c_B^T B^{-1} b + \bar{c}^T x$, in the hypothesis $\bar{c}^T \geq 0^T$ we have that $c^T x \geq c_B^T B^{-1} b$ for all $x \geq 0$ (and thus for all $x \in P$), where $c_B^T B^{-1} b = [c_B^T, c_F^T] \begin{bmatrix} B^{-1} b \\ 0 \end{bmatrix}$ is the value of the objective function corresponding to the basic feasible solution associated with B .

Definition 4.2.2

Given basis B and $z = c^T x$, system $\begin{cases} x_B = B^{-1} b - B^{-1} F x_F \\ z = c_B^T B^{-1} b + (c_F^T - c_B^T B^{-1} F) x_F \end{cases}$ is said to be in canonical form with respect to B .

Bland's rule

Whenever it is possible to choose, always choose the entering/leaving variable x_j with the smallest j . In particular, we have to:

- choose variable x_h to enter the basis defining $h := \operatorname{argmin}\{j : \bar{c}_j < 0\}$;
- among all rows t with $\bar{b}_t / \bar{a}_{th} = \theta$ that are eligible for the pivot operation, choose the one with minimum $\beta[t]$ so as to force the smallest-index variable $x_{\beta[t]}$ to leave the basis.

Theorem 4.2.2

Using Bland's rule, the simplex algorithm converges after, at most, $\binom{n}{m}$ iterations.

Proof:

Let us suppose by contradiction that the thesis is false and let us consider as counterexample the smallest LP problem for which there is no convergence. As already seen, in this case the simplex algorithm has to go through a cyclic sequence $B_1, B_2, \dots, B_k = B_1$ of bases. During this sequence, pivot operations are performed on all rows and all columns of the tableau, otherwise eliminating the not involved rows/columns we would obtain a smaller counterexample. It follows that all variables enter and leave, in turn, the current basis. Moreover, we must have $\bar{b}_t = 0$ for all rows $t \in \{1, \dots, m\}$, otherwise in the iteration in which the pivot operation is performed on row t , we would have $\theta > 0$, hence the value of the objective function would change- preventing cycling in the sequence of bases.

Consider now tableau T in which variable x_n leaves the current basis to let a given non-basic variable x_h enter the basis. Let us indicate with $x_{\beta[i]}$ the basic variable in row $i \in \{1, \dots, m\}$, and with t the row in which x_n is in the basis. The main features of this tableau are:

- $\bar{b}_i = 0$
- $\bar{c}_h < 0$ since x_h enters the basis
- $\bar{c}_{\beta[i]} = 0 \forall i$, the reduced cost of the basic variables
- $\bar{a}_{th} > 0$, the pivot element
- $\bar{a}_{ih} \leq 0 \forall i \neq t$, due to Bland's rule

Note that if there existed $i \neq t$ with $\bar{a}_{ih} > 0$, then Bland's rule would have certainly have preferred to let variable $x_{\beta[i]}$ leave the basis. Let us now consider tableau \tilde{T} in which x_n re-enters the basis; indicating with \tilde{c}_j the reduced costs in row 0 of \tilde{T} , we must have:

- $\tilde{c}_n < 0$ since x_n enters the basis

- $\tilde{c}_j \geq 0 \forall j \neq n$ due to Bland's rule.

Now, \tilde{T} has been obtained from T by means of a sequence of pivot operations, hence there exist appropriate multipliers μ_1, \dots, μ_m such that $[\text{row } 0 \text{ of } \tilde{T}] = [\text{row } 0 \text{ of } T] + \sum_{i=1}^m \mu_i [\text{row } i \text{ of } T]$, but then:

- $\tilde{c}_{\beta[t]} = \tilde{c}_n < 0, \bar{c}_{\beta[t]} = 0 \Rightarrow \tilde{c}_{\beta[t]} = \bar{c}_{\beta[t]} + \mu_t \Rightarrow \mu_t < 0$
- $\tilde{c}_{\beta[i]} \geq 0, \bar{c}_{\beta[i]} = 0 \Rightarrow \tilde{c}_{\beta[i]} = \bar{c}_{\beta[i]} + \mu_i \Rightarrow \mu_i \geq 0 \forall i \neq t$

hence there is a contradiction: $\underbrace{\tilde{c}_h}_{\geq 0} = \underbrace{\bar{c}_h}_{< 0} + \sum_{i \neq t} \underbrace{\bar{a}_{ih}}_{\leq 0} \underbrace{\mu_i}_{\geq 0} + \underbrace{\bar{a}_{th}}_{> 0} \underbrace{\mu_t}_{< 0} < 0$.

Duality in LP

Definition 5.1.1

Given a set $X \subseteq R^n$, an inequality $c^T x \geq c_0$ satisfied by all $x \in X$ is said to be valid for X .

Theorem 5.1.1 Farkas' Lemma

The inequality $c^T x \geq c_0$ is valid for the non-empty polyhedron $P := \{x \geq 0 : Ax = b\}$ if and only if $\mu \in R^m$ exists such that $c^T \geq \mu^T A$ and $c_0 \leq \mu^T b$.

Proof:

As already seen, the fact that the condition is sufficient is trivially true, given that for all $x \geq 0$ such that $Ax=b$ we have: $c^T x \geq \mu^T Ax = \mu^T b \geq c_0$. We will now prove that the condition is also necessary, i.e, that $c^T x \geq c_0$ valid for $P \neq \emptyset \Rightarrow \exists \mu \in R^m : c^T \geq \mu^T A, c_0 \leq \mu^T b$. By the hypothesis validity, we have that $c_0 \leq z^* := \min\{c^T x : Ax = b, x \geq 0\}$, which includes $z^* = -\infty$. Let then x^* be an optimal feasible solution found by the simplex algorithm applied to the aforementioned problem. This solution exists by the convergence property of the simplex algorithm. In addition, let B be an optimal basis associated with x^* , and let as usual $A = [B, F]$, $c^T = [c_B^T, c_F^T]$ and $x^* = (x_B^*, x_F^*)$ with $x_B^* = B^{-1}b$ and $x_F^* = 0$. We will prove that vector $u^T := c_B^T B^{-1}$ verifies the two conditions, hence that the thesis is valid. Recalling the reduced cost expression computed in correspondence of the optimal basis B we have: $\bar{c}^T := c^T - \underbrace{c_B^T B^{-1} A}_{u^T} \geq 0^T \Rightarrow c^T \geq u^T A$, and thus it satisfies the first condition. In addition,

for the optimization problem we have that: $c_0 \leq z^* = c^T x^* = c_B^T x_B^* + c_F^T x_F^* = c_B^T B^{-1}b = u^T b$, and thus also the second condition is verified.

Dual Problem

The considerations above show that, if the problem $\min\{c^T x : Ax = b, x \geq 0\}$ has an optimal finite value, then:

$$\begin{aligned} \min_x \{c^T x : Ax = b, x \geq 0\} &= \max_{c_0} \{c_0 : c^T x \geq c_0 \text{ valid for } P\} \\ &= \max_{c_0, u} \{c_0 : c_0 \leq u^T b, c^T \geq u^T A\} \text{ by Farkas' lemma} \\ &= \max_u \{u^T b : c^T \geq u^T A\} \text{ since, at the optimal solution } c_0 = u^T b. \end{aligned}$$

Proposition 5.3.1

The dual of the dual problem coincides with the primal problem.

Theorem 5.3.1 Strong Duality

Let $P := \{x \geq 0 : Ax \leq b\} \neq \emptyset$ with $\min\{c^T x : x \in P\}$ finite. Then $\min\{c^T x : Ax \leq b, x \geq 0\} = \max\{u^T b : c^T \geq u^T A, u \geq 0\}$

Proof:

The proof is the result of the considerations mentioned above.

Theorem 5.3.2 Weak Duality

Let $P := \{x \geq 0 : Ax \geq b\} \neq \emptyset$ and $D := \{u \geq 0 : c^T \geq u^T A\} \neq \emptyset$. For all pairs of points $\bar{x} \in P$ and $\bar{u} \in D$ we have that $\bar{u}^T b \leq c^T \bar{x}$.

Proof:

Given $\bar{x} \in P$ and $\bar{u} \in D$, we have $A\bar{x} \geq b, \bar{x} \geq 0, \bar{u} \geq 0$ and $c^T \geq \bar{u}^T A$, from which $\bar{u}^T b \leq \bar{u}^T A\bar{x} \leq c^T \bar{x}$.

Corollary 5.3.1

Consider the pair of primal $\min\{c^T x : Ax \geq b, x \geq 0\}$ and dual $\max\{u^T b : c^T \geq u^T A, u \geq 0\}$ problems. There are only 4 cases:

1. Both problems have a finite optimum, with $\min\{c^T x : Ax \geq b, x \geq 0\} = \max\{u^T b : c^T \geq u^T A, u \geq 0\}$;
2. The primal problem is unbounded and the dual is unfeasible;
3. The dual problem is unbounded and the primal is unfeasible;
4. Both problem are unfeasible

Proof:

Case 1 derives from the strong duality theorem, while 2 and 3 from the weak duality, given that the existence of a feasible solution in one of the two problems prevents the other from being unbounded. Case 4 is not excluded from any property and indeed occurs.

Theorem 5.5.1 Optimality Conditions

Two vectors $\bar{x} \in R^n$ and $\bar{u} \in R^m$ are optimal for problems $\min\{c^T x : Ax \geq b, x \geq 0\}$ and $\max\{u^T b : c^T \geq u^T A, u \geq 0\}$, respectively, if and only if the following optimality conditions hold:

- $A\bar{x} \geq b, \bar{x} \geq 0$, primal feasibility
- $c^T \geq \bar{u}^T A, \bar{u} \geq 0$, dual feasibility
- $\bar{u}^T (A\bar{x} - b) = 0$
- $(c^T - \bar{u}^T A)\bar{x} = 0$

The last two conditions are the complementary slackness conditions.

Bland's Rule(Dual)

Whenever it is possible to choose, always choose the entering/leaving variable x_j with the smallest index j

- Choose variable $x_{\beta[t]}$ leaving the basis with minimum index $\beta[t]$, i.e., the pivot row t with $\bar{b}_t < 0$ and minimum $\beta[t]$.
- among all variables x_h with $\bar{a}_{th} < 0$ and $\bar{c}_h/|\bar{a}_{th}| = \min\{\bar{c}_j/|\bar{a}_{tj}| : \bar{a}_{tj} < 0\}$ that are eligible for entering the basis, choose the one with minimum index h .

Integer Linear Programming

Definition 6.1.1

Given a set $S \subseteq R^n$, a convex hull of S is the smallest convex set $\text{conv}(S)$ that contains S.

Definition 6.2.1

An integer matrix A of size $m \times n$ with $m \leq n$ is said to be unimodular if for any of its $m \times m$ submatrices B, $\det(B) \in \{-1, 0, 1\}$.

Theorem 6.2.1

Let A be unimodular and b integer. Then the polyhedron $P := \{x \geq 0 : Ax = b\}$ has only integer vertices.

Proof:

Let x^* be any vertex of P. As already proven, there exists a basis B of A such that $x^* = (B^{-1}b, 0)$. Since B is a nonsingular $m \times m$ submatrix of A, we have $|\det(B)| = 1$, which implies the integrality of B^{-1} and of $B^{-1}b$ for all integers b.

Definition 6.2.2

An integer matrix A of size $m \times n$ is said to be totally unimodular (TUM) if $\det(Q) \in \{-1, 0, 1\}$ for any of its square submatrices Q of any order.

Theorem 6.2.2

Let A be totally unimodular and b integer. Then polyhedron $P := \{x \geq 0 : Ax \leq b\}$ has only integer vertices.

Proof:

Let x^* be any vertex of polyhedron P. First of all, we prove that $(x^*, s^* := Ax^* - b)$ is a vertex of polyhedron $P' := \{(x, s) \geq 0 : Ax - s = b\}$. If that was not the case, there would exist two distinct points (x^1, s^1) and (x^2, s^2) of P' such that $(x^*, s^*) = \lambda(x^1, s^1) + (1 - \lambda)(x^2, s^2)$ for some λ with $0 < \lambda < 1$. Note that x^1 and x^2 belong to P, given that $s^1 = Ax^1 - b \geq 0$ and $s^2 = Ax^2 - b \geq 0$. In addition $(x^1, s^1) \neq (x^2, s^2)$, hence $x^* = \lambda x^1 + (1 - \lambda)x^2$ cannot be a vertex of P. Since A is TUM, we have that $A' := [A, -I]$ is unimodular. By theorem 6.2.1, (x^*, s^*) is integer, hence x^* is integer as well.

Theorem 6.2.3

Let A be a matrix with $a_{ij} \in \{-1, 0, 1\} \forall i, j$. A is totally unimodular if the following conditions hold:

- 1 Every columns of A has no more than two non-zero elements
- 2 there exists a partition (I_1, I_2) of the rows of A such that each column with two non-zero elements has these two elements belonging to rows on different sets if and only if the two elements have the same sign.

Proof:

We have to prove that $\det(Q) \in \{-1, 0, 1\}$ for any submatrix Q of A of order k ($k=1, \dots, m$). The proof is by induction on k . If $k = 1$, then $Q = [a_{ij}]$ and hence $\det(Q) = a_{ij} \in \{-1, 0, 1\}$, as requested. Let us suppose now that $\det(Q') \in \{-1, 0, 1\}$ for any submatrix Q' of order k' , where $k' \geq 1$ is a fixed value. Let us consider any submatrix Q of order $k := k' + 1$. By condition (1) only three cases can occur:

- Q has one column of zeros: in this case $\det(Q)=0$.
- Q has a column with only one element different from zero: in this case, barring permutations of rows and/or

$$\text{columns, } Q \text{ is of the type } Q = \begin{bmatrix} \pm 1 & | & * \dots * \\ - & - & - \\ 0 & | & \\ \cdot & | & \\ \cdot & | & Q' \\ \cdot & | & \\ 0 & | & \end{bmatrix} \Rightarrow \det(Q) = \pm 1 \det(Q')$$

where $\det(Q') \in \{-1, 0, 1\}$ given that Q' has order $k-1=k'$.

- Each column of Q has exactly two non-zero elements: in this case, let $I(Q)$ be the set of rows of Q . We then have $\sum_{i \in I_1 \cap I(Q)} [\text{row } i \text{ of } Q] - \sum_{i \in I_2 \cap I(Q)} [\text{row } i \text{ of } Q] = [\text{null row}]$ given that in every column the two non-zero elements cancel out. It follows that the rows of Q are linearly dependent, hence $\det(Q)=0$.

In each of the three cases above we therefore have $\det(Q) \in \{-1, 0, 1\}$, hence this property applies to all submatrices Q of order $k=k'+1$. Applying the same reasoning, the result can be inductively extended to matrices Q of any order.

Proposition 6.2.1

Matrix A is TUM is and only if:

- A^T is TUM;
- matrix A' obtained from A permuting and/or changing sign to some columns and/or rows is TUM;

$$\bullet \text{ matrices } \begin{bmatrix} \pm 1 & | & \\ 0 & | & \\ \cdot & | & \\ \cdot & | & \\ \cdot & | & \\ 0 & | & \end{bmatrix} A \text{ and } \begin{bmatrix} 0 & | & \\ 0 & | & \\ \cdot & | & \\ \cdot & | & \\ \cdot & | & \\ 0 & | & \end{bmatrix} A \text{ are TUM.}$$

Definition 6.3.1

Given $x^* \in P$ a cut is an inequality of the type $\alpha^T x \leq \alpha_0$ such that:

1. $\alpha^T x \leq \alpha_0, \forall x \in X := P \cap Z^n$
2. $\alpha^T x^* > \alpha_0$

These two conditions imply that the inequality $\alpha^T x \leq \alpha_0$ is valid for X , but violated by point x^* .

Chvátal's inequalities

For all $r \in R$, let $\lfloor r \rfloor$ be the truncation of r to the nearest lower integer, i.e. $\lfloor r \rfloor := \max\{i \in Z : i \leq r\}$. Given system $Ax=b, x \geq 0$, choose randomly $u \in R^m$ and obtain by linear combination the valid equation for P : $u^T Ax = u^T b$. By means of truncation, we can obtain the "weakened" inequality $\alpha^T x := \lfloor u^T A \rfloor x \leq u^T b$, where coefficients α_j are defined as $\alpha_j := \lfloor u^T A_j \rfloor := \lfloor \sum_{i=1}^m u_i a_{ij} \rfloor$ for all $j=1, \dots, n$. At least, defining $\alpha_0 := \lfloor u^T b \rfloor$ we obtain the valid inequality for X but not necessarily for P $\alpha^T x \leq \alpha_0$.

It is possible to prove that, given a fractional vertex x^* of P , there always exists a vector $u \in R^m$ for which the corresponding inequality $\alpha^T x \leq \alpha_0$ is violated by x^* . With the Chvátal procedure it is therefore always possible to cut any fractional vertex of P , by choosing vector u appropriately.

However, it would be wrong to think that all valid inequalities for X are obtained directly starting from the initial system $Ax=b$: there exist valid inequalities for X that are not obtained as $\lfloor u^T A \rfloor x \leq \lfloor u^T b \rfloor$ for any $u \in R^m$. Indicating with $A^{(1)}x \leq b^{(1)}$ the families of inequalities $\lfloor u^T A \rfloor x \leq \lfloor u^T b \rfloor, u \in R^m$, we can define the first Chvátal's closure as $P_1 := \{x \geq 0 : Ax = b, A^{(1)}x \leq b^{(1)}\} \subseteq P$, where $P_1 = P$ if and only if P does not have fractional vertices. Although system $A^{(1)}x \leq b^{(1)}$, in principle, contains ∞^m inequalities, it is possible to prove that only a finite number of its inequalities is necessary to describe P_1 , hence P_1 is a polytope. The additional constraint $A^{(1)}x \leq b^{(1)}$ allows one to improve the initial formulation $Ax=b, x \geq 0$. If $\text{conv}(X)=P_1$, this formulation is ideal. Otherwise there exists fractional vertices of P_1 , that can be cut generating new Chvátal's inequalities. We keep going till we have a solution completely integer, i.e, we keep iterating till $P_1 \supset P_2 \supset \dots \supset P_K = \text{conv}(X)$.

Gomory's cut

The idea is to generate a cut using the information associated with the basis B corresponding to the optimal basic solution x^* of the current relaxation. If x^* is integer, it is not necessary to generate any cut. Otherwise, there exists at least one fractional component x_h^* . Variable x_h must then be a basic variable on a certain row t of the optimal tableau associated with x^* . The equation associated with row t of the optimal tableau is the following: $x_h + \sum_{j \in F} \bar{a}_{tj} x_j = \bar{b}_t (= x_h^*)$, where F is the set of indexes of non-basic variables. Row t is said to be the generating row for the cut. The previous equation is, by construction, valid for P . By applying Chvátal's truncation procedure starting from this equation, we immediately obtain Gomory's cut: $x_h + \sum_{j \in F} \lfloor \bar{a}_{tj} \rfloor x_j \leq \lfloor \bar{b}_t \rfloor$. By construction, this inequality is valid for X , but violated by x^* . Indeed, given that $x_j^* = 0$ for all $j \in F$ and that $\bar{b}_t = x_h^*$ is fractional, we have: $(x_h + \sum_{j \in F} \lfloor \bar{a}_{tj} \rfloor x_j)_{x=x^*} = x_h^* = \bar{b}_t > \lfloor \bar{b}_t \rfloor$. The previous equation is said to be the integer form of the cut. There exists an equivalent formulation, called fractional form, which is obtained by subtracting the two equations: $\sum_{j \in F} (\bar{a}_{tj} - \lfloor \bar{a}_{tj} \rfloor) x_j \geq \bar{b}_t - \lfloor \bar{b}_t \rfloor$, i.e., $\sum_{j \in F} \phi(\bar{a}_{tj}) x_j \geq \phi(\bar{b}_t)$, where for all $r \in R$ we indicate with $\phi(r) := r - \lfloor r \rfloor \geq 0$. Converting the first equation to standard form, we obtain $x_h + \sum_{j \in F} \lfloor \bar{a}_{tj} \rfloor x_j + s = \lfloor \bar{b}_t \rfloor$, where the slack variable $s \geq 0$ will be integer for all integer x , given that all coefficients are integer. We can then obtain the final equation as: $\sum_{j \in F} \phi(\bar{a}_{tj}) x_j + s = -\phi(\bar{b}_t), s \geq 0$ integer.

Graph Theory

Definition 7.3.1

The node-edge incidence matrix D of an undirected graph $G=(V,E)$ is the $|V| \times |E|$ matrix with elements $d_{ij} \begin{cases} 1 & \text{if the } j\text{-th edge } e \text{ is incident to node } i \\ 0 & \text{otherwise.} \end{cases}$

Definition 7.3.2

The node-edge incidence matrix D of a directed graph $G=(V,A)$ is the $|V| \times |A|$ matrix with elements $d_{ij} \begin{cases} 1 & \text{if the } j\text{-th leaves node } i \\ -1 & \text{if the } j\text{-th enters node } i \\ 0 & \text{otherwise.} \end{cases}$

Definition 7.3.3

The adjacency matrix Q of a simple graph $G = (V, E)$ is the symmetric $|V| \times |V|$ matrix with elements $q_{ij} \begin{cases} 1 & \text{if } [i, j] \in E \\ 0 & \text{otherwise.} \end{cases}$

Definition 7.3.4

The adjacency matrix Q of a simple directed graph $G = (V, A)$ is the symmetric $|V| \times |V|$ matrix with elements $q_{ij} \begin{cases} 1 & \text{if } (i, j) \in A \\ 0 & \text{otherwise.} \end{cases}$

Definition 7.4.1

Given two real functions f and g defined on the same domain, we have that $f(x)=O(g(x))$ if there exists two constants k_1 and k_2 such that $f(x) \leq k_1 g(x) + k_2$ for all x in the domain.

Definition 7.4.2

An algorithm is said to be polynomial if it need, in the worst case, a computing time $f(d) = O(d^k)$, where d is the size of the instance and k is a fixed constant.

Definition 7.4.3

A problem is said to be polynomial (or easy) if there exists a polynomial algorithm solving it.

Definition 7.4.4

A polynomial certificate is defined as an auxiliary piece of information that can be used to verify, in time polynomial in the size of the instance, the correctness of the answer for a given instance.

Definition 7.4.5

A decision problem belongs to the NP (Non-deterministic Polynomial) class if every of its instances with "yes" answer admits a polynomial certificate.

Definition 7.4.6

A decision problem belongs to the co-NP (NP complete) class if every of its instances with "no" answer admits a polynomial certificate.

Definition 7.4.7

Problem P_1 can be reduced in polynomial time to problem P_2 ($P_1 \propto P_2$) if there exists an algorithm to solve P_1 that:

- uses as black-box procedure any algorithm for P_2 ;
- is polynomial in the hypothesis that the algorithm for P_2 needs $O(1)$ time.

Definition 7.4.8

A problem $P_2 \in NP$ is said to be NP-complete if $P_1 \propto P_2$ for all $P_1 \in NP$.

Theorem 7.4.1

The problem: Given a directed graph $D=(N,A)$, does there exists an Hamiltonian circuit $D'=(N',A')$ in D ? is NP-complete.

Proof:

The problem belongs to the NP class, given that one of its instances with "yes" answer can be certified by exhibiting a Hamiltonian circuit in D .

Minimum Spanning Tree

$$\left\{ \begin{array}{l} \min \underbrace{\sum_{e \in E} c_e x_e}_{\text{total cost}} \\ \underbrace{\sum_{e \in E} x_e = n - 1}_{n-1 \text{ chosen edges}} \\ \underbrace{\sum_{e \in E} x_e \leq |S| - 1}_{\text{cycle elimination condition}} \quad \forall S \subseteq V : S \neq \emptyset \\ x_e \geq 0 \text{ integer, } e \in E \end{array} \right.$$

Algorithms

Simplex

Let $\beta[1], \dots, \beta[m]$ be the indexes of the columns of an initial feasible basis;
unbounded = false;
optimal = false;
while optimal = false and unbounded = false:
 let $B := [A_{\beta[1]} | \dots | A_{\beta[m]}]$ be the current feasible basis;
 compute B^{-1} and set $u^T := c_B^T B^{-1}$;
 compute the reduced cost $\bar{c}_h := c_h - u^T A_h$ of non-basic variables x_h ;
 if $\bar{c}_h \geq 0 \forall x_h$ non-basic then optimal = true;
 else
 choose a non-basic variable x_h with $\bar{c}_h < 0$;
 compute $\bar{b} := B^{-1}b$ and $\bar{A}_h := B^{-1}A_h$;
 if $\bar{a}_{ih} \leq 0 \forall i \in \{1, \dots, m\}$ then unbounded = true;
 else
 compute $t = \text{argmin}\{\bar{b}_i / \bar{a}_{ih}, i \in \{1, \dots, m\} : \bar{a}_{ih} > 0\}$;
 update the current basis by setting $\beta[t] = h$;

Pivot operation

save = $\bar{a}[t, h]$; #value of the pivot element;
for $j = 0$ to n : $\bar{a}[t, j] = \bar{a}[t, j] / \text{save}$;
for $i = 0$ to m :
 if $(i \neq t)$ and $(\bar{a}[i, h] \neq 0)$
 save = $\bar{a}[i, h]$;
 for $j = 0$ to n : $\bar{a}[i, j] = \bar{a}[i, j] - \text{save} * \bar{a}[t, j]$

Two phase method simplex

Let $\beta[1], \dots, \beta[m]$ be the indexes of the columns of an initial feasible basis;
Create the initial tableau $\bar{A} = \{\bar{a}[i, j] : i = 0, \dots, m; j = 0, \dots, n\}$, in canonical form w.r.t. the initial basis;
unbounded = false;
optimal = false;
while optimal = false and unbounded = false:
 if $\bar{a}[0, j] \geq 0 \forall j = 1, \dots, n$ then optimal = true;
 else
 choose a non-basic variable x_h with $\bar{a}[0, h] < 0$;
 if $\bar{a}[i, h] \leq 0 \forall i = 1, \dots, m$ then unbounded = true;
 else
 compute $t = \text{argmin}\{\bar{a}[i, 0] / \bar{a}[i, h] : i \in \{1, \dots, m\}, \bar{a}[i, h] > 0\}$;
 pivot(t, h);
 $\beta[t] = h$;

Results of the two phase method

1. $w^* > 0$. In this case, there exist no solution to the artificial problem with $y_i = 0 \forall i \in \{1, \dots, m\}$ (such solution would have cost $w=0$), hence the system $Ax=b, x \geq 0$, does not admit a solution: the original problem is hence infeasible.
2. $w^* = 0$. In this case we have $y^* = 0$, since $y^* \geq 0$ and $\sum_{i=1}^m y_i^* = 0$, hence x^* is a feasible solution of the original problem. With respect to the optimal tableau for the artificial problem, two sub cases may occur:
 - a All the artificial variables are non-basic variables: eliminating the columns of the tableau corresponding to the artificial variables, we have then a tableau for the original problem, already in canonical form w.r.t to a basis. Now the artificial objective function is converted in canonical form, and the original problem is solved by the simplex algorithm.
 - b There exists a basic variable y_h on a row i : in this case we have $y_h^* = 0$ (since $w^* = 0$), hence we are in the presence of degeneracy. If there exists a value $\bar{a}_{ij} \neq 0$ in correspondence to a variable $x_j (j \leq n)$, then it is sufficient to perform a pivot operation on element (i,j) . Since $\bar{b}_i = 0$, this operation is admissible even if $\bar{a}_{ij} < 0$ and does not entail variations to the objective function w . In this way, y_h leaves the basis. Repeating the procedure for all the basic y_h leads back to the non-degenerate case (a).
One last possibility remains: all values $\bar{a}_{i1}, \dots, \bar{a}_{in}$ are equal to zero. In this case, eliminating the artificial columns we obtain a tableaus equivalent to the original one, but with a zero row. This implies that the matrix A does not have full row rank m , and thus the i -th row of the current tableau may be eliminated without problems.

Cutting Plane Algorithm

solve the continuous relaxation $\min\{c^T x : Ax = b, x \geq 0\}$, and let x^* be the optimal basic solution found;
while x^* not integer:

solve the separation problem of x^* from X , identifying an inequality $\alpha^T x \leq \alpha_0$ valid for X but violated by x^* ;
add constraint $\alpha^T x \leq \alpha_0$ to the current continuous relaxation;
solve the current continuous relaxation (with dual simplex) and let x^* be the new optimal basic solution identified.

Branch and Bound

$m:=1$; $\text{parent}[1] := 0$; $q := \emptyset$; $z_{OPT} :=$ heuristic solution value (possibly $+\infty$)

solve the continuous relaxation $\min\{c^T x : Ax = b, x \geq 0\}$, and let x^* be the optimal solution found;

$LB[1] := c^T x^*$;

if $(x^*$ integer) and $(c^T x^* < z_{OPT})$:

$x_{OPT} = x^*$; $z_{OPT} := c^T x^*$

if $LB[1] < z_{OPT}$:

choose the fractional branching variable x_h^* ;

$vbranch[1] := h$; $value[1] := x_h^*$;

$Q := \{1\}$

while $Q \neq \emptyset$:

choose a node $t \in Q$, and set $Q := Q \setminus \{t\}$;

$h := vbranch[t]$; $val := value[t]$;

for child:=1 to 2:

$m:=m+1$;

if child = 1:

$parent[m] := t$;

else $parent[m] := -t$;

define problem PL_m associated with node m ;

solve problem PL_m and let x^* be the optimal solution found;

$LB[m] := c^T x^*$;

if $(x^*$ integer) and $(c^T x^* < z_{OPT})$:

$$\begin{aligned}
& x_{OPT} := x^*; z_{OPT} := c^T x^*; \\
& Q := Q \setminus \{j \in Q : LB[j] \geq z_{OPT}\} \\
& \text{if } LB[m] < z_{OPT}: \\
& \quad \text{choose the fractional branching variable } x_k^*; \\
& \quad vbranch[m] := k; value[m] := x_k^* \\
& \quad Q := Q \cup \{m\}
\end{aligned}$$