# Linear Programming

# 1 Linear Programming

Linear Programming problems are characterized by a linear *objective function* in *decision variables* and by *constraints* described by linear inequalities or equations. Wishing to give a geometric interpretation in the space of variables (as in the example we developed at the beginning of Chapter 1), constraints detect *half-spaces* (in the case of inequalities) or *hyperplanes* (in the case of equalities). "Level curves" of the objective function (i.e., points having equal value) detect hyperplanes.

A problem of this type is called a *Linear Programming* (*LP*) problem. Formally, a Linear Programming problem is an optimization (maximization or minimization) problem defined on $x$ variables taking values in $\boldsymbol{R}^n$ and characterized by the following properties:

*i*) the objective function $c(x)$: $\boldsymbol{R}^n \rightarrow \boldsymbol{R}$ is linear, i.e., it satisfies the relations $c(0)=0$, $c(\alpha x+\beta y) = \alpha c(x)+\beta c(y)$, $\forall x, y \in \boldsymbol{R}^n$, $\forall \alpha, \beta \in \boldsymbol{R}$;

*ii*) the feasible region is defined by a finite set of linear constraints of the type $h(x) = \gamma$, and/or $h(x) \leq \gamma$, and/or $h(x) \geq \gamma$, where $h(x)$: $\boldsymbol{R}^n \rightarrow \boldsymbol{R}$ is a linear function and $\gamma \in \boldsymbol{R}$.

A *LP* problem can be synthetically expressed by means of the following formulation, in which $A$ is a real matrix $m\times n$, $b\in\boldsymbol{R}^m$, $c$ and $x\in\boldsymbol{R}^n$:

$$\begin{aligned} &\max cx \\ &Ax \leq b \\ &x \geq 0. \end{aligned}$$

Below we will adopt the following notations:

- if $a$ and $b$ are vectors having the same dimensions, then $ab$ denotes their scalar product;
- if $A$ is a matrix $m$x$n$ and $b$ is a vector with dimension $n$ [$m$], then $Ab$ [$bA$] denotes the product of $A$ by $b$ [$b$ by $A$] considered as a column [row] matrix[(1)].

Writing the farmer problem of Chapter 1 in matricial form, we get:

$$\max\, [3000,\ 5000]\begin{bmatrix} x_L \\ x_P \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ 7 & 0 \\ 0 & 3 \\ 10 & 20 \end{bmatrix}\begin{bmatrix} x_L \\ x_P \end{bmatrix} \leq \begin{bmatrix} 12 \\ 70 \\ 18 \\ 160 \end{bmatrix},$$

$$x_L, x_P \geq 0.$$

Any *LP* problem can be easily related to the maximization form with constraints of smaller than or equal to and non-negative variables – or else to any other form – by means of the following equivalences we partly examined in Chapter 1:

$$\max \textstyle\sum_j c_j x_j \equiv -\min \textstyle\sum_j (-c_j)x_j$$

$$\textstyle\sum_j a_{ij}x_j = b_i \equiv \begin{cases} \sum_j a_{ij}x_j & \leq\ b_i \\ \sum_j (-a_{ij})x_j & \leq\ -b_i \end{cases}$$

$$\textstyle\sum_j a_{ij}x_j \geq b_i \equiv \sum_j (-a_{ij})x_j \leq -b_i$$

$$\textstyle\sum_j a_{ij}x_j \geq b_i \equiv \sum_j a_{ij}x_j - s_i = b_i, \qquad s_i \geq 0$$

$$\textstyle\sum_j a_{ij}x_j \leq b_i \equiv \sum_j a_{ij}x_j + s_i = b_i, \qquad s_i \geq 0$$

Variables $s_i$ are called *slack variables* because they provide the difference between the left side of the constraint and the known term.

---

(1) From now on, according to the context, we will consider a vector either as a row matrix or as a column matrix.

Moreover, a free variable $x$ can be replaced by non-negative variables by means of the transformation:

$$x = x^+ - x^-, \ x^+ \geq 0, \ x^- \geq 0.$$

From now on, when defining *LP* problems we will frequently use one of the two following forms:

$$\begin{array}{ll} \max \ cx & \min cx \\ \quad Ax \leq b & \quad Ax = b \\ & \quad x \geq 0 \end{array}$$

In some texts the first form is called *canonical* or *general*, whereas the second form is called *standard*.

Example: Shutters and Frames (mixed production)

A small concern in the Pavia lowland produces two types of shutters: an aluminum door and a wooden window. Taking into account production and staff costs, the profit brought by each door is 30 euros and the profit brought by each window is 50 euros. One week the three workers employed by the concern (a smith, a carpenter and an assembler) inform the proprietress of their availability to do 4, 12 and 18 extra working hours, respectively. Producing a door takes one hour's job by the smith and three hours' job by the assembler, whereas a window requires two hours' job by the carpenter and two hours' job by the assembler. The proprietress has to plan the extra production so as to maximize the profit.

Decision variables are $x_D$ and $x_W$ and respectively represent the number of doors and windows to be produced. Wishing to use a matricial representation, the variable vector is:

$$x = \begin{bmatrix} x_D \\ x_W \end{bmatrix},$$

whereas the coefficient vector of the objective function $c$ is

$$c = [30, 50].$$

In canonical form, the problem has the following matrix of constraint coefficients and the following vector of known terms:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 2 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \qquad b = \begin{bmatrix} 4 \\ 12 \\ 18 \\ 0 \\ 0 \end{bmatrix}$$

Note how the sign constraints of variables have been explicitly placed inside matrix $A$ ($-x_D \leq 0$, and $-x_W \leq 0$) because the canonical form does not comprise variables restricted in sign, and the possible sign constraints must be treated as explicit constraints of the problem.

Example: linear classifier

Given a series of $m$ surveys that were carried out (for example, the citrus fruits picked at several farms), each being characterized by a set of $n$ attributes (for example: weight, elliptical shape, color) and by a classification in two classes (for example Lemons and Oranges), we wish to find a criterion enabling to classify all future surveys, so that a machine can be "trained" to perform the automatic classification of citrus fruits and their packing. The problem can be related to Linear Programming. The data relative to surveys are contained in matrix $A$, where element $a_{ij}$ is the attribute $j$ of survey $i$. We intend to find a way to weigh the different attributes as well as a value representing a threshold separating one class from the other. Formally, we wish to find the weights $x_1, x_2, \ldots, x_n$ and a threshold $y$ such that:

$$\sum_{j=1}^{n} a_{ij}\, x_j \leq y \qquad \text{for each survey } i = 1, \ldots, m \text{ belonging to class 1}$$

$$\sum_{j=1}^{n} a_{ij}\, x_j > y \qquad \text{for each survey } i = 1, \ldots, m \text{ belonging to class 2.}$$

This is a set of $m$ linear constraints in $n$+1 unknowns. As soon as the *xs* and the $y$ are determined, we can to classify any other survey$*$ with $a_{*j}$ attributes by calculating:

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

and by assigning the survey $*$ to the first class if $w \leq y$ and to the second class if $w > y$. Actually, it should be noted that the second group of constraints we have imposed for determining the *xs* and the $y$ involve narrow inequalities: a fact that is not comprised by Linear Programming, where all constraints must include equality. In the case of linear classification, this inconvenience can be sidestepped by transforming the second group of constraints into the following inequalities:

$$\sum_{j=1}^{n} a_{ij}\, x_j \geq y + \varepsilon \qquad \text{for each survey } i = 1, \ldots, m \text{ belonging to class 2}$$

where ε is an appropriately small fixed value.

Note how in the case of the linear classifier we have indicated no objective function because the different feasible solutions (if any exist) are all equivalent to us.

**Exercise**

Develop an example of a linear classifier for a set containing some surveys having two attributes for which there are no feasible solutions.

**Exercise**

Transform the formulation of the linear classifier problem into standard form. Find out the possible objective functions.

## *1.1 Geometric representation*

When we are confronted with problems in canonical form having two (or three) variables, we can make use of geometric representation in order to gain greater intuitive understanding of the problem and even to solve it. Let us examine again the Shutters-and-Frames example. In Figure 1 the hatching indicates the *feasible region*, that is the set of all points satisfying the constraints. It corresponds to the intersection of the five half-spaces defined by constraint inequalities; a set of this type is called a *polyhedron* (or *polytope* in the specific case it is bounded, as in the considered example). We can immediately verify that, precisely because it results from the intersection of half-spaces, such set is *convex*; the term *convex polyhedron* is frequently used to characterize the geometry of the feasible region of a Linear Programming problem. A point belonging to the feasible set is called a *feasible solution*, whereas an external point is sometimes – and to some extent improperly – denoted as unfeasible solution.

In the figure constraints are pointed out by means of straight lines being the *locus* of the points satisfying the constraints as narrow equalities; such straight lines represent the boundary of half-spaces defined by the constraints themselves. The non-negativity constraints of variables, if treated as explicit constraints, have the boundary detected by Cartesian axes. In the figure, constraint gradients are indicated as well: they correspond to the red vectors "leaning" on straight lines and referring to constraint edges. As we have to do with linear constraints, constraint gradients are nothing but the different lines of matrix $A$. The gradient of a constraint indicates in which direction

the linear function on the left of the inequality grows. Since our problem implies constraints of smaller than or equal to, the feasible region is situated in the half-space opposite to the gradient.
Straight lines corresponding to constraints detect *vertices* and *faces* in the polyhedron: for example, points (2, 6) and (4, 3) are vertices and the segment joining them is a face. Later we will define the concepts of face and vertex more formally and will see how vertices can be considered as particular faces.

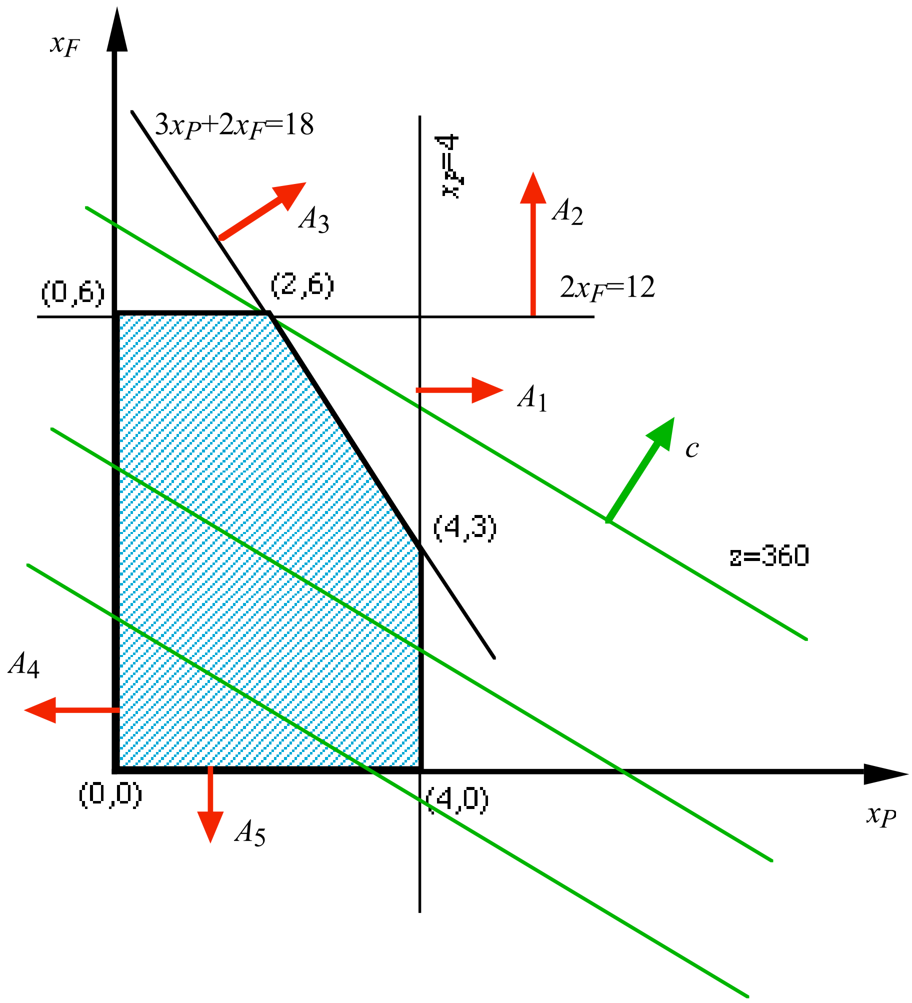


Fig. 1: representation and graphical solution of the Shutters-and-Frames problem

In Linear Programming vertices and faces play a particularly important role. Indeed, it is possible to prove that in a *LP* problem, if the optimal solution exists as finite, then at least one optimal solution is situated in a vertex; moreover, if a point inside a face is optimal solution to the problem, then all points of the face are optimal solutions. The truth of these properties is suggested by an examination of the figure. (It will be formally proved afterwards).
Let us try to solve the problem. Consider the straight line $30x_P+50x_F=z$. For each value of $z$ it defines the set of (not necessarily feasible) solutions having the value of the objective function equal to $z$. In the figure we indicate three of such straight lines, i.e., the ones corresponding to values 100, 200 and 360. As $z$ grows, straight lines are translated and move in the direction defined by vector $c$=[30, 50], which is gradient of the objective function and has been represented in green and slightly reduced.
Clearly, for each given value of $z$ it is possible to make a profit equal to that value if and only if the corresponding straight line has non-empty intersection with the feasible region; therefore, in order to find an optimal solution to our problem it suffices to move the straight line in the direction of the gradient as much as possible, imposing the constraint that the intersection with the feasible region remains non-empty. In the considered case the maximum value that can be assigned to $z$ is 360, and

for such value the intersection of the straight line with the feasible set is reduced to only one point: the vertex (2, 6), which is consequently the optimal solution.

In the problem we have just examined the feasible region is bounded. We might as well be confronted with problems in which, along some directions, the feasible region is unbounded. In such cases, depending on the specific objective function (i.e., the direction of its gradient), there may exist directions along which it is possible to move preserving the feasibility and (in maximization problems) making the value of the objective function grow without ever reaching a maximum limit. For example, in the Shutters-and-Frames problem, if the second and the third constraint were not there, we could let the value of $z$ grow to infinity without ever finding a value according to which the straight line $30_P+50x_F=z$ has empty intersection with the feasible region. In cases of this kind we speak of an *unbounded problem*, i.e., there exist feasible solutions, but there exists no maximum limit for the objective function (there merely exists an upper bound, i.e., $+\infty$). In practical applications a solution of this kind usually means that the model which was built represents the investigated reality in an improper or incomplete way.

A case somehow opposite to the preceding one is the case in which some of the constraints are incompatible with each other, so that the resulting feasible set is empty; in this case there are no solutions and the problem is called an *unfeasible problem*.

<u>Example: transformation into standard form and geometric representation</u>

Consider the following *LP* problem in canonical form:

$$
\begin{array}{lrcl}
\max & 3x_1 - x_2 & & \\
 & x_1+x_2 & \leq & 4 \\
 & -x_1+x_2 & \leq & 5 \\
 & -x_2 & \leq & 2
\end{array}
$$

By means of constraint transformations, and in particular by introducing slack variables, the problem can be transformed into the equivalent standard form:

$$
\begin{array}{ll}
\min & -3x_1 + x_2 \\
 & x_1 + x_2 + s_1 = 4 \\
 & -x_1 + x_2 + s_2 = 5 \\
 & -x_2 + s_3 = 2 \\
 & s_1, s_2, s_3 \geq 0
\end{array}
$$

Figure 2 shows the geometric representation of the problem. In particular, it should be noted that the feasible solutions of the problem belong to the hatched polygon (feasible region). Slack variables are associated with constraints and define the line supporting the side of the polygon ($s_i=0$) as well as the feasible half-plane corresponding to the constraint ($s_i \geq 0$). In the figure we indicate the gradient and the level curve of the objective function, i.e., the set of points such that, for a real datum $z$, $3x_1 - x_2 = z$. The optimal solution of the problem is given by the vertex $v$ shown in the figure, to which the following values correspond: $x_1 = 6$, $x_2 = -2$, $s_1 = 0$, $s_2 = 13$, $s_3 = 0$.

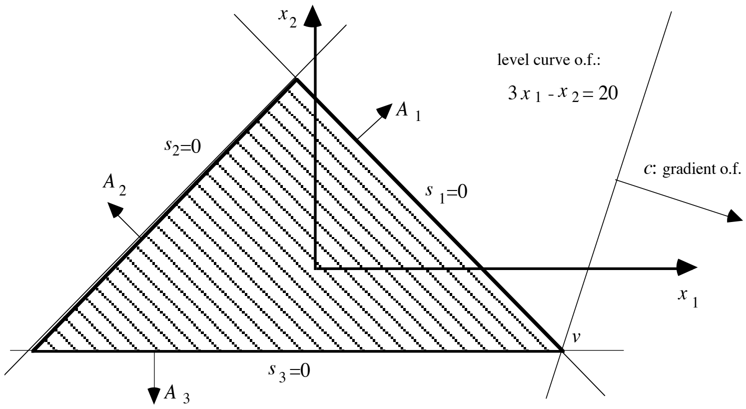


Fig. 2: geometric representation of a problem in standard form

**Exercise**

Build some more *LP* examples in the plane. In particular, formulate problems respectively resulting in: empty feasible region, non-finite optimum (objective function unbounded from above on the set of feasible solutions), non-unique optimal solution.

## *1.2 Some geometric aspects of Linear Programming

From the geometric point of view, the set of feasible solutions of a Linear Programming problem is the intersection of a certain number of half-spaces. In this paragraph we refer to some geometric concepts which allow to characterize such sets.

**Definition 1.1**

A set $S \subset \boldsymbol{R}^n$ is *convex* if for each $x, y \in S$ and each $\lambda \in [0,1]$ the convex combination $\lambda x + (1-\lambda)y \in S$.

From the geometric point of view, the union of all convex combinations of two points provides the segment joining them.

**Definition 1.2**

The set $C \subseteq \boldsymbol{R}^n$ is

- a *cone* if: $x \in C$, $\lambda \in \boldsymbol{R}_+ \Rightarrow \lambda x \in C$;
- a *convex cone* if: $x, y \in C$, $\forall \lambda, \mu \geq 0 \Rightarrow (\lambda x + \mu y) \in C$;
- a *polyhedral cone* if: $C = \{x: Ax \leq 0\}$;
- a *cone finitely generated* by $x^1,\ldots,x^m \in \boldsymbol{R}^n$ if: $C = cone\{x^1,\ldots,x^m\} = \{x = \lambda_1 x^1 + \ldots + \lambda_m x^m: \lambda_1,\ldots,\lambda_m \in \boldsymbol{R}_+\}$.

**Exercise**

Prove that the sets convex cone, polyhedral cone and finitely generated cone are convex sets.

**Definition 1.3**

$P \subseteq \boldsymbol{R}^n$ is a *convex polyhedron* if:

$$P = \{x \in \boldsymbol{R}^n: A_i x \leq b_i,\ i=1,\ldots,m\},$$

$P$ is the intersection of a finite number of *affine half-spaces*.

**Definition 1.4**

$Q \subseteq \boldsymbol{R}^n$ is a *convex polytope* if there exist $m$ vectors $v^1,\ldots,v^m$ such that:

$$Q = conv\{v^1,\ldots,v^m\} = \{x = \lambda_1 v^1 + \ldots + \lambda_m v^m: \lambda_1,\ldots,\lambda_m \geq 0, \sum_{i=1}^{m} \lambda_i = 1\},$$

$Q$ is the convex envelope of $v^1,\ldots,v^m$.

**Exercise**
Prove that the convex polyhedron and the convex polytope are convex sets.

**Exercise**
Prove that a polyhedral cone is a polyhedron.

As we informally saw in the first Chapter and in the Shutters-and-Frames example which was geometrically solved, it is intuitively clear that the vertices of the polyhedron enclosing the feasible region are important in order to limit the search for the feasible solutions. It is thus necessary to try to characterize them formally.

**Definition 1.5**
Let $P \subseteq \boldsymbol{R}^n$ be a convex set. A vector $x \in P$ is an *extreme point* of $P$ if it is not possible to find two vectors $y$, $z \in P$ being both different from $x$ and a scalar $\lambda \in [0,1]$ such that $x = \lambda y + (1-\lambda)z$.

Note that polyhedra have a finite set of extreme points.

**Exercise**
Give an example of a convex set having an infinite number of extreme points.

**Definition 1.6**
Let $P \subseteq \boldsymbol{R}^n$ be a convex polyhedron. A vector $x \in P$ is a *vertex* of $P$ if there exists a vector $c \in \boldsymbol{R}^n$ such that $cx > cy$, for each $y \in P$ and $y \neq x$.

According to the latter definition $x$ is a vertex if there exists a hyperplane (with gradient $c$) separating $x$ from the rest of the feasible region, as shown by the linear classifier example. Or else, there exists an appropriate objective function with gradient $c$ such that $x$ is the unique optimal solution of the problem defined on polyhedron $P$. Both definitions 1.5 and 1.6 detect what we informally called "vertex" of the polyhedron. Unfortunately, none of these two definitions are of any help from the algorithmic point of view.

**Definition 1.7**
Let us consider a polyhedron $P = \{x \in \boldsymbol{R}^n: A_i x \leq b_i,\ i = 1,\ldots,m\}$ and a point $x^* \in \boldsymbol{R}^n$; we define $I(x^*) = \{i: A_i x^* = b_i\}$ the set of *constraints active* in $x^*$.

**Definition 1.8**
Let $P = \{x \in \boldsymbol{R}^n: A_i x \leq b_i,\ i = 1,\ldots,m\}$ be a convex polyhedron. A vector $x$ is a *basic solution* if among vectors $A_i$ with $i \in I(x)$ there are $n$ of them being linearly independent. We have a basic feasible solution if $x \in P$.

Note that if $B \subseteq I(x)$ detects $n$ vectors $A_i$ being linearly independent (hence it is a basis), then the system

$$A_i x = b_i \quad i \in B$$

has a unique solution. This means that the intersection of hyperplanes being edges of the constraints in $B$ detects a unique point.

**Theorem 1.9**

Given a polyhedron $P=\{x\in\boldsymbol{R}^n: A_ix\leq b_i, i=1,\ldots,m\}$, definitions 1.5, 1.6 and 1.8 are equivalent.

Proof

Let us demonstrate the three implications.

**a)** $x$ vertex $\Rightarrow$ $x$ extreme point

Let $x$ be a vertex of $P$. Hence, according to definition 1.6, there exists $c\in\boldsymbol{R}^n$ such that $cx>cy$, for each $y\in P$ and $y\neq x$. Consider now any two points of $P$, $y$ and $z$, other than $x$, and a multiplier $\lambda\in[0,1]$; then by applying the definition we have $cx>cy$ and $cx>cz$, which implies $cx>c(\lambda y+(1-\lambda)z)$. It follows that $x\neq\lambda y+(1-\lambda)z$, which means that $x$ cannot be written as a convex combination of two other points $P$ being different from $x$ itself, that is to say that $x$ is an extreme point of $P$.

**b)** $x$ extreme point $\Rightarrow$ $x$ basic feasible solution

This time we suppose that $x$ is not a basic solution and prove that $x$ is not an extreme point. The fact that $x$ is not a basic solution means that in $\{A_i, i\in I(x)\}$ there do not exist $n$ linearly independent vectors. Then let $d\in\boldsymbol{R}^n$ be a non totally null vector orthogonal to each vector $A_i$, i.e., such that $A_id=0$ for each $i\in I(x)$.

We construct two points $y$ and $z$ starting from $x$: $y=x+\varepsilon d$ and $z=x-\varepsilon d$, where $\varepsilon$ is an appropriately small scalar. We observe that, due to the orthogonality of $d$, $A_iy=A_iz=A_ix=b_i$. This means that constraints active in $x$ are also active in $y$ and $z$. Because $x\in P$ we also have $A_ix<b_i$ for $i\in\{1,\ldots,m\}\backslash I(x)$, and by an appropriate choice of $\varepsilon$ also $A_iy$ and $A_iz$ are smaller than $b_i$ for $i\in\{1,\ldots,m\}\backslash I(x)$, then also $y$ and $z\in P$. But we can observe as well that $x=(y+z)/2$, which implies that $x$ cannot be an extreme point.

**c)** basic feasible solution $\Rightarrow$ vertex

Let $x$ be a basic feasible solution. We construct a vector $c$ as follows

$$c=\sum_{i\in B} A_i.$$

The result is:

$$cx=\sum_{i\in B} A_ix=\sum_{i\in B} b_i.$$

Moreover, for each $y\in P$ and each $i$ we obviously have $A_iy\leq b_i$ and

$$cy=\sum_{i\in B} A_iy\leq\sum_{i\in B} b_i. \tag{1.1}$$

This proves that $x$ is an optimal solution for the $LP$ problem defined on $P$ with $c$ as gradient of the objective function. Moreover, in (1.1) the equality is valid only if $A_iy=b_i$ for each $i\in I(x)$. Because $x$ is a basic solution, there are $n$ linearly independent constraints active in $x$, and $x$ is the unique solution of the system $A_ix=b_i$ for $i\in I(x)$. Hence, if $y\neq x$, $cy<cx$, and $x$ respects the definition of vertex.

Due to the transitivity of the implication, we have the equivalence of the three definitions. ♦

**Definition 1.10**

Two bases $B_1$ and $B_2$ are *adjacent* if they differ in only one component.

**Exercise**

Consider the polyhedron of Figure 1 and the indicated basic solution $x$. Detect the basic solutions adjacent to $x$.

The following theorem shows a fundamental property of convex polyhedra and helps us to find a formal demonstration of the reason why if in a Linear Programming problem there exists a finite optimal solution, this corresponds to at least one vertex.

**Theorem 1.11** [Motzkin, 1936] *Polyhedron decomposition*
*$P \subseteq \mathbf{R}^n$ is a polyhedron if and only if $P = Q + C$ where $Q$ is a polytope and $C$ is a polyhedral cone.*

Let $P = Q + C$, where $Q = conv\{x^1,\ldots,x^s\}$ and $C = cone\{y^1,\ldots,y^t\}$; we say that $P$ is generated by the *points* $x^1,\ldots, x^s$ and by the *directions* $y^1,\ldots, y^t$.
In case $conv\{x^1,\ldots,x^s\}$ is a minimal representation of polytope $Q$, vectors $x^1,\ldots,x^s$ coincide with *extreme points* $Q$. In case $cone\{y^1,\ldots,y^t\}$ is a minimal representation of cone $C$, vectors $y^1,\ldots,y^t$ are called *extreme rays* of $C$.

Example
Consider the polyhedron decomposition example illustrated in Figure 3, where:

$P = Q + C,\ Q = conv\{x^1,x^2,x^3\}$ and $C = cone\{y^1,y^2\}$.

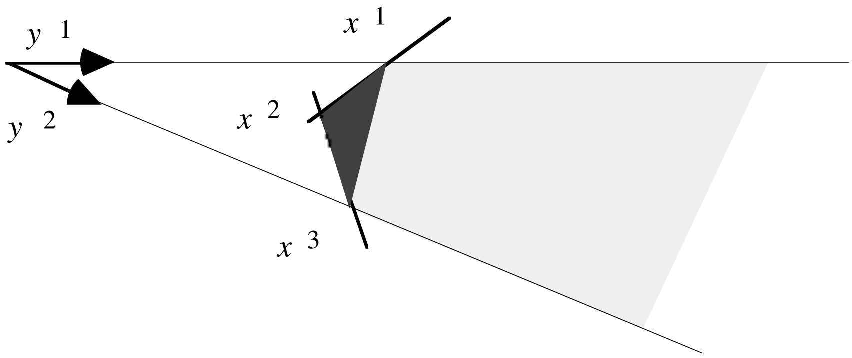


Fig. 3

The decomposition is minimal in the sense that the extreme points $x^1,x^2,x^3$ [the extreme rays $y^1,y^2$] are the minimum number of vectors for generating $Q$ [$C$].

**Theorem 1.12**
Let $P = \{x: Ax \leq b\} = Q + C$, where $Q = conv\{x^1,\ldots,x^s\}$ and $C = cone\{y^1,\ldots,y^t\}$. If the problem:

$$\begin{aligned} \max\ & cx \\ & Ax \leq b \end{aligned}$$

has finite optimum, then there exists a $k \in \{1,\ldots,s\}$ such that $x^k$ is an optimal solution.

Proof
The problem $\max\{cx: Ax \leq b\}$ is equivalent to:

$$\begin{aligned} \max\ & \sum_{i=1}^{s} \lambda_i cx^i + \sum_{j=1}^{t} \mu_j cy^j \\ & \sum_{i=1}^{s} \lambda_i = 1 \\ & \lambda_i \geq 0, \quad \mu_j \geq 0, \quad i=1,\ldots,s,\ j=1,\ldots,t. \end{aligned}$$

It clearly appears that the problem has finite optimum if and only if $cy^j \leq 0$, $j=1,\ldots,t$, and in case it has finite optimum there certainly exists an optimal solution in which $\mu_j = 0$, $j=1,\ldots,t$. Therefore, the problem can be rewritten as:

$$\max \quad \sum_{i=1}^{s} \lambda_i c x^i$$
$$\sum_{i=1}^{s} \lambda_i \;= 1$$
$$\lambda_i \geq 0, \quad i=1,\ldots,s,$$

which is equivalent to $\max\{cx^i: i=1,\ldots,s\}$, and this completes the demonstration. ♦

Now let us introduce a property of convex sets that will prove particularly useful afterwards.
**Theorem 1.13** (*of the separating hyperplane*)
Let $S$ be a closed convex set of $\boldsymbol{R}^n$ and let there be $c \notin S$. Then there exists a vector $x \in \boldsymbol{R}^n$ and $\varepsilon>0$ such that

$$cx \geq zx + \varepsilon \qquad \forall z \in S.$$

Proof
Let $z_0$ be the point of $S$ being nearest to $c$. Given that $c \notin S$, we have $||z_0 - c|| > 0$.
We prove that $(c - z_0)(z - z_0) \leq 0\ \forall z \in S$, which geometrically means that, by moving from $z_0$ towards $z$, we move far from $c$. To do this, we consider a convex combination of $z$ and $z_0$: $\lambda z + (1-\lambda)z_0$, $0 \leq \lambda \leq 1$.

$$\begin{aligned} ||c - z_0||^2 &\leq ||c - \lambda z - (1-\lambda)z_0||^2 \qquad \text{given that } z_0 \text{ is the nearest point to } c, \\ &= ||(c - z_0) + \lambda(z_0 - z)||^2 \\ &= ||c - z_0||^2 + 2\lambda(c-z_0)(z_0 - z) + \lambda^2||z_0 - z||^2 \end{aligned}$$

Simplifying, we obtain:

$$2\lambda\,(c - z_0)\,(z_0 - z) + \lambda^2||z_0 - z||^2 \geq 0$$

dividing by $\lambda > 0$, we have:

$$2(c - z_0)\,(z_0 - z) + \lambda||z_0 - z||^2 \geq 0.$$

If we make $\lambda$ tend to 0, we have:

$$(c - z_0)\,(z_0 - z) \geq 0.$$

Now we prove that $x = (c - z_0)$ is the gradient of the separating hyperplane we are looking for.
According to construction $x \neq 0$; we prove that $x$ separates $c$ from any $z \in S$, i.e., that $cx > zx$.
What follows holds for each $z \in S$ :

$$\begin{aligned} 0 &\leq (c - z_0)\,(z_0 - z) \\ &= (c - z_0)\,(c - c + z_0 - z) \\ &= (c - z_0)\,(c - z) + (c - z_0)\,(z_0 - c) \\ &= \quad x \quad\ (c - z) - ||c - z_0||^2 \end{aligned}$$

thus, we conclude that $cx \geq zx + \varepsilon$, where $\varepsilon = ||c - z_0||^2 > 0$. ♦

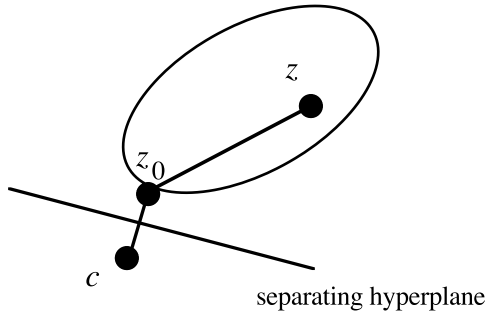


Fig. 4: separating hyperplane

Example: geometric interpretation of the linear classifier

Recalling the linear classifier example introduced previously, we can interpret its meaning geometrically. The $m$ available observations can be viewed as points in $\boldsymbol{R}^n$. The coefficients $x_1$, $x_2$, $\ldots$,$x_n$ and $y$ actually represent the gradient and the known term of a hyperplane separating the points belonging to the first class from those belonging to the second. Figure 5 shows a set of red and blue observations and the hyperplane separating them.

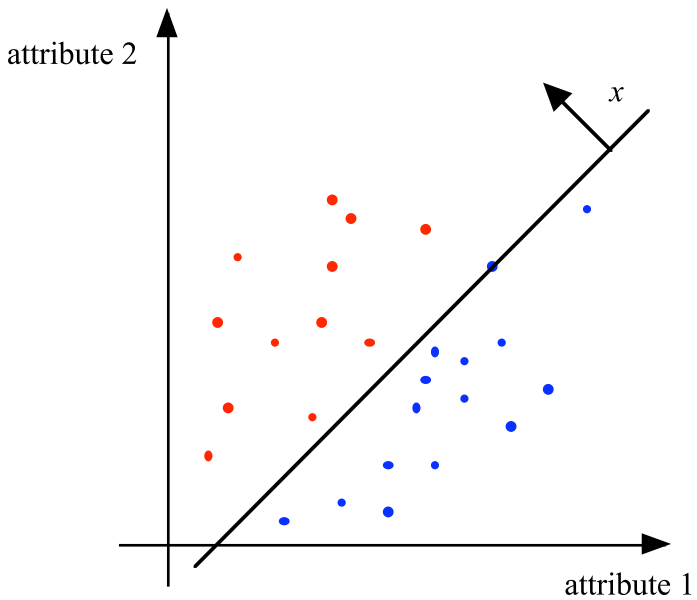


Fig. 5: linear classifier, separating hyperplane

## 2. Pairs of dual problems

As we saw for problems on a graph (shortest paths, maximum flow, minimum cost flow), in order to provide a solution algorithm it is necessary to characterize the optimal solution. With each Linear Programming problem we may associate another problem – called *dual* – providing valuable information concerning the solution. In order to understand its properties, we may try to formulate the dual problem in two fairly intuitive cases.

Example: a diet problem

In a chicken-farming concern the bird-feed consists in two types of cereals, $A$ and $B$. The daily ration must satisfy some nutritive requirements: it must contain at least 11 units of carbohydrates, 20 units of proteins and 9 units of vitamins per weight unit. The unit content of carbohydrates, proteins and vitamins in $A$ and $B$, the unit cost (in cents) as well as the minimum daily requirements are indicated in the following table:

| | *A* | *B* | min. daily requirements |
|---|---|---|---|
| carbohydrates | 2 | 2 | 11 |
| proteins | 4 | 2 | 20 |
| vitamins | 1 | 3 | 9 |
| unit cost | 12 | 16 | |

Let $x_1$ and $x_2$ be respectively the number of units of cereal *A* and *B* employed in the bird-feed. The number of units of carbohydrates contained in the feed is given by: $2x_1+ 2x_2$. Since the minimum requirement of carbohydrates is of 11 units, the result must be: $2x_1+2x_2\geq 11$. Similarly, for protein units the result must be $4x_1+2x_2\geq 20$ and for vitamin units it must be $x_1+3x_2\geq 9$. Obviously, to the three preceding constraints we must add the non-negativity conditions of decision variables $x_1$, $x_2\geq 0$. Finally, the cost of the decision $(x_1, x_2)$ is given by: $12x_1+16x_2$. Therefore, the minimum cost diet is given by a solution of the following *LP* problem:

$$\begin{array}{llcl}
\min & 12x_1 + 16x_2 & & \\
 & 2x_1 + 2x_2 & \geq & 11 \\
 & 4x_1 + 2x_2 & \geq & 20 \\
 & x_1 + 3x_2 & \geq & 9 \\
 & x_1, x_2 & \geq & 0
\end{array}$$

Observe how the examined problem is naturally associated with another problem that may be called the *problem of the chicken pill seller*. The purpose is to fix the selling prices of pills respectively containing carbohydrates, proteins and vitamins, so that the proceeds of the sale are maximized and the prices are competitive. In other words, the chicken breeder must think it is not disadvantageous to buy the pills instead of the cereals *A* and *B*. Let us suppose that each pill contains one unit of the corresponding nutritive element.

Let $y_1$, $y_2$ and $y_3$ be the unit selling prices of carbohydrate, protein and vitamin pills, respectively. Because the breeder must perceive the pill diet as being not more expensive than the cereal diet, the result must be: $2y_1+4y_2+y_3\leq 12$, i.e., the cost of the pills needed to be equivalent (from a nutritive point of view) to a ration of cereal *A* must not be greater than 12 cents. A similar reasoning is valid for cereal *B*: $2y_1+2y_2+3y_3\leq 16$. Selling prices must be non-negative and the proceeds of the sale are given by $11y_1+20y_2+9y_3$ (in fact, note that 11, 20 and 9 are the smallest number of carbohydrate, protein and vitamin pills needed to properly feed a chicken). Thus, the problem of the chicken pill seller is the following:

$$\begin{array}{llcl}
\max & 11y_1 + 20y_2 + 9y_3 & & \\
 & 2y_1 + 4y_2 + y_3 & \leq & 12 \\
 & 2y_1 + 2y_2 + 3y_3 & \leq & 16 \\
 & y_1, y_2, y_3 & \geq & 0
\end{array}$$

The two problems are summarized in Figure 6, which points out the relations between the variables of a problem and the constraints of another problem as well as between the known terms and the objective function.

$$\begin{array}{cccc}
 & \begin{array}{cc} x_1 & \quad x_2 \end{array} & & \max \\
\begin{array}{c} y_1 \\ y_2 \\ y_3 \end{array} & \boxed{\begin{array}{cc} 2 & \quad 2 \\ 4 & \quad 2 \\ 1 & \quad 3 \end{array}} & \geq & \boxed{\begin{array}{r} 11 \\ 20 \\ 9 \end{array}} \\
 & \leq & & \\
\min & \boxed{\begin{array}{cc} 12 & \quad 16 \end{array}} & &
\end{array}$$

Fig. 6

**Exercise**

Formulate the two problems by using the modeling language introduced previously (note how the data file is common to both problems). Compare the values of optimal solutions. Point out the relation between a problem's solution and the constraint slackness of the other problem. Try to vary, in turn, one of the coefficients of the known term vector by one unit and test by what amount the value of the optimal solution varies.

Example: a transportation problem

A chemical industry produces VCM in $n$ factories and, before selling it, stores it in $m$ warehouses. Let there be:

- $a_i$ the quantity of VCM produced by factory $i$ ($i$=1,…,$n$);
- $d_j$ the capacity of warehouse $j$ ($j$=1,…,$m$), assuming that $\sum_{i=1}^{n} a_i = \sum_{j=1}^{m} d_j$, i.e., the total production equals the total capacity of warehouses;
- $b_{ij}$ the transportation cost of a product unit from $i$ to $j$.

The purpose is to transport the produced VCM from factories to warehouses so as to minimize the transportation cost.

We denote by $y_{ij}$ the decision variables corresponding to the quantity of VCM transported from $i$ to $j$. In addition to the obvious non-negativity conditions of decision variables $y_{ij} \geq 0$, constraints impose that all produced VCM can be transported from factory $i$, that is: $\sum_{j=1}^{m} y_{ij} = a_i$, and that the VCM transported to warehouse $j$ can saturate its capacity, that is: $\sum_{i=1}^{n} y_{ij} = d_j$. The global transportation cost relative to a solution $y_{ij}$, $i$=1,…,$n$, $j$=1,…,$m$, is $\sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} y_{ij}$. The problem can be equally viewed as a specific flow problem in which the network is defined by as many nodes as factories and warehouses and in which there exist arcs going from each factory to each warehouse associated with cost $b_{ij}$. A node corresponding to a factory $i$ is associated with a flow supply $-a_i$, and a node associated with a warehouse $j$ is associated with a flow demand equal to $d_j$. Hence, the problem can be formulated as follows:

$$
\begin{array}{llll}
\min & \sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} y_{ij} & & \\
& -\sum_{j=1}^{m} y_{ij} = -a_i & & i=1,\dots,n, \\
& \sum_{i=1}^{n} y_{ij} = d_j & & j=1,\dots,m, \\
& y_{ij} \geq 0 & & i=1,\dots,n,\ j=1,\dots,m.
\end{array}
$$

In this case too it is possible to associate with the transportation problem data a further problem, that can be interpreted as a kind of "outsourcing" of the transportation process. A transportation company offers to buy the VCM from factory $i$ at unit price $\lambda_i$ and resell it to warehouse $j$ at unit price $\mu_j$. Purchase and sale prices become the decision variables of the transportation company problem. The company wishes to maximize the resulting earnings, that are given by $-\sum_{i=1}^{n} a_i \lambda_i + \sum_{j=1}^{m} d_j \mu_j$. Obviously, purchase and sale prices will have to be competitive in comparison with the transportation costs the chemical industry would have to bear. According to the offer made by the transportation company, the industry would pay the transportation of one VCM unit from $i$ to $j$ ($-\lambda_i + \mu_j$). Hence, the result must be $-\lambda_i + \mu_j \leq b_{ij}$. Therefore, the transportation company problem can be formulated as the following *LP* problem:

$$\begin{array}{llll} \max & -\sum_{i=1}^{n} a_i\lambda_i & +\sum_{j=1}^{m} d_j\mu_j & \\ & -\lambda_i & +\ \mu_j \ \leq\ b_{ij} & i=1,\dots,n,\ j=1,\dots,m. \end{array}$$

The relations between the two problems, assuming the case of two factories and three warehouses, are illustrated in Figure 7.

$$\begin{array}{c|cccccc|c|c|} & y_{11} & y_{12} & y_{13} & y_{21} & y_{22} & y_{23} & & \max \\ \lambda_1 & -1 & -1 & -1 & & & & & -a_1 \\ \lambda_2 & & & & -1 & -1 & -1 & & -a_2 \\ & & & & & & & = & \\ \mu_1 & 1 & & & 1 & & & & d_1 \\ \mu_2 & & 1 & & & 1 & & & d_2 \\ \mu_3 & & & 1 & & & 1 & & d_3 \\ & & & & \wedge\backslash & & & & \\ \min & b_{11} & b_{12} & b_{13} & b_{21} & b_{22} & b_{23} & & \end{array}$$

Fig. 7

Note that the matrix illustrated in the figure is simply the node-arc incidence matrix of the graph relative to the problem.

**Exercise**

Consider the shutter concern problem we discussed previously. Formulate the problem of a temporary work agency wishing to rent workers from the proprietress of the shutter enterprise. Point out the relations between variables of one problem and constraints of the other.

As it can be noticed, in the examples just described it is possible, in a quite natural way, to associate with the examined *LP* problem a new problem bearing a narrow relation with it. In matricial form, the problem of the chicken breeder who has to mix the bird-feeds becomes:

$$(2.1) \qquad P: \qquad \begin{array}{l} \min cx \\ Ax \geq b \\ x \ \geq 0 \end{array}$$

where vectors $c$ and $b$ are the feed costs and the minimum requirements of each nutritive element, whereas matrix $A$ provides, element by element, the nutritive supply of each type of feed. Formulated in matricial form as well, the chicken pill seller problem associated with the latter problem gives:

$$(2.2) \qquad D: \qquad \begin{array}{l} \max yb \\ yA \leq c \\ y \ \geq 0. \end{array}$$

This relation holds not only for the specific diet example, but for any *LP* problem. Problems $P$ and $D$ represent a *pair of dual problems* and, in such case, they are called *symmetric pair* because in both problems constraints are inequality constraints and in both problems variables are restricted in sign. $P$ is called the *primal* and $D$ the *dual*, even if their role is wholly interchangeable, as clearly shown by the following theorem.

**Theorem 2.1**

*The dual of the dual is the primal.*

Proof

Problem $D$ : $\max\{yb: yA \le c, y \ge 0\}$ can be equivalently[(1)] written – using the transformation rules summarized at the beginning of this chapter – as $\min\{-yb: -yA \ge -c, y \ge 0\}$.
The dual of this problem (see (2.1) and (2.2)) is $\max\{-cx: -Ax \le -b, x \ge 0\}$, which is equivalent to problem $P$. ♦

Now let us consider the example of the other pair of dual problems relative to the VCM transportation problem. The problem facing the owner of the factories and consisting in minimizing the transportation costs can be viewed in matricial form as follows:

$$\text{(2.3)}\qquad \begin{array}{l} \min\ yb \\ \quad yA = c \\ \quad y \ \ \ge 0. \end{array}$$

where $A$ is the node-arc incidence matrix of the complete bipartite graph between factories and warehouses, and the vector of coefficients $c_i$ corresponds to VCM supply with negative sign if index $i$ refers to a factory and to VCM demand if, otherwise, it refers to a warehouse.
The carrier problem in matricial form is:

$$\text{(2.4)}\qquad \begin{array}{l} \max\ cx \\ \quad Ax \le b \end{array}$$

where the vector of variables $x$ is obtained by juxtaposing the two vectors $\lambda$ and $\mu$, i.e., $x=[\lambda_i,\mu_i]$.
Therefore, we have the following *asymmetric pair*, thus called because in one problem we have free variables and inequality constraints, whereas in the other we have equality constraints and variables restricted in sign:

$$P: \begin{array}{l} \max\ cx \\ \quad Ax \le b \end{array} \qquad\qquad D: \begin{array}{l} \min\ yb \\ \quad yA = c \\ \quad y \ \ \ge 0. \end{array}$$

Actually, symmetric pair and asymmetric pair are equivalent. In fact, $P$ can be equivalently written as $\max\{c(x^+-x^-): A(x^+-x^-) \le b, x^+, x^- \ge 0\}$, and by applying the definition of symmetric pair we obtain the dual: $\min\{yb: yA \ge c, -yA \ge -c, y \ge 0\}$, which is equivalent to problem $D$ of the asymmetric pair.

**Exercise**
Prove the reverse, i.e., that the definition of symmetric pair can be obtained from the definition of asymmetric pair.

**Exercise**
Prove the theorem 2.1 for the asymmetric pair.

Given a $LP$ problem in any form, we can always derive its dual by referring to one of the four forms we saw through the equivalent transformations examined at the beginning of this chapter and by obtaining the corresponding dual in an "automatic" way.

Example: passage to the dual (1)
Consider the following $LP$ problem:

$$\begin{array}{lllll} \min & 3x_1 - 4x_2 & & + x_4 & \\ & 2x_1 - 2x_2 + 4x_3 - x_4 & \ge & -3 & \\ & -x_1 + x_2 + x_3 & = & 3 & \\ & x_2 \ge 0,\ x_4 \le 0 & & & \end{array}$$

---

(1) The equivalence between the two problems is preserved except for the value sign of the objective function; equivalent problems have the same feasible region and the same set of optimal solutions. This observation also applies to the next equivalence in the proof.

As a model we may choose, for instance, problem *P* of the asymmetric pair. Then we have to transform the objective function into maximum, all constraints into smaller than or equal to, and render sign constraints explicit. The equivalent problem is:

$$
\begin{array}{lllll}
-\max & -3x_1+4x_2 & & -x_4 & \\
 & -2x_1+2x_2-4x_3 & +x_4 & \leq & 3 \\
 & -x_1+x_2+x_3 & & \leq & 3 \\
 & x_1-x_2-x_3 & & \leq & -3 \\
 & -x_2 & & \leq & 0 \\
 & & x_4 & \leq & 0
\end{array}
$$

In order to write the dual we introduce 5 variables $(y_1, y_2, y_3, y_4, y_5)$ corresponding to the 5 constraints of the primal and write the dual respecting the scheme of the asymmetric pair:

$$
\begin{array}{llllllll}
\min & 3y_1 & +3y_2 & -3y_3 & & & & \\
 & -2y_1 & -y_2 & +y_3 & & & = & -3 \\
 & 2y_1 & +y_2 & -y_3 & -y_4 & & = & 4 \\
 & -4y_1 & +y_2 & -y_3 & & & = & 0 \\
 & y_1 & & & & +y_5 & = & 1 \\
 & \multicolumn{7}{l}{y_1, y_2, y_3, y_4, y_5 \geq 0}
\end{array}
$$

Alternatively – and much more rapidly, since no intermediate transformations are required – the dual of any *LP* problem can be written by applying the *P-D* correspondences indicated in the table of Figure 8, where $A_i$ indicates the *i*-th row of matrix *A*, and $A^i$ indicates the *i*-th column.

| min | max |
|---|---|
| variables | constraints |
| constraints | variables |
| cost vector $c$ | known term vector $b$ |
| known term vector $b$ | cost vector $c$ |
| $A_ix \geq b_i$ | $y_i \geq 0$ |
| $A_ix \leq b_i$ | $y_i \leq 0$ |
| $A_ix = b_i$ | $y_i \gtrless 0$ |
| $x_i \geq 0$ | $yA^i \leq c_i$ |
| $x_i \leq 0$ | $yA^i \geq c_i$ |
| $x_i \gtrless 0$ | $yA^i = c_i$ |

Fig. 8: Primal/Dual correspondence table

Example: passage to the dual (2)

Let us try to apply the correspondence table directly to the problem of the preceding example. Considering the constraints imposed on $x_2$ and $x_4$ in the form of sign constraints (being thus implicit), we have to introduce only two dual variables $(y_1, y_2)$ corresponding to the primal problem's constraints. The dual will have 4 constraints:

$$
\begin{array}{lllll}
\max & -3y_1 & +3y_2 & & \\
 & 2y_1 & -y_2 & = & 3 \\
 & -2y_1 & +y_2 & \leq & 4 \\
 & 4y_1 & +y_2 & = & 0 \\
 & -y_1 & & \geq & 1 \\
 & \multicolumn{4}{l}{y_1 \geq 0,\ y_2 \gtrless 0}
\end{array}
$$

In order to understand which is the type of constraint (≤,≥ or =), we must refer to the corresponding primal variable: $x_1$ and $x_3$ are free (first column corresponding to the minimization problem and last row in the table of Figure 8), so the corresponding constraints of the dual are equality constraints. The second constraint is of smaller than or equal to because $x_2$ takes non-negative values and, conversely, the fourth constraint is of greater than or equal to. Similar considerations apply to signs of dual variables, for which we have to refer to the constraint type of the primal.

The dual problem obtained from the application of the table is different from the dual we obtained by referring to one of the schemes relative to pairs of dual problems. However, by applying the transformation rules seen at the beginning of this chapter, we can easily prove that the two duals are equivalent. In particular, we can do so by undoubling the free variable, by introducing appropriate slack variables and by changing the direction of the objective function.

Example

Let us consider the following *LP* problem (where we imply that, if not explicitly remarked, a variable is not restricted in sign, such as, for instance, variable $x_2$).

$$\begin{array}{llll}
\max & 12x_1 + 7x_2 & & \\
 & 5x_1 + 7x_2 & = & 8 \\
 & 4x_1 + 2x_2 & \geq & 15 \\
 & 2x_1 + x_2 & \leq & 3 \\
 & x_1 & \geq & 0
\end{array}$$

Since it is a maximization problem, we take as reference the second column in the table of Figure 8. Treating the last constraint as an (implicit) sign constraint, the dual has three variables and two constraints, thus giving:

$$\begin{array}{llll}
\min & 8y_1 + 15y_2 + 3y_3 & & \\
 & 5y_1 + 4y_2 + 2y_3 & \geq & 12 \\
 & 7y_1 + 2y_2 + y_3 & = & 7 \\
 & y_2 \leq 0, y_3 \geq 0 & &
\end{array}$$

As it emerges from the examples described in the preceding paragraph, but also on the basis of the numerical results obtained in solving the diet problem, problems *P* and *D* are not merely "syntactically" related. The following theorem provides an initial relation between the objective function values of the two problems. If not differently indicated, from now on we will use the asymmetric duality form, but of course the obtained results are independent from the specific form being used.

**Theorem 2.2** (W*eak duality theorem*)

*If P and D admit the feasible solutions $\bar{x}$ and $\bar{y}$, respectively, then $c\bar{x} \leq \bar{y}b$.*

Proof

Let $\bar{x}$ and $\bar{y}$ be feasible solutions for *P* and *D*, respectively. We have:

$$\begin{array}{lll}
\bar{y}A = c & \Rightarrow & c\bar{x} = \bar{y}A\bar{x}, \\
A\bar{x} \leq b,\ \bar{y} \geq 0 & \Rightarrow & \bar{y}A\bar{x} \leq \bar{y}b.
\end{array}$$

Hence $c\bar{x} \leq \bar{y}b$. ♦

In general, if *P* and *D* are non-empty, it can be stated that:

$$\max\{cx: Ax \le b\} \le \min\{yb: yA=c, y \ge 0\}.$$

The weak duality theorem represents in itself a very useful tool to recognize an optimal solution. In fact, in case we are provided with a feasible solution for the primal and a feasible solution for the dual, both having the same value in the respective objective functions, we may be sure that none of the two can be improved. This intuitive consideration is formalized in the following corollary.

**Corollary 2.3**

*If $\bar{x}$ and $\bar{y}$ are feasible solutions for P and D, respectively, and $c\bar{x} = \bar{y}b$, then $\bar{x}$ and $\bar{y}$ are optimal solutions.*

Proof

For each solution $x$ feasible for $P$ and for each solution $y$ feasible for $D$, the result is:

$cx = \bar{y}Ax \le \bar{y}b = c\bar{x}$ hence $\bar{x}$ is optimal for $P$;

$yb \ge yA\bar{x} = c\bar{x} = \bar{y}b$ hence $\bar{y}$ is optimal for $D$. ♦

**Corollary 2.4**

*If P has unbounded optimality, then D is empty.*

In fact, even if $D$ contained only one solution, its value would represent a bound to the value of the solutions of $P$.

Example: Shutters and Frames

Let us go back to the Shutters-and-Frames concern problem. We already proved geometrically how vertex $x$=(2,6) is the optimal solution, with 360 as value of the objective function; now we can provide a formal proof of the optimality of (2, 6), by using the consequences of the weak duality theorem.

The problem's dual, in asymmetric form, is:

$$\begin{array}{llcccccl}
\min & 4y_1 & + & 12y_2 & + & 18y_3 & & \\
 & y_1 & + & & & 3y_3 & - y_4 & = 30 \\
 & & & 2y_2 & + & 2y_3 & - \quad y_5 & = 50 \\
 & \multicolumn{7}{l}{y_1, y_2, y_3, y_4, y_5 \ge 0.}
\end{array}$$

We consider the solution $y$=(0,15,10,0,0) (later we will see how to obtain it). We can immediately verify that such solution is feasible and that its value is given by: 12·15 + 18·10 = 360. The weak duality theorem and the corollary 2.3 enable us to state that solution $x$=(2,6) is optimal, and so is solution $y$.

**Exercise**

Prove the $P$-$D$ relations of Figure 8, starting from the definitions of symmetric pair and of asymmetric pair, respectively.

### *2.1 *Geometric interpretation of the dual problem*

Consider the asymmetric pair:

$$P: \begin{array}{l} \max\ cx \\ \quad Ax \le b \end{array} \qquad D: \begin{array}{l} \min\ yb \\ \quad yA = c \\ \quad y \ge 0. \end{array}$$

As we had occasion to see, the feasible region of $P$ is a convex polyhedron obtained by drawing the intersection of half-spaces defined by equality constraints. The feasible region of $D$ can also be easily interpreted from a geometric viewpoint, although the viewpoint must be changed. By writing the constraints of $D$ in full we get:

$$y_1A_1 + y_2A_2 + \ldots + y_mA_m = c$$

In practice, these constraints mean that we intend to express $c$ as linear combination of gradients of the constraints of $P$ ($A_1, A_2, \ldots, A_m$), using $y_1, y_2, \ldots, y_m$ as multipliers. Multipliers must be non-negative, and among all solutions we want the one minimizing the weighted sum of multipliers with coefficients $b_1, b_2, \ldots, b_m$.

<u>Example</u>

Consider the pair of dual problems:

$$\begin{array}{lrrrrl} P: & \max & 3x_1 & - x_2 & & \\ & & x_1 & + x_2 & \leq & 4 \\ & & -x_1 & + x_2 & \leq & 5 \\ & & & - x_2 & \leq & 2 \end{array} \qquad \begin{array}{lrrrrrl} D: & \min & 4y_1 & + 5y_2 & + 2y_3 & & \\ & & y_1 & - y_2 & & = & 3 \\ & & y_1 & + y_2 & - y_3 & = & -1 \\ & & & y_1, y_2, y_3 & & \geq & 0 \end{array}$$

In Figure 9 the pair is presented in tabular form, whereas in Figure 10 we provide the geometric representation of $P$.

| | $x_1$ | $x_2$ | | min |
|---|---|---|---|---|
| $y_1$ | 1 | 1 | | 4 |
| $y_2$ | -1 | 1 | $\leq$ | 5 |
| $y_3$ | 0 | -1 | | 2 |
| | || | | |
| max | 3 | -1 | | |

Fig. 9

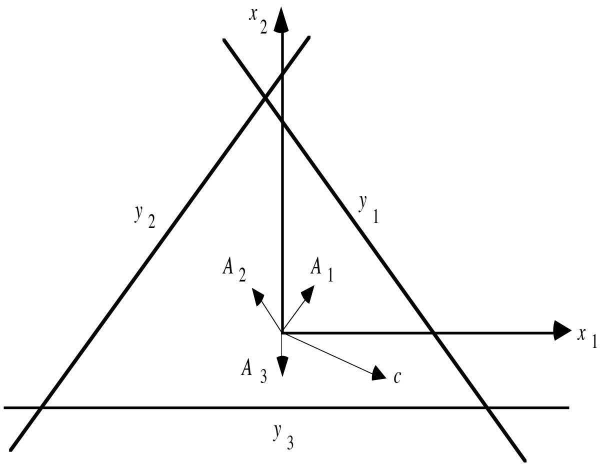


Fig. 10

We have to express vector $c$ as a non-negative linear combination of rows $A_1$, $A_2$ and $A_3$; evidently, such combination must use $A_1$ and $A_3$, in fact $c$ belongs to the cone generated by $A_1$ and $A_3$. Note as well that the optimal solution of $P$ falls on the lower right vertex of the triangle representing the feasible region and in that vertex the active constraints are precisely $A_1$ and $A_3$.

## 3 Solution of a Linear Programming problem

In this paragraph we approach the solution of a Linear Programming problem. The treatment will be quite informal and oriented towards the description of a simple algorithm. After introducing the algorithm's idea, we will start to go into details by pointing out the more formal aspects and also by means of a geometric interpretation.

### *3.1 Feasible growth directions*

We wish to solve a *LP* problem in the form:

$$P: \begin{array}{l} \max \; cx \\ \quad Ax \le b \end{array} \qquad\qquad D: \begin{array}{l} \min \; yb \\ \quad yA = c \\ \quad y \;\; \ge 0. \end{array}$$

Let $\bar{x} \in \mathbf{R}^n$ be a feasible solution for $P$; the question is whether such solution is optimal or it is possible to improve it. Taking advantage of the fact that the problem's feasible region is a convex set, notably a polyhedron, we can state that if there exists a point $x'$ better than $\bar{x}$, it must be possible to express it in the form:

$$x' = \bar{x} + \lambda\xi \tag{3.1}$$

where $\lambda>0$ is a scalar called *displacement step*, and $\xi$ is a vector of $\mathbf{R}^n$ called *displacement direction*. So, we can say that $\bar{x}$ is improvable if and only if there exists a direction $\xi$ for which, by an appropriate choice of the displacement step $\lambda>0$, point $x' = \bar{x} + \lambda\xi$ still falls in the interior of the feasible region and has objective function value $cx'>cx$. In other terms, $\bar{x}$ is optimal if and only if there exists no such direction $\xi$. Let us define this characteristic formally:

**Definition 3.1**

Considering a problem $P$: $\{\max cx\text{: } Ax \le b,\ x \in \mathbf{R}^n\}$ and a feasible point $\bar{x}$, we say that $\xi \in \mathbf{R}^n$ is a *feasible growth direction* if there exists a scalar $\lambda>0$ such that:

i) $c(\bar{x} + \lambda\xi) > c\bar{x}$

ii) $\bar{x} + \lambda\xi$ is feasible.

Let us concentrate on the search for a feasible growth direction $\xi$ for point $\bar{x}$. Imposing condition i) of definition 3.1 and recalling that $\lambda>0$, we obtain:

$$c(\bar{x} + \lambda\xi) > c\bar{x} \;\Rightarrow c\bar{x} + \lambda c\xi > c\bar{x} \Rightarrow c\xi > 0$$

The development of condition ii) implies a few distinctions. In fact, if point $\bar{x}$ were not adherent to any constraint of the polyhedron (hence $A\bar{x}<b$), we could evidently move in any direction – by an appropriate displacement step $\lambda>0$ – without risking going out of the feasible region. Otherwise, if $\bar{x}$ touched some constraint (hence, if the set of active constraint indexes $I(\bar{x})\neq\varnothing$), then even an infinitesimal displacement in the direction of one of the gradients of the constraints to which we keep adherent would be sufficient to lead us out of the feasible region. Hence, feasibility conditions for the choice of $\xi$ are:

$$A_I x' = A_I(\bar{x} + \lambda\xi) \le b_I$$

where $A_I$ and $b_I$ respectively indicate the submatrix of $A$ and the subvector of $b$ relative to active constraints $I(\bar{x})$. Developing the product and recalling the definition of active constraints, which precisely implies that $A_i\bar{x} = b_i$ for $i \in I(\bar{x})$, we have:

$$A_I(\bar{x} + \lambda\xi) \leq b_I \Rightarrow A_I\bar{x} + \lambda A_I\xi \leq b_I \Rightarrow A_I\xi \leq 0.$$

Summing up what we have saw up to now, we can say that optimality conditions for $P$ at point $\bar{x}$ are the **non**-existence of a solution for the following system.

$$\begin{array}{ll} & c\xi > 0 \\ (3.2) & \\ & A_I\xi \leq 0 \end{array}$$

If there exists such a direction $\xi$, we can move by a step $\lambda>0$. When moving along $\xi$, we must take care not to go out of the feasible region, and particularly we must keep control of constraints to which $\bar{x}$ is not adherent, i.e., constraints with indexes $\bar{I}(\bar{x})=\{1,\ldots,m\} \setminus I(\bar{x})$, called non-active constraints. In particular, we have to choose $\lambda$ in such a way that for these constraints it becomes:

$$A_i(\bar{x} + \lambda\xi) \leq b_i \quad \forall i \in \bar{I}(\bar{x})$$

hence

$$A_i\bar{x} + \lambda A_i\xi \leq b_i \quad \forall i \in \bar{I}(\bar{x}).$$

Therefore, for those indexes $i$ for which $A_i\xi \leq 0$ constraints will always be respected, whereas for those for which $A_i\xi > 0$ $\lambda$ a choice must be made such that:

$$\lambda \leq \frac{b_i - A_i\bar{x}}{A_i\xi} \qquad \forall i \in \bar{I}(\bar{x}),\ A_i\xi > 0\ .$$

Example: Shutters and Frames

Let us consider again the shutter concern example and let us analyze the problem of determining a feasible growth direction for the 4 points indicated in Figure 11.

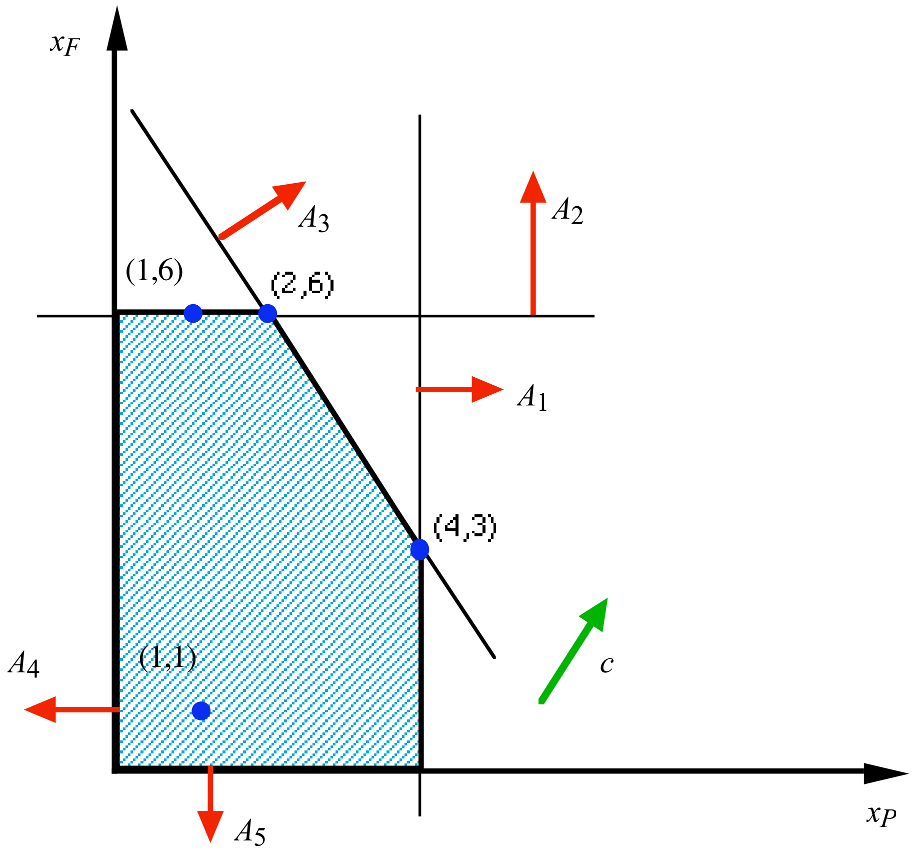


Fig. 11

For point $\bar{x} = \begin{bmatrix}1\\1\end{bmatrix}$ we have $I(\bar{x}) = \varnothing$, so we are free to choose $\xi$ as we judge it better, without risking going out of the feasible region by an appropriately large displacement step. Consequently, the choice of $\xi$ is merely influenced by the gradient of the objective function $c$. For example, if we take $\xi = \begin{bmatrix}2\\1\end{bmatrix}$, we have $c\xi = 2{\cdot}30 + 1{\cdot}50 > 0$, hence $\xi$ is a feasible growth direction and point $\bar{x}$ is not optimal. Wishing to determine a displacement step $\lambda>0$ guaranteeing the feasibility of the new point $\bar{x}+\lambda\xi$, we have to choose $\lambda$ in such a way that the first three constraints are respected; therefore:

$$\lambda \leq \frac{4 - [1,0]\bar{x}}{[1,0]\xi} \Rightarrow \lambda \leq \frac{3}{2}$$

$$\lambda \leq \frac{12 - [0,2]\bar{x}}{[0,2]\xi} \Rightarrow \lambda \leq 5$$

$$\lambda \leq \frac{18 - [3,2]\bar{x}}{[3,2]\xi} \Rightarrow \lambda \leq \frac{13}{8}$$

whereas we should not care about the fourth and the fifth constraint because $A_4\xi \leq 0$ and $A_5\xi \leq 0$. Any value of $\lambda \leq 3/2$ guarantees the feasibility of the new solution and, since it is a maximization and a feasible growth direction problem, we tend to choose $\lambda$ as great as possible.

At point $\bar{x} = \begin{bmatrix}1\\6\end{bmatrix}$ we have $I(\bar{x}) = \{2\}$, so in choosing $\xi$ we have to be careful that $[0,2]\xi \leq 0$. A feasible growth direction may be, for example, $\xi = \begin{bmatrix}1\\0\end{bmatrix}$. Note that, since $A_2\xi=0$, each displacement along $\xi$ guarantees that the new point keeps adherent to the second constraint. Otherwise, at point

$\bar{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ we have two active constraints, $I(\bar{x}) = \{1,3\}$. We can easily verify that $\xi = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ is a feasible growth direction.

Determining a feasible growth direction at point $\bar{x} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ seems to be not so easy; indeed, remembering that such solution is optimal, we should expect that none exists. Active constraints are $I(\bar{x}) = \{2,3\}$, and existence conditions for a feasible growth direction are given by the following system

$$\begin{array}{llcl} 30\xi_1 & + 50\xi_2 & > & 0 \\ & \phantom{+}2\xi_2 & \leq & 0 \\ 3\xi_1 & + 2\xi_2 & \leq & 0 \end{array}$$

which admits no solution.

**Exercise**

Compute the maximum value of step $\lambda$ for the feasible growth directions $\xi = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\xi = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ for the points $\bar{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$ and $\bar{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, respectively.

**Exercise**

Draw on Figure 11 the feasible growth directions and the new feasible points obtained after having applied the displacement steps determined in the preceding exercise.

At this point, it is intuitively clear how the crucial phase in the solution of a *LP* problem consists in determining a feasible growth direction, if one exists. So, we continue this rather informal development and examine in detail the problem of determining $\xi$ as formulated by system (3.2). Note that such system can be interpreted in terms of *LP*; particularly, the first constraint can be considered as an objective function to be maximized, whereas the other ones as real constraints. Therefore, we have:

$$P': \quad \begin{array}{l} \max\ c\xi \\ \quad A_I\xi \leq 0 \end{array}$$

Note that, if there exists any feasible solution $\xi'$ for $P'$ such that $c\xi'>0$, then $P'$ has unbounded optimality. In fact, if $\xi'$ is feasible, so is also $\alpha\xi'$ for each $\alpha\geq 0$, and by letting $\alpha$ grow to infinity the value of objective function $\alpha c\xi'$ grows to infinity as well.

As for any *LP* problem, we can write its dual: we introduce a dual variable $\eta_i$ for each $i\in I(\bar{x})$:

$$D': \quad \begin{array}{l} \min\ \eta 0 \\ \quad \eta A_I = c \\ \quad \eta \quad \geq 0. \end{array}$$

Note how the objective function of the dual $D'$ is useless because any feasible solution has value 0, and actually the problem is nothing but a feasibility problem. The primal $P'$ has a peculiar characteristic too: since the known term vector is composed of all 0, $\xi=0$ is always a feasible solution. Due to the weak duality theorem, if there exists a solution for $D'$, then $\xi=0$ is an optimal solution for $P'$. Conversely, if $P'$ has unbounded optimality, then $D'$ has no solution.

### *3.2 Algorithm scheme*

According to what we saw in the preceding paragraph, we intuitively understand how to build an algorithm for solving a *LP* problem having the form

$$P: \begin{array}{l} \max\ cx \\ \quad Ax \leq b \end{array}$$

starting from a feasible solution $\bar{x}$. The idea on which the algorithm is based can be summarized as follows.

As long as at current point $\bar{x}$ there exists a feasible growth direction $\xi$, we move along $\xi$ by a displacement step such as to guarantee the feasibility of the new point. When it is no longer possible to find a feasible direction, the algorithm stops and solution $\bar{x}$ is optimal.

Let us more deeply examine the latter aspect concerning the algorithm's stopping condition.

**Theorem 3.1**

Consider problem *P*, a feasible solution $\bar{x}$ of *P* and the problems related to the search for feasible growth directions *P'* and *D'*. If *D'* has a solution, then $\bar{x}$ is an optimal solution for *P*.

Proof

Let $\bar{\eta}$ be a feasible solution for *D'*. We consider the dual of *P*:

$$D: \begin{array}{l} \min\ yb \\ \quad yA = c \\ \quad y \geq 0. \end{array}$$

and build a solution $\bar{y}$ in the following way:

$$\bar{y}_i = \begin{cases} \bar{\eta} & \text{if } i \in I(\bar{x}) \\ 0 & \text{otherwise} \end{cases} .$$

Thus built, the solution is feasible for *D*, in fact $\bar{y} \geq 0$ and $\bar{y}A = \bar{\eta}A_I = c$. Now let us estimate the value of $\bar{y}$ in the objective function of *D*:

$$\bar{y}b = \sum_{i=1}^{m} \bar{y}_i b_i = \sum_{i \in I(\bar{x})} \bar{\eta}_i b_i = \sum_{i \in I(\bar{x})} \bar{\eta}_i A_i \bar{x} = c\bar{x}$$

where the last equality but one derives from the fact that for active constraints we have $A_i\bar{x}=b_i$ and the last one from the fact that $\bar{\eta}$ is a feasible solution for *D'*. In this way, we have built a dual feasible solution $\bar{y}$ having the same value as the primal feasible solution $\bar{x}$ and, as a consequence of the weak duality theorem, we can state that $\bar{x}$ is optimal. ♦

The critical point we have left aside, and whose proof we keep for the next in-depth paragraph, concerns the fact that either *P'* has unbounded solution (hence there exists a feasible growth direction) or there exists a solution of *D'*. This is demonstrated by the so-called Farkas's Lemma, a real "engine" of Linear Programming.

Example: Shutters and Frames

Consider again the shutter concern example and the points illustrated in Figure 11. We formulate problems *P'* and *D'* except for point $\bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ which (being in the interior of the feasible region) is clearly non-optimal and for which the choice of a feasible growth direction is totally free. In the case of $\bar{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$, set $I(\bar{x})$ only contains element 2, so we have $P'$ with only one constraint and $D'$ with only one variable and two equality constraints:

$$P': \quad \begin{array}{ll} \max 30\xi_1+50\xi_2 & \\ 2\xi_2 & \le 0 \end{array} \qquad D': \quad \begin{array}{ll} \min 0\eta_1 & \\ 0\eta_1 & = 30 \\ 2\eta_1 & = 50 \\ \eta_1 & \ge 0. \end{array}$$

The system of constraints of $D'$ is obviously inconsistent, hence $D'$ has no solution.

Otherwise, for $\bar{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ we have two active constraints, $I(\bar{x}) = \{1,3\}$. The problems are:

$$P': \quad \begin{array}{ll} \max 30\xi_1+50\xi_2 & \\ \xi_1 & \le 0 \\ 3\xi_1+2\xi_2 & \le 0 \end{array} \qquad D': \quad \begin{array}{ll} \min 0\eta_1+0\eta_2 & \\ \eta_1+3\eta_2 & = 30 \\ 2\eta_2 & = 50 \\ \eta_1,\ \eta_2 & \ge 0. \end{array}$$

Because we had previously detected a feasible growth direction for this point, $D'$ should have no solution.

**Exercise**

Formulate problems $P'$ and $D'$ for point $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$. Has $D'$ a solution?

The algorithm sketched above lacks a fundamental detail: assuming that $D'$ has no solution, how can we obtain a feasible growth direction $\bar{\xi}$? In building the direction we are guided by the information derived from problem $D'$.

For more simplicity, let us temporarily consider the case in which matrix $A_I$ is square and of full rank. Given such characteristics of matrix $A_I$, we can compute its inverse $A_I^{-1}$, so the system

$$\eta A_I = c$$

has a unique solution $\bar{\eta} = cA_I^{-1}$. If $\bar{\eta}$ is made of non-negative components, as we proved in theorem 3.1, then solution $\bar{x}$, that we took as reference, is optimal. Otherwise, if there is a $\bar{\eta}_h<0$, we can start precisely from such indication to build a feasible growth direction:

$$\bar{\xi} = - A_I^{-1}\, u_h \tag{3.3}$$

where $u_h$ denotes the $h$-th unit vector of order $n$, i.e., a vector with all 0 except one 1 in $h$-th position. In practice, (3.3) chooses as direction the $h$-th column of matrix $A_I^{-1}$ with changed sign.

Now let us prove that $\bar{\xi}$ really is a feasible growth direction. First of all

$$A_I\, \bar{\xi} = - A_I A_I^{-1}\, u_h = -u_h \le 0$$

hence feasibility conditions are respected, further

$$c\ \bar{\xi} = - c\, A_I^{-1}\, u_h = -\bar{\eta}_h > 0$$

hence growth conditions are also satisfied. Later on we shall be able to give a geometric interpretation of this choice as well, for the moment we say that, in practice, the direction we find is the one solving the system

$$A_I \, \bar{\xi} = - u_h.$$

Example: Shutters and Frames

We consider point $\bar{x}=\begin{bmatrix}0\\6\end{bmatrix}$ for the problem of the shutter concern: $I(\bar{x})=\{2,4\}$, and $A_I=\begin{bmatrix}0 & 2\\-1 & 0\end{bmatrix}$ is of full rank, so we can compute $A_I^{-1} = \begin{bmatrix}0 & -1\\1/2 & 0\end{bmatrix}$. The solution to the system of equality constraints of the dual $D'$ is $\bar{\eta} = [30,50]\, A_I^{-1} = [25,-30]$ which, because of the second negative component, does not respect sign constraints. A feasible growth direction is given by:

$$\bar{\xi} = -A_I^{-1} u_2 = - \begin{bmatrix}-1\\0\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}.$$

And in fact we can verify that $c\bar{\xi} = 30>0$, and $A_I\bar{\xi} = \begin{bmatrix}0\\-1\end{bmatrix} \le 0$.

Let us consider now the case in which the number of rows of $A_I$ is less than the number of columns. This implies that matrix $A_I$ cannot be inverted and system $\eta A_I = c$ has more equations than unknowns and might be inconsistent. If system $\eta A_I = c$ has a solution with components greater than or equal to 0, then we can state that for theorem 3.1 solution $\bar{x}$ is feasible; otherwise, if the system has no solution, or it has a solution with negative components, then we can build a feasible growth direction. A way of building such a direction is the following:

$$A_I \, \xi = 0$$
$$c \, \xi = 1.$$

In this way, the first set of constraints guarantees that we are keeping adherent to active constraints in $\bar{x}$, and in the possible new feasible point the cardinality of the set of active constraint indexes will grow by at least one unit. The second constraint guarantees that we are taking a growth direction. Note that value 1 of the known term was arbitrarily fixed, and any other choice, provided it is positive, supplies a growth direction.

Example: Shutters and Frames

Consider point $\bar{x}=\begin{bmatrix}1\\6\end{bmatrix}$. $I(\bar{x})=\{2\}$. As previously observed, the system $\{0\eta_1=30,\ 2\eta_1=50\}$ admits no solutions. The system to be solved in order to get a feasible growth direction is:

$$\begin{aligned} 2\xi_2 &= 0 \\ 30\xi_1 + 50\xi_2 &= 1 \end{aligned}$$

whose solution is $\bar{\xi}=\begin{bmatrix}1/30\\0\end{bmatrix}$. Note that if, instead of having a known term 1, we had indicated, say, 30, we would have obtained $\bar{\xi}=\begin{bmatrix}1\\0\end{bmatrix}$. However, it is evident that such choice does not influence the vector direction but just its length, and nothing changes except for the choice of the displacement step.

**Exercise**

Compute the maximum displacement steps along the direction $\bar{\xi}=\begin{bmatrix}1/30\\0\end{bmatrix}$ as well as along the direction $\bar{\xi}=\begin{bmatrix}1\\0\end{bmatrix}$.

We still have to examine the case in which the number of rows of $A_I$ is greater than the number of columns. Such case, known in the literature as *degeneracy*, is far from being rare in reality. Yet, it requires greater notational care and a closer analysis, that will be the object of next in-depth paragraphs.

Now we just have to formally summarize the solution algorithm of *LP* problems as it arises from previous considerations. The algorithm, known as *Primal-Dual Simplex*, takes as input matrix *A* with *m* rows and *n* columns, known term vector *b*, cost vector *c*, a feasible solution *x*, and returns as output the optimal solution *x* of *P* and the optimal solution *y* of *D*, if *P* and *D* admit a finite optimum as well as the logical variable *unbounded* having value "true" if *P* is unbounded (and *D* is empty).

```
Procedure Simplex_Primal_Dual(A,b,c,x,unbounded,y);
  begin
     optimal:=false; unbounded:=false;
     I:={i: Aix=bi};
     if I=∅ then grow_along(c,x,I,unbounded);
     while not optimal or not unbounded do
     begin
       if {ηAI=c} has no solution then begin compute ξ: AIξ=0, cξ=1;
                                             grow_along(ξ,x,I,unbounded);
                                           end;
       else if {ηAI=c} has a solution and ∃h: ηh<0 then begin compute ξ: AIξ=uh;
                                                               grow_along(ξ,x,I,unbounded);
                                                            end;
       else optimal:=true;          {the dual system has a solution: ηAI=c, η≥0}
     end;
  end.
```

Fig. 12

If in the first phase of the procedure the given point *x* is in the strict interior of all constraints ($I(x)=\emptyset$), we are free to choose any growth direction. The procedure chooses $\xi=c$, which is the direction providing the greatest increment of the objective function per displacement unit, equal to $\|c\|^2$.

The procedure **grow_along**($\xi$,*x*,*I*,unbounded), called inside the algorithm, computes the displacement step along the direction $\xi$ that was taken as input and, if the displacement step reveals to be finite, returns the new solution *x* and the new set of active constraint indexes.

**Procedure** grow_along($\xi$,x,I,unbounded);
  **begin**
    $\lambda:=\min\{(\frac{b_i - A_ix}{A_i\xi}: A_i\xi>0,\ i\in\{1,\ldots,m\}\setminus I),\ +\infty\};$
    **if** $\lambda=+\infty$ **then** unbounded:=true;
    **else begin**
      $x:=x+\lambda\xi;$
      $I:=\{i:\ A_ix=b_i\};$
    **end**;
  **end**.

Fig. 13

Note that the problem is unbounded if there exists no constraint blocking the growth along the direction $\xi$; in that case $\lambda$ is unbounded.

Example: Shutters and Frames

Let us try to apply the Primal-Dual Simplex algorithm to the usual problem of the shutter concern, starting from point $x=\begin{bmatrix}3\\1\end{bmatrix}$. In this point there are no active constraints, so the procedure begins moving along the gradient of objective function $c$.

We compute what are the constraints blocking the growth along $c$, i.e., those for which $A_ic>0$:

$$A_1c = [1,0]\begin{bmatrix}30\\50\end{bmatrix} = 30 > 0$$

$$A_2c = [0,2]\begin{bmatrix}30\\50\end{bmatrix} = 100 > 0$$

$$A_3c = [3,2]\begin{bmatrix}30\\50\end{bmatrix} = 190 > 0$$

$$A_4c = [-1,0]\begin{bmatrix}30\\50\end{bmatrix} = -30$$

$$A_5c = [0,-1]\begin{bmatrix}30\\50\end{bmatrix} = -50$$

Hence, the displacement step is given by

$$\lambda = \min\{\frac{b_1 - A_1x}{30}, \frac{b_2 - A_2x}{100}, \frac{b_3 - A_3x}{190}\} = \min\{\frac{1}{30}, \frac{10}{100}, \frac{7}{190}\} = \frac{1}{30}$$

which corresponds to a displacement reaching the edge of the first constraint. The new point is given by:

$$x = \begin{bmatrix}3\\1\end{bmatrix} + \frac{1}{30}\begin{bmatrix}30\\50\end{bmatrix} = \begin{bmatrix}4\\8/3\end{bmatrix}$$

in the new point $I(x) = \{1\}$. Note that the new set of active constraint indexes is given by the previous one together with all constraints that determined the maximum feasible value of $\lambda=1/30$. At this moment, the algorithm performs a second iteration. The dual system we are trying to solve is:

$$\eta_1 = 30$$
$$0\eta_1 = 50$$

which is clearly inconsistent. Therefore, we obtain a feasible growth direction by solving the system:

$$\xi_1 = 0$$
$$30\xi_1 + 50\xi_2 = 1$$

so the displacement direction is $\xi = \begin{bmatrix}0\\1/50\end{bmatrix}$, whereas – once we have verified that the only constraints blocking the growth are the second and the third – the displacement step is given by:

$$\lambda = \min\{\frac{b_2 - A_2x}{1/25}, \frac{b_3 - A_3x}{1/25}\} = \min\{\frac{20/3}{1/25}, \frac{2/3}{1/25}\} = \frac{50}{3}.$$

The new point is given by:

$$x = \begin{bmatrix} 4 \\ 8/3 \end{bmatrix} + \frac{50}{3} \begin{bmatrix} 0 \\ 1/50 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}.$$

In the new point, active constraints are those active in the starting point, i.e., the first, with the addition of the one that determined the value of $\lambda$, i.e., the third, hence $I(x) = \{1,3\}$. Let us solve the dual system relative to these active constraints:

$$\begin{aligned} \eta_1 + 3\eta_2 &= 30 \\ 2\eta_2 &= 50 \end{aligned}$$

whose solution is $\eta_1 = -45$, $\eta_2 = 25$. Because the constraints on the signs of dual variables are not satisfied, the solution is not optimal and we can build a feasible growth direction by solving the system:

$$\begin{aligned} \xi_1 &= -1 \\ 3\xi_1 + 2\xi_2 &= 0 \end{aligned}$$

hence $\xi = \begin{bmatrix} -1 \\ 3/2 \end{bmatrix}$. The maximum displacement along this direction is given by

$$\lambda = \min\{\frac{b_2 - A_2x}{3}, \frac{b_4 - A_4x}{1}\} = \min\{\frac{6}{3}, \frac{4}{1}\} = 2.$$

The new point is given by:

$$x = \begin{bmatrix} 4 \\ 3 \end{bmatrix} + 2 \begin{bmatrix} -1 \\ 3/2 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

where active constraints are $I(x)=\{2,3\}$. The dual system is:

$$\begin{aligned} 3\eta_2 &= 30 \\ 2\eta_1 + 2\eta_2 &= 50 \end{aligned}$$

which, as we saw, has solution $\eta_1 = 15$, $\eta_2 = 10$. Once arrived here, the algorithm ends.

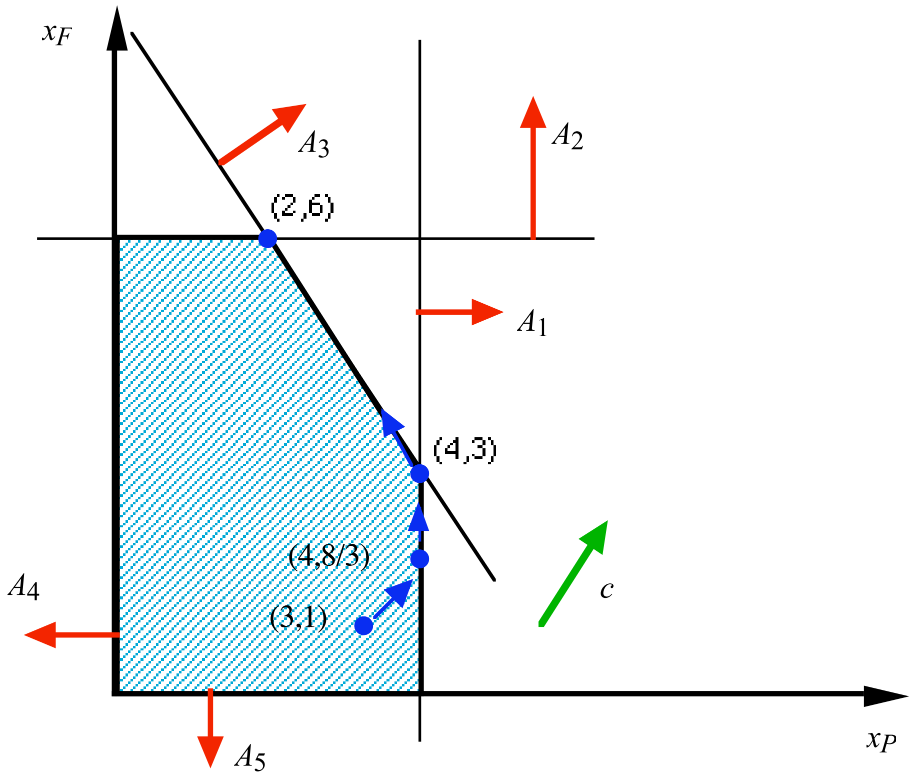


Fig. 14: evolution of the algorithm and indication of growth directions

### *3.3 *Linear Programming on cones*

We have seen how the heart of the solution algorithm of *LP* problems implies that at each iteration we are confronted with Linear Programming problems being particularly simple and whose feasible region can be represented by polyhedral cones:

$$P': \quad \max\ c\xi \qquad\qquad D': \quad \min \eta\, 0$$
$$A\xi \le 0 \qquad\qquad \eta A = c$$
$$\eta \ge 0.$$

Below we assume that, with no loss of generality, matrix $A$, $m \times n$, is of maximum rank and vector $c$ is non-null. The set of inequalities $A\xi \le 0$ obviously defines a polyhedral cone. But also the feasible region of the dual refers to a cone, because the set of points $C=\{z=\eta A,\ \eta \ge 0\}$ is exactly the cone finitely generated by rows of $A$, that is $C=cone\{A_1, A_2, \ldots, A_n\}$. Consequently, the dual problem consists in testing whether vector $c$ belongs to cone $C$.

As previously observed, $P'$ always admits at least one solution: the null solution ($\xi=0$) is feasible and has null value of the objective function. If there exists a feasible solution $\bar{\xi} \neq 0$ with $c\bar{\xi} > 0$, then also the solution $\alpha\bar{\xi}$ is feasible for any real $\alpha \ge 0$; in this case $P'$ is unbounded from above, and $\bar{\xi}$ detects an unbounded feasible growth direction for the objective function. In fact, by appropriately choosing the value of parameter $\alpha$, we can build feasible solutions with value of the objective function being arbitrarily large. The following theorem holds:

**Theorem 3.2**

*A problem of Linear Programming on cones either has an optimum in the origin or is unbounded.*

It can be easily verified that, in case $P'$ is unbounded, $D'$ cannot have any solution and, conversely, if $D'$ admits a solution, then $P$ cannot be unbounded. This is a direct consequence of the weak duality theorem applied to the particular case of *LP* on cones.

If $P'$ is unbounded, there exists a vector $\bar{\xi}$ such that we get:

$$A\bar{\xi} \leq 0,$$

$$c\bar{\xi} > 0,$$

whereas if $D'$ is feasible, there exists a $\bar{\eta}$ such that:

$$\bar{\eta}A = c,$$

$$\bar{\eta} \geq 0.$$

Now let us prove the *Fundamental Theorem of Linear Inequalities*, also known as *Farkas's Lemma* [Farkas 1884].

**Theorem 3.3**

*Let A be a matrix m×n of maximum rank and let there be* $c\in\boldsymbol{R}^n$. *Then the two systems:*

$$i)\ \begin{cases}\eta A &= c\\ \eta \geq 0\end{cases} \qquad ii)\ \begin{cases}A\xi \leq 0\\ c\xi > 0\end{cases}$$

*are mutually exclusive, i.e., either system i) has a solution or system ii) has a solution.*

Proof

Suppose there exists a solution for system i), that is a vector $\bar{\eta}\geq 0$ with $\bar{\eta}A=c$. We prove that a solution $\bar{\xi}$ with $c\bar{\xi}>0$ and $A\bar{\xi} \leq 0$ cannot exist. The proof is by contradiction. If there existed such a $\bar{\xi}$, we would have:

$$0 < c\bar{\xi} = \bar{\eta}A\bar{\xi} \leq 0,$$

where the first equality derives from $c = \bar{\eta}A$, whereas $\leq$ derives from $A\bar{\xi} \leq 0$ and $\bar{\eta}\geq 0$. Hence we have a contradiction.

Now let us prove that if there exists no solution for system i), there must exist a solution for system ii). If there exists no $\bar{\eta}\geq 0$ with $\bar{\eta}A=c$, this means that $c$ cannot be written as a non-negative linear combination of constraint gradients, thus it does not belong to the cone generated by vectors $A_i$, $i$=1, …,$m$. Such cone is given by $S = \{z: z = \eta A, \eta\geq 0\}$. Since $c$ does not belong to the closed and convex set $S$, then, according to the separating hyperplane theorem, we know that there exists a hyperplane separating $c$ from all points of $S$. Hence, there exists the hyperplane with gradient $\bar{\xi}$ such that

$$c\bar{\xi} > z\bar{\xi} \quad \forall z\in S.$$

Replacing $z$ by its definition, we have

$$c\bar{\xi} > \eta A\bar{\xi} \qquad \forall\eta\geq 0.$$

Choosing $\eta$=0 it follows that $c\bar{\xi} > 0$. Given that $\eta$ may also take arbitrarily large values (and grow to infinity), whereas $c\bar{\xi}$ has a finite value, then it must necessarily be $A\bar{\xi} \leq 0$. So we have found that there exists a solution of system ii). ♦

<u>Example</u>
Consider vector $c$=[-1,4] and matrix

$$A=\begin{bmatrix} 1 & -3 \\ 2 & 0 \\ 1 & 1 \end{bmatrix}.$$

The problem $P'$ has the following geometric representation:

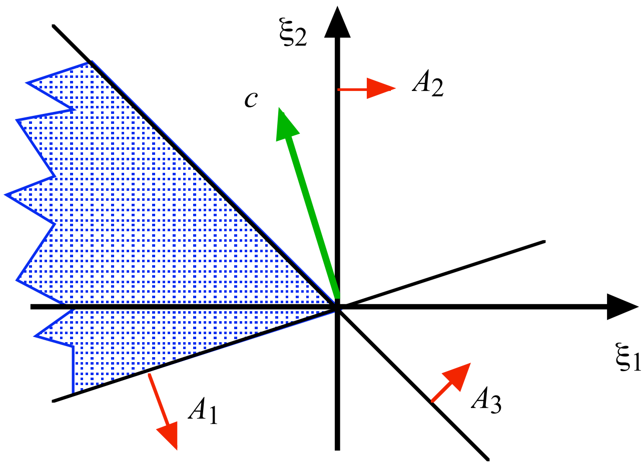


Fig. 15: the polyhedral cone of the primal space. The solution is unbounded.

Trying to solve the problem $P'$ geometrically, we see that there exists no finite optimal solution. The geometric representation of the dual is the following:

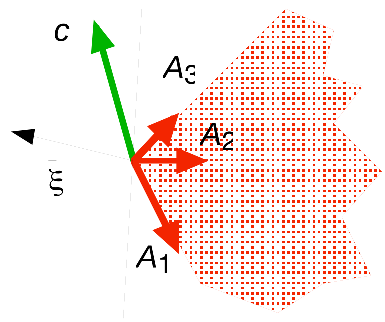


Fig. 16: space of the dual given by the cone finitely generated by constraint gradients. Separating hyperplane with gradient $\bar{\xi}$.

Also in this case it is easy to verify geometrically that vector $c$ does not belong to the cone finitely generated by $A_1$, $A_2$ and $A_3$.

**Exercise**
Represent problems $P'$ and $D'$ graphically with matrix $A$ of the preceding example and with vector $c$=[2,-1] and provide the solution obtained geometrically for both problems.

Now let us pose the problem of solving $P'$ analytically. As it follows from the above considerations, to do this we just need to find a feasible vector $\bar{\xi}$ such that $\bar{\xi}$>0 or, alternatively, a vector $\bar{\eta}\geq 0$ with

$\overline{\eta}A=c$. In the first case, the unboundedness of the problem would be proved, whereas in the second case we might conclude that the origin is an optimal solution. We will consider three separate cases.

*i) m=n.*

In this case matrix $A$ is square and, since it is of full rank, the system of equations of problem $D'$ has a unique solution given by $\overline{\eta}=cA^{-1}$. If $\overline{\eta}\geq 0$, then the optimal solution of $P'$ is $\overline{\xi}=0$. Otherwise, if $\overline{\eta}\not\geq 0$, i.e., if there exists an index $k$ for which $\overline{\eta}_k<0$, then denoting by $u_k$ the $k$-th unit vector we have $\overline{\eta}$ $))_k=cA^{-1}u_k$, hence setting $\overline{\xi}=-A^{-1}u_k$ we would get $c\overline{\xi}>0$, thus determining an unbounded growth direction for the objective function.

In practice, we have proved that:

- either there exists $\overline{\eta}\geq 0$ such that $\overline{\eta}A=c$
- or there exists $\overline{\xi}\in \boldsymbol{R}^n$ and an index $k$ for which

$$A_i\overline{\xi}\begin{cases} = 0 & \text{if } i\neq k \\ = -1 & \text{if } i=k \end{cases}$$

with $\quad c\overline{\xi}>0$.

From the geometric point of view, vector $\overline{\xi}=-A^{-1}u_k$, detects an *unbounded feasible growth direction*; in fact, we can move along such direction, letting the objective function grow and without ever going out of the cone; such direction is the one running along the line being intersection of all hyperplanes defining the cone except for the $k$-th: it is the *locus* of the points $\xi=\alpha\overline{\xi}$, with $\alpha\geq 0$, hence of the solutions of the system defined by the constraints

$$A_i\xi\begin{cases} = 0, & i\neq k, \\ < 0, & i=k. \end{cases}$$

*ii) m<n.*

In this case, since $A$ is of maximum rank, the system $\eta A=c$ either has no solution or has only one solution. Apart from a possible reordering of the columns of $A$, we can set $A=[A',A'']$, with $A'$ square matrix of order $m$, and $\det(A')\neq 0$; correspondingly, we decompose $c$ into $(c',c'')$. Then we have that if $\eta A=c$ admits a solution, such solution will be $\overline{\eta}=c'A'^{-1}$.

There are three possible cases:

*a)* $\overline{\eta}A=c$, with $\overline{\eta}\geq 0$, then, according to what stated above, the origin is optimal solution of $P'$.

*b)* $\overline{\eta}A=c$, with $\overline{\eta}_k=c'A'^{-1}u_k<0$, for any index $k$. Then vector $\overline{\xi}=(\xi',\xi'')$, with $\xi'=-A'^{-1}u_k$ and $\xi''=0$ is a solution of $P'$, and $c\overline{\xi}>0$; in fact

$$A\overline{\xi} = A'\xi' + A''\xi'' = -A'A'^{-1}u_k = -u_k \leq 0,$$

$$c\overline{\xi} = c'\xi' + c''\xi'' = \overline{\eta}A'\xi' = -\overline{\eta}A'A'^{-1}u_k = -\overline{\eta}_k > 0.$$

*c)* The system $\eta A=c$ has no solution, that is $\overline{\eta}A''=c'A'^{-1}A''\neq c''$. Thus there exists an index $k$ such that either $c''_k>c'A'^{-1}A''u_k$, or $c''_k<c'A'^{-1}A''u_k$.

Consider the first case (the other one can be similarly treated): we will have that vector $\overline{\xi}=(\xi', \xi'')$, with $\xi' = -A'^{-1}A''u_k$ and $\xi'' = u_k$, is feasible for $P'$, with $c\overline{\xi}>0$; in fact:

$$A\overline{\xi} = A'\xi' + A''\xi'' = -A'A'^{-1}A''u_k + A''u_k = 0,$$

$$c\overline{\xi} = c'\xi' + c''\xi'' = -c'A'A'^{-1}A''u_k + c''u_k = c''_k - c'A'^{-1}A''u_k > 0.$$

*iii*) $m>n$.

Let us try to refer to case *i*) with square matrix. Let $A_B$ be a non-singular square submatrix of $A$, with $B\subset\{1,...,m\}$*; i.e., $A_B$ is a *basis* for $\boldsymbol{R}^n$. Below we will call $B$ a *basis* and $A_B$ a *basis* (sub)*matrix* of $A$, and we will denote by $h=B(j)$, $j=1,\ldots,n$, the $j$-th index contained in $B$ (in practice, set $B$ is considered to be implemented by means of a list, and $B(j)$ is the $j$-th element of the list). Since we assumed that the rank of $A$, $r(A)$, is equal to $n$, then at least one basis $B$ always exists. According to what we saw at point *i*), we have that:

- either there exists $\overline{\eta}_B \geq 0$ such that $\overline{\eta}_B A_B=c$

  Then vector $(\overline{\eta}_B, \overline{\eta}_N)$, with $\overline{\eta}_N=0$ and $N=\{1,\ldots n\}\backslash B$, is a feasible solution for $D'$, hence $\xi=0$ is the only optimal solution of $P'$. In this case we say that $B$ is a *dual feasible* basis.

- or there exists $\overline{\xi}\in\boldsymbol{R}^n$ and $h\in B$ for which with $c\overline{\xi}>0$

  $$A_i\overline{\xi}\begin{cases} = \ 0 & \text{if } i\in B\backslash\{h\} \\ = -1 & \text{if } i=h \end{cases}$$

  Then, if for the remaining constraints we get $A_N\overline{\xi}\leq 0$, since $\overline{\xi}$ is feasible for $P'$, we have proved that $P'$ is unbounded and we have determined a growth direction, $\overline{\xi}=-A_B^{-1}u_k$, with $h=B(k)$; otherwise ($\overline{\eta}_B \ngeq 0$ and $A_N\overline{\xi}\nleq 0$), we still can say nothing neither about $P'$ nor about $D'$, so we have to seek another basis $B$.

The algorithm $\boldsymbol{S}$ of Figure 17 takes as input an initial basis $B$, in addition to matrix $A$ and to cost vector $c$, and modifies it at each step, until we get a basis which either is dual feasible or detects a decrease direction for primal $P'$. The procedure returns as output a variable indicating in which case of Farkas's lemma we are and, according to the case, also indicates a direction $\xi$ or a dual solution $\eta$. The procedure also returns the index of the last element that was removed from the basis.

---

* given a matrix $A$ and a set of indexes $B$, $A_B$ is the submatrix composed of rows $A_i$, $i\in B$. Similarly, $a_B$ denotes the subvector of vector $a$ containing the components whose indexes are in $B$.

```
Procedure S (A,c,B,η̄,ξ̄,case,h):
  begin
  1  termination:=false;
     repeat
  2    η̄B:=cAB^-1;
  3    if η̄B≥0   then begin termination:=true; case:=1; η̄=[η̄B,0] end
                else begin
  4                  h:=min{i∈B: η̄i<0};
  5                  ξ̄:=−AB^-1 uk;                                   {k : B(k)=h}
  6                  if Aξ̄≤0 then begin termination:=true; case:=2 end
                       else begin
  7                        s:=min{i: Aiξ̄>0};
  8                        B:=B∪{s}\{h}
                         end
                  end
     until termination
  end.
```

Fig. 17: algorithm for choosing the basis

If the algorithm returns case=1, then basis $B$ is dual feasible and the origin (i.e., the cone vertex) is the optimal solution of $P'$; otherwise, when we have case=2, index $h$ detects a decrease feasible direction for the objective function, i.e., the direction $-A_B^{-1}u_k$, with $h=B(k)$.

### Theorem 3.4

*The procedure **S** terminates in a finite number of steps, yielding a pair of vectors* $(\bar{\xi}, \bar{\eta})$*, such that it gives:*

$$\text{either} \quad \begin{cases} \bar{\eta}A = c \\ \bar{\eta} \geq 0 \end{cases} \qquad \text{or} \quad \begin{cases} A\bar{\xi} \leq 0 \\ c\bar{\xi} > 0. \end{cases}$$

The proof of the correctness of procedure ***S*** follows from the considerations formulated above. However, its termination in a finite number of steps is not so easy to prove and is based on the choice of indexes of the elements that must enter and leave the basis (instructions 4 and 7). This rule, that every time chooses the lowest index, is known as *Bland's anticycle rule*.

The procedure of Figure 18 summarizes the process for solving linear programming problems on cones.

```
Procedure Simplex_on_Cones (A,c,B,η̄,ξ̄,case,h):
   begin
      if |B|=n then S(A,c,B,η̄,ξ̄,case,h)
        else begin
           Choose A': det(A')≠0;                              {A=[A',A"]}
           η̄:=c'A'^-1;
           if η̄A"=c" then                                {the system has a solution}
             if η̄≥0   then case:=1                        {case i}
                        else begin
                           let h: η̄h<0;
                           ξ̄:= [ -A'^-1 eh ]
                                [    0      ]; case:=2       {case ii}
                        end
           else    begin                                  {the system has no solution}
                     let k: c"k−η̄A"uk≠0;
                     α:=sign(c"k−η̄A"uk);
                     ξ̄:=α [ -A'^-1 A"uk ]
                           [     uk      ]; case:=2; h:=0  {case ii}
                   end
         end
   end.
```

Fig. 18: algorithm for solving *LP* problems on cones

Observe that, for notational simplicity, we have assumed that matrix $A'$ is composed of the first columns of $A$. Further, parameter $h$ has a meaning only when we get the solution $\bar{\xi}$ of the second system (case=2), whereas $h = 0$ when $\bar{\eta}A=c$ has no solution.

It should be noted that, given a basis $B$, the scalar $-cA_B^{-1}u_k = -\bar{\eta}_h$, where $h=B(k)$ is the constraint index leaving the basis, is the unit increment of the objective function of $P'$ when we move far from the edge of the $h$-th constraint along direction $\bar{\xi}=-A_B^{-1}u_k$, i.e., when we increase the slack variable of the $h$-th constraint by one unit, letting slack variables of the other basis constraints (i.e., of constraints whose indexes are in $B$) keep value 0. For this reason, the value $-cA_B^{-1}u_k$ is usually defined as the *reduced cost* of the slack variable relative to the $h$-th constraint (or, more simply, the reduced cost of the $h$-th constraint).

If now we consider the pair of dual problems $P'$ and $D'$ relative to linear programming on cones, which was introduced at the beginning of this paragraph, then from theorem 3.3 it follows that $P'$ has a finite optimal solution **if and only if** $D'$ has a solution. In the following corollary we summarize this important result applying it to the different pairs of dual problems on cones.

### Corollary 3.5

*P' has a finite optimal solution if and only if D' has solution:*

(3.4a) $\quad P: \begin{array}{l} \max c\xi \\ A\xi \le 0 \end{array} \qquad D: \begin{array}{l} \min \eta\, 0 \\ \eta A = c \\ \eta \ \ge 0 \end{array}.$

(3.4b) $\quad P: \begin{array}{l} \max 0\xi \\ A\xi = b \\ \xi \ \ge 0 \end{array} \qquad D: \begin{array}{l} \min \eta b \\ \eta A \ge 0 \end{array};$

(3.4c) $P$: $\max 0\xi$, $A\xi \leq b$ $\qquad$ $D$: $\min \eta b$, $\eta A = 0$; $\eta \geq 0$

$$P:\ \begin{array}{l}\max 0\xi\\ A\xi \leq b\end{array} \qquad D:\ \begin{array}{l}\min \eta b\\ \eta A = 0\,;\\ \eta \ \geq 0\end{array} \tag{3.4c}$$

$$P:\ \begin{array}{l}\max 0\xi\\ A\xi \leq b\\ \xi \ \geq 0\end{array} \qquad D:\ \begin{array}{l}\min \eta b\\ \eta A \geq 0\,.\\ \eta \ \geq 0\end{array} \tag{3.4d}$$

**Exercise**
Prove corollary 3.5 for the pairs (3.4b), (3.4c), (3.4d) starting from (3.4a).

### *3.4 *Strong duality theorem and its consequences*

Let us formalize one of the main results of linear programming we already saw informally but can only prove after having examined the fundamental theorem of linear inequalities. Consider the pair of dual problems

$$P:\ \max\,\{cx\colon Ax \leq b\} \qquad\qquad D:\ \min\,\{yb\colon yA = c,\, y \geq 0\}$$

**Theorem 3.6** (*Strong duality theorem*)
*If P and D admit feasible solutions, then* $\max\{cx\colon Ax\leq b\} = \min\{yb\colon yA=c,\, y\geq 0\}$.

Proof
If $c$=0, then $\max\{cx\colon Ax\leq b\}$=0, whereas $y$=0, being feasible solution for $D$, is also optimal; in fact, $yb\geq 0$ for each feasible $y$ ($yb\geq yAx$=$0x$). Hence, in this case the theorem is banally true.

Consider now the more interesting case in which $c\neq 0$. Let $\bar{x}$ be an optimal solution for $P$; we denote by $I$ and $\bar{I}$ the set of active constraint indexes for $\bar{x}$ and its complement. Observe that $|I| \geq 1$. In fact, if $I = \emptyset$ then there exists $\lambda > 0$ such that $x' = \bar{x}+\lambda c$ is a feasible solution for $P$. In fact

$$Ax' = A\bar{x}+\lambda Ac \leq b \qquad \forall\lambda\in[0,\bar{\lambda}],$$

where

$$\bar{\lambda} = \begin{cases} \min\{\dfrac{b_i - A_i\bar{x}}{A_ic}\colon A_ic > 0\}, & \text{if } \exists\ i \in \{1,\ldots,m\}\colon A_ic > 0,\\ +\infty, & \text{if } A_ic \leq 0\ \forall\ i \in \{1,\ldots,m\}. \end{cases}$$

Since $cx' = c\bar{x}+\lambda\|c\|^2$, we have $cx' > c\bar{x}$, $\forall\lambda\in(0,\bar{\lambda}]$. Hence, if $I = \emptyset$ then $c$ is a feasible growth direction, contradicting the hypothesis that $\bar{x}$ is the feasible solution. So we may conclude that $I \neq \emptyset$.
Let $\xi\in\boldsymbol{R}^n$ be a feasible direction ($A_I\xi \leq 0$) and let:

$$\bar{\lambda} = \begin{cases} \min\{\dfrac{b_i - A_i\bar{x}}{A_i\xi}\colon i\in\bar{I} \text{ and } A_i\xi>0\}, & \text{if } \exists\ i\in\bar{I}\colon A_i\xi>0,\\ +\infty, & \text{if } A_i\xi\leq 0\ \ \forall\ i\in\bar{I}; \end{cases}$$

since $\bar{x}$ is an optimal solution, we must get $c\xi\leq 0$, otherwise we get $c\bar{x}+\bar{\lambda}c\xi > c\bar{x}$, which contradicts the hypothesis that $\bar{x}$ is optimal. Hence, the system:

$$\begin{cases} A_I\xi \leq 0\\ c\xi \ > 0 \end{cases}$$

has no solution and by theorem 3.3 the system

$$\begin{cases} \eta A_I = c \\ \eta \geq 0 \end{cases}$$

has solution $\bar{\eta}$. Then we have that $\bar{y} = (\bar{y}_I, \bar{y}_{\bar{I}}) = (\bar{\eta}, 0)$ is a feasible solution for $D$.

Further, we get $\bar{y}b = \bar{y}_I b_I = \bar{y}_I A_I \bar{x} = c\bar{x}$, hence by the weak duality theorem $\bar{y}$ is optimal. ♦

An immediate consequence of theorem 3.6 is a method for detecting if we are in front of an optimal solution. In fact, $\bar{x}$ is optimal for $P$ if and only if $c$ belongs to the cone generated by active constraint gradients.

<u>Example: Shutters and Frames</u>

Let us consider again the shutter concern problem and let us verify the optimality of point $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$ by merely formulating some considerations of a geometrical nature. In this point, active constraints are the second and the third. As it clearly appears from the figure, $c$ belongs to the cone generated by $A_2$ and $A_3$. Non-optimality can be easily verified for point $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ as well, in fact $c$ does not belong to the cone generated by $A_1$ and $A_3$, that are active constraint gradients.

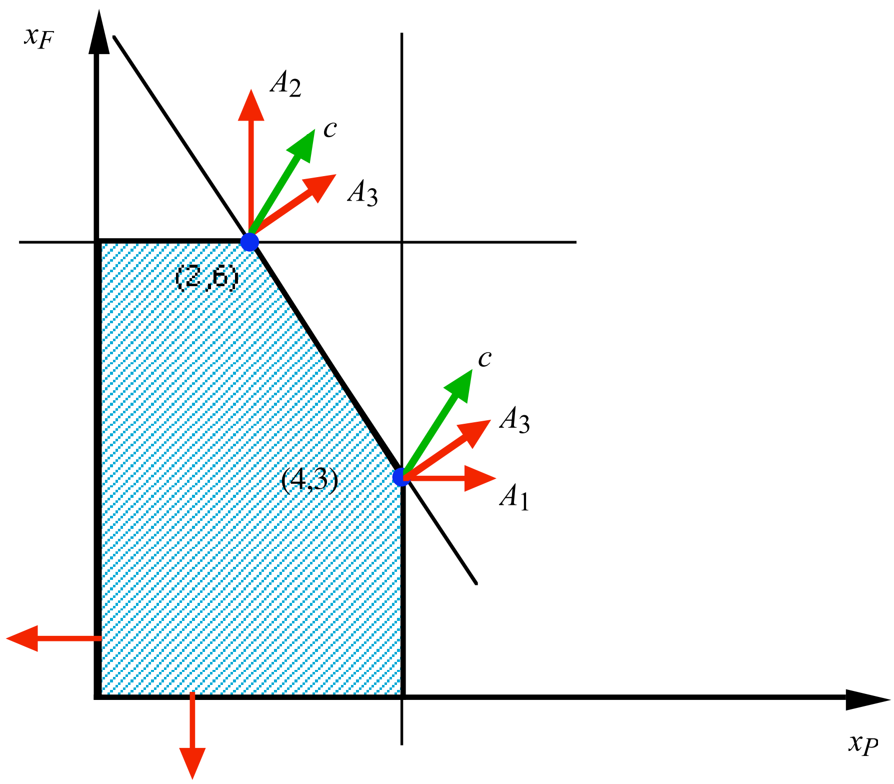


Fig. 19: optimality verification

Let us complete the characterization of the pairs of dual problems $P$ and $D$.

**Theorem 3.7**

If $P$ has finite optimum, then $D$ has finite optimum.

<u>Proof</u>

We prove the theorem for the pair $P$: $\max\{cx: Ax \leq b\}$, $D$: $\min\{yb: yA = c, y \geq 0\}$, but the result can be extended to any pair of dual problems.

Let $\bar{x}$ be an optimal solution for $P$.

From the weak duality theorem (theorem 2.2) follows that $D$ is bounded from below.

If $\{y: yA = c, y \geq 0\} = \emptyset$ then, by theorem 3.3, $\exists\, z: Az \leq 0$ and $cz > 0$, which implies the existence of a feasible solution that is better than $\bar{x}$: $\bar{x} + z \in \{x: Ax \leq b\}$, $c\,(\bar{x} + z) > c\bar{x}$. Which is a contradiction because $\bar{x}$ is an optimal solution for $P$. ♦

Hence, given a pair of dual problems $P$ and $D$, all cases that may occur are summarized in the following table.

| $D$ \ $P$ | finite optimum | unbounded | empty |
|---|---|---|---|
| finite optimum | * | | |
| unbounded | | | * |
| empty | | * | * |

Fig. 20: relations between $P$ and $D$. * indicates cases that may occur

Actually, the case in which both problems are empty is still to be verified, but this can be easily proved with a simple example that must be interpreted geometrically:

$$\max\{x_1: -x_1 - x_2 \leq -1,\ x_1 + x_2 \leq -1\},\ \min\{-y_1 - y_2: -y_1 + y_2 = 1, -y_1 + y_2 = 0,\ y_1, y_2 \geq 0\}.$$

### *3.5 Primal-Dual Simplex Algorithm

In the light of the results presented so far, we can describe the primal-dual simplex algorithm more formally, including the treatment of degeneracy as well, i.e., of those cases in which the number of active constraints in a point is greater than or equal to $n$.

As usual, consider the pair $(P, D)$:

$$P:\ \max\ cx \quad Ax \leq b \qquad\qquad D:\ \min\ yb \quad yA = c,\ y \geq 0$$

Let $\bar{x}$ be a feasible solution for $P$. Let us assume, with no loss of generality, that $\bar{x}$ belongs to the boundary of the feasible set of $P$, i.e., that the set $I$ of active constraints in $\bar{x}$ is non-empty.

A point $\bar{x}$ for which we get $I \neq \emptyset$ can be easily obtained unless $P$ is unbounded from above and consequently $D$ is empty: starting from any feasible point, we just need to move towards the direction indicated by vector $c$ until we meet the boundary.

Below we denote by **Determine**($\bar{x}$,$I$,*Dempty*) the procedure that either returns a feasible solution $\bar{x}$ such that $|I| \geq 1$ or determines that $D$ is empty (*Dempty*=true).

We can define the two following complementary problems:

- the *Restricted Primal* (*RP*):

$$\exists\ \xi: \begin{cases} A_I \xi \leq 0 \\ c\xi > 0; \end{cases}$$

- and the *Restricted Dual* (*RD*):

$$\exists\ \eta: \begin{cases} \eta A_I = c \\ \eta \geq 0. \end{cases}$$

Only one of them admits a solution. In order to determine which of them admits a solution, it suffices to call the procedure **Simplex_on_Cones**($A_I$,$c$,$B$,$\eta$,$\xi$,case,$h$) choosing $B$ in such a way that $B \subseteq I$ is a maximum cardinality set such that the rows of $A_B$ are linearly independent.

By **Choose**($B$,$I$) we will indicate the procedure that returns $B$ defined in this way.

If *RP* has solution $\xi$, then we have to consider solutions of type $x(\lambda)=\bar{x}+\lambda\xi$, with $\lambda \geq 0$. It can be immediately verified that:

- $\xi$ is a growth direction. In fact it gives

$$cx(\lambda) = c\bar{x} + \lambda c\xi > c\bar{x} \qquad \text{for } \lambda > 0;$$

- $\xi$ is a feasible direction. For each $\lambda \in [0,\bar{\lambda}]$, where

$$\bar{\lambda} = \begin{cases} \min\{\frac{b_i - A_i\bar{x}}{A_i\xi} : i\in\bar{I} \text{ and } A_i\xi > 0\}, & \text{if } \exists\ i\in\bar{I} \text{ and } A_i\xi > 0, \\ +\infty, & \text{if } \{i\in\bar{I}: A_i\xi > 0\} = \emptyset, \end{cases}$$

$x(\lambda)$ is a feasible solution. Further, if $\bar{\lambda}=+\infty$ then *P* is unbounded from above and *D* is empty.

To sum up, if *RP* has a solution, then either *D* is empty or we determine the solution $x(\bar{\lambda})$ such that $cx(\bar{\lambda}) > c\bar{x}$. Indeed, it should be observed that, according to construction, $\bar{\lambda} > 0$.

If *RD* has solution $\bar{\eta}$, then $\bar{x}$ and $\bar{y} = (\bar{\eta}, 0)$ are optimal solutions for *P* and *D*, respectively.

The algorithm summarized in the procedure of Figure 21 takes as input a feasible solution $\bar{x}$ and the set of active constraints *I*.

```
Procedure Simplex_Primal_Dual (A,b,c,I,x̄,ȳ,Dempty):
  begin
    optimal:=false; Dempty:=false;
    if |I|=0 then Determine(x̄,I,Dempty);
    Ī={1,..,m}\I ;
    if not Dempty then Choose (B,I);
    while not (Dempty or optimal) do
      begin
        Simplex_su_Cones(A_I,B,c,η,ξ,case,h);
        if case=1
          then begin ȳ_I:=η; ȳ:=[ȳ_I,0]; optimal:=true end
          else if {i∈Ī: A_iξ>0}=Ø then Dempty:=true
               else begin
                    λ̄:=min { (b_i−A_i x̄)/(A_i ξ) : i∈Ī and A_iξ>0};
                    x̄:=x̄+λ̄ξ; I:={i: A_i x̄=b_i};
                    if |B|=n
                      then begin k:=min{i: i∈I∩Ī}; B:=B\{h}∪{k} end
                      else Choose(B,I);
                    Ī:={1,..,m}\I
                  end
      end
  end.
```

Fig. 21: Primal-Dual Simplex Algorithm

By $I_r$ we indicate the set of active constraints at the iteration *r* of the algorithm and by $B_r$ a maximal subset of the set corresponding to linearly independent rows of matrix $A_{I_r}$; consider the case $r(A_{I_r})<n$, that is $|B_r|<n$. The procedure **Simplex_on_Cones** either returns $\eta\geq 0$ (*case*=1), thus causing the termination of the procedure, or returns $\xi$: $A_{I_r}\xi \leq 0$, $c\xi > 0$ (*case*=2), depending on the following cases:

a) $\eta A_{I_r} = c$ has a solution and $\eta_h < 0$. In this case the result is $A_h\xi < 0$, $A_i\xi = 0$, $i\in I_r\setminus\{h\}$.

Then it can be immediately verified that $h\notin I_{r+1}$, in fact $A_h\bar{x} + \bar{\lambda}A_h\xi < b_h$ because $\bar{\lambda} > 0$. The procedure **Choose**$(B,I)$ is called, which returns a set $B_{r+1}$ with cardinality greater than or equal to the cardinality of $B_r$. In fact, the displacement of $\bar{\lambda}$ along the feasible growth direction $\xi$ makes active at least one index $k$ – previously non-active – such that row $A_k$ is not linearly dependent from the rows of $A_{B_r\setminus\{h\}}$. If not, there would exist a non-null vector $\gamma$ such that:

$$A_k = \sum_{i\in B_r\setminus\{h\}} \gamma_i A_i.$$

Multiplying by $\xi$ we get $A_k\xi = \sum_{i\in B_r\setminus\{h\}} \gamma_i A_i\xi = 0$, but $A_k\xi > 0$ since $k\notin I_r$. Contradiction.

Note that we get $|B_{r+1}|>|B_r|$ in case there exist two or more indexes $k_1,\ldots, k_s \in I_{r+1}\setminus I_r$ such that matrix $A_{B_r\setminus\{h\}\cup\{k_1,\ldots,k_s\}}$ is composed of linearly independent rows.

b) $\eta A_{I_r} = c$ has no solution; **Simplex_on_Cones** approaches a problem having a number of rows smaller than the number of columns ($|I_r|<n$) admitting no solutions, thus it returns a feasible growth direction $\xi$ and the index $h$=0. Since $A_{I_r}\xi=0$ we get $I_{r+1}\supset I_r$; therefore there exists at least one index $k\in I_{r+1}\setminus I_r$ for which the rows of $A_{B_r\cup\{k\}}$ are linearly independent, as it can be verified in a way similar to the preceding case. Hence, the procedure **Choose**$(B,I)$ returns a set $B_{r+1}$ having cardinality strictly greater than the cardinality of $B_r$.

When $|B_r|=n$, since **Simplex_on_Cones** returns $\xi$ and $h$ such that

$$A_i\xi \begin{cases} = 0, & \text{if } i\in B_r\setminus\{h\}, \\ < 0, & \text{if } i=h, \end{cases}$$

we can immediately verify that the rows of $A_{B_r\setminus\{h\}\cup\{k\}}$, for each $k \in I_{r+1}\setminus I_r$ are linearly independent.

**Theorem 3.8**

*The procedure* ***Simplex_Primal_Dual*** *is correct.*

<u>Proof</u>

Clearly, if the procedure terminates, it yields the solution we were looking for; so we prove that it terminates in a finite number of iterations. At each iteration the system $A_{I_r}x = b_{I_r}$ has one and only one solution and further $cx^{(r+1)} > cx^{(r)}$, where $x^{(r)}$ indicates the feasible solution obtained at iteration $r$. Hence, there cannot exist two indexes $p$ and $q$, with $p > q$, such that $I_p=I_q$. The thesis follows from the fact that the distinct sets $I_r$, $r$=1, 2,.., $m$, are in a finite number.♦

## *3.6 Determination of a feasible solution*

So far we have always solved *LP* problems for which we knew a feasible solution that could be taken as starting point. In order to conclude the treatment of the subject and be able to solve any problem even in case no starting solution is available, we have to study how to determine a feasible solution or how to establish whether any exist.

Consider the problem:

$$P: \quad \max\ cx$$
$$Ax \le b$$

If $b \geq 0$, then $\bar{x}=0$ is banally a feasible solution for $P$. In case there exist some negative known terms, we need to solve an auxiliary problem in order to achieve a feasible starting solution.
Let $J_+ = \{i: b_i \geq 0\}$ and $J_- = \{i: b_i < 0\}$. We build the *auxiliary problem*

$$
\begin{array}{llcl}
AP: \max & & -ev & \\
 & A_{J_+}x & & \leq b_{J_+} \\
 & A_{J_-}x & -v & \leq b_{J_-} \\
 & & -v & \leq 0.
\end{array}
$$

It can be immediately verified that $x = 0$, $v = -b_{J_-}$ is a feasible solution for $AP$.

The **Simplex_Primal_Dual** algorithm allows to determine an optimal solution $(\bar{x},\bar{v})$ of $AP$.
Observe that an optimal solution exists because it was assumed that the objective function is bounded from above, and in fact it can have no positive values.

There are two possible cases:

i) $\bar{v} \neq 0$. In this case problem $P$ is not feasible. In fact, to a feasible solution $x$ of $P$ would correspond the solution $(x, 0)$ of $AP$ having null value of the objective function, consequently better than the optimum which was determined.

ii) $\bar{v} = 0$. In this case $\bar{x}$ is a feasible solution for $P$.

Thus, the solution of a $LP$ problem needs that the simplex algorithm is applied twice. For this reason it is frequently referred to as *two-phase method*.

<u>Example</u>
Let us consider the problem:

$$
\begin{array}{lrrcr}
\max & 3x_1 & +x_2 & & \\
 & x_1 & & \leq & 4 \\
 & -x_1 & -\ x_2 & \leq & -1 \\
 & -x_1 & +2x_2 & \leq & 2 \\
 & x_1 & +2x_2 & \leq & 14 \\
 & -x_1 & & \leq & 0 \\
 & & -x_2 & \leq & 0
\end{array}
$$

The second constraint has a negative known term excluding the origin from the feasible region. In order to verify if there exists a feasible solution, we formulate the auxiliary problem:

$$
\begin{array}{lrrrcr}
\max & & & -v_1 & & \\
 & x_1 & & & \leq & 4 \\
 & -x_1 & -\ x_2 & -v_1 & \leq & -1 \\
 & -x_1 & +2x_2 & & \leq & 2 \\
 & x_1 & +2x_2 & & \leq & 14 \\
 & -x_1 & & & \leq & 0 \\
 & & -x_2 & & \leq & 0 \\
 & & & -v_1 & \leq & 0
\end{array}
$$

A feasible solution is given by $x_1=0$, $x_2=0$, $v_1=1$, and from this solution we can apply the primal-dual simplex algorithm. The set of active constraints is $I=\{2,5,6\}$. The dual system is:

$$\begin{aligned} -\eta_1 - \eta_2 \quad &= 0 \\ -\eta_1 \quad - \eta_3 &= 0 \\ -\eta_1 \quad &= -1 \end{aligned}$$

having solution $\eta_1 = 1$, $\eta_2=-1$, $\eta_3=-1$, evidently unfeasible because of negative values. The feasible growth direction is obtained by solving the primal system:

$$\begin{aligned} -\xi_1 - \xi_2 - \xi_3 &= 0 \\ -\xi_1 \quad &= -1 \\ -\xi_2 \quad &= 0 \end{aligned}$$

having solution $\xi_1=1$, $\xi_2=0$, $\xi_3=-1$. The displacement step along this direction is by $\lambda=1$. The new point is: $x_1=1$, $x_2=0$, $v_1=0$, whose optimality can be easily verified. Since $v=0$ we can consider solution $x_1=1$, $x_2=0$ as starting point for solving the original problem.

**Exercise**

Consider the problem of the preceding example with known term vector $b$=(4,-10,2,14,0,0). Determine whether there exists a feasible solution or whether the feasible region is empty.

## 4 Complementary slackness and economic interpretation

Consider the following pair of dual problems:

$$P: \quad \max\ cx \qquad\qquad D: \quad \min\ yb$$
$$Ax \leq b \qquad\qquad yA = c$$
$$y \geq 0.$$

The following theorem holds; it is called complementary slackness theorem because it relates the constraint slackness of a problem to the variables of the other problem.

**Theorem 4.1**

*Let $\bar{x}$ and $\bar{y}$ be feasible solutions for P and D, respectively. The following properties are equivalent:*

*i)* $\bar{x}$ *and* $\bar{y}$ *are optimal solutions;*

*ii)* $c\bar{x} = \bar{y}b$;

*iii)* $\bar{y}(b - A\bar{x}) = 0$.

Proof

By the strong duality theorem *i*) and *ii*) are equivalent.

To complete the proof we just need to show that (*ii*) implies (*iii*) and that (*iii*) in its turn implies (*ii*), i.e.,:

$$c\bar{x} = \bar{y}b \quad \Rightarrow \quad \bar{y}A\bar{x} = \bar{y}b \quad \Rightarrow \quad \bar{y}(b - A\bar{x}) = 0,$$

$$\bar{y}(b - A\bar{x}) = 0 \quad \Rightarrow \quad \bar{y}A\bar{x} = \bar{y}b \quad \Rightarrow \quad c\bar{x} = \bar{y}b. \ \blacklozenge$$

**Definition 4.1**

Solutions $\bar{x}$ *and* $\bar{y}$ are said to be *complementary* if conditions *iii*) of theorem 4.1, called *complementary slackness* conditions, hold.

Given a pair of solutions $\bar{x}$ and $\bar{y}$ feasible for *P* and *D*, respectively, by theorem 4.1 such solutions are optimal if and only if *complementary slackness* conditions hold, component by component:

(4.1) $\quad \bar{y}_i\,(b_i - A_i\bar{x}) = 0, \ i = 1,\dots,m.$

In fact, since $\bar{y} \geq 0$ and $(b-A\bar{x}) \geq 0$ the scalar product in *iii)* is null if and only if (4.1) holds, element by element. Further, observe that (4.1) implies:

$$
(4.2) \qquad
\begin{array}{lcl}
\bar{y}_i > 0 & \Rightarrow & A_i\bar{x} = b_i, \\
A_i\bar{x} < b_i & \Rightarrow & \bar{y}_i = 0,
\end{array}
$$

that is: given $\bar{x}$ and $\bar{y}$ optimal for $P$ and $D$, if the $i$-th component of $\bar{y}$ is strictly positive, then the $i$-th constraint of $Ax \leq b$ is satisfied as equation; conversely, if the $i$-th constraint of $Ax \leq b$ is satisfied as strict inequality, then the $i$-th component of $\bar{y}$ is null. Obviously, if $\bar{x}$ and $\bar{y}$ are feasible and (4.1) holds, then $\bar{x}$ and $\bar{y}$ are optimal.

<u>Example: application of complementary slackness</u>

Consider the following pair of dual problems:

$$
P: \max\ x_1 + 2x_2 \qquad\qquad D: \min\ 5y_1 + 4y_2 + 3y_3
$$

$$
\begin{array}{rcl}
x_1 + x_2 & \leq & 5 \\
x_1 & \leq & 4 \\
x_2 & \leq & 3 \\
-x_1 & \leq & 0 \\
-x_2 & \leq & 0
\end{array}
\qquad\qquad
\begin{array}{l}
y_1 + y_2 - y_4 = 1 \\
y_1 + y_3 - y_5 = 2 \\
y_i \geq 0, \quad i = 1,\ldots,5.
\end{array}
$$

We consider $\bar{x}=\begin{bmatrix}2\\3\end{bmatrix}$ and verify if such solution is optimal by making use of the complementary slackness theorem.

By complementary slackness property we have:

$$
\begin{array}{ll}
\bar{x}_1 < 4 & \Rightarrow \bar{y}_2 = 0, \\
-\bar{x}_1 < 0 & \Rightarrow \bar{y}_4 = 0, \\
-\bar{x}_2 < 0 & \Rightarrow \bar{y}_5 = 0.
\end{array}
$$

Therefore, non-null components of a complementary solution $\bar{y}$ of $D$ must satisfy:

$$
\begin{cases}
y_1 = 1 \\
y_1 + y_3 = 2.
\end{cases}
$$

Hence, $\bar{y} = (1,0,1,0,0)$ is the dual complementary solution. Since $\bar{y} \geq 0$ we may conclude that $\bar{x}$ and $\bar{y}$ are optimal.

Otherwise, if we examine the point $\bar{x}=\begin{bmatrix}4\\0\end{bmatrix}$ and we control which constraints hold with the strict minimum, we can draw information about dual variables:

$$
\begin{array}{ll}
\bar{x}_1 + \bar{x}_2 < 5 & \Rightarrow \bar{y}_1 = 0, \\
\bar{x}_2 < 3 & \Rightarrow \bar{y}_3 = 0, \\
-\bar{x}_1 < 0 & \Rightarrow \bar{y}_4 = 0.
\end{array}
$$

Thus, the dual complementary solution is: $\bar{y}=(0,1,0,0,-2)$, which implies that $\bar{x}$ is not optimal.

**Exercise**

Compute the dual complementary solution for solution $\bar{x}=\begin{bmatrix}2\\6\end{bmatrix}$ in the shutter concern problem.

The complementary slackness theorem can be extended to the symmetric pair. Note that the conditions relate the constraint slackness of the primal to the variables of the dual but, differently from the asymmetric case, here we also have conditions relating the constraint slackness of the dual to the variables of the primal. This is due to the fact that in the symmetric pair both problems have inequality constraints, whereas in the asymmetric pair the dual has equality constraints, hence their slackness will always be null for any dual feasible solution.

**Theorem 4.2**
*Let:*

$$P: \quad \max\ cx \qquad\qquad D: \quad \min\ yb$$
$$\quad Ax \leq b \qquad\qquad\qquad yA \geq c$$
$$\quad x \geq 0 \qquad\qquad\qquad y \geq 0.$$

*And let $\bar{x}$ and $\bar{y}$ be feasible solutions for P and D, respectively;*
*then $\bar{x}$ and $\bar{y}$ are optimal solutions if and only if:*

$$\bar{y}_i\,(b_i - A_i\bar{x}) = 0, \quad i=1,\ldots,m,$$
$$(\bar{y}A^j - c_j)\bar{x}_j = 0, \quad j=1,\ldots,n.$$

Proof
Consider two solutions $\bar{x}$ and $\bar{y}$ feasible for $P$ and for $D$. Due to the feasibility of the two solutions we have:

$$c\bar{x} \leq \bar{y}A\bar{x} \leq \bar{y}b.$$

If $\bar{x}$ and $\bar{y}$ are optimal we have $c\bar{x} = \bar{y}b$, and the inequalities in the above expression hold as equalities. Therefore, considering the first equality $c\bar{x} = \bar{y}A\bar{x}$, we have:

$$(\bar{y}A - c)\bar{x} = 0$$

whereas, considering the second equality $\bar{y}A\bar{x} = \bar{y}b$, we have:

$$\bar{y}(b - A\bar{x}) = 0.$$

By repeating the same considerations, but in reverse, we prove that if complementary slackness conditions hold, then $c\bar{x} = \bar{y}b$ and solutions are optimal. ♦

### *4.1 Economic interpretation of dual variables*

By using the result of complementary slackness, we can give an economic interpretation to dual variables. Consider the pair of problems:

$$P: \quad \max\ cx \qquad\qquad D: \quad \min\ yb$$
$$\quad Ax \leq b \qquad\qquad\qquad yA = c$$
$$\qquad\qquad\qquad\qquad\qquad y \geq 0.$$

$P$ can be interpreted as a problem of utilizing the available resources (vector $b$) distributing them among a given set of activities (variables $x$) so as to maximize profit. Each column of matrix $A$ corresponds to an activity, and the value of the corresponding variable provides the activity level. Figure 22 illustrates a possible example (the optimal solution $\bar{x}$ is indicated as well).

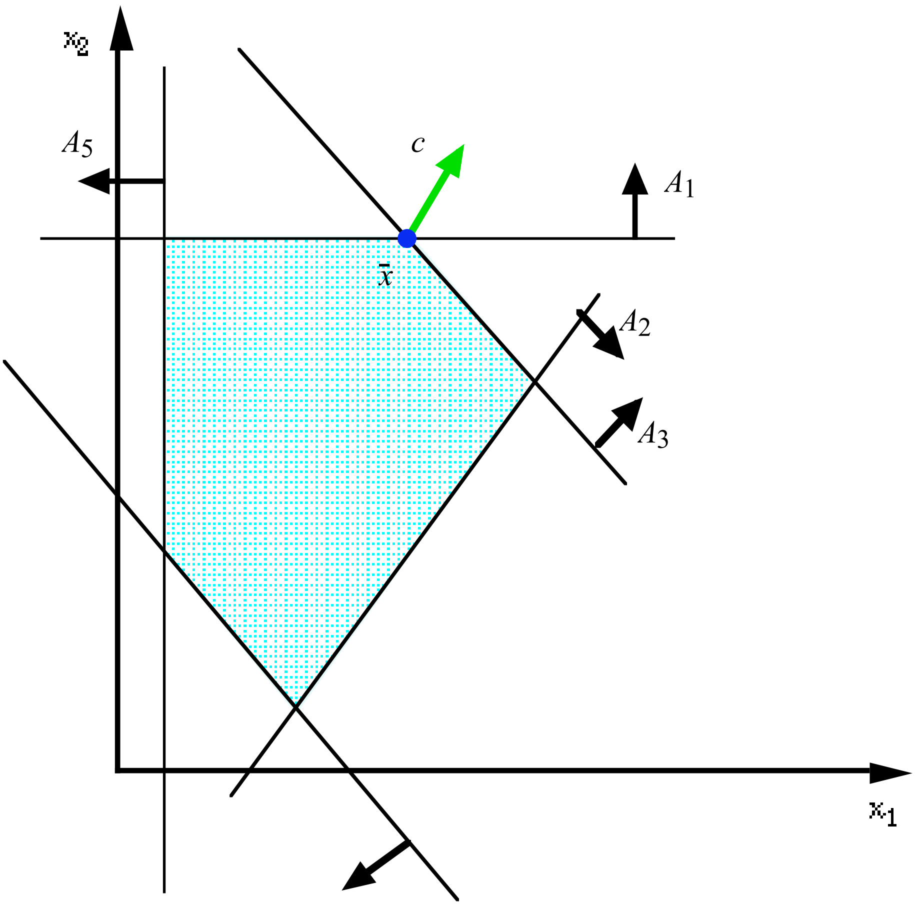


Fig. 22: example of optimal resource utilization problem

Let $\bar{y}$ be a complementary solution of $D$. Consider a slight variation of vector $b$, i.e., a variation of resource availability. We assume that such variation is small enough for optimal solution $\bar{x}$ to remain in the intersection of the hyperplanes where it was previously (thus, in the example of Figure 22 it must remain in the intersection of the first and the third constraint). We denote by $\bar{b}(\varepsilon)=b+\varepsilon$ the new value taken by vector $b$, where $\varepsilon\in\boldsymbol{R}^m$ is the vector of variations. We denote by $P(\varepsilon)$ and $D(\varepsilon)$ the primal and dual problems relative to the new vector $b$.

Clearly, having modified the known terms has no effect on the feasibility of $\bar{y}$ for $D(\varepsilon)$, in fact the constraints of the dual do not involve vector $b$. Obviously, known term variations influence the coordinates of the optimal solution of $P(\varepsilon)$ that we indicate by $\bar{x}(\varepsilon)$.

Observe that, according to previous considerations, $\bar{x}(\varepsilon)$ defines, for a ε being small enough, defines the same active constraints as $\bar{x}$. An immediate consequence of this is that the conditions of the complementary slackness theorem continue to hold also for the pair of solutions $\bar{x}(\varepsilon)$ and $\bar{y}$ of $P(\varepsilon)$ and $D(\varepsilon)$. In other words, $\bar{x}(\varepsilon)$ and $\bar{y}$ are feasible and complementary solutions, and the variation of the primal problem had no effect on the optimal solution of the dual. But such variation does have an effect on the value of such solution; in fact, the value of the objective function becomes:

$$c\bar{x}(\varepsilon)=\bar{y}(b+\varepsilon)=\bar{y}b+\bar{y}\varepsilon.$$

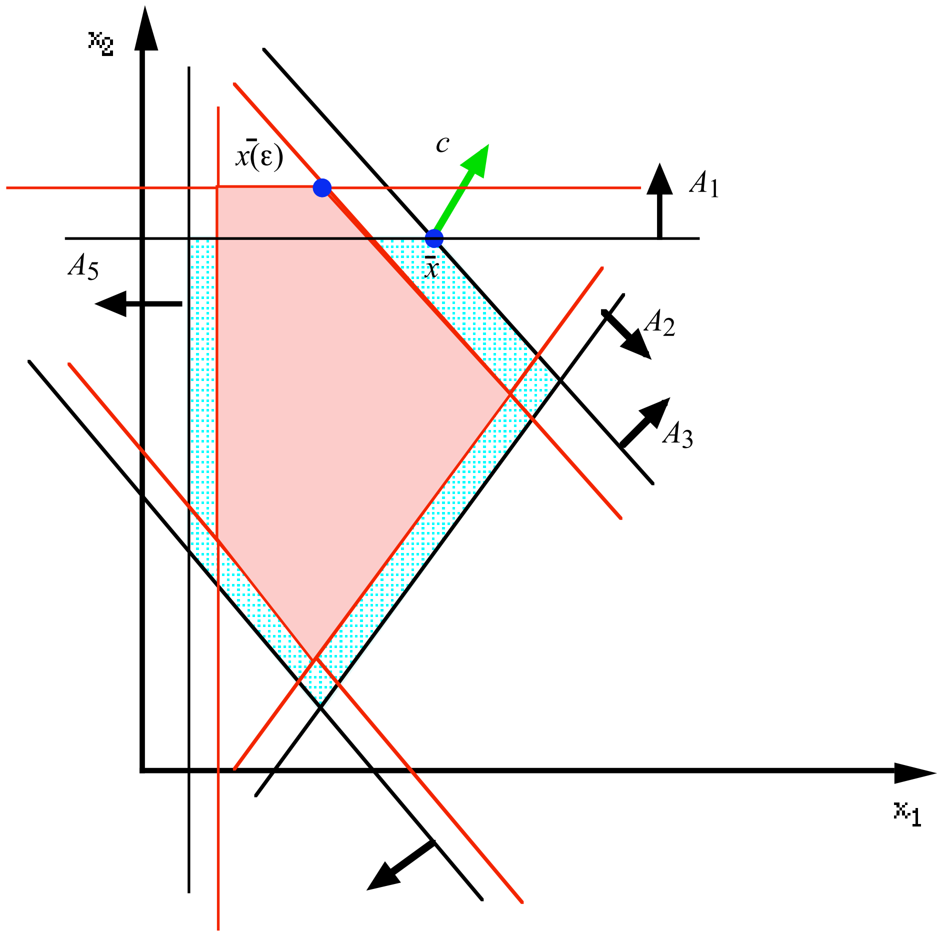


Fig. 23: resource variation and new optimal solution

Therefore, in the formulated hypotheses vector $\bar{y}$ represents the gradient of the optimal value of the objective function expressed according to the variation ε of $b$ and computed in the origin (ε=0). The single component $\bar{y}_i$ provides the variation of the optimal value of the objective function per unit variation[1] of the value of the $i$-th resource, thus indicating the maximum value that can be reasonably paid for an additional unit of such resource. In this perspective, we say that the optimal values of dual variables provide the *marginal values* (or *shadow prices*) of resources.
Shadow prices provide an estimate of the relative value assigned to different resources according to how they are used in the production process defined by problem $P$. On the basis of this interpretation, it also becomes clear why, in the optimal solution, dual variables relative to resources not utilized to the maximum are at zero.

Example: Shutters and Frames

In order to exemplify the concepts just presented, let us resume the shutter concern problem. More than once we have verified that point $\bar{x}=\begin{bmatrix}2\\6\end{bmatrix}$ is optimal and the corresponding dual complementary solution is $\bar{y}$=(0, 15, 10,0,0), with objective function value equal to 360. The question is: what happens if we increase the second resource by one unit, i.e., if we increase the time availability of the smith, bringing it from 12 to 13 working hours? Graphically, this corresponds to translating upwards by one unit the face corresponding to constraint $2x_2 \leq 12$. The vertex being intersection of

[1] Obviously, just in case we assume that such variation does not alter the optimal solution of the dual.

the faces corresponding to the first and to the third constraint moves to point $\bar{x}(\varepsilon)=\begin{bmatrix} 5/3 \\ 13/2 \end{bmatrix}$. This vertex keeps having the second and the third constraint as active constraints, so the dual solution $\bar{y}$ ))=(0, 15, 10,0,0) keeps being optimal. What changes is the value of the objective function, that now is 375. Observe that the increment is given by the value of the dual variable corresponding to the second constraint (15) by the increment of the second resource (1).

### **4.2 Primal-dual simplex and complementary slackness theorem*

Consider the asymmetric pair of problems ($P$, $D$) and let $\bar{x}$ be a feasible solution for $P$. From the complementary slackness theorem follows that $\bar{x}$ is an optimal solution if and only if there exists $\bar{y}$, a feasible solution of $D$ for which

$$\bar{y}(b - A\bar{x}) = 0$$

holds. Denoting by $I$ the set of active constraint indexes and by $\bar{I}= \{1,..,m\}\setminus I$, the preceding relation is equivalent to

$$\bar{y}_{\bar{I}}(b_{\bar{I}} - A_{\bar{I}}\bar{x}) = 0 \text{ which implies } \bar{y}_{\bar{I}} = 0.$$

Hence $\bar{x}$ is an optimal solution if and only if ($\bar{y}_I$, 0) is a feasible solution for the dual problem $D$, that is:

$$\begin{cases} \bar{y}_I A_I = c, \\ \bar{y}_I \quad \geq 0. \end{cases}$$

Observe that, according to how the algorithm builds the dual solution, at each iteration of **Simplex_Primal_Dual**, the pair of solutions $\bar{x}$ and $\bar{y}$ generated at each iteration of the algorithm respects the complementary slackness conditions.

### **4.3 Variation in the problem data: the case of vector c*

Remember that, in order to build the *LP* model which was solved, approximations had to be made because, for example, non-linear phenomena have been considered as linear, or else because of assumptions made in estimating parameters being not precisely known. Therefore, it is useful to know how stable the detected solution is, i.e., how much it is sensitive to small variations of data. In this case we speak of *sensitivity analysis*. Some of the data may occasionally be considered as function of one or more parameters, so the problem consists in determining the optimal value of the objective function as function of the parameters themselves; in this case we speak of *parametric analysis*. We partly examined this problem while studying how the variation of the known term can influence the value of the objective function. Now let us examine the variations of the vector of objective function $c$ and let us verify for which variations the optimal solution of the original problem remains optimal.

Consider the pair of problems $P$: max $\{cx: Ax \leq b\}$, $D$: min $\{yb: yA = c, y \geq 0\}$, with $\bar{x}$ and $\bar{y}$ as optimal solutions, respectively. Suppose that cost vector $c$ is replaced by the new vector $c'$. In order to keep having $\bar{x}$ as optimal solution, we must have that the dual complementary solution $y'$ relative to new vector $c'$ must remain optimal. Hence, if $I$ is the set of indexes active in $\bar{x}$, we must have:

$$y'A_I = c'$$
$$y'\geq 0$$

If vector $c'$ is function of a parameter $\alpha$, we can extract the optimality conditions to be imposed on the parameter.

Example

Consider the shutter concern problem, whose optimal solution, as we saw, is $\bar{x} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$. Now the question is for which door price values ($c_1=\alpha$) such solution remains optimal. Active constraints in $\bar{x}$ are the second and the third. We extract the dual solution $\bar{y}$ expressed in function of parameter $\alpha$:

$$3\, y_3 = \alpha$$
$$2\, y_2 + 2\, y_3 = 50$$

hence $\bar{y}=(0,\frac{\alpha}{3},50-\frac{2}{3}\alpha,0,0)$. In order to have the guarantee of optimality, we need to impose the non-negativity conditions on the different components of $\bar{y}$, hence we get:

$$\alpha \geq 0 \text{ and } \alpha \leq 75.$$

This means that solution $\bar{x}$ remains optimal for door price values comprised between 0 and 75. This analysis can be easily interpreted also from the geometric point of view. In fact, we have to extract for which values of $\alpha$ the vector $c$ continues to be included in the cone generated by $A_2$ and $A_3$. In order to do this it suffices to draw the horizontal line passing through $c$.

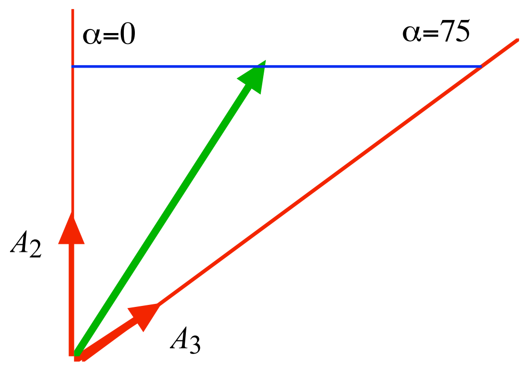


Fig. 24: geometric interpretation of the parametric analysis of $c$

**Exercise**

Compute (in the shutter concern problem) for which values of parameter $c_2=\alpha$ solution $\bar{x} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ remains optimal.

### *4.4 *Solution of problems of large extent: use of complementary slackness*

At the end of the chapter concerning flows we introduced a multi-terminal flow problem. The problem is specified by a network flow $G=(N,A)$, with costs $c_{ij} \geq 0$ and capacity $u_{ij}$ on arcs, as well as by a set of pairs of nodes $K$ and a matrix of demands for flow routing $d_{s_k t_k}$ between each pair of nodes $(s_k,t_k)$, $k \in K$. To formulate the problem we introduced the flow variables $x_p$ for each useful path $p$. $P$ denotes the set of all paths useful for solving the problem, and $P_k \subset P$ denotes – for each origin/destination pair $(s_k,t_k)$ – all paths going from $s_k$ to $t_k$. The cost of a path $p$ is given by the sum of the costs of the arcs composing it:

$$c_p = \sum_{(i,j) \in p} c_{ij}.$$

The path formulation of the problem is:

$$R\text{: } \min \sum_{k\in K}\sum_{p\in P_k} c_p x_p$$
$$\sum_{p\in P_k} x_p = d_{s_k t_k} \qquad \forall k\in K$$
$$\sum_{p\in P:(i,j)\in p} x_p \le u_{ij} \qquad \forall (i,j)\in A$$
$$x_p \ge 0 \qquad \forall p\in P$$

whose dual is:

$$RD\text{: } \max \sum_{k\in K} d_{s_k t_k} y_k + \sum_{(i,j)\in A} u_{ij} z_{ij}$$
$$y_k + \sum_{(i,j)\in p} z_{ij} \le c_p \qquad \forall k\in K,\ \forall p\in P_k$$
$$z_{ij} \le 0 \qquad \forall (i,j)\in A$$

This formulation, considering that the number of elementary paths between each pair of nodes of a graph can be exponential, requires a huge number of variables, which makes the problem difficult to solve. Yet, in order to solve the problem it is not necessary to make use of all paths and of the variables associated with them. In fact, knowing the set of paths $P' \subset P$ ($P'_k \subset P_k$, $k\in K$) that bring flow other than zero to the optimal solution, we might reduce the formulation as follows.

$$M\text{: } \min \sum_{k\in K}\sum_{p\in P'_k} c_p x_p$$
$$\sum_{p\in P'_k} x_p = d_{s_k t_k} \qquad \forall k\in K$$
$$\sum_{p\in P':(i,j)\in p} x_p \le u_{ij} \qquad \forall (i,j)\in A$$
$$x_p \ge 0 \qquad \forall p\in P'$$

But unfortunately, before obtaining the optimal solution, we are unable to know which is the set of paths $P'$ to which the attention must be limited. The idea contained in the solution method we propose is to consider any subset $P'$ ensuring the existence of a feasible solution of problem $M$ (called *master*). In comparison with the original problem, the master problem considers only one subset of variables but has the same set of constraints. Let us formulate the dual problem of $M$. We introduce a variable $y_k$ for each constraint imposed on the demand and a variable $z_{ij}$ for each capacity constraint. The dual is:

$$DM\text{: } \max \sum_{k\in K} d_{s_k t_k} y_k + \sum_{(i,j)\in A} u_{ij} z_{ij}$$
$$y_k + \sum_{(i,j)\in p} z_{ij} \le c_p \qquad \forall k\in K,\ \forall p\in P'_k$$
$$z_{ij} \le 0 \qquad \forall (i,j)\in A$$

It should be noted that $DM$ has exactly the same set of variables as $RD$, and indeed the latter variables refer to the pairs of $K$ and to the graph arcs and are not influenced by the choice of the subset of paths. What changes, with respect to the dual of $R$, is the set of constraints which, referring to the set of paths $P'$, is reduced. The optimal solution of $M$ has value greater than or equal to the optimal solution of $R$, in fact we have restricted the feasible region in $M$ by limiting the choice of variables. As it contains less constraints, $DM$ too has optimal value greater than or equal to $RD$. The question arising now is how can we understand whether the optimal solution of $M$ is optimal, i.e., whether the restricted set of paths $P'$ we considered is sufficient or must be extended.

Let us consider $\bar{x}$, optimal solution of *M*, and its complementary dual $(\bar{y},\bar{z})$. Wishing to verify the optimality of $\bar{x}$ for the original problem *R*, we should verify whether the dual solution $(\bar{y},\bar{z})$ is feasible for *RD* or there exist a pair $(s_k,t_k)$ and a path $p \in P_k$ for which

$$(4.3) \qquad \bar{y}_k + \sum_{(i,j)\in p} \bar{z}_{ij} > c_p.$$

Recalling that $c_p$ is given by the sum of the costs of the arcs composing *p*, the problem of seeking a path between $s_k$ and $t_k$ that makes (4.3) true can be related to the search for a longest path on graph *G* in which the cost of each arc $(i,j)$ is given by $\bar{z}_{ij} - c_{ij}$. If the length of the longest path is greater than or equal to $-\bar{y}_k$, we have found a constraint of *RD* violated by solution $(\bar{y},\bar{z})$. In this case we can add the detected path to the set of paths $P'_k$ and iterate the process. Otherwise, if for each $k \in K$ we cannot find such a path, we can say that no constraint of *RD* is violated and the pair of solutions $\bar{x}$ and $(\bar{y},\bar{z})$ is optimal.
Note that cost coefficients leading the search for the longest path are all smaller than or equal to zero, since variables $\bar{z}_{ij} \le 0$ and costs $c_{ij} \ge 0$. This means that we can easily determine the longest path by using one of the algorithms presented in the preceding chapter.

## *5 Complementary bases and simplex algorithm

Consider the pair of dual problems:

$$P: \quad \max\ cx \quad Ax \le b \qquad\qquad D: \quad \min\ yb \quad yA = c \quad y \ge 0$$

and let $B \subseteq \{1,\ldots,m\}$ be a set of indexes such that:

$$|B| = n,$$
$$det\,(A_B) \ne 0.$$

We associate with matrix $A_B$ the two vectors $\bar{x} \in \mathbf{R}^n$ and $\bar{y} \in \mathbf{R}^n$ defined as follows:

$$(5.1) \qquad \begin{aligned} &\bar{x} = A_B^{-1} b_B \\ &\bar{y} = (\bar{y}_B, \bar{y}_N) \text{ with } \bar{y}_B = cA_B^{-1}, \bar{y}_N = 0 \text{ and } N = \{1,\ldots,m\} \setminus B. \end{aligned}$$

Such vectors are called *basic solutions* for problems *P* and *D* and are associated with the *basis matrix* $A_B$. They are said to be *feasible*, *unfeasible*, *degenerate* and *non-degenerate* according to the conditions presented in the following table:

| | $\bar{x}$ | $\bar{y}$ |
|---|---|---|
| feasible | $A_N\bar{x} \le b_N$ | $\bar{y}_B \ge 0$ |
| unfeasible | $\exists\, j \in N$: $A_j\bar{x} > b_j$ | $\exists\, j \in B$: $\bar{y}_j < 0$ |
| degenerate | $\exists\, j \in N$: $A_j\bar{x} = b_j$ | $\exists\, j \in B$: $\bar{y}_j = 0$ |
| non-degenerate | $\forall\, j \in N$: $A_j\bar{x} \ne b_j$ | $\forall\, j \in B$: $\bar{y}_j \ne 0$ |

Observe that $\bar{x}$ and $\bar{y}$ satisfy complementary slackness conditions. In fact:

$$\bar{y}(b - A\bar{x}) = [\bar{y}_B, \bar{y}_N] \begin{bmatrix} b_B - A_B\bar{x} \\ b_N - A_N\bar{x} \end{bmatrix} = 0.$$

Vectors $\bar{x}$ and $\bar{y}$ are also called *pair of complementary basic solutions* associated with basis matrix (or just *basis*) $A_B$.

<u>Example</u>

Consider the pair of dual problems:

$$\begin{array}{lrrrl} P\text{: max} & x_1 & +3x_2 & & \\ & -2x_1 & +x_2 & \le & 1 \\ & x_1 & -2x_2 & \le & -4 \\ & x_1 & +x_2 & \le & 14 \\ & x_1 & & \le & 8 \\ & & -x_2 & \le & -4 \end{array}$$

$$\begin{array}{lrrrrrcl} D\text{: min} & y_1 & -4y_2 & +14y_3 & +8y_4 & -4y_5 & & \\ & -2y_1 & +y_2 & +y_3 & +y_4 & & = & 1 \\ & y_1 & -2y_2 & +y_3 & & -y_5 & = & 3 \\ & \multicolumn{7}{l}{y_1, y_2, y_3, y_4, y_5 \ge 0} \end{array}$$

The geometric representation of $P$ is illustrated in Figure 24.

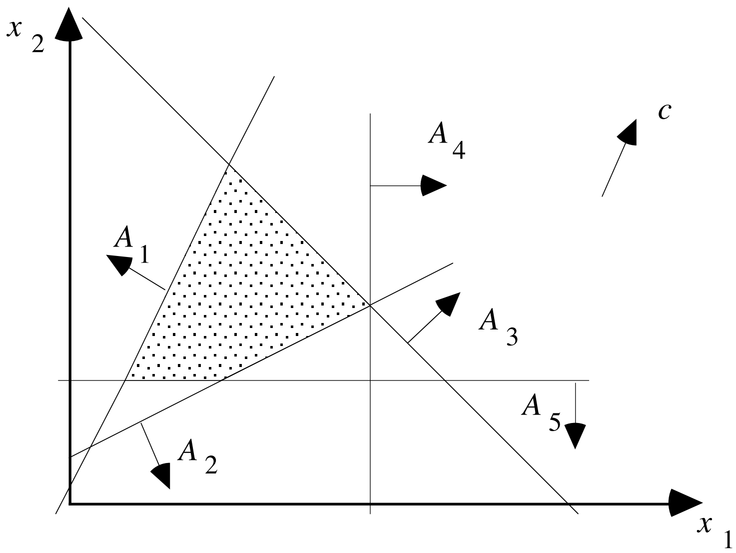


Fig. 24

Consider the basis $B$={2,5}. We have:

$$A_B = \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}, \quad A_B^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & -1 \end{bmatrix}, \quad \bar{x} = A_B^{-1}\begin{bmatrix} -4 \\ -4 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \qquad \begin{array}{l} A_1\bar{x} = -4 < 1 \\ A_3\bar{x} = 8 < 14 \\ A_4\bar{x} = 4 < 8 \end{array}$$

Hence, $\bar{x}$ is a basic feasible solution. The solution $\bar{y}$ associated with $A_B$ is given by:

$$\bar{y} = [\bar{y}_B, \bar{y}_N], \qquad \bar{y}_B = cA_B^{-1} = [1,-5], \quad \bar{y}_N = 0.$$

The solution $\bar{y}$ is unfeasible, $\bar{y}_5 = -5 < 0$. Further, note that $\bar{x}$ and $\bar{y}$ are non-degenerate solutions.

**Exercise**

Consider the basis $B$={2,4} for the problem of the preceding example. Verify optimality and degeneracy of the primal and dual solutions.

Let us consider now the problem of determining whether a given primal solution $\bar{x}$ is a basic solution. Remember that $I$ is the set of active constraints: $I = \{i: A_i\bar{x} = b_i\}$.

If $r(A_I) = n$, then $\bar{x}$ is a basic solution.

In fact, $r(A_I) = n \Rightarrow \exists\, B \subseteq I$: $|B| = n$, $det(A_B) \neq 0$; hence $\bar{x} = A_B^{-1} b_B$.

If $|I| = n$, then $B = I$ and $\bar{x}$ is a basic non-degenerate solution. In this case there exists only one vector $\bar{y}$ such that $(\bar{x},\bar{y})$ is a pair of complementary basic solutions. In fact, the basis matrix $A_B$ associated with $\bar{x}$ is univocally determined, and consequently vector $\bar{y}$ is univocally determined.

If $|I| > n$, then $\bar{x}$ is a degenerate solution. In this case, more than one basis matrices may correspond to $\bar{x}$, and consequently more than one basic solutions of $D$ can form a pair of complementary solutions with $\bar{x}$. Consider the following example.

Example

With reference to the problem presented in the example of Figure 24, consider $\bar{x} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}$ with $c\bar{x} = 26$.

We have:

$$\begin{array}{l} A_1\bar{x} = -10 < 1 \\ A_2\bar{x} = -4 \\ A_3\bar{x} = 14 \\ A_4\bar{x} = 8 \\ A_5\bar{x} = -6 \;\; < -4 \end{array}$$

Solution $\bar{x}$ is feasible and satisfies the constraints 2, 3, 4 as an equation, that is $I$ = {2,3,4}. $\bar{x}$ is a basic degenerate solution to which the following bases correspond: $B_1$={2,3}, $B_2$={2,4}, $B_3$={3,4},

$$\bar{x} = A_{B_1}^{-1}\begin{bmatrix} -4 \\ 14 \end{bmatrix} = A_{B_2}^{-1}\begin{bmatrix} -4 \\ 8 \end{bmatrix} = A_{B_3}^{-1}\begin{bmatrix} 14 \\ 8 \end{bmatrix} = \begin{bmatrix} 8 \\ 6 \end{bmatrix}.$$

As it can be verified, dual complementary solutions corresponding to $\bar{x}$ are given by:

$$\bar{y}^1 = [0,\tfrac{-2}{3},\tfrac{5}{3},0,0], \quad \bar{y}^2 = [0,\tfrac{-3}{2},0,\tfrac{5}{2},0], \qquad \bar{y}^3 = [0,0,3,-2,0],$$

which are unfeasible solutions for $D$ and $c\bar{x} = \bar{y}^1 b = \bar{y}^2 b = \bar{y}^3 b = 26$.

Now let us consider the problem of determining if a given dual solution $\bar{y}$ is a basic solution. Let $J = \{j: \bar{y}_j \neq 0\}$.

> If $A_j$, $j \in J$ are linearly independent,
> then $\bar{y}$ is a basic solution.

In fact, if $|J| = n$, then the basis matrix corresponding to $\bar{y}$ is $A_B = A_J$. In this case there exists only one vector $\bar{x}$ such that $(\bar{x},\bar{y})$ is a pair of complementary basic solutions. In fact, the basis matrix $A_B$ associated with $\bar{y}$ is univocally determined, and consequently vector $\bar{x}$ is univocally determined.
If $|J| < n$, then $\bar{y}$ is a degenerate solution; with this solution correspond more than one basis matrices obtained by adding (to matrix $A_J$) $n - |J|$ rows of $A$ (not being in $A_J$) so that the resulting matrix $A_B$ has non-null determinant. Observe that, since we have assumed $r(A)=n$, such a matrix exists. Consequently, more than one basic solutions of $P$ can form a pair of complementary solutions with $\bar{y}$. Consider the following example.

<u>Example</u>
Consider the example of Figure 24, but with gradient of the objective function $c = [1,1]$.
Let $\bar{y} = [0,0,1,0,0]$ and $\bar{y}b = 14$. $\bar{y}$ is a basic degenerate feasible solution for $D$. In fact, we get $J=\{3\}$. We can easily verify that the bases for $\bar{y}$ are $B_1 = \{1,3\}$, $B_2 = \{2,3\}$, $B_3 = \{3,4\}$, $B_4 = \{3,5\}$, so we get:

$$\bar{y}_{B_1} = cA^{-1}_{B_1} = \bar{y}_{B_2} = cA^{-1}_{B_2} = [0,1], \qquad \bar{y}_{B_3} = cA^{-1}_{B_3} = \bar{y}_{B_4} = cA^{-1}_{B_4} = [1,0].$$

As it can be verified, primal complementary solutions corresponding to $\bar{y}$ are given by:

$$\bar{x}^1 = A^{-1}_{B_1}\begin{bmatrix}1\\14\end{bmatrix} = \begin{bmatrix}\frac{13}{3}\\ \frac{29}{3}\end{bmatrix}, \qquad \text{basic feasible solution,}$$

$$\bar{x}^2 = A^{-1}_{B_2}\begin{bmatrix}-4\\14\end{bmatrix} = A^{-1}_{B_3}\begin{bmatrix}14\\8\end{bmatrix} = \begin{bmatrix}8\\6\end{bmatrix}, \qquad \text{basic degenerate feasible solution,}$$

$$\bar{x}^3 = A^{-1}_{B_4}\begin{bmatrix}14\\-4\end{bmatrix} = \begin{bmatrix}10\\4\end{bmatrix}, \qquad \text{basic unfeasible solution.}$$

The following theorem, which is a direct consequence of the definition of complementary basic solutions, holds.

**Theorem 5.1**
*Given a basic solution $\bar{x}$ for P [$\bar{y}$ for D], there exists at least one basic solution $\bar{y}$ for D [$\bar{x}$ for P] forming with $\bar{x}$ [$\bar{y}$] a pair of complementary basic solutions.*

### *5.1 Optimality conditions*

Since basic solutions are characterized by the basis defining them, we directly extract feasibility and optimality conditions on the basis matrix.
A basis matrix $A_B$ is called:

*primal feasible basis matrix* if: $b_N - A_N A_B^{-1} b_B \geq 0$;

*dual feasible basis matrix* if: $cA_B^{-1} \geq 0$.

Observe that a primal feasible basis matrix $A_B$ detects the basic solution $\bar{x} = A_B^{-1} b_B$ feasible for $P$. Similarly, a dual feasible basis matrix $A_B$ detects the solution $\bar{y} = [\bar{y}_B, \bar{y}_N] = [cA_B^{-1}, 0]$ feasible for $D$.

We can formalize the optimality conditions in the following theorem which is a consequence of previous considerations and of the complementary slackness theorem.

**Theorem 5.2** *Optimality conditions*

*Let $\bar{x}$ [$\bar{y}$] be a basic feasible solution for P [D] and $A_B$ a corresponding basis matrix.*

*Then $\bar{x}$ [$\bar{y}$] is an optimal solution if $cA_B^{-1} \geq 0$ [$b_N - A_N A_B^{-1} b_B \geq 0$].*

If $\bar{x}$ [$\bar{y}$] is non-degenerate, then, as we previously observed, the complementary solution $\bar{y}$ [$\bar{x}$] is univocally determined. It follows that the optimality condition is necessary and sufficient.
In the case of degenerate solutions, the condition is just sufficient, as shown by the following example.

<u>Example</u>

For the problem of Figure 25 consider the basic solution $\bar{x}$ such that $I = \{1,2,3\}$ and the bases $B_1 = \{1,2\}$, $B_2 = \{1,3\}$, $B_3 = \{2,3\}$.

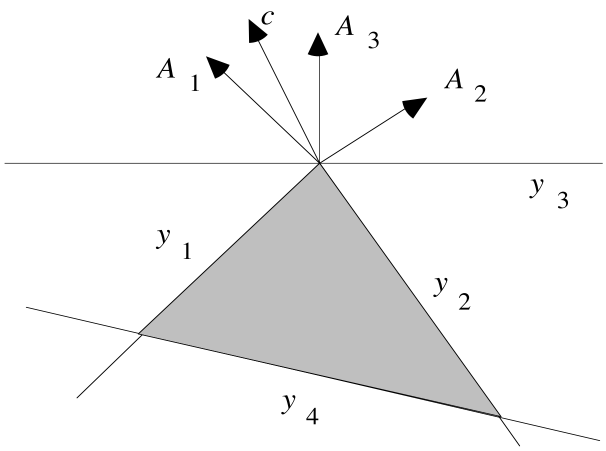


Fig. 25

$c \in cone(A_1, A_2) \Rightarrow \bar{y}_{B_1} \geq 0$;
$c \in cone(A_1, A_3) \Rightarrow \bar{y}_{B_2} \geq 0$;
$c \notin cone(A_2, A_3) \Rightarrow \bar{y}_{B_3} \ngeq 0$;

because in the last case we have $\bar{y}_2 < 0$ and $\bar{y}_3 > 0$. Hence, the basis matrix $A_{B_3}$ does not satisfy the optimality condition although $\bar{x}$ is an optimal solution.

We still have to prove that if $P$ and $D$ have finite optimum, then there exists a basis matrix, and therefore also a pair of optimal complementary basic solutions.

**Theorem 5.3**

*If P and D have finite optimum, then there exists at least one pair of complementary basic solutions, $(\bar{x},\bar{y})$, such that $\bar{x}$ is optimal solution of P and $\bar{y}$ is optimal solution of D.*

Proof

If it ends with $|B|=n$, the algorithm **Simplex_Primal_Dual** yields an optimal basis matrix $A_B$; otherwise, we have a case of dual degeneracy; thus, by performing further iterations starting from the non-basic optimal solution $x'$ we can determine a basic solution $\bar{x}$ such that $cx' = c\bar{x}$. ♦

**Exercise**

Build examples of P and D problems representable in $R^2$ with (both primal and dual) degenerate optimal solution and detect a basis matrix that does not satisfy optimality conditions.

### *5.2 Primal Simplex Algorithm*

The simplex method represents the first computationally efficient approach for solving Linear Programming problems. Originally proposed by G. B. Dantzig [1951] from a starting idea of J. von Neumann, the simplex method has been developed in several versions and is the basis of the most widely used Linear Programming codes.

Previously we introduced the primal-dual simplex algorithm. For the sake of completeness, now we introduce a variant called *Primal Simplex*. In practice, the primal simplex can be considered as a particular case of the primal-dual simplex in which from the very beginning we have a feasible basis, that is $|B|=n$. The peculiar property of the primal simplex of working on basic solutions allows, without making use of the procedure **Simplex_on_Cones**, the reorganization and simplification of some operations in the procedure **Simplex_Primal_Dual** as well as the elimination of operations relative to non-basic solutions.

The primal simplex algorithm is represented in Figure 26. In this algorithm we suppose we take as input a feasible basis $B$ and the inverse of the basis matrix, $A_B^{-1}$. In a finite number of steps, the algorithm yields an optimal basis matrix, or else it recognizes $D$ as empty, and consequently $P$ as unbounded from above.

```
Procedure Simplex_Primal (A,b,c,B,A_B^-1,Dempty,x̄,ȳ):
   begin
   1  optimal:=false; Dempty:=false;
      repeat
   2     x̄:=A_B^-1 b_B; ȳ_B:=cA_B^-1; ȳ_N:=0;
   3     if ȳ_B≥0 then optimal:=true
           else begin
   4          h:=min{i: ȳ_i<0}; r:=B^-1(h)(1);
   5          ξ:=−A_B^-1 e_r;
   6          if A_iξ≤0 ∀i∈N then Dempty:=true
                else begin
   7               x(λ):=x̄+λξ;
   8               θ:=max{λ: A_i x(λ)≤b_i, ∀i∈N, A_iξ>0}};
   9               k:=min{i∈N: A_i x(θ)=b_i};
   10              B:=B∪{k}\{h};
   11              update(A_B^-1)
                 end
              end
      until Dempty or optimal
   end
```

Fig. 26: Primal Simplex Algorithm.

<u>*Observations*</u>

The proposed algorithm coincides with the primal-dual simplex algorithm applied starting from a basic solution. In fact, as soon as it reaches a basis, the primal-dual simplex moves on adjacent bases. Thus, the correctness of the algorithm has been proved.

With update($A_B^{-1}$) we indicate a procedure determining the inverse of basis matrix $A_{B\cup\{k\}\setminus\{h\}}$ starting from the knowledge of the inverse $A_B^{-1}$.

The determination of θ at step 8 is achieved through *minimum ratio* evaluations:

$$\theta = \min\{\frac{b_i - A_i\bar{x}}{A_i\xi}: A_i\xi > 0, \forall i\in N\}.$$

The criterion for choosing the negative component $\bar{y}_B$ with lowest index $h$ and the choice of $k$ as lowest index detecting the minimum ratio represent Bland's anticycle rules, on which the proof of the algorithm convergence is based. In particular, the choice of $h$ corresponds to the detection of a growth direction. It can be experimentally proved that the choice of these directions produces a rather inefficient algorithm. Since in real problems the algorithm, even in the absence of anticycle rules, does not cycle, it is preferable to choose directions that (experimentally) guarantee a greater efficiency of the procedure.

For example, instruction 4 can be replaced by $h$:=max$\{|y_i|: y_i<0\}$. Observe that the value of the objective function $cx(\lambda)=c\bar{x}-\lambda y_h$. Therefore, the examined direction is the one providing the maximum unit increment of the objective function.

---

(1) $B^{-1}(h)$ provides the position in $B$ of index $h$ ($h=B(r) \Leftrightarrow r=B^{-1}(h)$).