Lecture notes on Linear Programming

February 7, 2025

Contents

1 Modeling

Example 1.1 *The problem of localization of fire stations.*

- *Objective: determine in which city to build fire stations in order to minimize the number of stations, knowing that every city must be able to be reached from at least one fire point in a maximum of 15 min.*

Example with 5 cities and the times between each city below:

1. Em qual cidade devemos construir um posto de bombeiros para
2. Minimizar o nº de postos
3. Sabendo que toda cidade deve ser alcançada por pelo menos 1 "fire point" = posto (?) em no máx 15min

1) 1, 2, 4
2) 1, 2, 5
3) 3,5
4) 1,4
5) 2,3,5

	C_1	C_2	C_3	C_4	C_5
C_1	0	10	25	15	30
C_2	10	0	20	50	10
C_3	25	20	0	30	10
C_4	15	50	30	0	70
C_5	30	10	10	70	0

- x_i *takes the value* 1 *if a station is built in city* i *and 0 otherwise.*
- *We can reach city 1 in a maximum of 15 minutes from 1, 2 or 4. To reach city 1 in a maximum of 15min, we then have to build a station in 1, 2 ou 4, i.e.,*

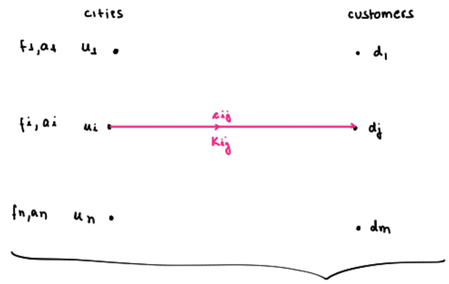
$$x_1 + x_2 + x_4 \geq 1.$$

Reasoning in the same way for the other cities, we obtain the PLI model:

$$\begin{cases} \min \ \sum_{i=1}^{5} x_i \\ x_1 + x_2 + x_4 \geq 1, \ x_1 + x_2 + x_5 \geq 1 \\ x_3 + x_5 \geq 1, \ x_1 + x_4 \geq 1 \\ x_2 + x_3 + x_5 \geq 1 \\ x_1, \, x_2, \, x_3, \, x_4, \, x_5 \in \{0,1\}. \end{cases}$$

Example 1.2 *Problem of locating plants and managing transport between plants and customers:*

- *n candidate cities.* u_i
- *Fix cust of placing a plant in city i: f_i.*
- *Daily production capacity of a plant located in city i: a_i.*
- *m customers: customer j with daily demand d_j.*
- *Transport cost between city i and customer j: c_{ij}.*
- *Maximum daily capacity that can be transported between i and j: K_{ij}.*



1) Var. Decisão : $y_i = \begin{cases} 1, \text{ se usina } i \text{ construída na cidade } i \\ 0, \text{ c.c} \end{cases}$

x_{ij} = qtd de produto transportada entre a usina i e o cliente j

2) Função Objetivo : $\min \underbrace{\sum_{i=1}^{n} f_i y_i}_{\text{custo de construção}} + \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij}}_{\text{custo de transporte}}$

3) Restrições : $0 \leq x_{ij} \leq K_{ij} y_i$

$\sum_{j=1}^{m} x_{ij} \leq a_i y_i, \ \forall i = 1, \ldots, n$

$\sum_{i=1}^{n} x_{ij} \geq d_j, \ \forall j = 1, \ldots, m$

Decision variables:

- *$y_i = 1$ if a plant is built in city i.*
- *x_{ij}: quantity transported between city i and customer j.*

The plant location model is given by:

$$\begin{cases} \overset{\text{min}}{\cancel{\max}} \ \sum_{i=1}^{n} f_i y_i + \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} x_{ij} \\ 0 \leq x_{ij} \leq y_i \, K_{ij} \\ \sum_{i \,|\, (i,j) \in E} \ x_{ij} \geq d_j, \ \forall \, j = 1, \ldots, m, \\ \sum_{j \,|\, (i,j) \in E} \ x_{ij} \leq a_i, \ \forall \, i = 1, \ldots, n, \\ y_i \in \{0,1\}, \ \forall \, i = 1, \ldots, n. \end{cases}$$

Example 1.3 *Investment problem.*

- *We have 4 projects. We want to decide which projects to invest in based on the following data*

Projeto	*Custo*	*VPLN*
P_1	5000	8000
P_2	7000	11000
P_3	4000	6000
P_4	3000	4000

→ "retorno"

and knowing that we have a capital of 14,000 to invest today.

(*) dividir por 1000 ñ muda a solução ótima. só multiplicar por 1000 depois

Model for choosing projects: using the variables x_i , with value equal to 1 if we invest in project i and 0 otherwise, the model is written

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14 \\ x_i \in \{0,1\},\ i = 1, \ldots, 4. \end{cases}$$

- No caso particular de variáveis binárias $x_i \in \{0,1\}$
  - ↳ podemos supor $c \geq 0$, $c_i \geq 0$
    se $c_i < 0$, fazemos uma mudança de variáveis
    $y_i = 1 - x_i \in \{0,1\}$
    $c_i x_i = c_i(1 - y_i) = \overbrace{c_i - c_i y_i}^{>0}$ c/ $y_i \in \{0,1\}$
  - ↳ $x_i^2 = x_i$ (passa ser linear, que são os problemas + simples)
  - ↳ se a função objetivo não é linear, mas tem termos como $x_1 x_2$, podemos transformar o problema de PLI, trocando $x_1 x_2$ por y com as restrições: * podemos linearizar

$$\begin{cases} x_1 + x_2 - y \leq 1 \\ -x_1 - x_2 + 2y < 0 \\ x_1, x_2, y \in \{0,1\} \end{cases}$$

$\to x_1 = 0$ ou $x_2 = 0$

We consider an integer linear programming (ILP) problem:

$$\begin{cases} \max\ c^\top x \\ Ax = b \\ x \geq 0 \\ x \in \mathbb{Z}^n \end{cases} \quad ou \quad \begin{cases} \min\ c^\top x \\ Ax = b \\ x \geq 0 \\ x \in \mathbb{Z}^n. \end{cases}$$

In the particular case of binary variables: $x_i \in \{0,1\}$:

- *We can assume $c \geq 0$. If $c_i < 0$, we make the change of variables $y_i = 1 - x_i \in \{0,1\}$.*
- *Observe that $x_i^2 = x_i$.*
- *If the objective function is not linear but has terms of the type $x_1 x_2$, we can transform the problem into a PLI exchanging $x_1 x_2$ by y with the restrictions*

$$\begin{aligned} x_2 + x_3 - y &\leq 1 \\ -x_2 - x_3 + 2y &\leq 0 \\ x_2,\ x_3,\ y &\in \{0,1\}. \end{aligned}$$

Example 1.4 *A Warehousing Problem.* *Consider the problem of operating a warehouse, by buying and selling a certain commodity which can be stored in the warehouse, in order to maximize profit over a certain time window. The warehouse has a fixed capacity C and there is a cost r per unit for holding stock for one period. The price, p_i of the commodity is known to fluctuate over a number of time periods—say months, indexed by i. In any period the same price holds for both purchase or sale. The warehouse is originally empty and is required to be empty at the end of the last period. To formulate this problem, variables are introduced for each time period. In particular, let x_i denote the level of stock in the warehouse at the beginning of period i. Let u_i denote the amount bought during period i, and let s_i denote the amount sold during period i. If there are n periods, the problem is*

<u>Armazém</u>: compra e vende um commodity que pode ser armazenado

- quer maximizar o retorno em um certo período de tempo
- tem uma capacidade máx. C
  e um custo (r) por unidade p/ guardar "stock" por um período
- o preço do commodity (p_i) muda de acordo c/ o nº de períodos (meses) i.
- Em qualquer período, o preço é o mesmo p/ comprar e p/ vender

- O armazém começa vazio e deve terminar vazio no último período

- P/ formular esse problema, variáveis são introduzidas p/ cada período:

Em particular: seja x_i o nível de "stock" no armazém no início do período i
u_i a qtd. comprada em i
s_i a qtd. vendida em i
} são n períodos

preço, qtd. vendida, qtd. comprada, custo de armazenamento, qtd. armazenada

$$\begin{array}{ll} \text{maximize} & \sum_i^n \left(p_i(s_i - u_i) - r x_i\right) \\ \text{subject to} & x_{i+1} = x_i + u_i - s_i \quad i = 1, 2, \ldots, n-1 \\ & 0 = x_n + u_n - s_n \\ & x_i + z_i = C \quad i = 2, \ldots, n \\ & x_1 = 0,\ x_i \geq 0,\ u_i \geq 0, \quad s_i \geq 0,\ z_i \geq 0, \end{array}$$

- Variáveis de Decisão: u_i, s_i e x_i
- F. Objetivo: max lucro - gastos

Slack variable p/ poder usar o "=" (folga)

* alguns solvers trabalham c/ restrições de igualdade, por isso às vezes é bom colocar z_i

$$\begin{cases} \max \sum_{i=1}^{n} p_i(s_i - u_i) - r x_i \\ x_1 = 0 \\ x_{i+1} = x_i + u_i - s_i \ , \ i = 1, \dots, n-1 \\ 0 = x_n + u_n - s_n \\ x_i \leq C \quad \text{ou} \quad x_i + z_i = C \\ x_i, u_i, s_i \geq 0 \ , \ z_i \geq 0 \end{cases}$$

$-1 \quad -1 \quad -1$

where z_i is a slack variable. If the constraints are written out explicitly for the case $n = 3$, they take the form

$-x_1 - u_1 + s_1$	$+\ x_2$		$= 0$
	$-x_2 - u_2 + s_2$	$+\ x_3$	$= 0$
	$x_2 + z_2$		$= C$
		$x_3 + u_3 - s_3$	$= 0$
		$x_3 + z_3$	$= C$

Note that the coefficient matrix can be partitioned into blocks corresponding to the variables of the different time periods. The only blocks that have nonzero entries are the diagonal ones and the ones immediately above the diagonal. This structure is common for problems involving time.

Example 1.5 Manufacturing Problem. *Suppose we own a facility that is capable of manufacturing n different products, each of which may require various amounts of m different resources. Each product j can be produced at any level $x_j \geq 0$, $j = 1, 2, \ldots, n$, and each unit of the jth product can sell for p_j dollars and needs a_{ij} units of the i-th resource, $i = 1, 2, \ldots, m$. Assuming linearity of the production costs, if we are given a set of m numbers $b_1, b_2, \ldots, b_m$ describing the available quantities of the m resources, and we wish to manufacture products to maximize the revenue, we need to maximize*

· Fábrica
↳ produz n produtos diferentes
↳ cada produto j pode ser produzido a qualquer nível $x_j \geq 0$, $j = 1, 2, \ldots, n$ e cada unidade pode ser vendida por $\$p_j$. Também precisa de a_{ij} unidades do recurso $i = 1, \ldots, m$

$$p_1 x_1 + p_2 x_2 + \ldots + p_n x_n$$

subject to the resource constraints

$$a_{i1} x_1 + a_{12} x_2 + \ldots + a_{in} x_n \leq b_i, \quad i = 1, \ldots, m$$

↳ Supondo linearidade dos custos de produção e $b_1, \ldots, b_m$ = qtd. disponível dos recursos

and the nonnegativity constraints on all production variables.

↳ Queremos maximizar o retorno

Example 1.6 Linear Classifier and Support Vector Machine. *Suppose several d-dimensional data points are classified into two distinct classes. For example, two-dimensional data points may be grade averages in science and humanities for different students. We also know the academic major of each student, as being in science or humanities, which serves as the classification. In general, we have vectors $a_i \in \mathbb{R}^d$ for $i = 1, 2, \ldots, n_1$ and vectors $b_j \in \mathbb{R}^d$ for $j = 1, 2, \ldots, n_2$. We wish to find a hyperplane that separates the a_i's from the b_j's. Mathematically, we wish to find $y \in \mathbb{R}^d$ and a number β such that*

$$a_i^T y + \beta \geq 1 \quad \text{for all } i,$$
$$b_j^T y + \beta \leq -1 \quad \text{for all } j,$$

where $\{x \ : \ x^T y + \beta = 0\}$ is the desired hyperplane, and the separation is defined by the values $+1$ and -1. This is a linear program.

· Suponha que diversos pontos d-dimensionais são classificados em 2 classes dist.

- **Network flow problems:**

Ex: pts 2d, médias de notas em ciências e humanidades
↳ sabemos o curso do estudante (C ou H) serve como classificação

Em geral, temos vetores $a_i \in \mathbb{R}^d$ p/ $i = 1, 2, \ldots, n_1$ e vetores $b_j \in \mathbb{R}^d$ $j = 1, 2, \ldots, n_2$
Queremos achar o hiperplano que separa a_i de b_j
Matematicamente, queremos achar $y \in \mathbb{R}^d$ e um nº β t.q.

$$\left.\begin{aligned} a_i^T y + \beta &\geq 1, \ \forall i \\ b_j^T y + \beta &\leq -1, \ \forall j \end{aligned}\right\}$$

onde $\{x : x^T y + \beta = 0\}$ é o hiperplano desejado e a separação é definida pelos valores +1 e −1

1.5) produzir n produtos a partir de m ingredientes

Estoque do ingrediente i : b_i

Produto j tem a_{ij} unidades do ingrediente i

Preço do produto j : p_j

1) Variáveis de Decisão : x_j = qtd. do produto j produzida , $x_j \in \mathbb{N}$

2) Função Objetivo : $\max \; p_1 x_1 + \cdots + p_n x_n$

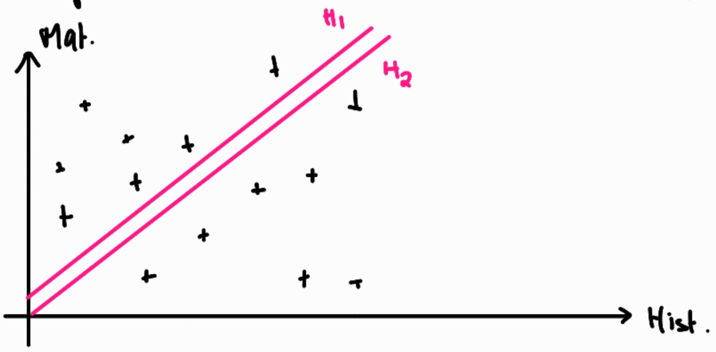
3) Restrições

• de Recurso:

$$a_{i1} p_1 + a_{i2} p_2 + \cdots + a_{in} p_n \leq b_i \;, \; i = 1, \ldots, m$$

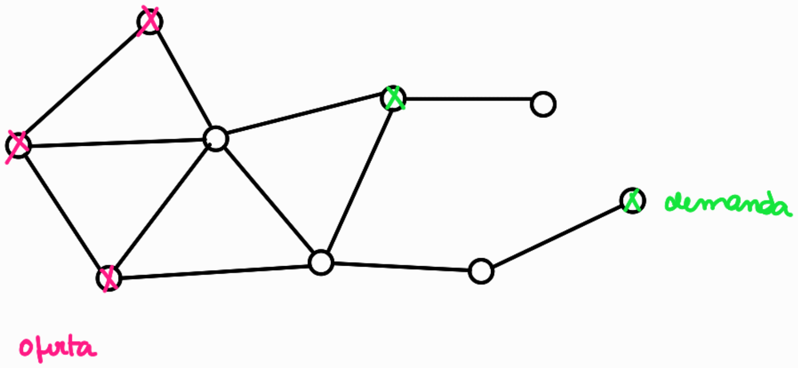
1.6) S.V.M - support Vector Machine - problema de viabilidade

* função objetivo irrelevante
* procuro um x que cumpra as restrições



$$\boxed{\begin{array}{l} a_i^T y + \beta \geq 1 \\ b_j^T y + \beta \leq -1 \end{array}}$$

1.7) Modelo de Transporte



1.8) Transportation Problem

Network Flow Problems:

Example 1.7 The general network flow problem: *Network-flow problems deal with the distribution of a single homogeneous product from plants (origins) to consumer markets (destinations). The total number of units produced at each plant and the total number of units required at each market are assumed to be known. The product need not be sent directly from source to destination, but may be routed through intermediary points reflecting warehouses or distribution centers. Further, there may be capacity restrictions that limit some of the shipping links. The objective is to minimize the variable cost of producing and shipping the products to meet the consumer demand. The sources, destinations, and intermediate points are collectively called nodes of the network, and the transportation links connecting nodes are termed arcs.*

	Urban transportation	Communication systems	Water resources
Product	Buses, autos, etc.	Messages	Water
Nodes	Bus stops, street intersections	Communication centers, relay station	Lakes, reservoirs, pumping stations
Arcs	Street (lanes)	Communication channels	Pipelines, canals, rivers.

Table 1: Examples of Network Flow Problems

Example 1.8 A special network flow problem: The Transportation Problem. *Quantities $a_1, a_2, \ldots, a_m$, respectively, of a certain product are to be shipped from each of m locations and received in amounts $b_1, b_2, \ldots, b_n$, respectively, at each of n destinations. Associated with the shipping of a unit of product from origin i to destination j is a shipping cost c_{ij}. It is desired to determine the amounts x_{ij} to be shipped between each origin–destination pair $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$; so as to satisfy the shipping requirements and minimize the total cost of transportation. To formulate this problem as a linear programming problem, we set up the array shown below:*

$$\begin{array}{|cccc|c}
x_{11} & x_{12} & \ldots & x_{1n} & a_1 \\
x_{21} & x_{22} & \ldots & x_{2n} & a_2 \\
\vdots & & & \vdots & \vdots \\
x_{m1} & x_{m2} & \ldots & x_{mn} & a_m \\
\hline
b_1 & b_2 & \ldots & b_n &
\end{array}$$

The ith row in this array defines the variables associated with the ith origin, while the jth column in this array defines the variables associated with the jth destination. The problem is to place nonnegative variables x_{ij} in this array so that the sum across the ith row is a_i, the sum down the jth column is b_j, and the weighted sum $\sum_{j=1}^{n} \sum_{i=1}^{m} c_{ij} x_{ij}$, representing the transportation cost, is minimized. Thus, we have the linear program-

ming problem:

$$
\begin{array}{lll}
\textit{minimize} & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} & \\
\textit{subject to} & \sum_{j=1}^{n} x_{ij} = a_i & \textit{for } i = 1,2,\ldots,m \qquad (1.1)\\
& \sum_{i=1}^{m} x_{ij} = b_j & \textit{for } j = 1,2,\ldots,n \qquad (1.2)\\
& x_{ij} \geq 0 & \textit{for } i = 1,2,\ldots,m, \quad j = 1,2,\ldots,n.
\end{array}
$$

For constraints (1.1) and (1.2) to be consistent, we must, of course, assume that $\sum_{i=1}^{m} a_i = \sum_{j=1}^{m} b_j$ *which corresponds to assuming that the total amount shipped is equal to the total amount received.*

Let us consider a simple example. A compressor company has plants in three locations: Cleveland, Chicago, and Boston. During the past week the total production of a special compressor unit out of each plant has been 35, 50, and 40 units respectively. The company wants to ship 45 units to a distribution center in Dallas, 20 to Atlanta, 30 to San Francisco, and 30 to Philadelphia. The unit production and distribution costs from each plant to each distribution center are given in the is presented in Table 2. What is the best shipping strategy to follow?

	Distribution centers				
Plants	1 Dallas	2 Atlanta	3 San Francisco	4 Philadelfia	Availability (units)
1 Cleveland	8	6	10	9	35
2 Chicago	9	12	13	7	50
3 Boston	14	9	16	5	40
Requirements (units)	45	20	30	30	125

Table 2: Unit Production and Shipping Costs

The linear-programming formulation of the corresponding transportation problem is:

Minimize $z = 8x_{11}+6x_{12}+10x_{13}+9x_{14}+9x_{21}+12x_{22}+13x_{23}+7x_{24}+14x_{31}+9x_{32}+16x_{33}+5x_{34}$,

subject to

$$
\begin{array}{cccccccccccccc}
x_{11} & +x_{12} & +x_{13} & +x_{14} & & & & & & & & & = & 35,\\
& & & & x_{21} & +x_{22} & +x_{23} & +x_{24} & & & & & = & 50,\\
& & & & & & & & x_{31} & +x_{32} & +x_{33} & +x_{34} & = & 40,\\
-x_{11} & & & & -x_{21} & & & & -x_{31} & & & & = & -45,\\
& -x_{12} & & & & -x_{22} & & & & -x_{32} & & & = & -20,\\
& & -x_{13} & & & & -x_{23} & & & & -x_{33} & & = & -30,\\
& & & -x_{14} & & & & -x_{24} & & & & -x_{34} & = & -30,
\end{array}
$$

$$x_{ij} \geq 0 \quad (i = 1,2,3;\ j = 1,2,3,4).$$

Example 1.9 A special network flow: The Assignment Problem. *n people are to be assigned to n jobs and that c_{ij} measures the performance of person i in job j. If we let*

$$x_{ij} = \begin{cases} 1 & \text{if person } i \text{ is assigned to job } j, \\ 0 & \text{otherwise,} \end{cases}$$

we can find the optimal assignment by solving the optimization problem:

$$\text{Maximize } z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij},$$

subject to:

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad (i = 1, \ldots, n)$$ mostra que cada pessoa tem só 1 trabalho

se mult. por (-1), o modelo seria de usual network interpretation

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad (j = 1, \ldots, n)$$ cada trabalho pode ser atribuído a só 1 pessoa

$$x_{ij} = 0 \quad \text{or} \quad 1 \quad (i = 1, \ldots, n; \; j = 1, \ldots, n)$$

The first set of constraints shows that each person is to be assigned to exactly one job and the second set of constraints indicates that each job is to be performed by one person. If the second set of constraints were multiplied by minus one, the equations of the model would have the usual network interpretation. As stated, this assignment problem is formally an integer program, since the decision variables x_{ij} are restricted to be zero or one. However, if these constraints are replaced by $x_{ij} \geq 0$, the model becomes a special case of the transportation problem, with one unit available at each source (person) and one unit required by each destination (job). Network-flow problems have integer solutions, and therefore formal specification of integrality constraints is unnecessary. Consequently, application of the simplex method, or most network-flow algorithms, will solve such integer problems directly.

Example 1.10 A special network flow: The Shortest-Path Problem. *Given a network with distance c_{ij} (or travel time, or cost, etc.) associated with each arc, find a path through the network from a particular origin (source) to a particular destination (sink) that has the shortest total distance. A number of important applications can be formulated as shortest- (or longest-) path problems where this formulation is not obvious at the outset. Further, the shortest-path problem often occurs as a subproblem in more complex situations, such as the subproblems in applying decomposition to traffic-assignment problems or the group-theory problems that arise in integer programming. In general, the formulation of the shortest-path problem is as follows:*

$$\text{Minimize } z = \sum_{i} \sum_{j} c_{ij} x_{ij}$$

• Var. Decisão:

$x_{ij} = \begin{cases} 1, \text{ se } (i,j) \text{ pertence ao caminho mais curto de } s \text{ até } t \\ 0, \text{ c.c} \end{cases}$

• tendo uma rede c/ uma distância (ou travel time ou custo, ...) associada a cada arco (aresta), encontre um caminho que saia de uma origem em particular (source / fonte) para um destino (sink) que tenha a menor distância total.

Shortest Path pode ser interpretado como network-flow: mandar 1 un. da source p/ a sink c/ custo mínimo. Source produz 1 un. sink demanda 1 un. } e os outros nós ñ adicionam nem tiram nada

$\min \sum_{i,j} c_{ij} x_{ij}$

$1 = \sum_{j \mid (s,j) \in E} x_{sj}$

$1 = \sum_{j \mid (j,t) \in E} x_{jt}$

$\sum_{j \mid (j,i) \in E} x_{ji} = \sum_{j \mid (i,j) \in E} x_{ij}$

$x_{ij} \in \{0,1\}, \forall i,j$

subject to:

$$\sum_j x_{ij} - \sum_k x_{ki} = \begin{cases} 1 & \textit{if } i = s \textit{ (source)}, \\ -1 & \textit{if } i = t \textit{ (sink)} \\ 0 & \textit{otherwise}, \end{cases}.$$

$$x_{ij} \geq 0 \textit{ for all arcs } i - j \textit{ in the network}$$

We can interpret the shortest-path problem as a network-flow problem very easily. We simply want to send one unit of flow from the source to the sink at minimum cost. At the source, there is a net supply of one unit; at the sink, there is a net demand of one unit; and at all other nodes there is no net inflow or outflow.

- **Linear integer programming problems:** Consider the problem

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

subject to

$$\begin{aligned} &\sum_{j=1}^{n} a_{ij} x_j = b_i \quad (i = 1, 2, \ldots, m), \\ &x \geq 0 \quad (j = 1, 2, \ldots, n) \\ &x_j \text{ integer} \qquad \text{(for some or all } j) \end{aligned} \tag{1.3}$$

Programa Inteiro Misto: nem todas as variáveis precisam ser inteiras

Programa Inteiro Puro: todas as var.'s precisam ser int.

This problem is called the (linear) integer-programming problem. It is said to be a mixed integer program when some, but not all, variables are restricted to be integer, and is called a pure integer program when all decision variables must be integers.

Example 1.11 Capital Budgeting. *In a typical capital-budgeting problem, decisions involve the selection of a number of potential investments. The investment decisions might be to choose among possible plant locations, to select a configuration of capital equipment, or to settle upon a set of research-and-development projects. Often it makes no sense to consider partial investments in these activities, and so the problem becomes a go–no-go integer program, where the decision variables are taken to be $x_j = 0$ or 1, indicating that the j-th investment is rejected or accepted. Assuming that c_j is the contribution resulting from the j-th investment and that a_{ij} is the amount of resource i, such as cash or manpower, used on the j-th investment, we can state the problem formally as:*

- seleção de potenciais investimentos
- As decisões podem ser investir ou não $x_j = 0$ ou 1
- c_j = retorno do investimento j
- a_{ij} = qtd de recurso ($ ou manpower) utilizado no inv. j

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

subject to

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

The objective is to maximize total contribution from all investments without exceeding the limited availability of any resource. One important special scenario for the capital-budgeting problem involves cash-flow constraints. In this case, the constraints

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

→ os fundos necessários p/ investimento deve ser menor ou igual aos fundos gerados nos investimentos anteriores + fundos exógenos aplicados

1.11) Problema de investimentos em projetos:

1) Var. Decisão: $x_j = \begin{cases} 1, & \text{se investimos no projeto } j \\ 0, & \text{c.c} \end{cases}$

Dados: c_j = retorno do projeto j

b_i = qtd de recurso i disponível

a_{ij} = qtd do recurso i consumido pelo projeto j

2) Função Objetivo:

Modelo é

$$\begin{cases} \max \sum_{j=1}^{n} c_j x_j \quad (\text{max. o lucro}) \\ \sum_{j=1}^{n} \underbrace{a_{ij} x_j}_{\text{qtd de recurso } i \text{ cons. pelo projeto } j} \leq b_i, \ \forall i \\ x_j \in \{0,1\}, \ \forall j \end{cases}$$

Modelagem de 2 restrições sobre os projetos:

A) Se investir no projeto i, temos que investir no projeto j

Ou seja $x_i = 1 \Rightarrow x_j = 1$

A restrição que modela A é $\boxed{x_j \geq x_i}$

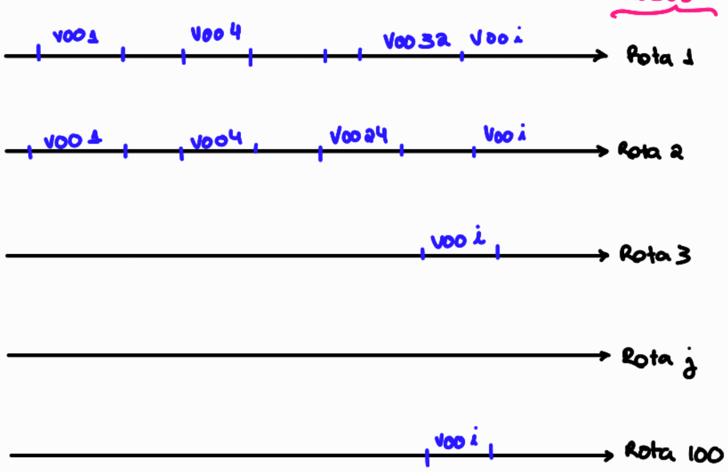
B) Entre os projetos 1, 2, 3 e 4, podemos inv. em apenas 1 deles

pode ser modelada por $x_1 + x_2 + x_3 + x_4 \leq 1$

1.12) Scheduling Problems - Problemas de Roteamento

* tenho que estipular 1º o conj. das rotas

dados



* mts n serão escolhidas

* n pode escolher 2 rotas c/ o mesmo voo

↳ pq n posso atribuir 2 tripulações p/ o mesmo voo

Queremos escolher as rotas, i.e., atribuir tripulações às rotas.

Dados: c_j = custo de atribuir uma tripulação à rota j

$a_{ij} = \begin{cases} 1, & \text{se o voo } i \text{ está na rota } j \\ 0, & \text{c.c} \end{cases}$

Ex: $a_{i1} = a_{i2} = a_{i3} = a_{i100} = 1$

$a_{11} = a_{12} = 1$

1) Variáveis de Decisão: $x_j \begin{cases} 1, & \text{se a rota } j \text{ é escolhida} \\ 0, & \text{c.c} \end{cases}$

3) Restrições:

* p/ todo voo eu vou atribuir apenas 1 tripulação

$$a_{i1} x_1 + a_{i2} x_2 + \cdots + a_{in} x_n = \sum_{j=1}^{n} a_{ij} x_j = 1, \ \forall i$$

$a_{i1} x_1 + a_{i2} x_2 + a_{i3} x_3 + a_{i100} x_{100} = 1$

2) Função Objetivo: $\max \sum_{j=1}^{n} c_j x_j$

a_{ij} = net cash flow from investment j in period i

se a_{ij} precisa de mais \$ investido no período i, $a_{ij} > 0$

se a_{ij} gera mais \$ no período i, $a_{ij} < 0$

b_i = cash flows exógenos incrementais

se fundos adicionais são adquiridos no período i, $b_i > 0$

se tiver que investir + \$, $b_i < 0$

reflect incremental cash balance in each period. The coefficients a_{ij} represent the net cash flow from investment j in period i. If the investment requires additional cash in period i, then $a_{ij} > 0$, while if the investment generates cash in period i , then $a_{ij} < 0$. The righthand-side coefficients $b_i$represent the incremental exogenous cash flows. If additional funds are made available in period i, then $b_i > 0$, while if funds are withdrawn in period i, then $b_i < 0$. These constraints state that the funds required for investment must be less than or equal to the funds generated from prior investments plus exogenous funds made available (or minus exogenous funds withdrawn).

The capital-budgeting model can be made much richer by including logical considerations. Suppose, for example, that investment in a new product line is contingent upon previous investment in a new plant. This contingency is modeled simply by the constraint

$$x_j \geq x_i,$$

which states that if $x_i = 1$ and project i (new product development) is accepted, then necessarily $x_j = 1$ and project j (construction of a new plant) must be accepted. Another example of this nature concerns conflicting projects. The constraint

$$x_1 + x_2 + x_3 + x_4 \leq 1,$$

for example, states that only one of the first four investments can be accepted. Constraints like this commonly are called multiple-choice constraints. By combining these logical constraints, the model can incorporate many complex interactions between projects, in addition to issues of resource allocation.

A particular case: the 0–1 knapsack problem. It is stated as

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

subject to

$$\sum_{j=1}^{n} a_j x_j \leq b,$$

$$x_j = 0 \quad \text{or} \quad 1 \qquad (j = 1, 2, \ldots, n)$$

This problem models the situation when a hiker must decide which goods to include on his trip. Here c_j is the "value" or utility of including good j, which weighs $a_j > 0$ pounds; the objective is to maximize the "pleasure of the trip," subject to the weight limitation that the hiker can carry no more than b pounds. The model can be altered by allowing more than one unit of any good to be taken, by writing $x_j \geq 0$ and x_j integer in place of the 0–1 restrictions on the variables.

Example 1.12 Scheduling problems: *Consider the scheduling of airline flight personnel. The airline has a number of routing "legs" to be flown, such as 10 A.M. New York to Chicago, or 6 P.M. Chicago to Los Angeles. The airline must schedule its personnel crews on routes to cover these flights. One crew, for example, might be* *scheduled to fly a route containing the two legs just mentioned. The decision variables, then, specify the scheduling of the crews to routes:*

$$x_j = \begin{cases} 1 & \text{if a crew is assgined to route } j, \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$a_{ij} = \begin{cases} 1 & \text{if leg } i \text{ is included on route } j, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$c_j = \text{Cost for assigning a crew to route } j$$

The coefficients a_{ij} define the acceptable combinations of legs and routes, taking into account such characteristics as sequencing of legs for making connections between flights and for including in the routes ground time for maintenance. The model becomes:

$$\begin{aligned} &\text{maximize } \textstyle\sum_{j=1}^{n} c_j x_j \\ &\text{subject to} \\ &\textstyle\sum_{j=1}^{n} a_{ij} x_j = 1 \quad (i = 1, 2, \ldots, m), \\ &x_j = 0 \ \text{ or } \ 1 \quad (j = 1, 2, \ldots, n) \end{aligned} \tag{1.4}$$

The ith constraint requires that one crew must be assigned on a route to fly leg i. An alternative formulation permits a crew to ride as passengers on a leg. Then the constraints (1.4) become:

$$\sum_{j=1}^{n} a_{ij} x_j \geq 1 \quad (i = 1, 2, \ldots, m)$$

If, for example,

$$\sum_{j=1}^{n} a_{1j} x_j = 3$$

then two crews fly as passengers on leg 1, possibly to make connections to other legs to which they have been assigned for duty.

Example 1.13 The traveling salesman problem. *Starting from his home, a salesman wishes to visit each of $(n-1)$ other cities and return home at minimal cost. He must visit each city exactly once and it costs c_{ij} to travel from city i to city j. What route should he select? If we let*

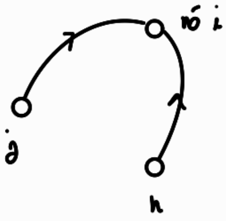
$$x_{ij} = \begin{cases} 1 & \text{if he goes from city } i \text{ to city } j, \\ 0 & \text{otherwise.} \end{cases}$$

1.13) Problema do Cacheiro Viajante

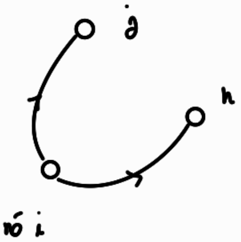
Seja $G = (X, E)$ um grafo. Dado um nó N e custos c_{ij} na aresta $(i,j) \in E$, quero encontrar a partir de N o caminho de custo mínimo que passe por todos os nós pelo menos 1 vez e volte p/ o nó original N.

1) Variáveis de Decisão: $x_{ij} = \begin{cases} 1, & \text{se passa pela aresta } (i,j) \\ 0, & \text{c.c} \end{cases}$ ("se o arco (i,j) pertence ao caminho")

3) Restrições:



$$x_{ji} = 1 \Rightarrow x_{ki} = 0 \quad \forall k \neq j, \; (k,i) \in E$$



$$x_{ij} = 1 \Rightarrow x_{ik} = 0 \quad \forall k \neq j, \; (i,k) \in E$$

2) Função Objetivo:

$$\min \sum_{(i,j) \in E} c_{ij} x_{ij}$$

We may formulate this problem as the assignment problem:

$$\text{Minimize } z = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij},$$

subject to:

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad (j = 1, \ldots, n)$$

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad (i = 1, \ldots, n)$$

$$x_{ij} \geq 0 \quad (i = 1, \ldots, n; \; j = 1, \ldots, n)$$

The constraints require that the salesman must enter and leave each city exactly once. Unfortunately, the assignment model can lead to infeasible solutions (see Figure 1). It is possible in a six-city problem, for example, for the assignment solution to route the salesman through two disjoint subtours of the cities instead of on a single trip or tour (see Figure 1). In this particular example, we can avoid the subtour solution of Figure 1 by including the constraint:

$$x_{14} + x_{15} + x_{16} + x_{24} + x_{25} + x_{26} + x_{34} + x_{35} + x_{36} \geq 1.$$

This inequality ensures that at least one leg of the tour connects cities 1, 2, and 3 with cities 4, 5, and 6. In general, if a constraint of this form is included for each way in which the cities can be divided into two groups, then subtours will be eliminated. The problem with this and related approaches is that, with n cities, $(2^n - 1)$ constraints of this nature must be added, so that the formulation becomes a very large integer-programming problem. For this reason the traveling salesman problem generally is regarded as difficult when there are many cities. The traveling salesman model is used as a central component of many vehicular routing and scheduling models. It also arises in production scheduling. For example, suppose that we wish to sequence $(n-1)$ jobs on a single machine, and that c_{ij} is the cost for setting up the machine for job j, given that job i has just been completed. What scheduling sequence for the jobs gives the lowest total setup costs? The problem can be interpreted as a traveling salesman problem, in which the "salesman" corresponds to the machine which must "visit" or perform each of the jobs. "Home" is the initial setup of the machine, and, in some applications, the machine will have to be returned to this initial setup after completing all of the jobs. That is, the "salesman" must return to "home" after visiting the "cities."

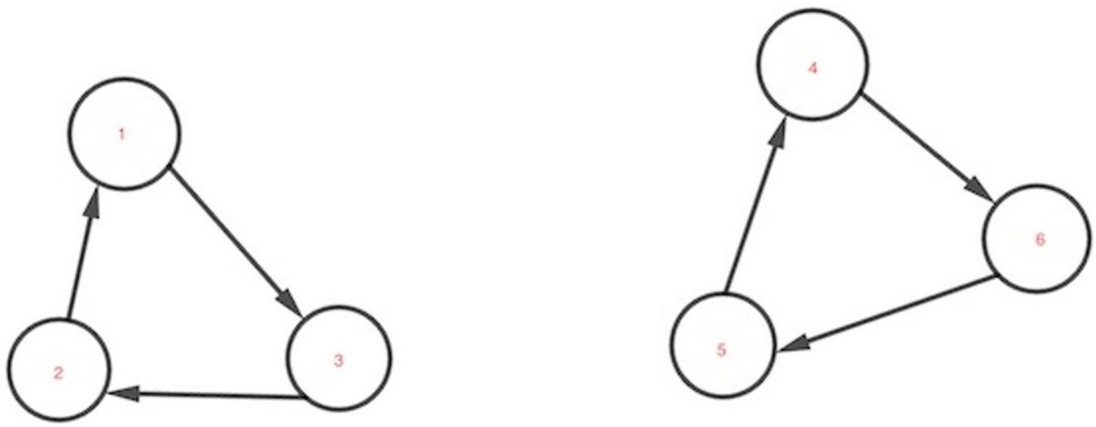


Figure 1: Disjoint subtours.

2 Affine spaces

* exercício de espaço afim p/ a prova

2.1 Definitions and affine hulls

Definition 2.1 (Affine space) *A subset E of $\mathbb{R}^n$ is an affine space if for every $x_1, x_2 \in E$ the line*

$$\{tx_1 + (1-t)x_2 : t \in \mathbb{R}\}$$

that passes through x_1 and x_2 is contained in E.

saber def. de espaço afim e Aff(X)

Definition 2.2 *An affine combination of vectors $x_1, \ldots, x_m \in \mathbb{R}^n$ is any point of form $\sum_{i=1}^m t_i x_i$ with $t_i \in \mathbb{R}$ and $\sum_{i=1}^m t_i = 1$.*

Remark 2.3 *An affine combination of $x_1, \ldots, x_m$ can be written as $x_1 + \sum_{i=1}^m t_i(x_i - x_1)$ where $t_i \in \mathbb{R}$.*

We obtain the following characterization of affine spaces:

Proposition 2.4 *A subset E of $\mathbb{R}^n$ is an affine space if and only if it contains all its affine combinations.*

All we have to show is that if E is an affine space then it contains all its affine combinations. We show the result by induction on the number m of points in the affine combination. If $m = 2$ the result follows by definition of an affine space. Now assume the result holds for m and take an affine combination $\sum_{i=1}^{m+1} t_i x_i$ of $m+1$ points $x_1, \ldots, x_{m+1}$ from E. Note that at least one the t_i is not equal to one. Without loss of generality assume that $t_1 \neq 1$. Note that due to the induction hypothesis, the point $x = \sum_{i=2}^{m+1} \theta_i x_i$ with $\theta_i = \frac{t_i}{1-t_1}$ belongs to E because $\sum_{i=2}^{m+1} \theta_i = 1$. Then

$$\sum_{i=1}^{m+1} t_i x_i = t_1 x_1 + (1-t_1)x \in E$$

since $x_1, x \in E$. We can see an affine subspace E as a translation of a vector space (a vector space being a special case of affine space) which is uniquely defined by E:

Proposition 2.5 *Let E be a nonempty affine space in $\mathbb{R}^n$. Then*

$$V = \{x - y : x, y \in E\}$$

is a vector space and for any $x_0 \in E$ we have

$$E = x_0 + V = \{x_0 + v : v \in V\}. \tag{2.5}$$

Let us show that V is a vector space. Take $v \in V$ and $\alpha \in \mathbb{R}$. We have $v = x - y$ for some $x, y \in E$. Then

$$\alpha v = \underbrace{x}_{\in E} - \underbrace{((1-\alpha)x + \alpha y)}_{\in E} \in V.$$

Take $v_1 = x_1 - y_1, v_2 = x_2 - y_2 \in V$ for some $x_1, x_2, y_1, y_2 \in E$. Then

$$v_1 + v_2 = \underbrace{x_1}_{\in E} - \underbrace{(y_1 + y_2 - x_2)}_{\in E} \in V.$$

Therefore V is a vector space.

Take $x \in E$. Then x can be written $x = x_0 + v$ with $v = x - x_0 \in V$. Therefore $E \subset x_0 + V$. Now if $x \in x_0 + V$ then $x = x_0 + x_1 - x_2$ for some $x_1, x_2 \in E$ and by Proposition 2.4 $x \in E$, which shows (2.5). The proposition above offers an equivalent characterization of affine spaces.

Proposition 2.6 *An intersection of affine spaces is an affine space. More precisely, for the family of affine spaces $E_\alpha = x_\alpha + V_\alpha, \alpha \in I$ where $V_\alpha = \{x - y : x, y \in E_\alpha\}$, if the intersection $E = \bigcap_{\alpha \in I} E_\alpha$ is nonempty, then*

$$E = x_0 + \bigcap_{\alpha \in I} V_\alpha \tag{2.6}$$

where x_0 is any point in E.

Let $x_1, x_2 \in E$. Then $x_1, x_2 \in E_\alpha$ for all $\alpha \in I$ and therefore the line passing through x_1, x_2 is contained in E_α for all $\alpha \in I$ and consequently is also contained in E which shows that E is an affine space.

Now let $x \in E$. We can write $x = x_0 + x - x_0$ with $x, x_0 \in E$. Therefore $x - x_0 \in V_\alpha$ for all α, $x - x_0 \in \cap_{\alpha \in I} V_\alpha$ and $x \in x_0 + \bigcap_{\alpha \in I} V_\alpha$. On the other hand, if $x \in x_0 + \bigcap_{\alpha \in I} V_\alpha$ then $x = x_0 + x_\alpha - y_\alpha$ for $x_\alpha, y_\alpha \in E_\alpha$. Since $x_0 \in E_\alpha$ for all α, we have that $x \in E_\alpha$ for all $\alpha \in I$ and therefore $x \in E$. Let X be a nonempty subset of $\mathbb{R}^n$. We denote by $\mathrm{Aff}(X)$ the smallest affine space of $\mathbb{R}^n$ that contains X which means that if S is an affine space of $\mathbb{R}^n$ that contains X then $\mathit{Aff}(X) \subset S$. We have the following characterization of $\mathit{Aff}(X)$:

Proposition 2.7 *We have* $\mathrm{Aff}(X) = \bigcap_{S \text{ affine space and } X \subset S} S.$

$$\ell(\lambda)$$

Let $S(X) = \bigcap_{S \textit{ affine space and } X \subset S} S$. By Proposition 2.6, $S(X)$ is an affine space which containts X. Therefore $\mathrm{Aff}(X) \subset S(X)$ by definition of $\mathit{Aff}(X)$. On the other hand, $\mathit{Aff}(X)$ is an affine space that contains X and therefore by definition of $S(X)$ we also have $S(X) \subset \mathit{Aff}(X)$. The set $\mathit{Aff}(X)$ is called the affine hull of X. To handle constraints of form $y \in \mathit{Aff}(X)$, we now obtain an algebraic representation of $\mathit{Aff}(X)$. Let x_0 be an arbitrary point in X. By definition of $\mathit{Aff}(X)$ this point x_0 also belongs to $\mathit{Aff}(X)$ and using Proposition 2.5, the set $\mathit{Aff}(X)$ can be written as

$$\mathit{Aff}(X) = x_0 + V_X$$

where V_X is a vector space. We now need an algebraic representation of this set V_X. The relation $X \subset \mathit{Aff}(X)$ can be re-written $X \subset x_0 + V_X$ which means that the vector space V_X must contain the set $X - x_0$. Moreover, if $V_1 \subset V_2$ then $x_0 + V_1 \subset x_0 + V_2$. Therefore, V_X is the smallest vector space which contains the set of vectors $X - x_0$ and we know that this smallest vector space is the linear span of $X - x_0$, i.e., the set of all linear combinations of vectors from $X - x_0$: $V_X = \mathit{Span}(X - x_0)$. We come to the representation

$$\mathit{Aff}(X) = x_0 + \mathit{Span}(X - x_0)$$

where x_0 is any point in X. Knowing that

$$\mathit{Span}(X - x_0) = \left\{ \sum_{i=1}^{k} \mu_i (x_i - x_0) : k \geq 1, \mu_i \in \mathbb{R}, x_i \in X \text{ for } i \geq 1 \right\},$$

we see that

$$\mathit{Aff}(X) = \left\{ x_0 + \sum_{i=1}^{k} \mu_i (x_i - x_0) : k \geq 1, \mu_i \in \mathbb{R}, x_i \in X \text{ for } i \geq 1 \right\}, \tag{2.7}$$

where x_0 is any point in X. From this characterization we obtain the following algebraic representation of $\mathit{Aff}(X)$:

Proposition 2.8 $\mathrm{Aff}(X)$ *is the set of all affine combinations of points from* X*:*

$$\mathrm{Aff}(X) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \geq 1, x_i \in X, \lambda_i \in \mathbb{R}, i \geq 1, \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

Using (2.7), if x is in $\mathit{Aff}(X)$, we can write

$$x = (1 - \sum_{i=1}^{k} \mu_i) x_0 + \sum_{i=1}^{k} \mu_i x_i = \sum_{i=0}^{k} \lambda_i x_i$$

for point $x_i \in X, i \geq 0$, reals μ_i, $\lambda_0 = 1 - \sum_{i=1}^{k} \mu_i$, and $\lambda_i = \mu_i, i \geq 1$. Since $\sum_{i=0}^{k} \lambda_i = 1$, x is indeed an affine combination of points from X. On the other hand, if x is an affine combination of points from X, we can write $x = \sum_{i=0}^{k} \lambda_i x_i$ with $x_i \in X, i \geq 0$ and reals λ_i summing up to one. Therefore

$$x = x_0 + \sum_{i=1}^{k} \lambda_i (x_i - x_0)$$

and we know from (2.7) that such x belongs to $\mathit{Aff}(X)$. From the above characterization of $\mathit{Aff}(X)$, we see that the affine hull of an affine space is this affine space.

2.2 Affinely spanning sets, affinely independent sets, affine basis and affine dimension

Just as vector spaces have basis, we would like to define basis for affine spaces, called affine basis, such that every vector in the affine space can be uniquely written as an affine combination of vectors from this space. Recall that an affine space Y is of form

$$Y = y_0 + V$$

where y_0 is any point in Y and V is the vector space $V = \{y_2 - y_1 : y_1, y_2 \in Y\}$. Assume that the dimension of V is k. To obtain an affine basis of Y it suffices to take a basis $(v_1, \ldots, v_k)$ of V. The set of vectors $(y_0, y_0 + v_1, y_0 + v_2, \ldots, y_0 + v_k)$ is then an affine basis of Y. Indeed, denoting $y_i = y_0 + v_i, i \geq 1$, any vector $y \in Y$ can be written as $y = y_0 + v$ for some $v \in V$ and thefore $y = y_0 + \sum_{i=1}^{k} \alpha_i v_i$ for some reals α_i, or equivalently,

$$y = \sum_{i=0}^{k} \lambda_i y_i \text{ with } \lambda_0 = 1 - \sum_{i=1}^{k} \alpha_i, \lambda_i = \alpha_i, i \geq 1.$$

Since in the expression above $\sum_{i=0}^{k} \lambda_i = 1$, any point $y \in Y$ can indeed be written as an affine combination of $(y_0, y_1, \ldots, y_k)$. We can also check that the coefficients $\lambda_i, i \geq 0$, in the decomposition of $y \in Y$ as an affine combination of $(y_0, y_1, \ldots, y_k)$ is uniquely determined by y. Indeed, assume that there are two sets of coefficients $\lambda_i, i \geq 0$, and $\lambda_i', i \geq 0$ summing up to one, such that $y = \sum_{i=0}^{k} \lambda_i y_i = \sum_{i=0}^{k} \lambda_i' y_i$. Then

$$0 = (\lambda_0 - \lambda_0') y_0 + \sum_{i=1}^{k} (\lambda_i - \lambda_i')(y_0 + v_i) = \underbrace{(\sum_{i=1}^{k} \lambda_i - \sum_{i=1}^{k} \lambda_i')}_{=1-1=0} y_0 + \sum_{i=1}^{k} (\lambda_i - \lambda_i') v_i.$$

Since vectors v_i are linearly independent the above relation implies that $\lambda_i = \lambda'_i = 0, i \geq 1$. Moreover, $\lambda_0 + \sum_{i=1}^n \lambda_i = \lambda'_0 + \sum_{i=1}^n \lambda'_i = 1$ which gives $\lambda_0 = \lambda'_0$. This shows that any point in Y can be uniquely written as an affine combination of $(y_0, y_1, \ldots, y_k) = (y_0, y_0 + v_1, y_0 + v_2, \ldots, y_0 + v_k)$. Therefore we have checked that for any $y_0 \in Y$ and any basis $(v_1, \ldots, v_k)$ of V the set of vectors $(y_0, y_0 + v_1, y_0 + v_2, \ldots, y_0 + v_k)$ is an affine basis of Y.

Reciprocally, consider a basis $(y_0, y_1, \ldots, y_n)$ of vectors from Y. Let us show the following property:

(P) this basis must have $n + 1 = k + 1$ vectors and that $(y_1 - y_0, \ldots, y_n - y_0)$ is a basis of V.

Before this discussion, we need to characterize vectors $(y_0, y_1, \ldots, y_n)$ such that the coefficients λ_i in the affine combination

$$y = \sum_{i=0}^{n} \lambda_i y_i \tag{2.8}$$

are uniquely determined by y. Assume that this is not the case, that is to say that we have

$$y = \sum_{i=0}^{n} \lambda_i y_i = \sum_{i=0}^{n} \lambda'_i y_i$$

for two different sets of coefficients $\lambda_i, i \geq 0$ and $\lambda'_i, i \geq 0$ summing up to one. Then

$$\sum_{i=0}^{n} \theta_i y_i = 0 \text{ for } \theta_i = \lambda_i - \lambda'_i, \sum_{i=0}^{n} \theta_i = \sum_{i=0}^{n} \lambda_i - \sum_{i=0}^{n} \lambda'_i = 1 - 1 = 0,$$

where coefficients θ_i are not all null. Therefore, to ensure that coefficients λ_i in the affine combination (2.8) are uniquely determined by y, we need the following:

Definition 2.9 (Affinely independent vectors) *Vectors* $(y_0, \ldots, y_n)$ *are affinely independent if*

$$\sum_{i=0}^{n} \lambda_i y_i = 0, \sum_{i=0}^{n} \lambda_i = 0 \Rightarrow \lambda_i = 0, i = 0, \ldots, n.$$

From our discussion above, if vectors $(y_0, y_1, \ldots, y_n)$ are affinely independent then coefficients λ_i in the affine combination (2.8) are indeed uniquely determined by y. From the definition of an affine basis, vectors forming an affine basis must be affinely independent.

Let us now show (P). Let $v \in V$. The vector $y_0 + v$ belongs to Y and therefore there are $\lambda_i, i \geq 0$ summing up to one such that $y_0 + v = \sum_{i=0}^n \lambda_i y_i$ and therefore $v = \sum_{i=1}^n \lambda_i (y_i - y_0)$. It follows that the set of vectors $(v_1, \ldots, v_n)$ with $v_i = y_i - y_0$ spans V and therefore $n \geq k$. Equivalently we say that $(y_0, y_1 - y_0, \ldots, y_n - y_0)$ affinely spans Y:

Definition 2.10 (Affinely spanning set) *We say that the set of vectors* $(y_0, y_1, \ldots, y_n)$ *affinely spans* Y *if* $\mathrm{Aff}(\{y_0, y_1, \ldots, y_n\}) = Y$ *or equivalently if* $\mathrm{Span}(\{y_1 - y_0, \ldots, y_n - y_0\}) = V$.

On the other hand, if reals $\lambda_i, i \geq 1$ satisfy $\sum_{i=1}^{n} \lambda_i(y_i - y_0) = 0$ then we have $\sum_{i=0}^{n} \theta_i y_i = 0$ with $\theta_i = \lambda_i, i \geq 1, \theta_0 = -\sum_{i=1}^{n} \lambda_i$, $\sum_{i=0}^{n} \theta_i = 0$, and since $(y_0, y_1, \ldots, y_n)$ are affinely independent this implies $\theta_i = \lambda_i = 0, i \geq 1$. This shows that vectors $(v_1, \ldots, v_n)$ in V with $v_i = y_i - y_0$ are linearly independant. Therefore $n \leq k$. We have shown that $n = k$ and that in fact $(v_1, \ldots, v_n)$ is a basis of V.

Summarizing our observations, we have shown that all basis of $Y = y_0 + V$ with V of dimension k are of the form $(y_0, y_1, \ldots, y_k)$ where y_0 is any point in Y and $(y_1 - y_0, \ldots, y_k - y_0)$ is a basis of V. Equivalently, an affine basis for Y is a set of affinely independent vectors in Y that affinely spans Y.

The affine dimension of the affine space $Y = y_0 + V$ is the dimension of V. The number of elements in all affine basis of Y is the maximal number of affinely independent vectors in Y or equivalently the minimal number of vectors that affinely span Y.

For any $y \in Y$ and any affine basis $(y_0, y_1, \ldots, y_k)$ of Y there exists a unique set of coefficients $\lambda_i, i \geq 0$ summing to one such that

$$y = \sum_{i=0}^{k} \lambda_i y_i.$$

These coefficients are called the barycentric coordinates of y in the basis.

2.3 Outer description of affine and vector spaces

So far we have give a direct way of constructing affine spaces. Namely given a set of vectors Y, all affine combinations of vectors from Y define an affine space, which we denoted *Aff*(Y), the smallest affine space that contains Y. Similarly, the set of all linear combinations of vectors in X is a vector space. It is possible to obtain outer representations of vector and affine spaces expressing them as solution sets to linear systems of equations. More precisely, we have the following:

Proposition 2.11 *V is a vector space if and only if V is the solution of a system of linear equations of form*

$$a_i^T x = 0, i = 1, \ldots, m. \tag{2.9}$$

Moreover, the minimal number of equations in the system above is the dimension of $V^\perp$.

Let V be a vector space and let $(a_1, \ldots, a_m)$ be a set of vectors spanning $V^\perp$: $[a_1, \ldots, a_m] = V^\perp$. Then V is the set of solutions to the system of linear equations (2.9). Indeed, if x is in V then x is orthogonal to all vectors $a_i, i = 1, \ldots, m$, and therefore x is a solution to (2.9). On the other hand, if x is a solution to (2.9) then x is orthogonal to each $a_i, i = 1, \ldots, m$ and therefore belongs to $[a_1, \ldots, a_m]^\perp = V$. Reciprocally, the solution set to (2.9) is a vector space.

Proposition 2.12 *A set is an affine space if and only if it is the solution of a system of linear equations of form*

$$a_i^T x = b_i, i = 1, \ldots, m, \tag{2.10}$$

where $a_i^T a = b_i, i = 1, \ldots, m$. Moreover, the minimal number of equations in the system above is the dimension of $V^\perp$.

Let $Y = a + V$ be an affine space. Take $(a_1, \ldots, a_m)$ such that $[a_1, \ldots, a_m] = V^\perp$ and define b such that $a_i^T a = b_i, i = 1, \ldots, m$. Then Y is the solution set to (2.10). Indeed, if y is in Y then $y = a + v$ for some $v \in V$ with v orthogonal to each $a_i, i = 1, \ldots, m$. Therefore, $y^T a_i = a^T a_i = b_i, i = 1, \ldots, m$, and y is a solution to (2.10). On the other hand, if y is a solution to (2.10) then $y - a \in [a_1, \ldots, a_m]^\perp = V$ and $y \in Y$. Reciprocally, we know that the solution set to (2.10) is an affine space. Propositions 2.11 and 2.12 offer simple ways of checking if a given vector belongs or does not belong to a vector space or to an affine space: it suffices to check if this vector satisfies the corresponding linear equations given by these propositions.

2.4 Examples of affine spaces

Affine space of dimension 0 in $\mathbb{R}^n$ are singletons in $\mathbb{R}^n$, translations by a given vector a of the vector space $\{0\}$. They are solution to a system of n linear equations with n unknown of form (2.10) with invertible matrix $[a_1; a_2; \ldots; a_n]$.

Affine spaces of dimension 1 in $\mathbb{R}^n$ are lines:

$$\{a + td : t \in \mathbb{R}\}.$$

An affine basis is given by two different points y_1, y_2 on the line. It is the set of solutions of a system of linear equations of form (2.10) with n vaiables and $n - 1$ equations.

Affine spaces of dimension $n - 1$ in $\mathbb{R}^n$ are called hyperplanes. They are the solution set of a single linear equation $a^T x = b$ where $a \neq 0$.

The largest possible affine space in $\mathbb{R}^n$, of dimension n, is $\mathbb{R}^n$ itself.

3 Convexity

3.1 Basic definitions and properties.

We start with the definition of a convex set in $\mathbb{R}^n$.

Definition 3.1 *A nonempty set $C \subset \mathbb{R}^n$ is convex if, for any $x_1, x_2 \in C$ and $\alpha \in [0, 1]$, it holds $\alpha x_1 + (1 - \alpha) x_2 \in C$.*

Example 3.2 *The following sets in $\mathbb{R}^n$ are convex:*

- *An affine space (the set of solutions of a linear system of equations $C := \{x \in \mathbb{R}^n \mid Ax = b\}$ where A is some $m \times n$ matrix and $b \in \mathbb{R}^m$);*

- *the solution set to an arbitrary number of linear inequalities:* $a_\alpha^T x \le b_\alpha, \alpha \in I$. *In particular the solution set the finite linear inequalities* $Ax \le a$ *for some* $m \times n$ *matrix* A *is convex and called a polyhedron;*

- *an ellipsoid:*
$$\{x : (x - x_0)^T Q(x - x_0) \le r^2\}$$
where Q *is definite positive;*

- *the* ε*-fattening* X^ε *of a convex set* X *given by*
$$X^\varepsilon = \{y : \inf_{x \in X} \|y - x\| \le \varepsilon\};$$

- *the balls* $\mathbb{B}(x_0, r) = \{x : \|x - x_0\| \le r\}$*;*

- *the* m*-dimensional simplex*
$$S = \Big\{ \sum_{i=0}^{m} \lambda_i x_i : \lambda_i \ge 0, i \ge 1, \sum_{i=0}^{m} \lambda_i = 1 \Big\} \tag{3.11}$$
for vectors $x_0, x_1, \ldots, x_m$*, affinely independent. This is the convex hull of points* $x_1, \ldots, x_m$*. As a special case, the unit simplex is the* n*-dimensional simplex obtained taking convex hull of 0, and vectors* $e_1, e_2, \ldots, e_n$*, of the canonical basis. It can be expressed as* $\{x : x \ge 0, \sum_{i=1}^n x_i \le 1\}$*. The probability simplex is the* $(n-1)$*-dimensional simplex given by unit vectors* $e_1, \ldots, e_n$*. It is the set of vectors satisfying* $x \ge 0$ *and* $\sum_{i=1}^n x_i = 1$ *which correspond to the set of possible discrete probability distributions with* n *elements in the support.*

  The m*-dimensional simplex* S *given by* (3.11) *can be seen as a polyhedron given by the solution set to linear equalities and inequalities. Indeed, if vectors* $x_0, x_1, \ldots, x_m$ *are affinely independent then*
$$A = \begin{bmatrix} x_0 & x_1 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$
has full column right and therefore has a left inverse B*. If* e *is a vector in* $\mathbb{R}^{m+1}$ *with all components equal to one, we then have*
$$\begin{aligned} y \in C \quad &\Leftrightarrow \quad \exists \theta \ge 0 : A\theta = \begin{pmatrix} y \\ 1 \end{pmatrix}, e^T\theta = 1, \\ &\Leftrightarrow \quad \exists \theta \ge 0 : \theta = B \begin{pmatrix} y \\ 1 \end{pmatrix}, e^T\theta = 1, \\ &\Leftrightarrow \quad B \begin{pmatrix} y \\ 1 \end{pmatrix} \ge 0, e^T B \begin{pmatrix} y \\ 1 \end{pmatrix} = 1. \end{aligned}$$

We have shown in the previous section that a set is affine if and only if it is an affine combination of its members. Similarly we can show that a set is convex if and only if it is a convex combination of its members, knowing that convex combination is an affine combination with nonnegative weights:

Definition 3.3 *A convex combination of vectors $x_1, \dots, x_m$ is a vector of form*

$$x = \sum_{i=1}^{m} \lambda_i x_i$$

where

$$\lambda_i \geq 0, i \geq 1, \sum_{i=1}^{m} \lambda_i = 1.$$

We have the following analogue of Proposition 2.4 (we skip the proof which is similar to the proof of Proposition 2.4):

Proposition 3.4 *X is convex if and only if every convex combination of points from X belongs to X.*

We have seen that a nonempty intersection of affine spaces is an affine space. A nonempty intersection of convex sets is also convex:

Proposition 3.5 *A nonempty set which is the intersection of a (possibly infinite) set of convex sets is convex.*

For any nonempty convex set X, the convex hull of X is the smallest convex set which contains X. It is denoted by $\mathrm{Conv}(X)$ and due to the previous proposition it is the intersection of all convex sets that contain X. We can derive an algebraic representation of $\mathrm{Conv}(X)$:

Proposition 3.6 *Let X be a nonempty set. Then $\mathrm{Conv}(X)$ is the set of all convex combinations of points from X:*

$$\mathrm{Conv}(X) = \Big\{ \sum_{i=1}^{k} \lambda_i x_i : k \geq 1, \lambda_i \geq 0, x_i \in X, i \geq 1, \sum_{i=1}^{k} \lambda_i = 1 \Big\}.$$

Let

$$S(X) = \Big\{ \sum_{i=1}^{k} \lambda_i x_i : k \geq 1, \lambda_i \geq 0, x_i \in X, i \geq 1, \sum_{i=1}^{k} \lambda_i = 1 \Big\}.$$

It is straightforward to check that $S(X)$ is a convex set that contains X. Therefore $Conv(X) \subset S(X)$. Reciprocally, if $x \in S(X)$, then $x = \sum_{i=1}^{k} \lambda_i x_i$ for $x_i \in X$ and nonnegative coefficients λ_i summing to one. Since $x_i \in X \subset Conv(X)$ and since $Conv(X)$ is convex, x which is a convex combination of points from convex set $Conv(X)$ belongs to $Conv(X)$. Therefore $S(X) \subset Conv(X)$, which achieves the proof.

3.2 Operations preserving convexity

The following proposition is straightforward:

Proposition 3.7 *The following operations preserve convexity:*

- *Intersection: if $C_\alpha, \alpha \in I$ are convex sets then $\bigcap_{\alpha \in I} C_\alpha$ is convex.*
- *Image of a convex set under an affine function: if $C \subset \mathbb{R}^n$ is convex and $f : \mathbb{R}^n \to \mathbb{R}^m$ is affine then $f(C)$ is convex. Examples: if C is convex then $C + x_0$ (translation of C) and αC (scaling) are convex.*
- *Inverse image of a convex set under an affine function: if $f : \mathbb{R}^n \to \mathbb{R}^m$ is affine and C is convex then*
$$f^{-1}(C) = \Big\{x : f(x) \in C\Big\}$$
is convex.
- *Projection: the projection of a convex set onto some of its coordinates is convex: if $C \subset \mathbb{R}^{m+n}$ is convex then*
$$\Pi = \Big\{x_1 \in \mathbb{R}^m : \exists x_2 \in \mathbb{R}^n, (x_1, x_2) \in C\Big\}$$
is convex.
- *Sum of two sets: if C_1 and C_2 are convex then $C_1 + C_2 = \{c_1 + c_2 : c_1 \in C_1, c_2 \in C_2\}$ is convex.*
- *Cartesian product of two sets:if C_1 and C_2 are convex then $C_1 \times C_2$ is convex.*

Application: The ellipsoid
$$C = \Big\{x : (x - x_0)^T Q^{-1}(x - x_0) \leq 1\Big\}$$
is convex. It is the image of the unit ball $\{u : \|u\|_2 \leq 1\}$ under the affine mapping $f(u) = Q^{1/2}u + x_0$. It is also the inverse image of the unit ball under the affine mapping $f(x) = Q^{-1/2}(x - x_0)$.

We define the perspective function $P : \mathbb{R}^{n+1} \to \mathbb{R}^n$ with domain $\mathbb{R}^n \times \{t \in \mathbb{R} : t > 0\}$ by $P(x, t) = \frac{x}{t}$.

Proposition 3.8 *If $C \subset dom(P)$ is convex then $P(C)$ is convex.*

Let $x, y \in \mathbb{R}^{n+1}$ with $x_{n+1}, y_{n+1} > 0$. Then for $0 \leq \theta \leq 1$ we have
$$P(\theta x + (1 - \theta)y) = tP(x) + (1 - t)P(y)$$
where
$$t = \frac{\theta x_{n+1}}{\theta x_{n+1} + (1 - \theta)y_{n+1}} \in [0, 1].$$
It follows that $P([x, y]) = [P(x), P(y)]$ and therefore if C is convex then $P(C)$ is convex too.

Proposition 3.9 *If $C \subset \mathbb{R}^n$ is convex then $P^{-1}(C)$ is convex.*

Let $(x_1, t_1), (x_2, t_2) \in P^{-1}(C)$ and $0 \leq \theta \leq 1$. We want to show that

$$\frac{\theta x_1 + (1-\theta)x_2}{\theta t_1 + (1-\theta)t_2} \in C.$$

This follows from the fact that

$$\frac{\theta x_1 + (1-\theta)x_2}{\theta t_1 + (1-\theta)t_2} = \mu \frac{x_1}{t_1} + (1-\mu)\frac{x_2}{t_2}$$

where

$$\mu = \frac{\theta t_1}{\theta t_1 + (1-\theta)t_2}$$

and $\frac{x_1}{t_1} \in C, \frac{x_2}{t_2} \in C$.

Proposition 3.10 *Let $g : \mathbb{R}^n \to \mathbb{R}^{m+1}$ given by*

$$g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}.$$

Let $f : \mathbb{R}^n \to \mathbb{R}^{m+1}$ be the linear fractional function given by

$$f(x) = \frac{Ax+b}{c^T x + d}$$

where $dom(f) = \{x : c^T x + d > 0\}$. From the previous propsition, we obtain that if $C \subset dom(f)$ is convex then $f(C)$ is convex and if $C \subset \mathbb{R}^m$ is convex then $f^{-1}(C)$ is convex.

We also have the following:

Proposition 3.11 *Let C be a convex set and X a random variable which belongs to C with probability one. Then $\mathbb{E}[X] \in C$.*

Definition 3.12 *The closed convex hull of a nonempty set $A \subset \mathbb{R}^n$ is defined as $\overline{Conv}(A) := \bigcap_{C \in \mathcal{B}(A)} C$ where $\mathcal{B}(A) = \{C \,|\, C \text{ is closed convex and } A \subset C\}$.*

Remark 3.13 *For future use, notice, with the notations above, that simple calculations yields $\mathcal{B}(\{y\} + A) = \{y\} + \mathcal{B}(A)$. Also, for any family of sets $\mathcal{B}$, it holds $\bigcap_{C \in \mathcal{B}} (\{y\} + C) = \{y\} + \bigcap_{C \in \mathcal{B}} C$. Combining the two above relations, we conclude that that*

$$\overline{Conv}(\{y\} + A) = \{y\} + \overline{Conv}(A).$$

Example 3.14 *Let $X = \{x_1, \ldots, x_m\} \subset \mathbb{R}^n$. We claim that*

$$\overline{Conv}(X) = \Big\{ \sum_{i=1}^{k} \lambda_i x_i : \lambda_i \geq 0, \ \sum_{i=1}^{k} \lambda_i = 1 \Big\}$$

which is a compact set. Notice that, in this case, the set $Conv(X)$ is the image of the compact set $\Delta = \Big\{\lambda \ : \ \lambda_i \geq 0, \ \sum_{i=1}^{k} \lambda_i = 1\Big\}$ by the continous function $\phi : \mathbb{R}^n \to \mathbb{R}$, $\phi(\lambda) = \sum_{i=1}^{m} \lambda_i x_i$ and therefore it is also compact. Then, since $Conv(X)$ is compact, it concides with $\overline{Conv}(X)$, and the claim follows.

3.3 Projection onto a convex set

Proposition 3.15 *Given a nonempty closed and convex set $C \subset \mathbb{R}^n$ and $\hat{x} \notin C$ there exists a unique point $a \in C$ such that $\|\hat{x} - a\| = \min\{\|\hat{x} - y\| \mid y \in C\}$.*

Take $x_0 \in C$ and define $\delta = \|\hat{x} - x_0\|$. Since $D = B(\hat{x}, \delta) \cap C$ is a nonempty compact set and the function $f(x) = \|\hat{x} - x\|$ is continuous in $\mathbb{R}^n$, it follows that f achieves its minimum value in D, at some point $a \in D$. Then, by definition, it holds

$$\|\hat{x} - a\| = \min\{\|\hat{x} - x\| \mid x \in D\} \leq \delta \leq \|\hat{x} - y\| \quad \forall y \in C - D = C \cap D^c,$$

which clearly shows that $\|\hat{x} - a\| = \min\{\|\hat{x} - y\| \mid y \in C\}$. Let us denote $d(\hat{x}, C) = \min\{\|\hat{x} - y\| \mid y \in C\}$ and assume that there exists another point $\hat{a} \in C$ such that $\|\hat{x} - \hat{a}\| = d(\hat{x}, C)$. Convexity of C implies that $(a + \hat{a})/2 \in C$, which combined with the Parallelogram identity

$$\|x + y\|^2 + \|x - y\|^2 = 2\Big[\|x\|^2 + \|y\|^2\Big]$$

with

$$x = \frac{\hat{x} - a}{2}, \quad y = \frac{\hat{x} - \hat{a}}{2},$$

yields

$$d(\hat{x}, C)^2 \leq \left\|\hat{x} - \frac{(a + \hat{a})}{2}\right\|^2 = \frac{\|\hat{x} - a\|^2}{2} + \frac{\|\hat{x} - a\|^2}{2} - \frac{\|a - \hat{a}\|^2}{4} = d(\hat{x}, C)^2 - \frac{\|a - \hat{a}\|^2}{4}$$

which implies that $a = \hat{a}$ and ends the proof.

Under the assumptions of the previous proposition, the point a will be called the projection of x onto C and it will be denoted by $P_C(x)$. Also, we will denote $d(x, C) = \|x - PC(x)\|$.

Proposition 3.16 *Consider a nonempty closed and convex set $C \subset \mathbb{R}^n$ and $\hat{x} \notin C$. Point $x \in C$ is the projection onto C of $\hat{x}$ if and only if*

$$\langle \hat{x} - x, y - x \rangle \leq 0 \quad \forall y \in C. \tag{3.12}$$

Let $a = P_C(\hat{x})$. By the convexity of C and the definition of the projection, it follows that, for any $y \in C$ and $t \in [0, 1]$

$$\|\hat{x} - (a + t(y - a))\|^2 \geq \|\hat{x} - a\|^2,$$

which yields

$$t\|y - a\|^2 \geq 2\langle \hat{x} - a, y - a \rangle.$$

Now, taking $t \to 0$, if follows that (3.12) holds.

Now assume now that, for some $x \in C$, it holds

$$\langle \hat{x} - x, y - x \rangle \leq 0 \quad \forall y \in C.$$

Relation

$$\|\hat{x} - y\|^2 = \|\hat{x} - x\|^2 + \|y - x\|^2 - 2\langle \hat{x} - x, y - x \rangle \geq \|\hat{x} - x\|^2 \quad \forall y \in C$$

shows that $x = P_C(\hat{x})$, which ends the proof.

4 Cones

Definition 4.1 *A nonempty set C is conic if for every $x \in C$ and every $t \geq 0$ we have $tx \in C$.*

Definition 4.2 *A cone is a convex conic set.*

We deduce:

Proposition 4.3 *A nonempty set C is a cone if and only if it satisfies the following two properties:*

(i) for every $x \in C$ and every $t \geq 0$ we have $tx \in C$;

(ii) for every $x, y \in C$ we have $x + y \in C$.

If C satisfies (i), (ii), for every $x, y \in C$ for every $0 \leq t \leq 1$ we have $tx \in C, (1-t)y \in C$, and therefore the sum $tx + (1-t)y \in C$. Reciprocally, if C is a cone then C is conic and for every $x, y \in C$ we have

$$x + y = 2\, \underbrace{(\frac{1}{2}x + \frac{1}{2}y)}_{\in C \text{ by convexity}} \quad \in C.$$

Definition 4.4 *A conic combination of points $x_1, \ldots, x_m$ is a vector of form $\sum_{i=1}^m \lambda_i x_i$ where $\lambda_i \geq 0, i \geq 1$.*

We deduce from Proposition 4.3 that C is a cone if and only if it contains all conic combinations of points from C.

Example 4.5 *The set of solutions to the (possibly infinite) set of inequalities $a_\alpha^T x \leq 0, \alpha \in I$ is a cone. In particular, the solution set to a homogeneous finite system of m homogeneous linear inequalities $Ax \leq b$ where A is an $m \times n$ matrix is a cone, called polyhedral cone.*

Similarly to affine spaces and convex sets, a nonempty intersection of cones is a cone. For any nonempty set X we can also define the smallest cone that contains X, the conic hull of X denoted by Cone(X). The conic hull of X is the intersection of all cones that contain X and we obtain the following representation of the conic hull (the proof is similar to the proof of Proposition 3.6):

Proposition 4.6 *Let X be a nonempty set. Then*

$$\mathrm{Conic}(X) = \Big\{ \sum_{i=1}^k \lambda_i x_i : k \geq 1, x_i \in X, \lambda_i \geq 0, i \geq 1 \Big\}.$$

Example 4.7 *The following sets are cones:*

- *The solution set C to an homogeneous system of linear inequalities $C = \{x : a_i^T x \leq 0, i = 1, \ldots, m\}$. This cone is called polyhedral.*

- *The norm cone: $C = \{(x,t) : \|x\| \leq t\}$. For $\|\cdot\| = \|\cdot\|_2$, we obtain the second-order cone or Lorentz cone or ice-cream cone $C = \{(x,t) \in \mathbb{R}^{n+1} : \|x\|_2 \leq t\}$.*

- *The set of semidefinite positive matrices.*

Also, for any nonempty set X, the closed conical hull of of a set X is the smallest closed cone containing X. It is denoted by $\overline{Conic}(X)$.

Proposition 4.8 *Let $X = \{x_1, \ldots, x_m\}$. Then,*

$$\overline{\text{Conic}}(X) = \Big\{ \sum_{i=1}^{m} \lambda_i x_i \ : \ \lambda_i \geq 0, \ i = 1, \ldots, m \Big\}.$$

Let $K = Conic(X) = \Big\{ \sum_{i=1}^{m} \lambda_i x_i \ : \ \lambda_i \geq 0, \ i = 1, \ldots, m \Big\}$. We will show that K is closed. First, notice that any linear combination of elements of X, with nonnegative linear coefficients, can be rewritten as a linear combination of all the elements of X, also with nonnegative linear coefficients. Next, take a sequence $\{x^k\} \subset K$ such that $x^k \to \hat{x}$ when $k \to +\infty$. We want to show that $\hat{x} \in K$. If $\hat{x} = 0$ then $\hat{x} \in K$. Now assume $\hat{x} \neq 0$. Without loss of generality, we can assume that $x^k \neq 0$ for all $k \geq 1$. Now, for each $k \geq 1$, we have that $x^k = \sum_{i=1}^m \lambda_i^k x_i$ where all λ_i^k are nonnegative and at least one of them is different from zero. Define $L_k = \sum_{i=1}^m \lambda_i^k$ and $\bar{x}^k = (1/L_k)\, x^k = (1/L_k) \sum_{i=1}^m \lambda_i^k x_i$ for all $k \geq 1$. Since $\{\bar{x}^k\} \subset \text{Conv}(X)$ which is compact, there exists a subsequence $\{\bar{x}^k : k \in J \subset \mathbb{N}\} \subset \{x^k\}$ converging to some $\bar{x} \in \text{Conv}(X)$. We have two cases to consider:

i) $0 \notin \text{Conv}(X)$. In this case, we have that $\bar{x} \neq 0$ and therefore $L_k = \left\|x^k\right\| / \left\|\bar{x}^k\right\| \to \bar{L} = \|\hat{x}\| / \|\bar{x}\| \in \mathbb{R}$ when $k \to +\infty, k \in J$, and it follows that $\hat{x} = \bar{L}\bar{x} \in K$.

ii) $0 \in \text{Conv}(X)$. Let us write $0 = \sum_{i=1}^m \bar{\lambda}_i x_i$ with $\bar{\lambda}_i \geq 0$ $i = 1, \ldots, m$ and $\sum_{i=1}^m \bar{\lambda}_i = 1$, and let us define $I = \left\{ i \in \{1, \ldots, m\} : \bar{\lambda}_i > 0 \right\} \neq \emptyset$. Take $x = \sum_{i=1}^m \lambda_i x_i \in K$ with $\lambda_i \geq 0$ $i = 1, \ldots, m$ and $x \neq 0$, and define $\delta = \min\left\{ \lambda_i / \bar{\lambda}_i : i \in I \right\} \geq 0$ and $I_1 = \left\{ i \in I : \delta = \lambda_i / \bar{\lambda}_i \right\} \neq \emptyset$. Simple calculations show that

$$x = \sum_{i=1}^{m} (\lambda_i - \delta \bar{\lambda}_i) x_i \quad \lambda_i - \delta \bar{\lambda}_i \geq 0 \ \ i = 1, \ldots, m, \quad \lambda_\ell - \delta \bar{\lambda}_\ell = 0 \ \forall\, \ell \in I_1$$

which means that $x \in \text{Conic}(X')$ where $X' = \{x_i : i \in J = \{1 \ldots, m\} - I_1\}$. This argument shows that it holds

$$K = \bigcup_{i \in \{1, \ldots, m\}} \text{Conic}(X_i) \text{ where } X_i = X - \{x_i\}$$

Notice, by the previous argument in case i), that, if $0 \notin \mathrm{Conv}(X_i)$ for all $i \in I$, then $\mathrm{Conic}(X_i)$ is a closed set for all $i \in I$ and, therefore, the above realtion shows that K is closed. If $0 \in \mathrm{Conv}(X_i)$ for some $i \in I$, then for this index i we could apply the process described above and, in this way, decompose the set $\mathrm{Conic}(X_i)$, where $X_i = X - \{x_i\}$, as the union of conic sets of subsets of X_i, each of these subsets with one less elements than X_i. Since X is a finite set, this process, after a finite number of steps, will lead to a decomposition of K as the union of a finite family of closed sets, and the claim will follow.

Proposition 4.9 *If C is a closed cone and $\hat{x} \notin C$, then it holds*

$$\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle = 0, \quad \langle \hat{x} - P_C(\hat{x}), \hat{x} \rangle > 0 \quad \textit{and} \quad \langle \hat{x} - P_C(\hat{x}), y \rangle \leq 0 \quad \forall\, y \in C,$$

Using Proposition 3.16 and the fact that C is a cone, we have that

$$\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle = \langle \hat{x} - P_C(\hat{x}), 2P_C(\hat{x}) - P_C(\hat{x}) \rangle \leq 0$$

and, also,

$$\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle = -2 \left\langle \hat{x} - P_C(\hat{x}), \frac{1}{2} P_C(\hat{x}) - P_C(\hat{x}) \right\rangle \geq 0$$

which yield the first claim of the lemma. To prove the second claim of the lemma, note that it holds

$$\begin{aligned} \langle \hat{x} - P_C(\hat{x}), \hat{x} \rangle &= \langle \hat{x} - P_C(\hat{x}), \hat{x} - P_C(\hat{x}) + P_C(\hat{x}) \rangle = \|\hat{x} - P_C(\hat{x})\|^2 + \langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle \\ &= \|\hat{x} - P_C(\hat{x})\|^2 + 2 \left\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) - \frac{1}{2} P_C(\hat{x}) \right\rangle \geq \|\hat{x} - P_C(\hat{x})\|^2 > 0 \end{aligned}$$

where the before-last inequality follows from Proposition 3.16 and the fact that $\frac{1}{2} P_C(\hat{x}) \in C$. Using again Proposition 3.16 and the fact that C is a cone, we have for any $t > 0$ and $y \in C$,

$$\langle \hat{x} - P_C(\hat{x}), y \rangle = \frac{\langle \hat{x} - P_C(\hat{x}), ty - P_C(\hat{x}) \rangle}{t} + \frac{\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle}{t} \leq \frac{\langle \hat{x} - P_C(\hat{x}), P_C(\hat{x}) \rangle}{t}.$$

Taking $t \to +\infty$ in the relation above, yields the third part of the claim.

5 Separation theorems

Definition 5.1 *We say that a point $x \in C$ is in the relative interior of the set $C \subset \mathbb{R}^n$ if there exists $r > 0$ such that $B(x, r) \cap \mathrm{Aff}(C) \subset C$. We denote the set of such points by $ri(C)$ and call this set the relative interior of C.*

Definition 5.2 *An hyperplane in $\mathbb{R}^n$ is an affine space $H \subset \mathbb{R}^n$ of dimension equal to $n-1$.*

Proposition 5.3 *A set $H \subset \mathbb{R}^n$ is an hyperplane if and only if there exists a nonzero vector $a \in \mathbb{R}^n$ and a real constant c such that $H = \{x \in \mathbb{R}^n \mid \langle x, a\rangle = c\}$.*

Any subspace V of $\mathbb{R}^n$ of dimension $n-1$ is of the form $V = \{x \in \mathbb{R}^n \mid \langle a, x\rangle = 0\}$ for some nonzero vector $a \in \mathbb{R}^n$. Hence, it follows that $H = V + \{x_0\}$ where V is a linear subspace of dimension $n-1$ and $H = \{x \in \mathbb{R}^n \mid \langle a, x\rangle = 0\} + \{x_0\}$ for some vector $a \neq 0$. Now, taking $c = \langle a, x_0\rangle$ the result follows.

Definition 5.4 *Given an hyperplane $H = \{x \in \mathbb{R}^n \mid \langle a, x\rangle = c\}$ we define the associated positive and negative closed semi-spaces $H_+ := \{x \in \mathbb{R}^n \mid \langle a, x\rangle \geq c\}$ and $H_- := \{x \in \mathbb{R}^n \mid \langle a, x\rangle \leq c\}$, respectively.*

Definition 5.5 *A convex polytope $H \subset \mathbb{R}^n$ is the set obtained by the intersection of a finite number of closed semi-spaces of $\mathbb{R}^n$, that is, $H = \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{c} \langle a_1, x\rangle \leq b_1 \\ \langle a_2, x\rangle \leq b_2 \\ \vdots \\ \langle a_m, x\rangle \leq b_m \end{array} \right\}$. A polyhedral set is a nonempty bounded and convex polytope.*

Theorem 5.6 *Let $C \subset \mathbb{R}^n$ be a convex set and $y \in \overline{C}^c$. Then, there exists a vector $a \in \mathbb{R}^n$ such that $\langle a, y\rangle > \sup_{x \in C} \langle a, x\rangle$.*

In view of Proposition 3.16, we have, for any $x \in C$,

$$0 \leq \langle y - P_{\overline{C}}(y), P_{\overline{C}}(y) - x\rangle = \langle y - P_{\overline{C}}(y), P_{\overline{C}}(y) - y + y - x\rangle = \langle y - P_{\overline{C}}(y), y - x\rangle - \|y - P_{\overline{C}}(y)\|^2 .$$

Hence, defining $a = y - P_{\overline{C}}(y)$, it follows that

$$\langle a, y\rangle \geq \langle a, x\rangle + d(y, \overline{C})^2 \quad \forall x \in C,$$

which, in view of the fact that $d(y, \overline{C}) > 0$, implies the claim.

Remark 5.7 *Under the assumptions of Theorem 5.6, any hyperplane*

$$H = \{x : \langle a, x\rangle = b\}$$

with $\langle a, y\rangle > b > \sup_{x \in C} \langle a, x\rangle$ strictly separates y from C, in the sense that y and C are contained in the respective interiors of the negative and positive semi-spaces H_- and H_+. For instance, we can take $H = \left\{ x : \langle a, x\rangle = \left\langle a, \frac{y + P_{\overline{C}}(y)}{2} \right\rangle \right\}$.

Theorem 5.8 *Let $C \subset \mathbb{R}^n$ be a convex set and y a boundary point of C, that is $y \in \overline{C} \backslash \mathrm{int}(C)$. Then, there exists a vector $a \in \mathbb{R}^n$ such that $\langle a, y\rangle \geq \sup_{x \in C} \langle a, x\rangle$.*

Take a sequence $\{y^k\} \subset \overline{C}^c$ such that $y^k \to y$ when $k \to +\infty$. From Theorem 5.6, it follows that, for each $k \in \mathbb{N}$, there exists a nonzero $a^k \in \mathbb{R}^n$ such that

$$\langle a^k, y^k \rangle > \langle a^k, x \rangle \quad \forall x \in C.$$

Let a_* be an accumulation point of the sequence $\{a^k / \|a^k\|\}$. It follows from the above relation, dividing each side of the inequality by $\|a^k\|$ and taking limits when $k \to +\infty$, that

$$\langle a_*, y \rangle \geq \langle a_*, x \rangle \quad \forall x \in C$$

which clearly implies the claim and ends the proof.

Definition 5.9 *An hyperplane H such that a set C is contained in one of its related semi-spaces and contains a point y of the boundary of C is called a supporting hyperplane of C.*

Theorem 5.10 *Let C and D be two closed convex sets such that $C \cap D = \emptyset$. Then, there exists an hyperplane that separates C and D, that is, there exists a nonzero vector $a \in \mathbb{R}^n$ such that $\inf_{c \in C} \langle a, c \rangle \geq \sup_{d \in D} \langle a, d \rangle$.*

Define set $E = D - C = \{d - c : d \in D, c \in C\}$ which, in view of Proposition 3.7, is convex. Notice that $0 \notin E$. We have two options, $0 \in \overline{E}^c$ or $0 \in \overline{E}$. In the first case by Theorem 5.6 and in the second case by Theorem 5.8 (notice that 0 is not a point of the interior of E since $0 \notin E$ hence in the second case we have $0 \in \overline{E} \backslash \mathrm{int}(E)$ and the assumptions of Theorem 5.8 apply), it follows that there exists $a \in \mathbb{R}^n$ such that

$$0 = \langle a, 0 \rangle \geq \sup_{u \in E} \langle a, u \rangle = \sup_{d \in D, c \in C} \langle a, d - c \rangle \geq \langle a, d \rangle - \langle a, c \rangle \quad \forall\, c \in C,\, d \in D,$$

which clearly implies the claim.

Theorem 5.11 *Let C and D be two closed convex sets such that C is bounded and $C \cap D = \emptyset$. Then, there exists an hyperplane that strictly separates C and D, that is, there exists a nonzero vector $a \in \mathbb{R}^n$ such that $\inf_{c \in C} \langle a, c \rangle > \sup_{d \in D} \langle a, d \rangle$.*

Define the convex set $E = D - C$. Assume that $0 \in \overline{E}$. Then, there exist sequences $\{x^k\} \subset C$ and $\{y^k\} \subset D$ such that $u^k = x^k - y^k \to 0$ when $k \to +\infty$. Since C is compact, there exists $\hat{x}$ an accumulation point of $\{x^k\}$, which, in view of the above relation, is also an accumulation point of $\{y^k\}$. Since C and D are closed, it follows that $\hat{x} \in C \cap D$ which is a contradiction. Hence $0 \notin \overline{E}$ and from Theorem 5.6 it follows that there exist nonzero vector $a \in \mathbb{R}^n$ and a real scalar b such that,

$$0 = \langle a, 0 \rangle > b > \sup_{u \in E} \langle a, u \rangle \geq \langle a, d \rangle - \langle a, c \rangle \quad \forall\, c \in C,\, d \in D.$$

The above relation implies that

$$0 > b \geq \sup_{d \in D} \langle a, d \rangle - \inf_{c \in C} \langle a, c \rangle$$

which yields the claim.

Theorem 5.12 *Let C and D be two convex sets such that $\mathrm{ri}(C) \cap \mathrm{ri}(D) = \emptyset$. Then, there exists an hyperplane that separates C and D, that is, there exists a nonzero vector $a \in \mathbb{R}^n$ such that $\sup_{d \in D} \langle a, d \rangle \leq \inf_{c \in C} \langle a, c \rangle$.*

Define the convex set $E = D - C$. We have two possibilities, $0 \notin \overline{E}$ or $0 \in \overline{E}$. In the second case, in view of the assumptions, we have that $0 \notin \mathrm{ri}(E) = \mathrm{ri}(D) - \mathrm{ri}(C)$ and therefore $0 \notin \mathrm{int}(E)$ (otherwise, if we had $0 \in \mathrm{int}(E)$, since $\mathrm{int}(E) \subset \mathrm{ri}(E)$, we would also have $0 \in \mathrm{int}(E)$). Therefore, 0 is a boundary point of $\overline{E}$. Thus, in the first case by Theorem 5.6 and in the second case by Theorem 5.8, it follows that there exists $a \in \mathbb{R}^n$ such that

$$0 = \langle a, 0 \rangle \geq \sup_{u \in E} \langle a, u \rangle \geq \langle a, d \rangle - \langle a, c \rangle \quad \forall\, c \in C,\, d \in D,$$

which clearly implies that $a \in \mathbb{R}^n$ satisfies the claim.

Definition 5.13 *Given a convex set C, a point $x \in C$ is an extreme point of C if there are no $x_1, x_2 \in C$, $x_1 \neq x_2$, and $\alpha \in (0,1)$ such that $x = \alpha x_1 + (1-\alpha)x_2$. That is, x is not a convex combination of points in C. We will denote the set of extreme points of C by ext(C).*

Example 5.14

i) If $C = [a, b]$ then ext$(C) = \{a, b\}$

ii) If $C = (a, b)$ then ext$(C) = \emptyset$.

iii) If $C = [a, +\infty)$, then ext$(C) = \{a\}$.

iv) If $C = B(a, r) = \{y \in \mathbb{R}^n \,:\, \|y - a\|_2 \leq r\}$, then ext$(C) = \{x \,:\, \|x - a\|_2 = r\}$.

Lemma 5.15 *Every bounded and closed convex set has at least one extreme point.*

Let C be a bounded closed convex set, $\delta = \max\left\{\|x\|^2 \,:\, x \in C\right\}$ and $\hat{x} \in$C such that $\|\hat{x}\|^2 = \delta$. We will prove that $\hat{x}$ is an extreme point of C. Assume that $\hat{x} = \alpha x_1 + (1-\alpha)x_2$ for some $x_1, x_2 \in C$ and $\alpha \in (0,1)$. The relation

$$\begin{aligned}\delta &= \|\alpha x_1 + (1-\alpha)x_2\|^2 = \alpha \|x_1\|^2 + (1-\alpha)\|x_2\|^2 - \alpha(1-\alpha)\|x_1 - x_2\|^2 \\ &\leq \alpha\delta + (1-\alpha)\delta - \alpha(1-\alpha)\|x_1 - x_2\|^2 = \delta - \alpha(1-\alpha)\|x_1 - x_2\|^2\end{aligned}$$

implies that $x_1 = x_2$. In view of the assumptions, we conclude that $x_1 = x_2 = \hat{x}$, which yields the claim.

Lemma 5.16 *Let C be a convex set, H a supporting hyperplane of C, and $T = C \cap H$. Then, ext$(T) \subset$ ext(C).*

Let us write $H = \{x \mid \langle a, x\rangle = c\}$ and let us assume that $C \subset H_+ = \{x \mid \langle a, x\rangle \geq c\}$. Let x_0 be an extreme point of T. Assume that $x_0 = \alpha x_1 + (1-\alpha)x_2$ for some $x_1, x_2 \in C$ and $\alpha \in (0,1)$. If we have $\max\{\langle a, x_1\rangle, \langle a, x_2\rangle\} > c$ then since $\langle a, x_2\rangle, \langle a, x_1\rangle \geq c$, we would have $\langle a, x_0\rangle > c$, which contradicts the fact that $x_0 \in H$. Therefore, it follows that $\langle a, x_1\rangle = \langle a, x_2\rangle = c$ which means that $x_1, x_2 \in T$. Recalling that x_0 is an extreme point of T, the relation $x_0 = \alpha x_1 + (1-\alpha)x_2$ for $x_1, x_2 \in T$ is only possible if $x_1 = x_2 = x_0$. Hence, we conclude that x_0 is an extreme point of C.

Lemma 5.17 *Let $C \subset \mathbb{R}^n$ be a convex set and $y \in \mathbb{R}^n$ a given point. Then, $ext(C) + \{y\} = ext(C + \{y\})$. That is, a is an extreme point of the set C if and only if $a + y$ is an extreme point of the set $\hat{C} = C + \{y\}$.*

Assume that $a + y \notin \text{ext}(C) + y$, i.e., a is not an extreme point of C. Then, there exist $x_1, x_2 \in C$ and $\alpha \in (0,1)$ such that $a = \alpha x_1 + (1-\alpha)x_2$. Hence, it follows that $a + y = \alpha(x_1 + y) + (1-\alpha)(x_2 + y)$, which shows that $a + y$ is not an extreme point of $\hat{C}$. Similarly if $a + y$ for $a \in C$ is not an extreme point of $\hat{C}$ then $a + y = \alpha(a_1 + y) + (1-\alpha)(a_2 + y)$ for $a_1 \neq a_2 \in C$ and $a = \alpha a_1 + (1-\alpha)a_2$ is not an extreme point of C and $a + y \notin \text{ext}(C) + y$.

Theorem 5.18 *Let $C \subset \mathbb{R}^n$ be a bounded closed and convex set. Then, $C = \overline{Conv}(ext(C))$. That is, C is the closed convex hull of the set of its extreme points.*

We will prove the result using induction in n, the dimension of the space. When $n = 1$, the bounded closed and convex sets of $\mathbb{R}$ are points and closed and bounded intervals, and then the claim holds trivially. Next, assume that the claim holds for all bounded closed and convex sets of $\mathbb{R}^m$. Let C be a bounded closed and convex set of $\mathbb{R}^{m+1}$ and $K = \overline{\text{Conv}}(\text{ext}(C))$. Under the assumptions on C, it is easy to see that $K \subset C$ (every point in K is a limit of points of $\text{Conv}(\text{ext}(C))$ and therefore, by convexity, of points of C and such limit belongs to C due to the fact that C is closed). Assume that $C \neq K$ and take $y \in C \backslash K$. Since K is closed and convex, it follows that there exists an hyperplane H that strictly separates y and K, that is, there exist a nonzero vector $a \in \mathbb{R}^{m+1}$ such that

$$\langle a, y\rangle < \inf\{\langle a, k\rangle \mid k \in K\}.$$

Take $c_0 = \inf\{\langle a, x\rangle \mid x \in C\} < +\infty$, and $x_0 \in C$ such that $c_0 = \langle a, x_0\rangle$, which are well defined since C is compact. It follows that the hyperplane $H_1 = \{x \in \mathbb{R}^{m+1} \mid \langle a, x\rangle = c_0\}$ is a supporting hyperplane for C. In addition, note that for $x \in K$, we have $\langle a, x\rangle \geq \inf\{\langle a, k\rangle \mid k \in K\} > \langle a, y\rangle \geq c_0$ and therefore $x \notin H_1$, i.e., $H_1 \cap K = \emptyset$. Now define $T = C \cap H_1$ and notice that $T \neq \emptyset$ since $x_0 \in T$ and it holds trivially that T is a bounded, closed and convex set. Since $\dim(H_1) = m$, it follows by the induction assumption that T contains extreme points that, by Lemma 10.2 are extreme points of C. Thus we have found extreme points of C not in K, which is a contradiction Therefore, it holds $C = K$ which ends the proof.

Corollary 5.19 *A polyhedron is the convex combination of its extreme points.*

6 General theorem on the alternative.

In this section, we will consider the system of inequalities

$$(\mathcal{S}:)\begin{cases} a_i^T x > b_i, \ i = 1, \dots, p, \\ a_i^T x \geq b_i, \ i = p+1, \dots, m \end{cases} \tag{6.13}$$

in variable $x \in \mathbb{R}^n$. We associate with $\mathcal{S}$ the two systems of linear equations and inequalities in variable $\lambda \in \mathbb{R}^m$:

$$\mathcal{T}_I : \begin{cases} \lambda \geq 0, \\ \sum_{i=1}^m \lambda_i a_i = 0, \\ \sum_{i=1}^p \lambda_i > 0, \\ \sum_{i=1}^m \lambda_i b_i \geq 0. \end{cases} \qquad \mathcal{T}_{II} : \begin{cases} \lambda \geq 0, \\ \sum_{i=1}^m \lambda_i a_i = 0, \\ \sum_{i=1}^m \lambda_i b_i > 0. \end{cases} \tag{6.14}$$

The next result relates the infeasibility of system $\mathcal{S}$ with the feasibility of systems $\mathcal{T}_I$ and $\mathcal{T}_{II}$.

Theorem 6.1 (General Theorem on Alternative (GTA)) *The system $\mathcal{S}$ is infeasible if and only if either $\mathcal{T}_I$ or $\mathcal{T}_{II}$, or both of these systems have a solution.*

To prove the GTA we will use the following results.

Lemma 6.2 (The homogeneous Farkas Lemma) *The system of homogeneous linear inequalities*

$$(F): \begin{cases} a^T x & < 0, \\ a_i^T x & \geq 0, \ i = 1, \dots, m. \end{cases}$$

is infeasible if and only if there exists $\lambda \geq 0$ such that $a = \sum_{i=1}^m \lambda_i a_i$.

Let us assume that there exist $\lambda \geq 0$ such that $a = \sum_{i=1}^m \lambda_i a_i$ and let x be a solution of system (F). Then, it follows that $a^T x = \sum_{i=1}^m \lambda_i a_i^T x \geq 0$ which is a contradiction. In summary, existence of such a λ implies the infeasibility of the system (F). Next, consider the set $X = \{a_1, \dots, a_m\}$, and let $K = \text{Conic}(X)$ which, by Proposition 4.8, is a closed cone. Notice that there exists $\lambda \geq 0$ satisfying $a = \sum_{i=1}^m \lambda_i a_i$ if and only if $a \in K$. Let us assume that $a \notin K$. Then, by Proposition 4.9, it follows that $x = P_K(a) - a$ satisfies

$$\langle a, x\rangle < 0 \text{ and } \langle y, x\rangle \geq 0 \quad \forall y \in K.$$

which means, taking $y = a_i$, $i = 1, \dots, m$, that system (F) is feasible and ends the proof.

Proposition 6.3 *The system of linear inequalities* (6.13) *has no solution if and only if this is the case for the following homogeneus system*

$$(\mathcal{S}^*): \begin{cases} -s < 0, \\ t - s \geq 0, \\ a_i^T x - b_i t - s \geq 0, \ i = 1, \dots, p, \\ a_i^T x - b_i t \geq 0, \ i = p+1, \dots, m. \end{cases} \tag{6.15}$$

Let $\hat{x}$ be a solution of the system (6.13) (system ($\mathcal{S}$)). Define $\hat{t} = 1$ and $\hat{s} = \min\{1, a_i^T\hat{x} - b_i \; i = \dots, p\}$. It follows that

$$\begin{aligned}
&\hat{s} > 0,\\
&\hat{t} - \hat{s} = 1 - \hat{s} \geq 0,\\
&a_i^T\hat{x} - b_i\hat{t} - \hat{s} = a_i^T\hat{x} - b_i - \hat{s} \geq 0 \quad i = 1, \dots, p,\\
&a_i^T\hat{x} - b_i\hat{t} = a_i^T\hat{x} - b_i \geq 0 \quad i = p+1, \dots, m,
\end{aligned}$$

which means that $(\hat{x}, \hat{s}, \hat{t})$ is a solution of (6.15) (system ($\mathcal{S}^*$)). Now, let $(\hat{x}, \hat{s}, \hat{t})$ be a solution of system ($\mathcal{S}^*$). Since $\hat{t} \geq \hat{s} > 0$ it follows that we can define $\bar{x} = \hat{x}/\hat{t}$ and we get

$$a_i^T\bar{x} - b_i = \frac{1}{\hat{t}}\left(a_i^T\hat{x} - \hat{t}b_i\right) \geq \frac{\hat{s}}{\hat{t}} > 0 \quad i = 1, \dots, p,$$

and

$$a_i^T\bar{x} - b_i = \frac{1}{\hat{t}}\left(a_i^T\hat{x} - \hat{t}b_i\right) \geq 0 \quad i = p+1, \dots, m,$$

which means that $\bar{x}$ is a solution of (6.13) (system ($\mathcal{S}$)). We have just proved that one system is feasible if and only if the other system is also feasible, which is equivalent to the claim.

Proof of the GTA Rewitting system (S^*) as follows

$$\begin{cases}
(0, -1, 0)^T(x, s, t) < 0,\\
(0, -1, 1)^T(x, s, t) \geq 0,\\
(a_i, -1, -b_i)^T(x, s, t) \geq 0, \; i = 1, \dots, p,\\
(a_i, 0, -b_i)^T(x, s, t) \geq 0, \; i = p+1, \dots, m.
\end{cases}$$

and using the Farkas Lemma, it follows that (S^*) is infeasible if, and only if, there exists $\bar{\lambda} = (\lambda_0, \lambda_1, \dots, \lambda_m) \geq 0$ such that

$$(0, -1, 0) = \lambda_0(0, -1, 1) + \sum_{i=1}^{p} \lambda_i(a_i, -1, -b_i) + \sum_{i=p+1}^{m} \lambda_i(a_i, 0, -b_i).$$

This means that $\bar{\lambda}$ satisfies

$$\bar{\lambda} \geq 0, \quad 0 = \sum_{i=1}^{m} \lambda_i a_i, \quad \sum_{i=1}^{p} \lambda_i = 1 - \lambda_0, \quad \sum_{i=1}^{m} \lambda_i b_i = \lambda_0 \tag{6.16}$$

Since $\sum_{i=1}^{p} \lambda_i \geq 0$ it follows that $\lambda_0 \in [0, 1]$. Therefore, $\lambda = (\lambda_1, \dots, \lambda_m)$ satisfies

$$\lambda \geq 0, \quad 0 = \sum_{i=1}^{m} \lambda_i a_i, \quad \sum_{i=1}^{p} \lambda_i = 1, \quad \sum_{i=1}^{m} \lambda_i b_i = 0 \;\text{ if }\; \lambda_0 = 0$$

or

$$\lambda \geq 0, \quad 0 = \sum_{i=1}^{m} \lambda_i a_i, \quad \sum_{i=1}^{p} \lambda_i \in [0, 1), \quad \sum_{i=1}^{m} \lambda_i b_i \in (0, 1] \;\text{ if }\; \lambda_0 \in (0, 1].$$

which means that in the first case above λ satisfies system $\mathcal{T}_I$ and in the second case it satisfies system $\mathcal{T}_{II}$. Now, let x be a solution to system (S) (6.13) and let $\lambda = (\lambda_1, \ldots, \lambda_m) \geq 0$ be a solution to either system $\mathcal{T}_I$ or system $\mathcal{T}_{II}$. If λ solves system $\mathcal{T}_I$, it follows that

$$\sum_{i=1}^{p} \lambda_i(a_i^T x) > \sum_{i=1}^{p} \lambda_i b_i \text{ and } \sum_{i=p+1}^{m} \lambda_i(a_i^T x) \geq \sum_{i=p+1}^{m} \lambda_i b_i.$$

Threfore, we have that

$$0 = \left(\sum_{i=1}^{m} \lambda_i a_i\right)^T x = \sum_{i=1}^{m} \lambda_i \left(a_i^T x\right) > \sum_{i=1}^{m} \lambda_i b_i^T \geq 0$$

and we obtain a contradiction. If λ solves system $\mathcal{T}_{II}$, then we have

$$0 = \left(\sum_{i=1}^{m} \lambda_i a_i\right)^T x = \sum_{i=1}^{m} \lambda_i(a_i^T x) \geq \sum_{i=1}^{m} \lambda_i b_i > 0$$

and we obtain a contradiction. Hence in both cases we obtain a contradiction, which shows that if either $\mathcal{T}_I$ or system $\mathcal{T}_{II}$ is solvable then system (S) is infeasible, and ends the proof.

7 Linear programming problems

7.1 Basic definitions and examples

The linear programming problem (LPP) consists in the minimization or maximization of a linear objective function subjected to a finite set of constraints defined by linear equalities and inequalities. The LPP, in its standard form, is formulated as follows:

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } Ax = b \text{ and } x \geq 0 \end{aligned} \tag{7.17}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and A is an $m \times n$ matrix.

Remark 7.1 *Notice that other formulations of the LPP can be cast in the form of* (7.17)*. For example,*

i) inequality constraints can be recast as equality constraints introducing slack or surplus nonnegative variables, i.e. $a^T x \leq d$ *can be reformulated as* $a^T x + s = b$, $s \geq 0$, *and* $a^T x \geq d$ *can be reformulated as* $a^T x - s = b$, $s \geq 0$, *respectively;*

ii) free variables, which means variables that are not required to be nonnegative, can be rewritten as the difference of two nonnegative variables.

Assumption 7.2 *The* $m \times n$ *matrix* A *satisfies the full rank assumption if* $m < n$ *and the* m *rows of* A *are linearly independent.*

7.2 Basic solutions and extreme points

Consider the system of linear equations

$$Ax = b \tag{7.18}$$

where $b \in \mathbb{R}^m$, and A is an $m \times n$ matrix that satisfies Assumption 7.2.

Definition 7.3 *Consider an ordered set of m indexes $J \subset \{1, 2, \ldots, n\}$ such that the matrix B, formed by the columns of A indexed by J, is invertible. The basic solution of the system (7.18) defined by B is the vector x with $x_J = B^{-1}b$ and $x_I = 0$ where $I = \{1, 2, \ldots, n\} \backslash J$. The components x_j, $j \in J$ are called basic variables of the basic solution. If one or more of the basic variables is equal to zero, we call the basic solution a degenerate basic solution.*

Remark 7.4 *To simplify the exposition of the results that will follow in these notes, usually we will denote a basic solution, related to a matrix B as in the definition above, by $x = (x_B, 0) = (B^{-1}b, 0)$. Formally, this corresponds to the matrix B being equal to the submatrix formed by the first m columns of the matrix A. By a simple reordering of the variables, we can reduce any case of basic solution to this case. Also, we will denote by B the set of indexes which index the vectors of the matrix B.*

Now consider the system of constraints

$$Ax = b, \quad x \geq 0 \tag{7.19}$$

where $b \in \mathbb{R}^m$, and $A \in M(m, n)$ satisfies Assumption 7.2.

Definition 7.5 *A vector satisfying (7.18) is said to be feasible for these constraints. A feasible solution to the system (7.19) that is also basic for the related linear system (7.18) is said to be a basic feasible solution; if this solution is also a degenerate basic solution, it is called a degenerate basic feasible solution.*

Definition 7.6 *An optimal solution to LPP (7.17) that is also basic for the related linear system (7.18) is said to be an optimal basic solution; if this solution is also a degenerate basic solution, it is called a degenerate optimal basic solution.*

Theorem 7.7 *Consider the LPP (7.17) where A satisfies the Assumption 7.2. Then,*

i) if there is a feasible solution to the system (7.19), there is a basic feasible solution;

ii) if there is an optimal feasible solution to the LPP (7.17), there is an optimal basic feasible solution.

First, we prove i). Notice that if $x = 0$ is a feasible solution to (7.19) then it is a degenerate basic feasible solution and the claim is proved. Next, consider $x = (x_1, x_2, \ldots, x_n) \neq 0$, a feasible solution for (7.19), and let $J = \{i \in \{2, \ldots, n\} \,|\, x_i > 0\}$. Let $B = \{a_j \,|\, j \in J\}$ denote the set of column vectors of A indexed by J. If B is a linear independent set, then we have a basic feasible solution and the claim is proved. Otherwise, consider the vector $y = (y_1, y_2, \ldots, y_n)$ with $\{y_j \,|\, j \in J\}$ satisfying

$$\sum_{j \in J} y_j a_j = 0$$

and $y_j = 0$ for all $j \notin J$. Clearly, we can assume that at least one component of y is positive. Notice that the vector $x_\varepsilon = x + \varepsilon y$ is a solution of the linear system (7.18) for any $\varepsilon \in \mathbb{R}$. In particular, considering

$$\varepsilon = \min \left\{ -\frac{x_i}{y_i} \,|\, y_i < 0, \right\} > 0$$

we obtain a new feasible solution to (7.19), since clearly $x_\varepsilon \geq 0$, and in addition $(x_\varepsilon)_j = 0$ for all $j \notin J$ and, at least for some $j \in J$, $(x_\varepsilon)_j$ became 0. This process shows how to reduce to zero at least one of the positive components of any nonzero and nonbasic feasible solution to (7.19). Clearly, after a finite number of steps of this procedure, we will obtain a feasible solution to (7.19), whose nonzero variables are associated to linear independent column vectors of the matrix A, in summary, a basic feasible solution. Then, the claim follows.

Now, we prove ii). Observe that we can proceed as in the proof of item i) of the theorem, but considering in addition that x is optimal for the LPP (7.17). In this case, let $\varepsilon > 0$ such that $x_1 = x_\varepsilon = x + \varepsilon y \geq 0$ is feasible (taking $\varepsilon \leq \min \left\{ -\frac{x_i}{y_i} \,|\, y_i < 0, \right\} > 0$)) and $x_2 = x - \varepsilon y \geq 0$ is feasible (taking $\varepsilon \leq \min \left\{ \frac{x_i}{y_i} \,|\, y_i > 0, \right\} > 0$) and x_1, x_2 are feasible solutions to (7.19). Next, if $c^T y > 0$ we have

$$c^T x_2 = c^T x - \varepsilon c^T y < c^T x$$

which is not possible while if $c^T y < 0$ we have

$$c^T x_1 = c^T x + \varepsilon c^T y < c^T x$$

which is also a contradiction. Therefore, $c^T y = 0$. Hence, $c^T x = c^T x_1 = c^T x_2$. Using the same procedure as in item i), we can obtain a basic feasible solution for system (7.19) with the value of the objective function equal to $c^t x$, that is, a basic optimal solution to (7.19), and the claim follows.

Theorem 7.8 (Equivalence of Extreme Points and Basic Solutions) *Consider the system (7.19) where A satisfies the Assumption 7.2. Let K be the convex polytope consisting of all feasible solutions to this system. The, a vector x is an extreme point of K if, and only if, x is a basic feasible solution to (7.19).*

Let x be a basic solution to (7.19) associated to a nonsingular submatrix B of A. Following the notations of Definition 7.3, consider the set of indexes J and I associated to B. We have that $x = (x_B, 0)$ and $x_I = 0$. Assume that $x = \alpha\hat{x} + (1-\alpha)\bar{x}$ for some $\alpha \in (0,1)$ and $\hat{x}, \bar{x} \in K$. Since $\hat{x}, \bar{x} \geq 0$, it follows that $\hat{x}_I = \bar{x}_I = 0$. These relations and the feasibility of $\hat{x}$ and $\bar{x}$ imply that $\hat{x}_J = \bar{x}_J = B^{-1}b$. Hence, we obtain that $x = \hat{x} = \bar{x}$ and we conclude that x is an extreme point of K. Next, let x be an extreme point of K. Define $J = \{i \in \{1, 2, \ldots, n\} \mid x_i > 0\}$ and $I = \{1, 2, \ldots, n\} \backslash J$, and let B be the submatrix of vectors of A indexed by J. If the column vectors of B are linearly independent, then x is a basic solution to (7.19) (a degenerated one if $|J| < m$). On the other hand, if these vectors are linearly dependent, then there exists a vector $u \in \mathbb{R}^{|J|}$ such that $Bu = 0$. Taking $y \in \mathbb{R}^n$ with $y_J = u$ and $y_I = 0$, and $\varepsilon > 0$ small enough, it is easy to see that the vectors $x_1 = x + \varepsilon y$ and $x_2 = x - \varepsilon y$ are feasible to (7.19), which, combined with the relation $x = (1/2)x_1 + (1/2)x_2$, shows that x is not an extreme point of K, yielding a contradiction and ending the proof.

7.3 Duality

In this subsection, we will consider the Linear Programming Problems (7.17)

Remark 7.9 *Other formulations of the LPP can be recast as instances of* (7.17)*. Notice that* $\{x \mid Ax = b\} = \{x \mid Ax \geq b, \ -Ax \geq -b\}$.

Definition 7.10 *The dual problem of the LPP* (7.17) *is the LPP*

$$\begin{aligned} &\textit{maximize } b^T y \\ &\textit{subject to } y^T A \leq c^T \end{aligned} \tag{7.20}$$

Remark 7.11 *Note that we can reformulate the above problem as an equivalent minimization LPP:*

$$\begin{aligned} &\textit{minimize } \ -b^T y \\ &\textit{subject to } y^T A \leq c^T \end{aligned} \tag{7.21}$$

Theorem 7.12 (Weak duality) *If x and y are feasible for* (7.17) *and* (7.20)*, respectively, then* $c^T x \geq y^T b$.

Simple calculations show that

$$y^T b = y^T (Ax) = (y^T A)x \leq c^T x$$

since $y^T A \leq c^T$ and $x \geq 0$.

The following result follows immediately from the previous theorem.

Corollary 7.13 *If x and y are feasible for (7.17) and (7.20), respectively, and $c^T x = y^T b$, then x and y are optimal for (7.17) and (7.20), respectively,*

Theorem 7.14 (Strong duality) *If either of the problems (7.17) or (7.20) has a finite optimal solution, so does the other, and the corresponding values of the objective functions are equal. If either problem has an unbounded objective, the other problem has no feasible solution.*

The second statement of the theorem follows trivially from Theorem 7.12 since any feasible solution of either of the problems provides a bound for the functional value of the other problem. To prove the first statement, let x_0 be an optimal feasible solution for the LPP (7.17), set $z_0 = c^T x_0$ and define the set $C = \left\{t(z_0 - c^T x, b - Ax) \,|\, t \geq 0,\ x \geq 0\right\}$. Observe that C is a closed convex cone. Also, it holds $(1,0) \notin C$, since this would imply that, for some $\hat{x} \geq 0$ and $\hat{t} > 0$, it holds $b - A\hat{x} = 0$, but in this case we must have $\hat{t}(z_0 - c^T\hat{x}) \leq 0$ since $\hat{x}$ would be feasible for the LPP problem (7.17). Now, by Theorem 5.6, it follows that there exists $a = (s, y)$ and a real scalar δ such that

$$s = (s,y)^T(1,0) < \delta < (s,y)^T\left(t(z_0 - cx, b - Ax)\right) = t\left(sz_0 - sc^T x + y^T b - y^T Ax\right) \ \forall\, t \geq 0,\ x \geq 0.$$

The above relation implies that $\delta \leq 0$ and, therefore, that $s < 0$. Notice that through a simple rescaling we can assume that $s = -1$. Dividing by t and making $t \to +\infty$, it follows that

$$0 \geq (z_0 - y^T b) + (-c + A^T y)^T x \quad \forall\, x \geq 0,$$

which implies that

$$0 \geq -c + A^T y \text{ and } 0 \geq z_0 - y^T b.$$

The first relation shows that y is feasible for the dual problem (7.20). The second relation, combined with this fact and Theorem 7.12, implies that $z_0 = y^T b$, whence it follows that y is an optimal solution for this problem and that both problems, (7.17) and (7.20), have the same optimal value. Next, notice that if the LPP (7.20) has a solution then its equivalent reformulation (7.11) has the same solution, and, in view of the first part of the proof, the the dual of this later problem also has an optimal solution with the same optimal value. It is easy to see the the dual of the LPP (7.11) is an equivalent reformulation of the LPP (7.17), which ends the proof.

We now present another proof of the duality theorem for LPs.

Consider the linear program

$$(LP) \quad c_* = \begin{cases} \min c^T x \\ a_i^T x \geq b_i,\ i = 1, \ldots, m. \end{cases} \tag{7.22}$$

A systematic way for bounding a function f on a set of form $\{x : g_i(x) \geq b_i,\ i = 1, \ldots, m\} = \{x : g(x) \geq b\}$ consists in using Lagrange duality. If x satisfies $g_i(x) \geq b_i, i = 1, \ldots, m$, then for any $\lambda \geq 0$ we have

$$\sum_{i=1}^m \lambda_i g_i(x) \geq \sum_{i=1}^m \lambda_i b_i.$$

Assuming that for all $x \in \mathbb{R}^n$ we have $f(x) \geq \sum_{i=1}^m \lambda_i g_i(x)$ then for any $\lambda \geq 0$ the quantity $\sum_{i=1}^m \lambda_i b_i$ is a lower bound for $f(x)$ for every x such that $g(x) \geq b$. Therefore the best lower bound that can be found on c_* with this technique is

$$\begin{array}{l} \max\ \sum_{i=1}^m \lambda_i b_i \\ \lambda \geq 0,\ \sum_{i=1}^m \lambda_i g_i(x) \leq f(x),\ \forall\, x \in \mathbb{R}^n. \end{array} \tag{7.23}$$

Let us look how the above problem writes when $f(x) = c^T x$ and $g_i(x) = a_i^T x$ are linear. In this situation the contraint $\sum_{i=1}^m \lambda_i g_i(x) \leq f(x),\ \forall\, x \in \mathbb{R}^n$, implies that $c = \sum_{i=1}^n \lambda_i a_i$, i.e., $c = A^T \lambda$ and therefore (7.23) becomes the linear problem

$$d_* = \left\{ \begin{array}{l} \max\ \lambda^T b \\ \lambda \geq 0,\ A^T \lambda = c \end{array} \right. \tag{7.24}$$

which is called the dual problem of (primal) problem (7.22). Therefore, we have the following:

Proposition 7.15 (Weak duality) *The optimal value of the dual problem is less than or equal to the optimal value of primal problem.*

Let us show that for linear problems, if the primal problem (7.22) is solvable then both the primal and the dual are solvable and have the same optimal value.

Indeed, a is a lower bound on c_* if and only if $c^T x \geq a$ for every x satisfying $Ax \geq b$, or equivalently if and only if the system of inequalities

$$\mathcal{S}_a : \quad Ax \geq b, \quad -c^T x > -a,$$

has no solution. Using GTA, $\mathcal{S}_a$ has no solution if and only if

$$\mathcal{T}_I : \left\{ \begin{array}{l} \lambda_0 > 0, \lambda_i \geq 0, i = 1, \ldots, m, \\ -\lambda_0 c + \sum_{i=1}^m \lambda_i a_i = 0, \\ -\lambda_0 a + \sum_{i=1}^m \lambda_i b_i \geq 0, \end{array} \right. \quad \text{or} \quad \mathcal{T}_{II} : \left\{ \begin{array}{l} \lambda_i \geq 0, i = 1, \ldots, m, \\ -\lambda_0 c + \sum_{i=1}^m \lambda_i a_i = 0, \\ -\lambda_0 a + \sum_{i=1}^m \lambda_i b_i > 0, \end{array} \right.$$

has a solution.

We now check that $\mathcal{S}_a$ has no solution if and only if $\mathcal{T}_I$ has a solution. We already know that if $\mathcal{T}_I$ has a solution then $\mathcal{S}_a$ has no solution. Let us now assume that $\mathcal{S}_a$ has no solution and that the system $Ax \geq b$ has a solution. We want to prove that $\mathcal{T}_I$ has a solution. By contradiction, if $\mathcal{T}_I$ has no solution, from our previous observations, necessarily, $\mathcal{T}_{II}$ must have a solution. Moreover, every solution to $\mathcal{T}_{II}$ satisfies $\lambda_0 = 0$, otherwise it is also a solution to $\mathcal{T}_I$. Therefore system $\mathcal{T}_{II}$ has a solution for any a, c (since $\lambda_0 = 0$), in particular, say, for $a = 0, c = -1$. Using GTA, this means that the system

$$-0^T x > -1, \quad Ax \geq b$$

has no solution, which is a contradiction with the fact that there exists x such that $Ax \geq b$. Therefore if $\mathcal{S}_a$ has no solution then $\mathcal{T}_I$ has a solution. If $\mathcal{T}_I$ has a solution, then it also has a solution with $\lambda_0 = 1$ (from a solution λ with $\lambda_0 \neq 0$, we can build a new solution

$(1, \lambda_1/\lambda_0, \ldots, \lambda_m/\lambda_0)$ since $\lambda_0 > 0$). Now, $(1, \lambda)$ is a solution to $\mathcal{T}_I$ if and only if $\lambda \in \mathbb{R}^m$ satisfies:

$$(SD): \quad \lambda \geq 0, \; c = A^T\lambda, \; a \leq \lambda^T b.$$

We have shown the following:

Proposition 7.16 *a is a lower bound on the optimal value of (7.22) if and only if there exists (λ, a) such that $\lambda \geq 0, \; c = A^T\lambda, \; a \leq \lambda^T b$.*

We loose nothing in looking for lower bounds a in (SD) of form $a = \lambda^T b$. In this situation, the best lower bound a on (7.22) given by a solution of (SD) is the optimal value of the dual problem

$$(D) \quad \left\{ \begin{array}{l} \max \; \lambda^T b \\ \lambda \geq 0, \; A^T\lambda = c. \end{array} \right.$$

We have now shown:

Proposition 7.17 *Whenever λ is feasible for (D) then $\lambda^T b$ is a lower bound on the optimal value of (7.22). If (7.22) is feasible then for every $a \leq c_*$, there exists a feasible solution λ to (D) such that $\lambda^T b \geq a$*

Theorem 7.18 (Duality Theorem in Linear Programming) *Consider linear program (LP) along with its dual (D) given by:*

$$(LP) \quad c_* = \left\{ \begin{array}{l} \min c^T x \\ a_i^T x \geq b_i, \; i = 1, \ldots, m. \end{array} \right. \quad (D) \quad \left\{ \begin{array}{l} \max \; \lambda^T b \\ \lambda \geq 0, \; A^T\lambda = c. \end{array} \right. \tag{7.25}$$

Then:

1) Duality is symmetric: the dual to the dual is the dual.
2) The value of the dual objective at every dual feasible solution is $\leq$ the value of the primal objective at every primal feasible solution.
3) The following 5 properties are equivalent:
(i) The primal is feasible and bounded below.
(ii) The dual is feasible and bounded above.
(iii) The primal is solvable.
(iv) The dual is solvable.
(v) Both the primal and the dual are feasible.
Whenever (i)-(v) hold both the primal and the dual are solvable and have the same optimal value.

1) Writing the dual as

$$\begin{array}{c} -\min \; -b^T y \\ \left[\begin{array}{c} A^T \\ -A^T \\ I_m \end{array} \right] y \geq \left[\begin{array}{c} c \\ -c \\ 0 \end{array} \right] \end{array}$$

the dual of the dual is the linear program

$$\begin{aligned}&-\max\ (\lambda_1-\lambda_2)^T c\\ &A(\lambda_1-\lambda_2)+\lambda_3=-b,\\ &\lambda_3\ge 0,\ \lambda_1,\lambda_2\ge 0\end{aligned}$$

which is equivalent to (LP) setting $x=\lambda_2-\lambda_1$.

2) is weak duality.

3) (i) $\Rightarrow$ (iv). Taking $a=c_*$ we get that there exists a dual feasible solution λ satisfying $\lambda^T b\ge c_*$. Together with 2) this implies $c_*=\lambda^T b$ and therefore $d_*=c_*$.

(iv) $\Rightarrow$ (ii) is clear.

(ii) $\Rightarrow$ (iii) comes from the primal-dual symmetry and (i) $\Rightarrow$ (iv).

(iii) $\Rightarrow$ (i) is clear.

Clearly (i)-(iv) imply (v). If (v) holds then the primal is feasible and bounded below by the value of the dual objective at any feasible dual point (see 2)).

We deduce the following necessary and sufficient optimality conditions for linear programs:

Theorem 7.19 *Consider primal problem (LP) and its dual (D) given by (7.25). Then a pair (x,λ) of primal-dual feasible solutions is optimal if and only if*

$$\lambda_i[Ax-b]_i=0,\ i=1,\dots,m.$$

The above condition can also be written $c^T x=\lambda^T b$ (no duality gap).

If (x,λ) is a primal-dual optimal solution then optimal values of the primal and the dual are equal, i.e., $c^T x=\lambda^T b$. On the other hand, if $c^T x=\lambda^T b$, then due to 2) in Theorem 7.18, x and λ are optimal primal and dual solutions. Finally, there is no duality gap iff $\lambda_i[Ax-b]_i=0,\ i=1,\dots,m$, because if (x,λ) is a primal-dual feasible solution then

$$\lambda^T(Ax-b)=x^TA^T\lambda-b^T\lambda=x^Tc-b^T\lambda.$$

Remark 7.20 *Notice that, starting with the primal-dual pair of LPPs given in (7.25) and taking into account Remark 7.9, we can obtain the following primal-dual pair of LPPs*

(7.26)

Primal problem (P)	$\max\ c^\top x$ $Ax=b$ $x\ge 0$	$\max\ c^\top x$ $Ax=b$ $Cx\le d$ $x\ge 0$	$\min\ c^\top x$ $Ax=b$ $x\ge 0$	$\min\ c^\top x$ $Ax=b$ $Cx\le d$ $x\ge 0$
Dual problem (D)	$\min\ \lambda^\top b$ $A^\top\lambda\ge c$	$\min\ \lambda^\top b+\mu^\top d$ $A^\top\lambda+C^\top\mu\ge c$ $\mu\ge 0$	$\max\ \lambda^\top b$ $A^\top\lambda\le c$	$\max\ \lambda^\top b+\mu^\top d$ $A^\top\lambda+C^\top\mu\le c$ $\mu\le 0$

For each one of these primal-dual pairs of LPPs a similar result to Theorem 7.18 also holds.

7.4 The SIMPLEX algorithm

7.4.1 Simplex algorithm: compact version, examples

The Simplex is an algorithm for solving linear programming problems. We sill consider the linear programming problem

$$\begin{cases} \max\ c^\top x = \sum_{i=1}^n c_i x_i \\ \sum_{j=1}^n A_{ij} x_j = b_i \quad i = 1, \dots, m_1, \\ \sum_{j=1}^n A_{ij} x_j \geq b_i, \quad i = m_1 + 1, \dots, m_1 + m_2, \\ \sum_{j=1}^n A_{ij} x_j \leq b_i, \quad i = m_1 + m_2 + 1, \dots, m, \\ x \geq 0 \end{cases} \tag{7.27}$$

where $m = m_1 + m_2 + m_3$. In (7.27), we can assume that $b_i \geq 0$ for all $i = 1, \dots, m$ (if for some i it holds $b_i < 0$, then we can re-write the corresponding equality or inequality by multiplying both left and right hand sides by -1).

We introduce (non-negative) slack variables e_i in order to re-write (7.27) as follows

$$\begin{cases} \max\ c^\top x = \sum_{i=1}^n c_i x_i \\ \sum_{j=1}^n A_{ij} x_j = b_i, \quad i = 1, \dots, m_1, \\ \sum_{j=1}^n A_{ij} x_j - e_i = b_i, \quad i = m_1 + 1, \dots, m_1 + m_2, \\ \sum_{j=1}^n A_{ij} x_j + e_i = b_i, \quad i = m_1 + m_2 + 1, \dots, m, \\ x \geq 0, \quad e \geq 0 \end{cases} \tag{7.28}$$

where $b \geq 0$. The Simplex algorithm has two phases summarized below.

Phase I of the Simplex The objective of the first phase of the simplex is to find out whether problem (7.28) has a solution. Furthermore, if (7.28) has a solution, Phase I provides a base matrix of constraints (set of independent columns of A), used to start phase II.

Phase II of the Simplex For a problem of the form (7.28) that has a solution, phase II finds an optimal solution based on the matrix of constraints provided by phase I. Phase II then allows a problem to be solved linear, given a basis of the constraints matrix.

Details of Phase I. Phase I consists of solving the problem

$$\begin{cases} \min \sum_{i=1}^{m_1+m_2} v_i \\ \sum_{j=1}^n A_{ij} x_j + v_i = b_i, \quad i = 1, \dots, m_1, \\ \sum_{j=1}^n A_{ij} x_j - e_i + v_i = b_i, \quad i = m_1 + 1, \dots, m_1 + m_2, \\ \sum_{j=1}^n A_{ij} x_j + e_i = b_i, \quad i = m_1 + m_2 + 1, \dots, m, \\ x \geq 0, \quad e \geq 0, \ v \geq 0, \end{cases} \tag{7.29}$$

which minimizes the sum of constraint violations. The problem above is always viable and has an optimal solution. As we have the evident base $B = (v, e_{m_1+m_2+1}, \dots, e_m))$ for this problem, we can apply Phase II of the Simplex to solve (7.29) from this base.

Proposition 7.21 *Problem (7.28) has a solution if and only if the optimal value of problem (7.29) is 0.*

If the optimal value of problem (7.29) is 0 and if there are artificial variablesv_i in the basic solution, we can replace each of these variables with variables from the initial problem x_i or e_i.

Details of Phase II. Here, without loss of generality, we will consider the LP problem (7.17). Starting with that base B determined in Phase I, we can reformulate this problem as follows

$$\begin{aligned} \min\ z = z_0 + \bar{c}_D^\top x_D \\ x_B + N x_D = \bar{b} \\ x_B \geq 0,\ x_D \geq 0. \end{aligned} \tag{7.30}$$

where, denoting $A = (B|D)$, we have $z_0 = c_B^T B^{-1} b$, $N = B^{-1} D$, $\bar{b} = B^{-1} b$ and $\bar{c}_D = c_D - N^T c_B$

To simplify the presentation we will assume that $\{1, 2, \ldots, m\}$ is the set of indexes of the columns of A which correspond to the vectors in B.

We will start with a previous definition and some previous reults that will be used to analyze the Simplex method.

Definition 7.22 *Let x be a basic solution associated to a basis matrix B and let c_B be the vector of costs of the basic variables. We define the vector of reduced costs $\bar{c}$ associated to x according to the formula*

$$\bar{c}^T = c^T - c_B^T B^{-1} A. \tag{7.31}$$

Remark 7.23 *Notice that $\bar{c}_B = c_B - \left(c_B^T B^{-1} B\right)^T = 0$, that is, the reduced costs corresponding to basic variables are equal to zero.*

The next result clarifies the role of the reduced costs in the behaviour of the objective function.

Lemma 7.24 *Let B be a basis for problem (1.5) and consider D and N as in (7.30). Take $\theta > 0$ and $r \in D$, and define $x_r^\theta = \theta$, $x_l^\theta = 0$ for $l \in D$, $l \neq r$, and $x_B^\theta = B^{-1} b - N x_D^\theta$. Then, for $x^\theta = (x_B^\theta, x_D^\theta)$, it holds $c^T x^\theta = z_0 + \theta \bar{c}_r$.*

Simple calculations show that

$$c^T x^\theta = c_B^T B^{-1} b + \left(c_D^T - c_B^T B^{-1} D\right) x_D = c_B B^{-1} b + \theta \left(c_r - c_B^T B^{-1} A_r\right) = c_B B^{-1} b + \theta \bar{c}_r.$$

Remark 7.25 *The previous lemma shows that, if the reduced cost corresponding to the non basic variable x_r is positive, then by increasing the value if the non-basic variable from 0 to $\theta > 0$ we can we can decrease the value of the objective function. In particular, if for some $\theta > 0$ we have that $x^\theta \geq 0$ with x^θ defined as above, then this vector would be a new feasible solution with a lower value of the objective function.*

Given a basis, the next result shows a procedure to construct a new basis by changing one vector of the former one, and presents formulas that relate the coordinates of a vector in both basis.

Lemma 7.26 *Let $B = \{b_1, \ldots, b_m\}$ be a basis of $\mathbb{R}^m$ and $u \in \mathbb{R}^m$ a vector satisfying $\left(B^{-1}u\right)_r \neq 0$. Then, it holds*

i) the set $\bar{B} = \{\bar{b}_i, \ldots, \bar{b}_m\}$ with $\bar{b}_i = b_i$ for all $i \in \{1, \ldots, m\} - \{r\}$ and $\bar{b}_r = u$ is also a basis of $\mathbb{R}^m$;

ii) for any vector $x = \sum_{i=1}^m x_i b_i$ it holds $x = \sum_{i=1}^m \bar{x}_i \bar{b}_i$ with $\bar{x}_i = x_i - x_r \left(\left(B^{-1}u\right)_i / \left(B^{-1}u\right)_r\right)$ for $i \in \{\ldots, m\} - \{r\}$ and $\bar{x}_r = x_r / \left(B^{-1}u\right)_r$.

First we prove item ii). Simple calculations show that $u = BB^{-1}u = \sum_{i=1}^m \left(B^{-1}u\right)_i b_i$ whence it follows that

$$b_r = \frac{1}{\left(B^{-1}u\right)_r}\left[u - \sum_{i=1,\, i\neq r}^{m} \left(B^{-1}u\right)_i b_i\right]$$

and therefore we have that

$$\begin{aligned} x = \sum_{i=1\, \imath\neq r}^{m} x_i b_i + x_r b_r &= \sum_{i=1\, \imath\neq r}^{m} x_i b_i + x_r \frac{x_r}{\left(B^{-1}u\right)_r}\left[u - \sum_{i=1,\, i\neq r}^{m} \left(B^{-1}u\right)_i b_i\right] \\ &= \sum_{i=1\, \imath\neq r}^{m} \left[x_i - \frac{x_r \left(B^{-1}u\right)_i}{\left(B^{-1}u\right)_r}\right] b_i + \frac{x_r}{\left(B^{-1}u\right)_r} u \end{aligned}$$

which end the proof of item ii). To prove item i) notice that item ii) shows that any vector $x \in \mathbb{R}^m$ can be written as a linear combination of the vectors of $\bar{B}$. Hence, $\bar{B}$ is also a basis.

The next result shows a procedure to, given a feasible basic solution, produce a new feasible basic solution. The next result shows how to select a non-basic variable and a value θ to be assigned to this variable in order to achieve this.

Lemma 7.27 *Let B be a basis for problem (1.5) such that the associated basic solution is feasible and non-degenerated, and consider D and N as in (7.30). In addition, assume that for some $r \in D$ and $l \in \{1, \ldots, m\}$ it holds $N_{lr} > 0$. Define $x_B = B^{-1}b$ and*

$$\theta = \min\left\{\frac{x_{B(i)}}{N_{ir}} \;:\; N_{ir} > 0\right\}$$

and let x^θ defined as in Lemma 7.24. Then, x^θ is feasible for the LPP (7.30). In addition, for any $s \in \{1, \ldots, m\}$ such that $\theta = x_{B(s)}/N_{sr}$, it holds $x^\theta_{B(s)} = 0$.

Notice that it holds $x_B = B^{-1}b \geq 0$. From the definition we have that $x^\theta_D \geq 0$. In addition, for any $i \in \{1, \ldots, m\}$, we have that

$$x^\theta_{B(i)} = (B^{-1}b)_i - (Nx^\theta)_i = x_{B(i)} - \theta N_{ir} \geq x_{B(i)} \geq \ 0 \text{ if } \ N_{ir} \leq 0$$

and, using the definition of θ,

$$x^\theta_{B(i)} = (B^{-1}b)_i - (Nx^\theta)_i = x_{B(i)} - \theta N_{ir} \geq \ 0 \text{ if } \ N_{ir} > 0. \tag{7.32}$$

Since it holds trivially that $Ax^\theta = b$, it follows the first part of the claim. The second part of the claim follows trivially from (7.32) and the definition of θ.

Notice that, under the assumptions of Lemma 7.27, and assuming additionally $\bar{c}_r > 0$, the procedure described in this lemma would produce a new basic feasible solution with a larger value of the objective function.

Finally, complementing Lemma 7.22, the next theorem shows that all the reduced costs of a feasible basic solution being non-negative is a necessary and sufficient for this solution being optimal.

Theorem 7.28 *Let x be a basic solution associated to a basis matrix B and $\bar{c}$ the corresponding vector of reduced costs. Then*

(a) if $\bar{c} \leq 0$, then x is optimal;

(b) if x is optimal and non degenerate, then $\bar{c} \leq 0$.

First we prove item (a). Note that $\lambda = \left(B^{-1}\right)^T c_B$ satisfies

$$\bar{c} = c - A^T\lambda.$$

Hence, if $\bar{c} \leq 0$, then λ is a feasible for its dual LPP (according to (7.26)). Additionally, we have that

$$\lambda^T b = bc_B^T B^{-1} b = c_b^T x_B = c^T x$$

which means, in view of Theorem 7.14, that λ and x are optimal solutions for problems (7.17) and (7.20), respectively. Item (b) follows straightforwardly from Lemma 7.24.

An iteration of the simplex method:

1. We start with a basis consisting of the basic columns $A_{B(1)}, \ldots.., A_{B(m)}$, and an associated basic feasible solution x, with value of the objective function equal to $z_0 = c_B B^{-1} b$.

2. Compute the reduced costs $\bar{c}_l = c_l - c_B^T B^{-1} A_l$, for all nonbasic indices l. If they are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates; else, choose some j for which $\bar{c}_j > 0$.

3. Compute $u = B^{-1} A_j$, If no component of u is positive, the problem is unbounded and the algorithm terminates, i.e., the optimal cost is $-\infty$.

4. If some component of u is positive, let

$$\theta_0 = \min\left\{\frac{x_{B(i)}}{u_i} \,:\, u_i > 0\right\} = \min\left\{\frac{(B^{-1}b)(i)}{u_i} \,:\, u_i > 0\right\}$$

5. Let l be such that $\theta_0 = x_{B(\ell)}/u_\ell$. Form a new basis by replacing $A_{B(l)}$ with A_j. If y is the new basic feasible solution, the values of the new basic variables are $y_j = \theta_0$ and $y_{B(i)} = x_{B(i)} - \theta_0 u_i \; i \neq l$, and the value fo the objective function at y is equal to $z_0 + (x_{B(\ell)}/u_\ell)\bar{c}_j$

Remark 7.29 *Note that we need to calculate $B^{-1}A_j$ where x_j is a nonbasic variable. If we consider all the nonbasic variables, we would need to perform $n-m$ of such calculations. In general, we can calculate $d = B^{-1}u$ in two different ways: 1) solving the linear system $Bd = u$ or 2) calculating the inverse B^{-1} and performing the multiplication*

Theorem 7.30 *Assume that the LPP (7.17) does not have degenerate basic solutions. Then, the Simplex method finds an optimal basic solution or determines that this problem is unbounded below in a finte number of iterations.*

At each iteration the Simplex method determines that the objective function is unbounded on the feasible set and stops or finds a new feasible basic solution. In addition, since along all the iterations (in the latter case) the objective function decreases, it follows that all the feasible basic solutions generated by the method are different. Notice that from the n columns of the matrix A we can select at most $C_m^n = n!/m!(n-m)!$ different sets of m vectors, then it follows that the number of feasible basic solutions of the system in (7.17) is finite. Then, the Simplex method determines the unboundedness of the objective function, or it stops at some iteration, otherwise it would generate more than C_m^n feasible basic solutions. Then, at the last iteration the Simplex method stops because it has established the unboundedness of the objective function or because the sopping criteria of optimality involving the reduced costs holds, in which case it has found an optimal feasible basic solution.

Given the basic basis $B = \left\{A_{B(1)}, \dots, A_{B(m)}\right\}$ (assuming to easy the presentation that $B = \{1, 2, \dots, m\}$) and the related reduced costs vector $\bar{c}$, we can rewrite problem (7.30) as follows

$$\begin{aligned} &\max\ z = c_B^T B^{-1} b + \bar{c}^T x = z_0 + (0, \bar{c}_D)^\top (x_B, x_D) \\ &A'x = b' \\ &x = (x_B; x_D) \geq 0. \end{aligned} \tag{7.33}$$

where

$$A' = B^{-1}A = \left[I \,|\, B^{-1}D\right] = \left[I \,|\, N\right], \quad b' = B^{-1}b, \quad z_0 = c_B^T B^{-1} b, \quad \bar{c}_D^T = c_D^T - c_B^T B^{-1} D. \tag{7.34}$$

The Simplex method for the formulation (7.33):

1. Along the iterations B is the identity matrix.

2. Calculate the reduced costs $\bar{c}_N^T = c_B^T - c_N^T N$. If $\bar{c}_N \leq 0$ the current basic feasible solution is optimal, and the algorithm terminates; else, choose some $r \in N$ for which $(\bar{c}_N)_r > 0$ (the variable x_r will enter the basis).

3. If all the components of N_r are non-positive, then the problem is unbounded and the algorithm terminates, i.e., the optimal cost is $+\infty$.

4. If some component of N_r is positive, let

$$\theta_0 = \min\left\{\frac{b_i}{(N_r)_i} \ : \ (N_r)_i > 0\right\}$$

5. Let s be such that $\theta_0 = b_s/(N_r)_s$. Form a new basis $\bar{B}$ by replacing e_s with N_r. and update the model

$$B = B \cup \{r\} - \{s\}, \quad N = N \cup \{s\} - \{r\},$$

and (update B, N and b)

$$B = I, \quad [N \,|\, b] = \bar{B}^{-1}[N \,|\, b] \text{ (}\leftarrow \text{ full tableu implementation does this, without calculating } \hat{B}^{-1}\text{)}$$

Naive implementation: In steps 2) and 3) of the implementation above: first, we calculate the so called *vector of multipliers* $p = c_B^T B^{-1}$ by solving the linear system $pB = c_B$, which requires $\mathcal{O}(m^3)$ arithmetic operations; then we calculate the reduced $n - m$ costs $\bar{c}_l = c_l - c_B^T B^{-1} A_l = c_l - p^T A_l$ for $j \notin B$, which requires $\mathcal{O}(mn)$ arithmetic operations, due to the calculation of $n = m$ inner products of vectors in $\mathbb{R}^n$; next, once we have determined the non basic variable x_j that will enter the base, we calculate vector $u = B^{-1}A_j$ by solving the linear system $Bu = A_j$, adding another $\mathcal{O}(m^3)$ arithmetic operations. Thus, in total, this naive implementation requires $\mathcal{O}(m^3 + mn)$ arithmetic operations.

Full tableau implementation of the Simplex method:

We can represent (7.33) using the tableau below

Basic variables	z	x_B	x_D	
x_B	0	I	$B^{-1}D$	$\bar{b}$
	1	0	$-\bar{c}_D$	z_0

(7.35)

with $x_B = \{x_1, \dots, x_m\}$, $x_D = \{x_{m+1}, \dots, x_n\}$, $\bar{c}_D = \{\bar{c}_{m+1}, \dots, \bar{c}_n\}$ and $\bar{b} = B^{-1}b$. A more detailed version of the tableau is presented below.

		Basic variables	z	x_1	$\dots$	x_r	$\dots$	x_m	x_{m+1}	$\dots$	x_s	$\dots$	x_n	
L_1		x_1	0											$\bar{b}_1$
		$\vdots$												$\vdots$
L_r	$\leftarrow$	x_r	0			I					N_{rs}			$\bar{b}_r$
		$\vdots$												$\vdots$
L_m		x_m	0											$\bar{b}_m$
L_{m+1}			1	0	$\dots$	0	$\dots$	0	$-\bar{c}_{m+1}$	$\dots$	$-\bar{c}_s$	$\dots$	$-\bar{c}_n$	z_0
											$\uparrow$			

(7.36)

Next, it follows a detailed presentation of the Simplex method iteration. If $\bar{c}_N \leq 0$, i.e., is $-\bar{c}_i \geq 0, i = m+1, \dots, m$, in the table (7.41), then we find the optimal solution $x_B = \bar{b}$ and $x_N = 0$, being z_0 the optimal value. Otherwise, we introduce the index s satisfying

$$-\bar{c}_s = \min\{-\bar{c}_i, \ i = m+1, \dots, n\}.$$

If $N_{is} \leq 0$ for $i = 1, \dots, m$, then the optimal value of (7.28) is $+\infty$. To see this, take $x_N(s) = \theta$ and $x_N(i) = 0$ otherwise. Take x_B such that $x_B + N x_N = b$. Then $x_N \geq 0$,

$x_B = b - Nx_N = b - \theta N_{\cdot s} \geq 0$ and the the objective at (x_B, x_N) is $z_0 + \theta c_s \to +\infty$ when $\theta \to +\infty$. Otherwise, we introduce the index r defined by

$$\frac{\bar{b}_r}{N_{rs}} = \min\left\{\frac{b_i}{N_{is}}\ :\ i = 1, \ldots, m,\ \text{e } N_{is} > 0\right\}.$$

The new basis $\bar{B}$ is $\bar{B} = B \cup \{s\} - \{r\}$ and we update the table (7.41) (i.e. we re-write s (7.28) using the base $\bar{B}$) with the following sequence of operations:

$$\begin{array}{lcl} L_r & \leftarrow & \dfrac{L_r}{N_{rs}} \\ L_i & \leftarrow & L_i - N_{is}L_r, \quad i = 1, \ldots, m,\ i \neq r, \\ L_{m+1} & \leftarrow & L_{m+1} + c_s L_r. \end{array}$$

The element N_{rs} is called pivot.

Let $\bar{B}$ be the new basis described above. In view of Lemma 7.26, and noticing that $N_{rs} = (B^{-1}A_s)_r$, it follows that the iteration of the Simplex method described above is performing the construction of $\bar{B}^{-1}[A\,|\,b] = \left[\left[I\,|\,\bar{B}^{-1}\bar{D}\right]\,|\,\bar{B}^{-1}b\right]$. That is, the matrix representation associated to the formulation (7.33) for the new basis $\bar{B}$.

In addition, notice that the value of the new basic variable according to the tableau is $\bar{b}_r/N_{rs}$. We will prove that the applying that the operation performed on the last row of the tableau computes the updated vector of reduced costs, that is, the vectors of reduced costs of the new basic solution associated to the basis $\bar{B}$, and also the updated value of the objective function at the new basic solution. For the transformation of the last row, we will multiply the r-th row for $\bar{c}_s/A_{sr}$ and then we will add it to the last row. The purpose here is to set the reduced cost of x_s (the new basic variable) equal to zero, which is the value of the reduced costs of the basic variables. Notice that the last row of the tableau has the form

$$L_{m+1} = \left(-\bar{c}^T, z_0\right) = \left(-c^T + c_B^T B^{-1}A, z_0\right)$$

where $z_0 = c_B^T B^{-1}b$, while the r-th row of the matrix $B^{-1}A$ has the expression

$$L_r = \left(e_r^T B^{-1}A, e_r^T B^{-1}b\right)$$

Hence, after applying the elementary operations we obtain the following new expression for the last row

$$\begin{aligned} \left(\frac{\bar{c}_s}{N_{rs}}\right) L_r + L_{m+1} &= \left(-c^T + \left(\left(\frac{\bar{c}_s}{N_{rs}}\right) e_r^T + c_B^T\right) B^{-1}A, \left(\frac{\bar{c}_s}{N_{rs}}\right)\left(B^{-1}b\right)_r + z_0\right) \\ &= \left(-c^T + p^T \bar{B}^{-1}A, \left(\frac{\bar{c}_s}{N_{rs}}\right)\left(B^{-1}b\right)_r + z_0\right) \end{aligned}$$

where $p^T = \left(\left(\bar{c}_s/N_{rs}\right) e_r + c_B\right)^T B^{-1}\bar{B}$. First, notice that Lemma 7.24 with $\theta = 1$ shows that, after the Simplex iteration, the value of the objective function changes from $z_0 = c_B^T B^{-1}b$ to $z_0 + (x_r/A_{sr})c_s = z_0 + \left((B^{-1}b)_r/A_{sr}\right) c_s$, which coincides with the value of the last component of the vector above. Next, notice that relations $B^{-1}\bar{B}_i = e_i$ for all $i \in B$ with $i \neq s$ and $B^{-1}\bar{B}_s = N_s$ imply that

$$p_i = \left(\left(\frac{\bar{c}_s}{N_{rs}}\right) e_r + c_B\right)_i = (c_B)_i = c_i \quad \forall\, i \in B \text{ with } i \neq r$$

and

$$p_s = \left(\left(\frac{\bar{c}_s}{N_{rs}}\right) e_r^T + c_B^T\right) N_s = \left(\frac{\bar{c}_s}{N_{rs}}\right) N_{rs} + c_B^T N_s = \bar{c}_s + c_B^T N_s = \bar{c}_s + c_B^T B^{-1} A_s = c_s.$$

Therefore, $p = c_{\bar{B}}$ and it holds $-c^T + p^T \bar{B}^{-1} A = -c + c_{\bar{B}}^T \bar{B}^{-1} A$, which is the vector of reduced costs corresponding to the new basis $\bar{B}$. Hence, we have that

$$\left(\frac{\bar{c}_s}{N_{rs}}\right) L_r + L_{m+1} = \left(-c + c_{\bar{B}} \bar{B}^{-1} A, \bar{B}^{-1} b\right),$$

and the claim follows.

In each iteration of the implementation discussed above we modify each row of the tableau by adding another row multiplied by a certain constant. Hence, it follows that this implementation requires $\mathcal{O}(mn)$ arithmetic operations, since nm is the number of elements of the tableau.

As a final comment, notice that in the case of a minimization problem, the only modification when applying the simplex method is that we check if any reduced cost is negative in order to choose the variable entering the basis (since we want to decrease the value of the objective function).

Example 7.31 *We are going to solve the following LPP by the Simplex method*

$$\begin{cases} \min\ z = 6x_1 + 7x_2 + 4x_3 \\ 4x_1 + 5x_2 + x_3 \leq 23 \\ 5x_1 + 6x_2 + 2x_3 \geq 28 \\ x_1 \geq 0, \quad x_2 \geq 0 \ \ x_3 \geq 0. \end{cases}$$

Introducing the slack and artificial variables we have the following reformulation of LPP above in the standard form.

$$\begin{cases} \min\ z = v \\ 4x_1 + 5x_2 + x_3 + e_1 = 23 \\ 5x_1 + 6x_2 + 2x_3 - e_2 = 28 \\ x_1 \geq 0, \quad x_2 \geq 0 \ \ x_3 \geq 0 \ \ e_1 \geq 0 \ \ e_2 \geq 0. \end{cases}$$

Introducing the artificial variables we have the following auxiliar LPP for the Phase I

$$\begin{cases} \min\ z = v \\ 4x_1 + 5x_2 + x_3 + e_1 = 23 \\ 5x_1 + 6x_2 + 2x_3 - e_2 + v = 28 \\ x_1 \geq 0, \quad x_2 \geq 0 \ \ x_3 \geq 0 \ \ e_1 \geq 0 \ \ e_2 \geq 0 \ \ v \geq 0. \end{cases}$$

Phase I: We start with the basic variables $x_B^T = (e_1, v) = (23, 28)^T$ *and the non-basic variables* $x_D^T = (x_1, x_2, x_3, e_2) = (0, 0, 0, 0)$. *Hence, we have that* $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *and* $N = \begin{pmatrix} 4 & 5 & 1 & 0 \\ 5 & 6 & 2 & -1 \end{pmatrix}$. *In addition the vector of reduced costs is* $\bar{c}_B = 0$ *and* $\bar{c}_N^T = c_N^T - c_B^T B^{-1} N = 0 - (0, 1)N = (-5, -6, -2, 1)$. *Therefore, we start the iterations with the tableau below*

$$
\begin{array}{ll|c|ccccccc|c}
 & \textit{Basic variables} & z & x_1 & x_2 & x_3 & e_1 & e_2 & v & b \\
\hline
L_1 & e_1 & 0 & 4 & 5 & 1 & 1 & 0 & 0 & 23 \\
L_2 \longleftarrow & v & 0 & 5 & 6 & 2 & 0 & -1 & 1 & 28 \\
\hline
L_3 & & 1 & 5 & 6 & 2 & 0 & -1 & 0 & 28 \\
 & & & & \uparrow & & & & &
\end{array}
\tag{7.37}
$$

$$
\begin{array}{ll|c|ccccccc|c}
 & \textit{Basic variables} & z & x_1 & x_2 & x_3 & e_1 & e_2 & v & b \\
\hline
L_1 & x_2 & 0 & 4/5 & 1 & 1/5 & 1/5 & 0 & 0 & 23/5 \\
L_2 \longleftarrow & v & 0 & 1/5 & 0 & 4/5 & -6/5 & -1 & 1 & 2/5 \\
\hline
L_3 & & 1 & 1/5 & 0 & 4/5 & -6/5 & -1 & 0 & 2/5 \\
 & & & & & \uparrow & & & &
\end{array}
\tag{7.38}
$$

$$
\begin{array}{ll|c|ccccccc|c}
 & \textit{Basic variables} & z & x_1 & x_2 & x_3 & e_1 & e_2 & v & b \\
\hline
L_1 & x_2 & 0 & 3/4 & 1 & 0 & 1/2 & 1/4 & -1/4 & 9/2 \\
L_2 & x_3 & 0 & 1/4 & 0 & 1 & -3/2 & -5/4 & 5/4 & 1/2 \\
\hline
L_3 & & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0
\end{array}
\tag{7.39}
$$

Al the reduced costs are non-negative, hence we have an optimal solution for the auxiliary LPP.

Next we perform the Phase II. Using the last tableau above we have following basic solution for the LPP $x_B = (x_2, x_3)^T = (9/2, 1/2)^T$ *and* $x_D = (x_1, e_1, e_2)^T = 0$ *with* $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *and* $N = \begin{pmatrix} 3/4 & 1/2 & 1/4 \\ 1/4 & -3/2 & -5/4 \end{pmatrix}$, *vector of reduced costs* $\bar{c}$ *with* $\bar{c}_B = 0$ *and* $\bar{c}_N = (1/4, -5/2, -13/4)^T$, *and value of the objective function* $z_0 = 67/2$.

$$
\begin{array}{ll|c|cccccc|c}
 & \textit{Basic variables} & z & x_1 & x_2 & x_3 & e_1 & e_2 & b \\
\hline
L_1 & x_2 & 0 & 3/4 & 1 & 0 & 1/2 & 1/4 & 9/2 \\
L_2 \longleftarrow & x_3 & 0 & 1/4 & 0 & 1 & -3/2 & -5/4 & 1/2 \\
\hline
L_3 & & 1 & 1/4 & 0 & 0 & -5/2 & -13/4 & 67/2 \\
 & & & \uparrow & & & & &
\end{array}
\tag{7.40}
$$

	Basic variables	z	x_1	x_2	x_3	e_1	e_2	b
L_1	x_2	0	0	1	-3	5	$19/16$	3
L_2	x_1	0	1	0	4	-6	-5	2
L_3		1	0	0	$-1/4$	-1	$-37/4$	33

(7.41)

Al the reduced costs are non-negative, hence we have an optimal solution for the auxiliary LPP.

The Revised Simplex method: We can improve the complexity estimation, in terms of number of arithmetic operations for iteration, presented in the analysis of the naive implementation above, in several ways. The revised Simplex method, presented below, has a similar complexity, in terms of arithmetical operations, than the full tableau implementation of the Simplex, but it is more efficient in terms of data storage. Notice that what the Simplex method essentially do is to update the formulation (7.33) while the basis are changing along the iterations. Note that last three relations in (7.34) can be rewritten as

$$\pi^t = c_B^T B^{-1}, \quad b' = B^{-1}b, \quad z_0 = c_B^T b', \quad \bar{c}_D^T = c_D^T - \pi^T D. \tag{7.42}$$

Relations in (7.42) contain all the information required to verify the optimality of the given basis B. The Revised Simple method works with the relations above, updating the basic matrix along the iterations, working only with the current basis, the corresponding basic solution, and the vector that will enter this basis to form a new one, In contrast, the full tableau implementation works with the whole matrix defining the linear restrictions of the LPP being solved.

The Revised Simplex method (version 1):

1. We start with a basis consisting of the basic columns $A_{B(1)}, \ldots, A_{B(m)}$, the associated basic feasible solution x, and the basis matrix B.

2. Compute the vector of multipliers $\pi^T = c_B^T B^{-1}$ by solving the the liner system $\Leftrightarrow$ $\pi^T B = c_B^T$.

3. Compute the reduced costs $\bar{c}_l = c_l - \pi^T A_l$, for the nonbasic indices l. If they are all non-positive, the current basic feasible solution is optimal, and the algorithm terminates; else, choose some r for which $\bar{c}_r > 0$ (the variable x_r will enter the basis).

4. Compute $u = B^{-1}A_r$ by solving the liner system $Bu = A_r$. If no component of u is positive, the problem is unbounded and the algorithm terminates, i.e., the optimal cost is $+\infty$.

5. If some component of u is positive, let $x_B = B^{-1}b$ ($\Leftrightarrow \; Bx_B = b$) and

$$\theta_0 = \min \left\{ \frac{x_{B(i)}}{u_i} \; : \; u_i > 0 \right\}$$

6. Let s be such that $\theta_0 = x_{B(s)}/u_s$. Form a new basis $\bar{B}$ by replacing $A_{B(s)}$ with A_r. If x is the new basic feasible solution, the values of the new basic variables are $x_r = \theta_0$ and $x_{B(i)} = x_{B(i)} - \theta_0 u_i$ $i \neq s$.

Remark 7.32 *Notice that the Revised Simplex method presented above works, in each iteration, with the current basis, the corresponding basic solution, and the vector that will enter this basis to form a new one. Only this information is updated along the iterations, in contrast with the full tableau implementation that works with the whole matrix defining the linear restrictions of the LPP being solved (considering the reformulation (7.33)).*

Next we solve the same example presented above using the variant of the Revised Simplex Method presented above.

Example 7.33 *Solve the following LPP by the Revised Simplex method*

$$\begin{cases} \min\ z = 6x_1 + 7x_2 + 4x_3 \\ 4x_1 + 5x_2 + x_3 \leq 23 \\ 5x_1 + 6x_2 + 2x_3 \geq 28 \\ x_1 \geq 0,\ \ x_2 \geq 0\ \ x_3 \geq 0. \end{cases}$$

Introducing the slack and artificial variables we have the following reformulation of LPP above in the standard form.

$$\begin{cases} \min\ z = 6x_1 + 7x_2 + 4x_3 \\ 4x_1 + 5x_2 + x_3 + e_1 = 23 \\ 5x_1 + 6x_2 + 2x_3 - e_2 = 28 \\ x_1 \geq 0,\ \ x_2 \geq 0\ \ x_3 \geq 0\ \ e_1 \geq 0\ \ e_2 \geq 0. \end{cases}$$

Introducing the artificial variables we have the following auxiliar LPP for the Phase I

$$\begin{cases} \min\ z = v \\ 4x_1 + 5x_2 + x_3 + e_1 = 23 \\ 5x_1 + 6x_2 + 2x_3 - e_2 + v = 28 \\ x_1 \geq 0,\ \ x_2 \geq 0\ \ x_3 \geq 0\ \ e_1 \geq 0\ \ e_2 \geq 0\ \ v \geq 0. \end{cases}$$

Phase I: We start with the basic variables $x_B^t = (e_1, v)$ and the non-basic variables $x_N^T = (x_1, x_2, x_3, e_2) = (0,0,0,0)$. Hence, we have that $c_B^T = (0,1)$, $c_D^T = (0,0,0,0)$, $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $N = \begin{pmatrix} 4 & 5 & 1 & 0 \\ 5 & 6 & 2 & -1 \end{pmatrix}$, $x_B = (23, 28)$ and the value of the objective function is $z_0 = 28$.

i) First, we find the vector of multipliers π by solving the linear system $\pi^T B^{-1} = c_B^T$. It follows that $\pi = (0,1)^T$.

ii) Note that $\bar{c}_1 = 0 - (0,1)(4,5)^T = -5 < 0$, hence we choose x_1 to enter the basis.

iii) Compute $u = B^{-1}A_1$ by solving the linear system $Bu = A_1 = (4,5)^T$. It follows that $u = (4,5)^T$. The components of u are positive, then we calculate

$$\theta_0 = \min\left\{\frac{23}{4}, \frac{28}{5}\right\} = \frac{28}{5}.$$

Hence, we choose $(x_B)_2 = v$ to exit the basis.

iv) The new basis is $B = \begin{pmatrix} 1 & 4 \\ 0 & 5 \end{pmatrix}$ and we have that $N = \begin{pmatrix} 5 & 1 & 0 & 0 \\ 6 & 2 & -1 & 1 \end{pmatrix}$. The new feasible solution is $x = (x_B; x_N)$ with $x_B = (e_1, x_1)^T$ where $x_1 = \theta_0 = 28/5$ and $e_1 = e_1 - \theta_0 u_1 = 23 - (28/5)4 = 3/5$, and $x_N = (x_2, x_3, e_2, v)^T$. Note that value of the objective function is $z_0 = 0$ (since v is not in the basis). Clearly we have an optimal solution of the auxiliary LPP. So we can end Phase I and start Phase II. Notice that the vector of multipliers $\pi = (0,0)$ is the solution of the linear system $\pi^T B = c_B^T = (0,0)^T$. Hence the vector of reduced costs is $\bar{c}^T = c^T - \pi^T N = (0,0,0,0,1) - 0 = (0,0,0,0,1)$ has nonnegative components, thus we have indeed an optimal solution.

Next we perform the Phase II. The new basic variables are x_1 and e_1, the initial basis is $B = \begin{pmatrix} 4 & 1 \\ 5 & 0 \end{pmatrix}$ and we have that $N = \begin{pmatrix} 5 & 1 & 0 \\ 6 & 2 & -1 \end{pmatrix}$. The initial feasible solution is $x = (x_B; x_N)$ with $x_B = (x_1; e_1) = B^{-1}b = (28/5, 3/5)^T (\Leftrightarrow Bx_B = b \Leftrightarrow x_B)$ and $x_N^T = (x_2, x_3, e_2, v) = (0,0,0,0)$, $c_B = (6,0)^T$.

i) First, we find the vector of multipliers π by solving the linear system $\pi^T B = c_B^T$. It follows that $\pi = (0, 6/5)^T$.

ii) Note that $\bar{c}_1 = 7 - (0, 6/5)(5,6)^T = 7 - 36/5 = -1/5 < 0$, hence we choose x_2 to enter the basis.

iii) Compute $u = B^{-1}A_2$ by solving the linear system $Bu = A_2 = (5,6)^T$. It follows that $u = (6/5, 1/5)^T$. The components of u are positive, then we calculate

$$\theta_0 = \min\left\{\frac{28/5}{6/5}, \frac{3/5}{1/5}\right\} = \frac{3/5}{1/5} = 3.$$

Hence, we choose $(x_B)_2 = e_1$ to exit the basis.

iv) The new basic variables are x_1 and x_2, the new basis is $B = \begin{pmatrix} 4 & 5 \\ 5 & 6 \end{pmatrix}$ and we have that $N = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$. The new feasible basic solution is $x = (x_B; x_N)$ with $x_B = (x_1, x_2)^T$ where $x_2 = \theta_0 = 3$ and $x_1 = x_1 - \theta_0 u_1 = (28/5) - 3(6/5) = 2$ and $x_N^T = (x_3, e_1, e_2) = (0,0,0)$.

v) We find the vector of multipliers π by solving the linear system $\pi^T B = c_B^T = (6,7)$. It follows that $\pi = (-1, 2)^T$.

vi) The vector of reduced costs is

$$\bar{c}_B = (0,0)^T, \quad \bar{c}_N^T = c_N^T - \pi^T N = (4,0,0) - (-1,2)\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} = (1,1,2),$$

and, since all its components are nonnegative components, then we have an optimal solution.

Remark 7.34 *Notice that, in steps 2) and 3) of the implementation of the revised method, presented above: the calculation of $p = c_B^T B^{-1}$ and $u = B^{-1}A_j$ requires $\mathcal{O}(m^3)$ arithmetic operations Also, the calculation of the reduced costs for the $(n-m)$ non basic variables requires $\mathcal{O}(mn)$ arithmetic operations; and, finally, updating the inverse of the matrix of the basis requires $\mathcal{O}(m^2)$ arithmetic operations. Hence, the cost of the implementation of revised variant of the Simplex requires $\mathcal{O}(m^3 + mn)$ arithmetic operations, which coincides with the estimation for the naive implementation showed above. These estimations can be improved if we have the matrix B^{-1} at our disposal, in each iteration. In this case we can reduce the operational cost of calculating p and u as above to $\mathcal{O}(m^2)$ arithmetic operations and the cost of the whole implementation to $\mathcal{O}(m^2 + mn)$ arithmetic operations.*

In view of the above remark, we could improve the complexity estimations of the naive implementation of the Simplex method and of the first variant of the Revised Simplex method presented above by making available at the beginning of the iteration the matrix B^{-1} and then devising an efficient way of calculating $\bar{B}^{-1}$ where $\bar{B}$ denotes the new base. In our case, this means calculating $\bar{B}^{-1}$ using at most $\mathcal{O}(m^2)$ arithmetic operations. Next we will present such a way or making the calculation of B^{-1}. Let

$$B = \left[A_{B(1)} \dots A_{B(m)}\right]$$

be the basis at the beginning of the iteration and let

$$\bar{B} = \left[A_{B(1)} \dots A_{B(l-1)} A_{B(j)} A_{B(l+1)} \dots A_{B(m)}\right]$$

the basis that results from x_j replacing x_l in the basis, at the end of the iteration. Notice that this implies that for $u = B^{-1}A_j$ it holds $u_l > 0$. Now, assuming that we have B^{-1}, we will present an efficient way of calculating $\bar{B}^{-1}$.

Definition 7.35 *Given a matrix, not necessarily square, the operation of adding a constant multiple of one row to the same row or to another row is called an elementary row operation.*

Lemma 7.36 *The elementary row operation of multiplying the ith row of a matrix A by a constant β and adding the result to the j-th row of A can be obtained by multiplying the matrix A by the nonsingulat matrix $Q_{ij} = I + D_{ij}$ where D_{ij} is the matrix with all entries equal to zero except the (i,j)-th entry which is equal to β.*

First, note that the determinant of Q_{ij} is equal to one, which proves that it is a nonsingular matrix. The proof of the second part of the claim follows straightforwardly from direct calculations, which ends the proof of the lemma.

Remark 7.37 *It follows trivially that any sequence of elementary operations applied to a matrix can be obtained by multiplying this matrix by the nonsingular matrix equal to the product of all matrices representing each elementary operation, in the same order that these operations are applied. Also, notice that performing one elementary operation on a $k \times p$ matrix requires, at most, $2p$ arithmetic operations.*

Under the initial assumptions, simple calculations show that

$$B^{-1}\bar{B} = [e_1|\ldots|e_{l-1}|\,u\,|e_{l+1}|\ldots|e_m]$$

where $\{e_1, e_2, \ldots, e_m\}$ denotes the canonical base of $\mathbb{R}^m$ and $u = B^{-1}A_j$ with $u_l > 0$. Now let Q be the nonsingular matrix corresponding to the sequence of elementary operations which transforms the vector u in the vector e_l. Note that that we can obtain this transformation performing m elementary operations on u as follows: first we multiply the l-th row of u by $1/u_l$ and then we add the l-th row of u multiplied by $-u_j$ to the j-th row of u, for all $j \neq i$. Now it follows

$$\begin{aligned} QB^{-1}\bar{B} &= Q\left([e_1|\ldots|e_{l-1}|\,u\,|e_{l+1}|\ldots|e_m]\right) = \left([Qe_1|\ldots|Qe_{l-1}|\,Qu\,|Qe_{l+1}|\ldots|Qe_m]\right) \\ &= \left([e_1|\ldots|e_{l-1}|\,e_l\,|e_{l+1}|\ldots|e_m]\right), \end{aligned}$$

since, the definition of Q and simple calculations show that $Qe_k = e_k$ for all $k \neq l$. The above relation implies that $QB^{-1} = \bar{B}^{-1}$ which means that applying the sequence of m elementary operations that defined Q to the matrix B^{-1} allows to compute $\bar{B}^{-1}$. Hence, given B^{-1}, we need $\mathcal{O}(m(2m)) = \mathcal{O}\left(m^2\right)$ to compute $\bar{B}^{-1}$. In practical terms we can use the following procedure

$$\left[B^{-1}\,|\,B^{-1}A_s\right] \underset{\text{elementary operations}}{\longrightarrow} \left[\bar{B}^{-1}\,|\,e_r\right]$$

where e_r denotes the r-th vector of the canonical basis, the elementary operations transform $B^{-1}A_s$ in e_r, and s is the index of the variable that enters the basis and r is the index of the variable that exits the basis.

The Revised Simplex method (version 2):

1. We start with a basis B consisting of the basic columns $A_{B(1)}, \ldots, A_{B(m)}$, the associated basic feasible solution x, and the inverse B^{-1} of the basis matrix.

2. Compute the vector of multipliers $\pi^T = c_B^T B^{-1}$.

3. Compute the reduced costs $\bar{c}_l = c_l - \pi^T A_l$, for a nonbasic indices l. If they are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates; else, choose some j for which $\bar{c}_j < 0$ (the variable x_j will enter the basis).

4. Compute $u = B^{-1}A_j$, If no component of u is positive, the problem is unbounded and the algorithm terminates, i.e., the optimal cost is $-\infty$.

5. If some component of u is positive, let $x_B = B^{-1}b$ and

$$\theta_0 = \min\left\{\frac{x_{B(i)}}{u_i}\ :\ u_i > 0\right\}$$

6. Let l be such that $\bar{\theta} = x_{B(l)}/u_l$. Form a new basis $\bar{B}$ by replacing $A_{B(l)}$ with A_j. If y is the new basic feasible solution, the values of the new basic variables are $y_j = \theta_0$ and $y_{B(i)} = x_{B(i)} - \theta_0 u_i \; i \neq l$.

7. Form the $m \times (m+1)$ matrix $[B^{-1}|u]$. Apply the elementary operations previously describe to transform the vector u in the vector e_l. The resulting matrix is $\left[\bar{B}^{-1} \,|\, e_l\right]$.

Next we solve the same example presented above using the second variant of the Revised Simplex Method presented above.

Example 7.38 *Solve the following LPP by the Revised Simplex method*

$$\begin{cases} \min\ z = 6x_1 + 7x_2 + 4x_3 \\ \begin{array}{ccccccc} 4x_1 & + & 5x_2 & + & x_3 & \leq & 23 \\ 5x_1 & + & 6x_2 & + & 2x_3 & \geq & 28 \end{array} \\ x_1 \geq 0, \;\; x_2 \geq 0 \;\; x_3 \geq 0. \end{cases}$$

Introducing the slack and artificial variables we have the following reformulation of LPP above in the standard form.

$$\begin{cases} \min\ z = v \\ \begin{array}{ccccccccc} 4x_1 & + & 5x_2 & + & x_3 & + & e_1 & = & 23 \\ 5x_1 & + & 6x_2 & + & 2x_3 & - & e_2 & = & 28 \end{array} \\ x_1 \geq 0, \;\; x_2 \geq 0 \;\; x_3 \geq 0 \;\; e_1 \geq 0 \;\; e_2 \geq 0. \end{cases}$$

Introducing the artificial variables we have the following auxiliary LPP for the Phase I

$$\begin{cases} \min\ z = v \\ \begin{array}{ccccccccccc} 4x_1 & + & 5x_2 & + & x_3 & + & e_1 & & & = & 23 \\ 5x_1 & + & 6x_2 & + & 2x_3 & - & e_2 & + & v & = & 28 \end{array} \\ x_1 \geq 0, \;\; x_2 \geq 0 \;\; x_3 \geq 0 \;\; e_1 \geq 0 \;\; e_2 \geq 0 \;\; v \geq 0. \end{cases}$$

Phase I: We start with the basic variables $x_B = (e_1, v)$ *and the non-basic variables* $x_N = (x_1, x_2, x_3, e_2) = (0,0,0,0)$. *Hence, we have that* $c_B = (0,1)$, $c_N = (0,0,0,0)$, $B = B^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ *and* $N = \begin{pmatrix} 4 & 5 & 1 & 0 \\ 5 & 6 & 2 & -1 \end{pmatrix}$, $x_B = (23, 28)$ *and the value of the objective function is* $z_0 = 28$.

i) First, we find the vector of multipliers π *by solving the linear system* $\pi^T B = c_B^T = (0,1)$. *It follows that* $\pi^T = (0,1)$.

ii) Note that $\bar{c}_1 = c_1 - \pi^T N_1 = 0 - (0,1)(4,5)^T = -5 < 0$, *hence we choose* x_1 *to enter the basis.*

iii) Compute

$$u = B^{-1} A_1 = A_1 = \begin{pmatrix} 4 \\ 5 \end{pmatrix}.$$

The components of u are positive, then we calculate

$$\theta_0 = \min\left\{\frac{23}{4}, \frac{28}{5}\right\} = \frac{28}{5}.$$

Hence, we choose $(x_B)_2 = v$ to exit the basis.

iv) The new feasible basic solution is $x = (x_B, x_D)$ with $x_B = (e_1, x_1)$ and $x_D = (x_2, x_3, e_2, v)$. The new basis is $\bar{B} = \begin{pmatrix} 1 & 4 \\ 0 & 5 \end{pmatrix}$ and we have that $\bar{D} = \begin{pmatrix} 5 & 1 & 0 \\ 6 & 2 & -1 \end{pmatrix}$. To calculate $\bar{B}^{-1}$ we do

$$\left[B^{-1} \,|\, A_6\right] = \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \mid \begin{pmatrix} 4 \\ 5 \end{pmatrix}\right] \underset{\textit{elementary operations}}{\longrightarrow} \left[\begin{pmatrix} 1 & -4/5 \\ 0 & 1/5 \end{pmatrix} \mid \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \left[\bar{B}^{-1} \,|\, e_2\right]$$

Note that value of the objective function is $z_0 = 0$ (since v is not in the basis). Clearly we have an optimal solution of the auxiliary LPP. So we can end Phase I and start Phase II.

Next we perform the Phase II. The new basic variables are x_1 and e_1 (in that order, notice that we have reordered the basic variables), the initial basis is $B = \begin{pmatrix} 4 & 1 \\ 5 & 0 \end{pmatrix}$ and we have that $D = \begin{pmatrix} 5 & 1 & 0 \\ 6 & 2 & -1 \end{pmatrix}$ and $B^{-1} = \begin{pmatrix} 0 & 1/5 \\ 1 & -4/5 \end{pmatrix}$ (after reordering the variables is easy to obtain this later matrix by reordering the elements of the matrix B^{-1} calculated at the end of Phase I). The initial feasible solution is $x = (x_B, x_D)$ with

$$x_B^T = (e_1, x_1)^T = B^{-1}b = \begin{pmatrix} 0 & 1/5 \\ 1 & -4/5 \end{pmatrix} \begin{pmatrix} 23 \\ 28 \end{pmatrix} = (28/5, 3/5)^T$$

i) Now, we find the vector of multipliers

$$\pi^T = c_B^T B^{-1} = (0, 6) \begin{pmatrix} 0 & 1/5 \\ 1 & -4/5 \end{pmatrix} = (0, 6).$$

ii) Note that $\bar{c}_1 = c_1 - \pi^T D_1 = 7 - (0,6)(5,6)^T = 7 - 36 = -29 < 0$, hence we choose x_2 to enter the basis.

iii) Compute

$$u = B^{-1}A_2 = \begin{pmatrix} 0 & 1/5 \\ 1 & -4/5 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 6/5 \\ 1/5 \end{pmatrix}$$

The components of u are positive, then we calculate

$$\theta_0 = \min\left\{\frac{28/5}{6/5}, \frac{3/5}{1/5}\right\} = \frac{3/5}{1/5} = 3.$$

Hence, we choose $(x_B)_2 = e_1$ to exit the basis.

iv) *The new basic variables are* x_1 *and* x_2*, the new basis is* $\bar{B} = \begin{pmatrix} 4 & 5 \\ 5 & 6 \end{pmatrix}$ *and we have that* $\bar{N} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$. *To calculate* $\bar{B}^{-1}$ *we do*

$$\left[B^{-1} \,|\, A_2\right] = \left[\begin{pmatrix} 0 & 1/5 \\ 1 & -4/5 \end{pmatrix} \,|\, \begin{pmatrix} 6/5 \\ 1/5 \end{pmatrix}\right] \underset{\textit{elementary operations}}{\longrightarrow} \left[\begin{pmatrix} -6 & 5 \\ 5 & -4 \end{pmatrix} \,|\, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] = \left[\bar{B}^{-1} \,|\, e_2\right]$$

The new feasible solution is $x = (x_B, x_D)$ *with*

$$x_B = (e_1, x_1)^T = B^{-1}b = \begin{pmatrix} -6 & 5 \\ 5 & -4 \end{pmatrix} \begin{pmatrix} 23 \\ 28 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

and $x_N = (x_2, x_3, e_2)^T = (0, 0, 0)^T$.

v) *We find the vector of multipliers*

$$\pi^T = c_B^T B^{-1} = \begin{pmatrix} -6 & 5 \\ 5 & -4 \end{pmatrix} \begin{pmatrix} 6 \\ 7 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

vi) *The vector of reduced costs is*

$$\bar{c}_B^T = (0, 0), \quad \bar{c}^T = c^T - \pi^T D = (4, 0, 0) - (-1, 2) \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} = (1, 1, 2),$$

and, since all its components are nonnegative components, then we have an optimal solution.

Remark 7.39

i) *Another option to improve the complexity estimations of the naive implementation of the Simplex method is to calculate the expressions of the form* $B^{-1}d$ *by solving the related system* $Bu = d$ *in a more efficient way. Notice that, given* LU *decomposition of the matrix* B*, solving the linear system above would require only* $\mathcal{O}(m^2)$ *arithmetic operations. A more efficient procedure would be the updating of the* LU *decomposition of the basis matrix to obtain a similar decomposition for the new basis matrix.*

ii) *In some cases, in particular when* n *is much larger than* m*, this new variant could provide some improvement over the tableau implementation previously discussed. Notice that when* n *is much larger than* m*, if we calculate the reduced costs of the non-basic variable sequentially and stop when we find the first negative value (for the minimization case) we could reduce, in practice, the term* mn *in the complexity estimation of Remark 7.34 to something smaller, for example,* m *multiplied by some fraction of* n*. The full tableau implementation of the Simplex method always calculates the whole vector of reduced costs.*

7.4.2 Viewing the Simplex method as a descent method

Definition 7.40 *Let $x = (x_B, x_N) = (x_B, 0)$ the basic solution associated to B. For $j \notin B$, the jth basic direction at x is the vector d defined by letting $d_j = 1$, $d_i = 0$ for $i \notin \{1, 2, \dots, m\}$ and $i \neq j$, and $d_B = -B^{-1}A_j$.*

Definition 7.41 *Let x be feasible for the linear system in (7.17). A vector $d \in \mathbb{R}^n$ is said to be a feasible direction at x, if there exists a positive scalar θ_0 for which $x + \theta d$ is feasible for the linear system in (7.17) for all $\theta \in [0, \theta_0]$.*

Remark 7.42 *Let d be the jth basic direction at x, θ_0 defined as in (7.43). For $\theta \in (0, \theta_0]$, let $x^\theta = x + \theta d = (x_B + \theta d_B, \theta d_D)$, which is feasible for the system in (7.17). Following the notations of (7.30), we have that*

$$c^T x^\theta = z_0 + \hat{c}_D^\top x_D^\theta = z_0 + (c_D - \left(B^{-1}D\right)^T c_B)(\theta d_D) = z_0 + \theta \left(c - \left(\left(B^{-1}A\right)^T c_B\right)\right)_j = z_0 + \theta \bar{c}_j$$

Hence, if the reduced cost corresponding to the non basic variable x_j is negative, we can decrease the objective function by moving along the jth basic direction at x. In particular, in this case, the basic feasible solution defined in Lemma 7.44, obtained from the basic solution x, strictly decreases the value of the objective function.

Lemma 7.43 *Let $x = (x_B, x_N) = (x_B, 0)$ be the basic solution associated to B and d be the jth basic direction at x. If x is nondegenerated, then d is a feasible direction at x.*

Simple calculations show that $Ad = Bd_B + Nd_N = B(-B^{-1}A_j) + A_j = 0$. Hence $A(x+\theta d) = Ax = b$ for all $\theta \in \mathbb{R}$. In addition, note that, for all $\theta > 0$,

$$(x + \theta d)_j = x_j + \theta d_j = \theta > 0, \quad \text{and} \quad (x + \theta d)_i = x_i = 0 \;\forall\, i \notin \{1, 2, \dots, m\} \text{ and } i \neq j.$$

Also, since $x_B > 0$, it holds, for $i \in B$, that

$$(x + \theta d)_i > 0 \text{ if } d_i \geq 0,$$

and

$$(x + \theta d)_i = x_i + \theta d_i \geq 0 \;\forall\, \theta \in [0, \theta_0] \quad \text{otherwise}$$

where

$$\theta_0 = \min \left\{ -\frac{x_l}{d_l} \;:\; d_l < 0,\, l \in B \right\} > 0. \tag{7.43}$$

In summary, we have proved that $x+\theta d$ is feasible for the linear system in (7.17) for $\theta \in [0, \theta_0]$, which proves the claim.

Lemma 7.44 *Let x be the basic solution associated to B. Assume that x is non degenerated and that for some $j \in B$ it holds that $\min\{d_i \; i = 1, \dots, m\} < 0$ where $d = B^{-1}A_j$. Let $i_0 \in \bar{x} = x + \theta_0 d$ where $\theta_0 = \min\{-x_i/d_i : d_i < 0, i \in B\}$. Then, $\bar{x}$ is a feasible basic solution to the system in (7.17)*

From the proof of Lemma 7.43, it follows that $\bar{x}$ is feasible for the system in (7.17). In addition, taking $i \in B$ such that $\theta_0 = -x_i/d_i$, it follows that $\hat{x}_i = 0$. Consider now the sets of indexes $\bar{B} = B \cup \{j\} - \{i\}$. We claim that the associated matrix $\bar{B} = \left[A_{\bar{B}(1)} \dots A_{\bar{B}(m)}\right]$ is nonsingular. Notice that $d = B^{-1}A_j$ is the vector of coordinates of A_j in the basis B. We have that $A_j = \sum_{l \in B} d_l A_l$ where A_k denotes the k-th column vector of A for all $k \in \{1, 2, \dots, m\}$; and, since $d_i \neq 0$, it follows that

$$A_i = \sum_{l \in B,\, l \neq i} \frac{-d_l}{d_i} A_l + \frac{1}{d_i} A_j$$

Now, take any $w \in \mathbb{R}^n$ and let $w = \sum_{l \in B} \alpha_l A_l$. From the above relation it follows that

$$A_i = \sum_{l \in B,\, l \neq i} \frac{-d_l}{d_i} A_l + \frac{1}{d_i} A_j$$

which implies that

$$w = \sum_{l \in B,\, l \neq i} \alpha_l A_l + \alpha_i \left(\sum_{l \in B\, l \neq i} \frac{-d_l}{d_i} A_l + \frac{-1}{d_i} A_j \right) = \left(\frac{-\alpha_i}{d_i} \right) A_j + \sum_{l \in B\, l \neq i} \left(\alpha_l - \frac{d_l}{d_i} \right) A_l$$

The above relation shows how to write any vector $w \in \mathbb{R}^n$ as a linear combination of the column vectors of $\bar{B}$, which, in particular, implies the claim and ends the proof of the lemma.

Remark 7.45 *The process described in the lemma above will be called variable x_j entering the basis.*

7.5 Duality and sensitivity analysis

Recall the primal-dual pairs of LPPs (see Remark 7.20) and the duality Theorem 7.14.

Interpretation of the dual variables. We consider the primal problem of profit maximization

$$\max\ c^\top x, \quad Cx \leq d,\ x \geq 0 \tag{7.44}$$

where d_i represents quantity of resource$i$ available. The dual problem is written

$$\min\ \mu^\top d, \quad C^\top \mu \geq c,\ \mu \geq 0. \tag{7.45}$$

We consider a perturbation of the amount of available resource in (7.44):

$$\max\ c^\top x, \quad Cx \leq d + \varepsilon,\ x \geq 0. \tag{7.46}$$

and the corresponding dual problem

$$\min\ \mu^\top (d + \varepsilon), \quad C^\top \mu \geq c,\ \mu \geq 0. \tag{7.47}$$

We assume that the solutions to the initial dual problem (7.45) are also solutions to the perturbed dual problem (7.47). In this case, denoting by μ^* a solution of the dual problem (7.45), the optimal value of the perturbed problem (7.46) is the optimal value of the initial problem plus $d^\top \mu^*$. In other words, μ_i^* represents the increase in profit when the quantity of resource i increases by one unit. This interpretation is valid for perturbations given by d_i. A solution of the dual problem (graphically or using simplex) allows us to determine for which values of the perturbations the set of solutions of the dual problem does not change.

Example 7.46 *We want to produce strawberry and rose jelly. The amount of sugar available is 8 tons and the capital is 40,000 reais. Strawberries are bought for 2 reais per kg and roses for 15 reais per kg. Strawberry jam (resp. rose) is sold for 4.50 reais per kg (resp. 12.60 reais per kg). Strawberry jam (resp. rose) is obtained by mixing 50% strawberries (resp. 40% roses) and 50% (resp. 60%) sugar. Model the problem and write the dual problem. Provide an interpretation of the dual solution and study the variation in profit when the amount of available sugar changes.*

Model the problem and write the dual problem. Provide an interpretation of the dual solution and study the variation in profit when the amount of available sugar changes.

Let x_1 and x_2 be the quantities of sugar destined respectively for the production of strawberry and rose jelly. The problem of maximizing the profit is modeled by the LPP

$$\left\{\begin{array}{llll} \max\ 7x_1 + 11x_2 & & & \\ x_1 \quad + \quad x_2 & \leq & 8000 \\ 2x_1 \quad + \quad 10x_2 & \leq & 40000 \\ x_1 \geq 0, \quad x_2 \geq 0. & & \end{array}\right.$$

If the amount of sugar were $8000 + \varepsilon$ the problem to be solved would be

$$\left\{\begin{array}{lll} \max\ 7x_1 + 11x_2 & & \\ x_1 \quad + \quad x_2 & \leq & 8000 + \varepsilon \\ 2x_1 \quad + \quad 10x_2 & \leq & 40000 \\ x_1 \geq 0, \quad x_2 \geq 0. & & \end{array}\right. \tag{7.48}$$

Let's find a solution to the perturbed problem (7.48) graphically and using the Simplex algorithm. We write the (7.48) problem in standard form:

$$\left\{\begin{array}{lll} \max\ 7x_1 + 11x_2 & & \\ x_1 \quad + \quad x_2 \quad + \quad e_1 & = & 8000 + \varepsilon \\ 2x_1 \quad + \quad 10x_2 \quad + \quad e_2 & = & 40000 \\ x_1 \geq 0,\ x_2 \geq 0,\ e_1 \geq 0,\ e_2 \geq 0. & & \end{array}\right.$$

We do not seed Phase I. We initiate Phase II with $x_B = \{e_1, e_2\}$:

		z	x_1	x_2	e_1	e_2	
	e_1	0	1	1	1	0	$8000 + \varepsilon$
$\leftarrow$	e_2	0	2	$\underline{10}$	0	1	40000
	z	1	-7	-11	0	0	0
				$\uparrow$			

Variable x_2 enter the basis and e_2 leaves the basis if $4000 \leq 8000 + \varepsilon$, i.e., if $\varepsilon \geq -4000$. In this case, in the first iteration we have

		z	x_1	x_2	e_1	e_2	
$\leftarrow$	e_1	0	$\underline{0.8}$	0	1	-0.1	$4000 + \varepsilon$
	x_2	0	0.2	1	0	0.1	4000
	z	1	-4.8	0	0	1.1	44000
			$\uparrow$				

In the second iteration, x_1 enter the basis and e_1 leaves the basis if $1.25(4000+\varepsilon) \leq 5 \times 4000$, i.e., if $\varepsilon \leq 12000$:

	z	x_1	x_2	e_1	e_2	
x_1	0	1	0	1.25	-0.125	$5000 + 1.25\varepsilon$
x_2	0	0	1	-0.25	0.125	$3000 - 0.25\varepsilon$
z	1	0	0	6	0.5	$68000 + 6\varepsilon$

We have found an optimal solution : $x_1 = 5000 + 1.25\varepsilon$ kg of sugar for the strawberry jelly, $x_2 = 3000 - 0.25\varepsilon$ kg of sugar for the rose jelly, and $x_3 = x_4 = 0$. The optimal value is $68000 + 6\varepsilon$. We can verify graphically this result

The dual problem of the perturbed problem (7.48) is

$$\begin{cases} \min\ (8000+\varepsilon)\mu_1 + 40000\mu_2 \\ \mu_1 + 2\mu_2 \geq 7 \\ \mu_1 + 10\mu_2 \geq 11 \\ \mu_1 \geq 0,\ \ \mu_2 \geq 0. \end{cases}$$

We can check graphically or using the simplex algorithm that the optimal value of this perturbed dual problem is

- $\mu_1^* = 6$ e $\mu_2^* = 0.5$ para $-4000 \leq \varepsilon \leq 12000$;
- $\mu_1^* = 0$ e $\mu_2^* = 3.5$ para $\varepsilon > 12000$;
- $\mu_1^* = 11$ e $\mu_2^* = 0$ para $\varepsilon < -4000$.

The dual variable μ_1^* represents the increase in the profit when the quantity of sugar increases by 1kg. Consequently:

- *If the initial amount of sugar is between* $8,000 - 4,000 = 4,000$ *and* $8,000 + 12,000 = 20,000$, $\mu_1^* = 6$, *i.e., the increase in profit for each kg of sugar added is 6 reais. Then it is interesting to buy sugar until the amount of 20,000kg of sugar is reached if the purchase price of the sugar is less than 6 reais per kg. Any amount of sugar added above 20,000 kg does not produce any additional profit, since in this case* $\mu_1^* = 0$ *(with the available capital we cannot purchase more than 20,000 kg of strawberries).*

- *If the initial amount of sugar is less than* $8,000 - 4,000 = 4,000$, *each kg of sugar added until the total amount reaches* 4000 *kg can be used to make rose jelly and produces a profit increase of 11 reais. So it is interesting buy sugar if the price of sugar is less than 11 reais per kg.*

8 Column generation for LP

Consider the linear program (P) and its dual (D) given by

$$(P)\left\{\begin{array}{l}\max c^T x\\ Ax=b, x\geq 0\end{array}\right. \quad (D)\left\{\begin{array}{l}\min \pi^T b\\ A^T\pi\geq c.\end{array}\right.$$

Assume that the primal problem is bounded from below and feasible. Then the optimal values of (P) and (D) are equal. Let f_* be this optimal value.

The column generation algorithm works as follows. At iteration $k \geq 1$, we select a set S_k of columns of A and denote by x_k the vector whose variables have indexes in $\mathcal{S}_k$. For $k=1$, we can take a known feasible solution of $Ax=b, x\geq 0$ (for instance given by a heuristic applied to (P) in several applications of column generation) and take the set S_1 to be the set of columns of A corresponding to positive entries of this solution. We denote by A_k (resp. c_k) the submatrix of A (resp. subvector of c) whose columns (resp. entries) are indexed by S_k. We solve the following relaxed problem P_k along with its dual D_k given by

$$(P_k)\left\{\begin{array}{l}\max c_k^T x_k\\ A_k x_k=b, x_k\geq 0\end{array}\right. \quad (D_k)\left\{\begin{array}{l}\min \pi^T b\\ A_k^T\pi\geq c_k.\end{array}\right.$$

The common (by the duality theorem for bounded feasible LPs) optimal value f_k of (P_k) and (D_k) satisfies $f_k \leq f_*$ (any feasible solution of (P_k) can be extended, by completing $x_i=0$ for $i\notin S_k$, to a feasible solution of (P) with the same value of the objective function). Therefore, if π_k is an optimal solution of (D_k) which is feasible for (D), i.e., $A^T\pi\geq c$, then we have both $\pi_k^T b\geq f_*$ (by feasibility of π_k for (D)) and $\pi_k^T b = f_k \leq f_*$, which shows that $\pi_k^T b = f_k = f_*$, and, therefore, that π_k is optimal for (D). It also follows that completing x_k to a feasible solution of D, x, in the manner described above, we obtain an optimal solution to (P) since x would be feasible for (P) and we would have that $c^T x = c_k^T x_k = f_k = f_*$. On the contrary, if π_k is not feasible for (D), which means that, for some $i\notin S_k$, it holds $c(i) - a_i^T\pi_k > 0$, where a_i is ith column of A, then we add one such column a_i to S_k to generate S_{k+1} (for instance the column providing the largest reduced cost) and continue the process, solving the relaxed problems (P_{k+1}) and (D_{k+1}). If A has a finite number of columns, this process stops in a finite number of iterations finding an optimal solution to (P). In practice, column generation is useful for problems with a large number of variables and such that a large number of variables are null at an optimum. Observe that since (P_1) is feasible and the feasible set of (P_k) contains the feasible set of (P_1) all problems (P_k) are bounded and feasible.

9 Branch & Bound

Solve the ILPP

$$\left\{\begin{array}{l}\max\ 8x_1+11x_2+6x_3+4x_4\\ 5x_1+7x_2+4x_3+3x_4\leq 14\\ x_i\in\{0,1\}\ i=1,\ldots,4.\end{array}\right.$$

1. Set $L = -\infty$. In N_1 we solve

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \le 14 \\ 0 \le x_i \le 1,\ i = 1, \dots, 4. \end{cases}$$

   We have $z^* = 22$, $x_1^* = 1$, $x_2^* = 1$, $x_3^* = 0.5$, $x_4^* = 0$. The solution is not feasible for the ILPP and $E[z^*] > L$, then N_1 is active.

2. In N_1, we choose the variable with the largest fractional for "branching": x_3.

   (2.a) In N_2 we solve

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \le 14 \\ 0 \le x_i \le 1,\ i \ne 3 \\ x_3 = 0. \end{cases}$$

   We have $z^* = 21.67$, $x_1^* = 1$, $x_2^* = 1$, $x_3^* = 0.5$, $x_4^* = 0.67$. The solution is not feasible for the ILPP and $E[z^*] > L$, then N_2 is active.

   (2.b) In N_3 we solve

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \le 14 \\ 0 \le x_i \le 1,\ i \ne 3 \\ x_3 = 1. \end{cases}$$

   We have $z^* = 21.86$, $x_1^* = 1$, $x_2^* = 0.71$, $x_3^* = 1$, $x_4^* = 0$. The solution is not feasible for the ILPP and $E[z^*] > L$, then N_3 is active

3. We choose an active leaf with the largest value z^*: N_3.

4. For N_3, we choose the variable with the largest fractional for "branching": x_2.

   (4.a) In N_4 we solve

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \le 14 \\ 0 \le x_1 \le 1,\ 1 \le x_4 \le 1 \\ x_3 = 1,\ x_2 = 0. \end{cases}$$

   We have $z^* = 18$, $x_1^* = 1$, $x_2^* = 0$, $x_3^* = 1$, $x_4^* = 1$. The solution is integer and $z^* = 18 > L$, then $L \leftarrow z^*$, i.e. $L = 18$. The solution is feasible for the ILPP, then N_4 is inactive.

   (4.b) In N_5 we solve

$$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \le 14 \\ 0 \le x_1 \le 1,\ 1 \le x_4 \le 1 \\ x_3 = 1,\ x_2 = 1. \end{cases}$$

   We have $z^* = 21.8$, $x_1^* = 0.6$, $x_2^* = 1$, $x_3^* = 1$, $x_4^* = 0$. The solution is not feasible for the ILPP and $L < E[z^*]$, then N_5 is active.

5. We choose an active leaf with the largest value z^*: N_5.

6. For N_5, we choose the variable with the largest fractional for "branching": x_1.

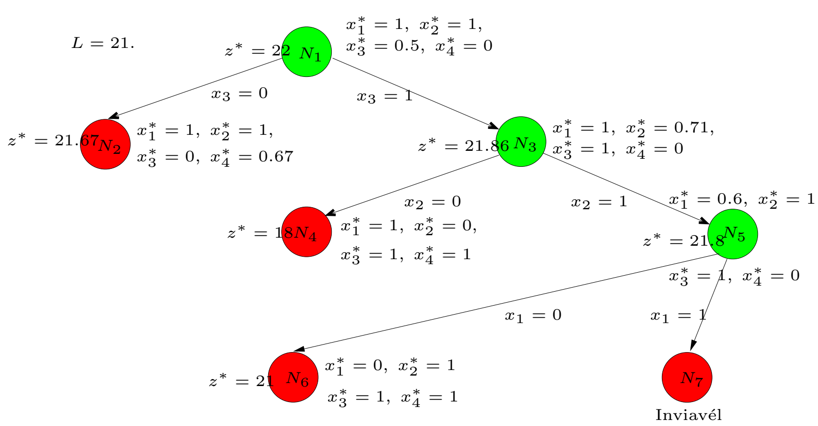
   (6.a) In N_6 we solve
   $$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14 \\ 0 \leq x_1 \leq 1,\ 1 \leq x_4 \leq 1 \\ x_3 = 1,\ x_2 = 1. \end{cases}$$
   We have $z^* = 21$, $x_1^* = 0$, $x_2^* = 1$, $x_3^* = 1$, $x_4^* = 1$. The solution is integer and $z^* = 21 > L$, then $L \leftarrow z^*$, i.e. $L = 21$. The solution is feasible for the ILPP, then N_6 is inactive.

   (6.b) In N_7 we solve
   $$\begin{cases} \max\ 8x_1 + 11x_2 + 6x_3 + 4x_4 \\ 5x_1 + 7x_2 + 4x_3 + 3x_4 \leq 14 \\ 1 \leq x_4 \leq 1\ x_1 = 1, \\ x_3 = 1,\ x_2 = 1. \end{cases}$$
   The problem is infeasible, then N_7 is inactive.

7. We choose an active leaf with the largest value z^*: N_2.

8. For N_2, $E[z^*] < L$, then N_2 is inactive.

9. All leaves are inactive. The algorithm stops. The optimal solution is $x_1^* = 0$, $x_2^* = 1$, $x_3^* = 1$, $x_4^* = 1$ with $z^* = 21$.

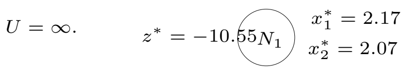


We do not have active leaves: we have found an optimal solution

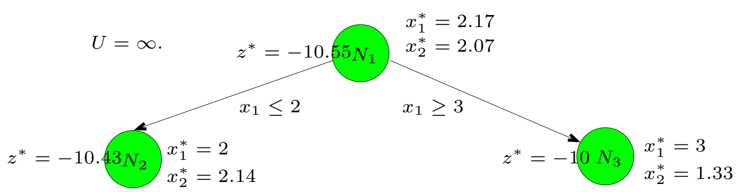
9.1 Example II

We will consider the minimization problem

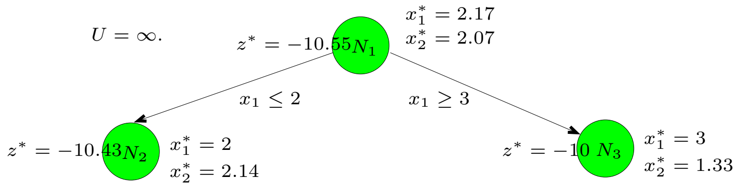
$$\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ x_1,\ x_2 \in \mathbb{Z}_+. \end{cases}$$



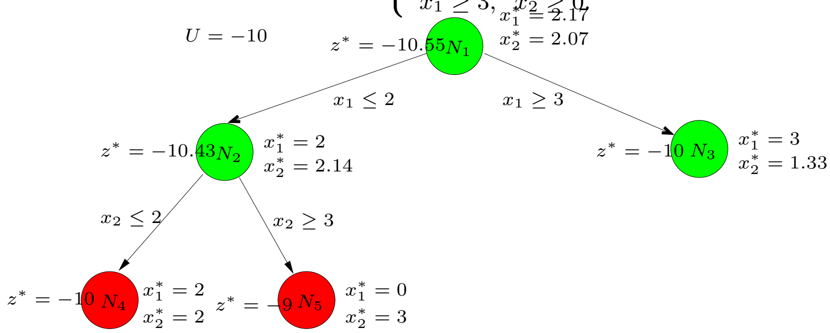
Em N_1, resolvemos $\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ x_1, x_2 \geq 0. \end{cases}$



Em N_2, resolvemos $\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ 0 \leq x_1 \leq 2,\ x_2 \geq 0. \end{cases}$



Em N_3, resolvemos $\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ x_1 \geq 3,\ \ x_2 \geq 0, \end{cases}$



Em N_4, resolvemos $\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ 0 \leq x_1 \leq 2,\ 0 \leq x_2 \leq 2. \end{cases}$

10 Large scale linear optimization

10.1 The Representation Theorem

Definition 10.1 *Consider a polyhedron P defined by linear equalities and inequality constraints, and let x^* be an element of $\mathbb{R}^n$.*

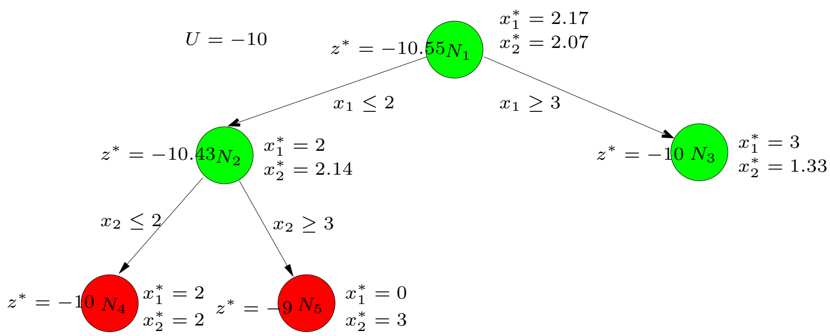
(a) The vector x^* is a basic solution if:

(i) All equality constraints are active;

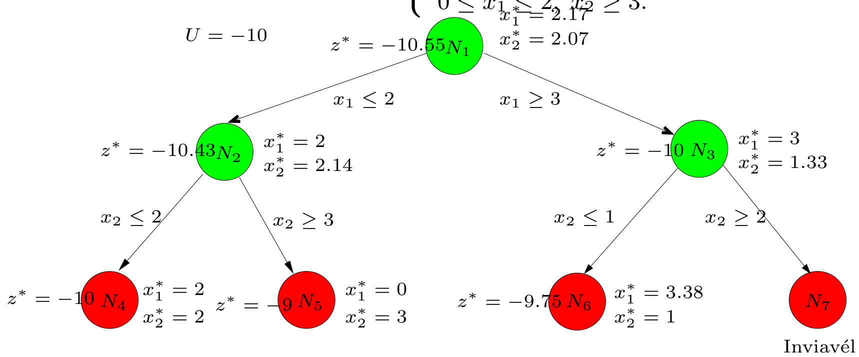
(ii) out of the constraints that are active at x^*, there are n of them that are linearly independent.

(b) If $x^*$is a basic solution that satisfies all of the constraints, we say that it is a basic feasible solution.

Lemma 10.2 *Suppose that the polyhedron $P = \{x \in \mathbb{R}^n \,|\, a_i^T x \geq b_i,\ i = \ \ldots, m\}$ is nonempty. If x^* is a basic feasible solution, then x^* is an extreme point of P*



$$\text{Em } N_5, \text{ resolvemos} \quad \begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ 0 \leq x_1 \leq 2,\ x_2 \geq 3. \end{cases}$$



$$\text{Em } N_6, \text{ resolvemos} \quad \begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \leq 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \leq 1 \\ 3 \leq x_1,\ 0 \leq x_2 \leq 1. \end{cases}$$

Let $I(x^*) = \{i \in \{1, \dots, m\} \mid a_i^T x^* = b_i\}$. Assume that $x^* = \alpha x_1 + (1-\alpha)x_2$ with $x_1,\ x_2 \in P$ and $\alpha \in (0,1)$. Since $a_i^T x_l \geq b_i$ for all $i \in I(x^*)$ and $l = 1,\ 2$, relation

$$0 = b_i - a_i^T x^* = b_i - \left(a_i^T x_1 + (1-\alpha)a_i^T{}_2\right) = \left(b_i - a_i^T x_1\right) + (1-\alpha)\left(b_i - a_i^T{}_2\right) \leq 0$$

implies that

$$a_i^T x_l = b_i \quad \forall i \in I(x^*),\ l = 1,\ 2.$$

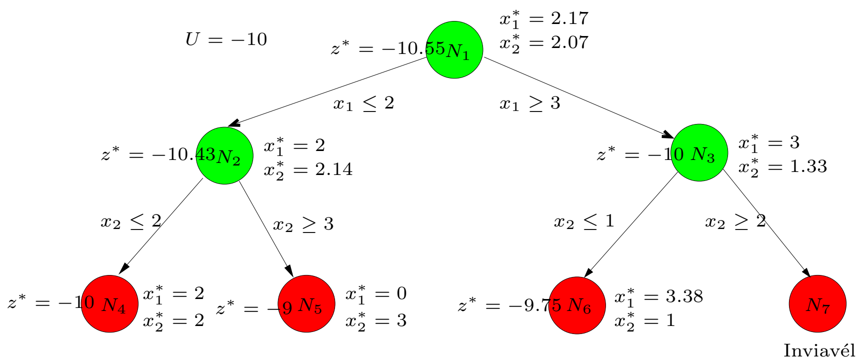
Taking $i_1,\ i_2, \dots,\ i_n \in I(x^*)$ such that $\{a_{i_1}, a_{i_2}, \dots, a_{i_n}\}$ is a l.i. set of vectors, it follows from the above relations and the definition of $I(x^*)$ that $x^* = x_1 = x_2$ which ends the proof.

Theorem 10.3 *Suppose that the polyhedron* $P = \{x \in \mathbb{R}^n \mid a_i^T x \geq b_i,\ i = \dots, m\}$ *is nonempty. Then, the following are equivalent:*

(a) The polyhedron P *has at least one extreme point.*

(b) The polyhedron P *does not contain a line.*

(c) There exist n *vectors out of the family* $a_1, \dots, a_m$ *which are linearly independent.*



Em N_7, resolvemos $\begin{cases} \min\ -2x_1 - 3x_2 \\ \frac{2}{9}x_1 + \frac{1}{4}x_2 \le 1 \\ \frac{1}{7}x_1 + \frac{1}{3}x_2 \le 1 \\ 3 \le x_1,\ 2 \le x_2. \end{cases}$

Não tem mais folhas ativas: achamos uma solução ótima
Uma solução ótima é (2;2) e o valor ótimo é -10.

(b)⇒(a). Let $x \in P$ and $I(x) = \{i \in \{1,\ldots,m\} \,|\, a_i^T x = b_i\}$. If n of the vectors $\{a_i,\ i \in I(x)\}$ are l.i. then it follows that x is a basic solution and therefore is an extreme point of P. If $\{a_i,\ i \in I(x)\}$ is a l.d. set of vectors then let $n(x)$ be the cardinality of the largest (in terms of number of vectors) l.i. set of vectors that can be extracted from $\{a_i,\ i \in I(x)\}$. It follows that there exists $d \neq 0$ such that $a_i^T d = 0$ for $i = 1,\ldots,m$. Consider the point $x(t) = x + td$. Clearly it holds $a_i^T x(t) = a_i^T x = b_i$ for $i = 1,\ldots,m$ and all $t \geq 0$. Now, let us define $I(x,d) = \{t \in \mathbb{R} \,|\, x + td \in P\}$ which bounded below or above in view of the assumptions. Let us assume that $t_1 = \sup\{t \in I(x,d)\} < +\infty$. Since P is closed we have that $t_1 \in I(x,d)$. Clearly, from this definition, it follows that at least one inequality will become active at $x(t_1)$. Hence, for some $j \notin I(x)$ it holds $a_j^T(x + t_1 d) < b_j$ which implies that $a_j^T d < 0$ since $a_j^T x \geq b_j$. Since $< \{a_i\ i \in I(x)\} > \subset < d >^\perp$ it follows that $a_j \notin < \{a_i\ i \in I(x)\} >$ which means that $n(x(t_1)) \geq n(x) + 1$. Now, repeating this process as long as required we will end up with some point in $\hat{x} \in P$ with $n(\hat{x}) = n$ which ends the proof in this case. The proof in the case when $t_1 = \inf\{t \in I(x,d)\} > -\infty\}$ is similar.

(a)⇒(c). Let $x \in P$ be an extreme point of P and $I(x) = \{i \in \{1,\ldots,m\} \,|\, a^T x = b_i\}$. Let us assume that $\{a_i,\ i \in I(x)\}$ is an l.d. set of vectors. Then it follows that there exists $d \neq 0$ such that $a_i^T d = 0$ for $i = 1,\ldots,m$. Consider the point $x(t) = x + td$ and notice that for $\varepsilon > 0$ small enough it holds $x(\varepsilon), x(-\varepsilon) \in P$. Since $x = (1/2)(x(\varepsilon) + x(-\varepsilon))$ we obtain a contradiction. Hence it follows that $\{a_i\ i \in I(x)\}$ is a l.i. set of vectors and ends the proof.

(c)⇒(b). Let us assume that for some $x \in P \subset \mathbb{R}^n$ and $d \neq 0$, the line $\{x + td \,|\, t \in \mathbb{R}\}$ is contained in P. It follows that $a_i^T(x + td) \geq b$ for all $t \in \mathbb{R}$ and $i = 1,\ldots,m$. This implies that $a_i^T d = 0$ for $i = 1,\ldots,m$. If we have a l.i. set of n vectors $\{a_{i_1}, a_{i_2}, \ldots, a_{i_n}\}$ it would follow that $d = 0$ which is a contradiction. This ends the proof.

Definition 10.4 *Given a nonempty convex set $P \subset \mathbb{R}^n$ and a point $x \in P$, we define the recession cone at x as the set of all directions x along which we can move indefinitely away from x, without leaving the set P. More formally, the recession cone at y is defined as the* *set* $\{w \in \mathbb{R}^n \,|\, x + td \in P,\ \forall t \geq 0\}$.

Lemma 10.5 *Given a nonempty polyhedron* $P = \{x \in \mathbb{R}^n \,|\, Ax \geq b\}$ *and a point* $x \in P$, *it holds that the recession cone at* x *is the closed and convex cone* $\{w \in \mathbb{R}^n \,|\, Aw \geq 0\}$.

Let w be a direction in the recession cone of P. Then, for any $x \in P$, we have that

$$A(x + tw) \geq b\ \forall\, t > 0 \Rightarrow Aw \geq \frac{-1}{t} Ax \underset{t \to +\infty}{\rightarrow} 0$$

Hence it holds $Aw \geq 0$. On the other hand for any $w \in \mathbb{R}^n$ with $Aw \geq 0$ it holds that

$$A(x + tw) = Ax + tAw \geq Ax \geq b\ \forall 1, t \geq 0,$$

which means that $x + tw \in P$ for all $t \geq 0$, that is, w is a recession direction for P. The convexity of the set $\{w \,|\, Aw \geq 0\}$ follows trivially from the definition and the fact that P is defined by linear inequalities. Also from the definition it follows trivially that is a cone. If $\{w^k\}$ is a sequence of directions in the recession cone converging to some $\hat{w}$, it follows that

$$0 \leq Aw^k \underset{k \to +\infty}{\rightarrow} A\hat{w} \Rightarrow 0 \leq A\hat{d}$$

which means that $\hat{w}$ is also in the recession cone and ends the proof.

Corollary 10.6 *Given a nonempty polyhedron* $P = \{x \in \mathbb{R}^n \,|\, Ax = b,\ x \geq 0\}$ *and a point* $x \in P$, *it holds that the recession cone at* x *is the closed and convex cone* $\{w \in \mathbb{R}^n \,|\, Aw = 0,\ w \geq 0\}$.

Note that $P = \{x \in \mathbb{R}^n \,|\, \hat{A}x \geq \hat{b}\}$ where $\hat{A} = \begin{bmatrix} A \\ -A \\ I \end{bmatrix}$ and $\hat{b} = \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$. Now, from the previous lemma, it follows that the recession cone of P is the set $\{w \in \mathbb{R}^n \,|\, \hat{A}w \geq 0\}$ which in view of the above definitions coincides with the set $\{w \in \mathbb{R}^n \,|\, Aw = 0,\ w \geq 0\}$.

Definition 10.7

(a) A nonzero element x *of a polyhedral cone* $C \subset \mathbb{R}^n$ *is called an extreme ray if there are* $n - 1$ *linearly independent constraints that are active at* x.

(b) An extreme ray of the recession cone associated with a nonempty polyhedron P *is also called an extreme ray of* P.

Example 10.8 *We will show that when the optimal cost of problem 7.25 is* $-\infty$, *the simplex method provides us at termination with an extreme ray. Let us recall that, in this case, at termination we have a basic matrix* B, *a nonbasic variable* x_j; *with negative reduced cost* $\hat{c}_j$ *and with corresponding column of the tableau* $B^{-1}A_j$ *which has no positive elements.*

Consider the vector $\hat{w} \in \mathbb{R}^n$ defined by $\hat{w}_j = 1$, $\hat{w}_l = 0$ for all $l \notin B$ with $l \neq j$ and $\hat{w}_B = -B^{-1}A_j$. Simple calculations show that

$$\hat{w} \geq 0, \quad A\hat{w} = B\hat{w}_B + A_j + \hat{N}\hat{w}_{\hat{N}} = -A_j + A_j + 0 = 0$$

where $\hat{N} = N - \{j\}$. Hence, according to Corollary 10.6, the vector $\hat{w}$ belongs to the recession cone of the feasible set of P. In addition, relation $A\hat{w} = 0$ means that $\hat{w}$ satisfies m linearly independent constraints with equality, and relations $\hat{w}_i = 0$ for $i \notin B$, $i \neq j$ means that $\hat{w}$ satisfies additional $n - m - 1$ l.i. linear constraints, totalizing $n - 1$ l.i. linear constraints. Hence, we conclude that $\hat{w}$ is an extreme ray of the feasible set. In addition, notice that simple calculations show that

$$c^T\hat{w} = -c_B B^{-1}A_j + c_j = \hat{c}_j < 0.$$

Theorem 10.9 *Consider the problem of minimizing c^Tx over a* **pointed** *polyhedral cone $C = \{x \in \mathbb{R}^n \,|\, a_i^T x \geq 0 \ i = 1, \ldots m\}$. The optimal cost is equal to $-\infty$ if and only if some extreme ray d of C satisfies $c^Td < 0$.*

Let d be an extreme ray of C satisfying $c^Td < 0$. It follows that $a + td \in C$ for all $t \geq 0$ and, since $c^T(a + td) = c^Ta + tc^Td \to -\infty$ when $t \to +\infty$, we conclude that the linear optimization problem considered above is unbounded below. Now assume that optimal cost of c^Tx over C is equal to $-\infty$. Then it follows that for some $\hat{x} \in C$ it holds $b := c^T\hat{x} < 0$. Now consider the (non empty) polyhedral $P = \{x \in R^n \,|\, a_i^Tx \geq 0,\ c^Tx = b\} \subset C$. Since C is pointed it follows that C and, therefore, P do not contain lines. By Theorem 10.3, it follows that P contains at least one extreme point $\bar{x}$ at which we have n active linear constraints whose associated coefficient vectors are l.i. Hence, it follows that at least $n-1$ constraints of the form $a_i^Tx \geq 0$ should be active at $\bar{x}$ which means that this point is an extreme ray of C and ends the proof.

Theorem 10.10 *Consider the problem of minimizing c^Tx subject to $Ax \geq b$, and assume that the feasible set has at least one extreme point. The optimal cost is equal to $-\infty$ if and only if some extreme ray d of the feasible set satisfies $c^Td < 0$.*

For the proof will denote the feasible set as $S = \{x \,|\, Ax \geq b\}$. Let d extreme ray of satisfying $c^Td < 0$. Since $x + td \in S$ for all $t > 0$ and $c^T(x + td) = c^Tx + tc^Td \to -\infty$ when $t \to +\infty$ the "if" part of the claim follows.

Next, assume that the optimal cost of the LLP mentioned above is $-\infty$. This means that the assciated dual problem

$$\begin{aligned} &\max \ b^Tp \\ &\quad s.t. \quad A^Tp = c, \quad p \geq 0. \end{aligned}$$

is unfeasible. Therefore the related LPP

$$\begin{aligned} &\max \ 0 \\ &\quad s.t. \quad A^Tp = c, \quad p \geq 0. \end{aligned}$$

is also unfeasible which implies that its dual,

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

is unbounded below. Since $\{x \mid Ax \geq b\}$ has at least one extreme point it follows that the rows of A spam $\mathbb{R}^n$ and therefore $\{x \mid Ax \geq 0\}$ is a pointed polyhedral cone. Now, by Theorem 10.9, it follows that there exists an extreme ray d of this set that satisfies $c^T d < 0$. Since the extreme rays of a polyhedral are the extreme rays of its recession cone the result follows.

Theorem 10.11 (Representation theorem) *Let $P = \{x \in \mathbb{R}^n \mid Ax \geq b\}$ be a nonempty polyhedron with at least one extreme point. Let $x^1, \ldots, x^k$ be the extreme points of P, and let $w^1, \ldots, w^r$ be a complete set of extreme rays of P. Let*

$$Q = \left\{ \sum_{i=1}^{k} \lambda_i x^i + \sum_{j=1}^{r} \theta_j w^j \mid \lambda_i \geq 0 \ , \theta_j \geq 0, \ i = 1, \ldots, k, \ j = 1, \ldots, r, \ \sum_{i=1}^{k} \lambda_k = 1 \right\}.$$

Then, $Q = P$.

First we prove that $Q \subset P$. Take any λ_i $i = 1, \ldots, k$ and θ_j $j = 1, \ldots, r$ as in the definition of Q. By convexity it follows that $\sum_{i=1}^k \lambda_i x^i \in P$ and by Lemma 10.5 it follows that $\sum_{j=1}^r \theta_j w^j$ is a direction in the recession cone. These properties imply that $\sum_{i=1}^k \lambda_i x^i + \sum_{j=1}^r \theta_j w^j \in P$ which ends the proof of the statement.

Next we prove that $P \subset Q$. Let us assume that there exists $z \in P - Q$. Note that Q is a closed and convex set since it is the sum of two closed and convex sets: the convex hull of the (finite) set of extreme points of P and the conic hull of the (finite) set of extreme rays of P. Then, by Theorem 5.10, there exists an hyperplane that strictly separate z from Q. That is, there exists $(p, q) \in \mathbb{R}^n \times \mathbb{R}$ such that $p^T z + q < 0$ and $p^T w + q \geq 0$ for all $w \in Q$. In particular, it follows that

$$p^T x^i \geq -q > p^T z \ \text{ for } \ i = 1, \ldots, k \tag{10.49}$$

and

$$p^T(\lambda w^j) + q \geq 0 \ \text{ for } \ j = 1, \ldots, r \ \text{ and all } \ \lambda > 0 \ \Rightarrow \ p^T w^j \geq 0 \ \text{ for } \ j = 1, \ldots, r. \tag{10.50}$$

Now consider the LPP

$$\begin{aligned} \min \ & p^T x \\ s.t. \quad & Ax \geq b \end{aligned}$$

which has z as a feasible point. If its optimal value is finite then there exists an extreme point x^i which is optimal, and it follows that $p^T x^i < p^T z$ which is a contradiction with (10.49). On the other hand, it its optimal value is $-\infty$ then, by Theorem 10.10, there exists an extreme ray w^j satisfying $p^T w^j < 0$ which is a contradiction with (10.50). Hence it follows that $Q \subset P$ which ends the proof of the theorem.

10.2 Dantzig-Wolfe decomposition.

Consider the LPP

$$\begin{array}{ll} \min & c^T x \\ \text{subject to} & D_1x_1 + D_2x_2 = b_0 \\ & \quad F_1x_1 = b_1 \\ & \quad F_2x_2 = b_2 \\ & \quad x_1, x_2 \geq 0. \end{array} \tag{10.51}$$

where $x = (x_1, x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^n$ and $b_0 \in R^{m_0}$. We will call the first block of constraints "coupling contraints" since they establish relations that should hold for the whole set of variables. The second block of constraints have a separable structure which as we shall see is very convenient. Next, consider the following reformation of the problem

$$\begin{array}{ll} \min c^T x & \\ \quad \text{subject to} & D_1x_1 + D_2x_2 = b_0 \\ & x_1 \in P_1 \\ & x_2 \in P_2. \end{array} \tag{10.52}$$

where $P_i = \{x \in \mathbb{R}^n \,|\, F_i x = b_i,\ x \geq 0\}$ $i = 1, 2$. For $i = 1, 2$, let x_i^j $j \in J_i$ be the extreme points of P_i, and let w_i^k $k \in K_i$ be a complete set of extreme rays of P_i. Then, any $x_i \in P_i$ can be written as

$$x_i = \sum_{j \in J_i} \lambda_i^j x_i^j + \sum_{k \in K_i} \theta_i^k w_i^k$$

where $\lambda_i^j \geq 0$,$\theta_i^k \geq 0$ for $i \in J_i$ $k \in K_i$ and $\sum_{i \in J_i} \lambda_i^j = 1$ for $i = 1, 2$. Hence, we can reformulate the LPP (10.52) as follows

$$\begin{array}{rl} \min \displaystyle\sum_{j \in J_1} \lambda_1^j c_1^T x_1^j + \sum_{k \in K_1} \theta_1^k c_1^T w_1^k + \sum_{i \in J_2} \lambda_2^j c_2^T x_2^j + \sum_{k \in K_2} \theta_2^k c_2^T x_2^k & \\ s.t. \quad \sum_{j \in J_1} \lambda_1^j D_1 x_1^j + \sum_{k \in K_1} \theta_1^k D_1 w_1^k + \sum_{j \in J_2} \lambda_2^j D_2 x_2^j + \sum_{k \in K_2} \theta_2^k D_2 w_2^k & = b_0 \\ \sum_{j \in J_1} \lambda_1^j & = 1 \\ \sum_{j \in J_2} \lambda_2^j & = 1 \\ \lambda_i^j,\ \theta_i^k \geq 0 & \forall\, i, j, k. \end{array} \tag{10.53}$$

The above problem will be called the master problem. Notice that we can rewrite the (m_0+2) equality linear restrictions as follows

$$\sum_{j \in J_1} \lambda_1^j \begin{bmatrix} D_1 x_1^j \\ 1 \\ 0 \end{bmatrix} + \sum_{j \in J_2} \lambda_2^j \begin{bmatrix} D_2 x_2^j \\ 0 \\ 1 \end{bmatrix} + \sum_{k \in K_1} \theta_1^j \begin{bmatrix} D_1 w_1^k \\ 0 \\ 0 \end{bmatrix} + \sum_{k \in K_2} \theta_2^j \begin{bmatrix} D_1 w_2^k \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} b_0 \\ 1 \\ 1 \end{bmatrix}$$

Let us assume that we have a feasible basic solution for the master problem (10.53) associated to a basis B. Consider the dual vector $p^T = c_B^T B^{-1} \in \mathbb{R}^{m_0+2}$ and let us denote

$p = (q, r_1, r_2)$ with $r_1, r_2 \in \mathbb{R}$ being related to the last two equality restrictions in (10.53). To check optimality we look at the reduced costs: for variable λ_i^j we have

$$c_i x_i^j - \begin{bmatrix} q^T \, r_1 \, r_2 \end{bmatrix} \begin{bmatrix} D_i x_i^j \\ e_i^T \end{bmatrix} = (c_i^T - q^T D_i) x_i^j - r_1$$

and for variable θ_i^k we have

$$c_i^T w_i^k - \begin{bmatrix} q^T \, r_1 \, r_2 \end{bmatrix} \begin{bmatrix} D_i w_i^k \\ 0 \\ 0 \end{bmatrix} = (c_i^T - q^T D_i) w_i^k$$

for $i = 1, 2$ where $e_1 = (1, 0)$ and $e_2 = (0, 1)$. Now, instead of evaluating all the expressions above we will solve the following LPP

$$\begin{array}{rl} \min & (c_i^T - q^T D_i)x \\ \text{subject to} & x \in P_i. \end{array}$$

which we will call subproblems, using the Simplex method.

We have the following options

(a) If the optimal cost in the subproblem is $-\infty$, then, upon termination, the simplex method provides us with an extreme ray $w_i^{\hat{k}}$ $\hat{k} \in K_i$ that satisfies $(c_i^T - q^T D_i) w_i^{\hat{k}} < 0$ (see Example 10.8). In this case, the reduced cost of the variable $\theta_1^{\hat{k}}$ is negative. At this point, we can generate the column

$$\begin{bmatrix} D_i w_i^{\hat{k}} \\ 0 \\ 0 \end{bmatrix}$$

associated with $\theta_i^{\hat{k}}$, and have it enter the basis in the master problem.

b) If the optimal cost in the subproblem is finite and smaller than r_i; then, upon termination, the simplex method provides us with an extreme point $x_1^{\hat{j}}$ $\hat{j} \in J_i$ that satisfies $(c_i^T - q^T D_i) x_i^{\hat{j}} < r_i$. In this case, the reduced cost of the variable $\lambda_i^{\hat{j}}$ is negative. At this point, we can generate the column

$$\begin{bmatrix} D_i x_i^{\hat{j}} \\ 0 \\ 0 \end{bmatrix}$$

associated with $\lambda_i^{\hat{j}}$, and have it enter the basis in the master problem.

c) Finally, if the optimal cost in the subproblem is finite and no smaller than r_i, this implies that $(c_i^T - q^T D_i) x_i^j \geq r_i$ for all extreme points x_i^j, and $(c_i^T - q^T D_i) w_i^k \geq 0$ for all extreme rays w_i^k. In this case, the reduced cost of every variable λ_i^j or w_i^k is nonnegative.

Next we present the resulting algoritm:

Dantzig-Wolfe decomposition algorithm with delayed column generation

1. A typical iteration starts with a total of m_0+2 extreme points and extreme rays of P_1 and P_2, which lead to a basic feasible solution to the master problem, the corresponding inverse basis matrix B^{-1} and the dual vector $p^T=(q,r_1,r_2)^T=c_B^TB^{-1}$.
2. Form and solve the two subproblems. If the optimal cost in the first subproblem is no smaller than r_1 and the optimal cost in the second subproblem is no smaller than r_2, then all reduced costs in the master problem are nonnegative, we have an optimal solution, and the algorithm terminates.
3. If the optimal cost in the i-th subproblem is $-\infty$, we obtain an extreme ray w_i^k associated with a variable θ_i^k whose reduced cost is negative; this variable can enter the basis in the master problem.
4. If the optimal cost in the ith subproblem is finite and less than r_i we obtain an extreme point x_i^j, associated with a variable λ_i^j whose reduced cost is negative; this variable can enter the basis in the master problem.
5. Having chosen a variable to enter the basis, generate the column associated with that variable, carry out an iteration of the revised simplex method for the master problem, and update B^{-1} and p.

Remark 10.12 *Note that the method above solves the Dantzig-Wolfe reformulation* (10.56) *by using an implementation of the Reduced Simplex method.*

Remark 10.13 *Note that we can also apply the Dantzig-Wolfe decomposition method to the case in which we do not have coupling constraints and therefore we do not separate the variables in blocks. For example, we can consider an LPP of the form*

$$\begin{array}{ll} \min & c^T x \\ \textit{subject to} & D_1 x = b_0 \\ & F_1 x = b_1 \\ & x \geq 0, \end{array} \tag{10.54}$$

which can be rewritten as

$$\begin{array}{ll} \min & c^T x \\ \textit{subject to} & D_1 x = b_0 \\ & x \in P_1 \end{array} \tag{10.55}$$

where $P_1=\{x\in\mathbb{R}^n \,|\, F_1x=b_1,\ x\geq 0\}$ *and then apply a similar process to the one above considering the representation of* P_1 *in terms of its extreme points and rays.*

Remark 10.14 *Let us assume that at iteration l we have (what we call) the restricted master problem:*

$$\begin{aligned} \min \sum_{j\in J_1^l} \lambda_1^j c_1^T x_1^j + \sum_{k\in K_1^l} \theta_1^j c_1^T w_1^k + \sum_{j\in J_2^l} \lambda_2^j c_2^T x_2^j + \sum_{k\in K_2^l} \theta_2^k c_2^T x_2^k & \\ s.t. \quad \sum_{j\in J_1^l} j \in J_1^l \lambda_1^j D_1 x_1^j + \sum_{k\in K_1^l} \theta_1^k D_1 w_1^k + \sum_{j\in J_2^l} \lambda_2^j D_2 x_2^j + \sum_{k\in K_2^l} \theta_2^k D_2 w_2^k & = b_0 \\ \sum_{j\in J_1^l} \lambda_1^j & = 1 \\ \sum_{j\in J_2^l} \lambda_2^j & = 1 \\ \lambda_i^j,\ \theta_i^k \geq 0 & \quad \forall\, i, j, k. \end{aligned} \tag{10.56}$$

where $J_i^l \subset J_i$ and $K_i^l \subset K_i$ for $i = 1, 2$ and l. That is, the restricted master problem is defined by considering only a subset of the extreme points and extreme rays of the sets P_1 and P_2. Then, an alternative method of based on the Dantzig-Wolfe decomposition is to replace the step 1) above for the resolution of this restricted master problem and then add a new extreme point and/or a new extreme ray to define a new restricted master problem (with more variables), more closely related in some sense to the original problem. In this case we obtain the following formulation:

Dantzig-Wolfe decomposition algorithm with column generation

1. *Solve the restricted master problem. If the optimal cost is $-\infty$ then the master problem (10.56), and therefore the LP (10.52), are also unbounded below and the algorithm stops. Hence, will assume that the (10.52) has solution.*

2. *Take the optimal dual vector $p^T = (q, r_1, r_2)$ and form and solve the two subproblems. If the optimal cost in the first subproblem is no smaller than r_1 and the optimal cost in the second subproblem is no smaller than r_2, then all reduced costs in the master problem are nonnegative, we have an optimal solution, and the algorithm terminates.*

3. *If the optimal cost in the i-th subproblem is $-\infty$, we obtain an extreme ray w_i^k associated with a variable θ_i^k whose reduced cost is negative. Since we have already solved the restricted master problem in step 1), it follows that $w_i^k \notin K_i^l$ and we can add this variable to the set of extreme rays in order to define the new restricted master problem with $K_i^{l+1} = K_i^l \cup \{w_i^k\}$. In addition, this variable can enter the basis in this new restricted master problem.*

4. *If the optimal cost in the ith subproblem is finite and less than r_i we obtain an extreme point x_i^j associated with a variable λ_i^j whose reduced cost is negative. Since we have already solved the restricted master problem in step 1), it follows that $x_i^j \notin J_i^l$ and we can add this variable to the set of extreme points in order to define the new restricted master problem with $J_i^{l+1} = J_i^l \cup \{x_i^l\}$. In addition, this variable can enter the basis in this new restricted master problem.*

Example 10.15 *Solve the LPP*

$$\begin{aligned} min \ & -4x_1 - x_2 - 6x_3 \\ subject\ to\ & 3x_1 + 2x + 4x_3 = 17 \\ & 1 \le x_1 \le 2 \\ & 1 \le x_2 \le 2 \\ & 1 \le x_3 \le 2 \end{aligned}$$

Let us consider, in the notations of LPP (10.52), $D_1 = [3\ 2\ 4]$, $b_0 = 17$, and $P_1 = \{x = (x_1, x_2, x_2) \in \mathbb{R}^3 \,|\, 1 \le x_1 \le 2,\ 1 \le x_2 \le 2,\ 1 \le x_3 \le 2\}$. In this particular case, P is a cube, then it has 8 extreme points which are its vertices. In addition, since P is bounded it does not have extreme rays. Denoting the extreme points of P by x^j $j = 1, 2, \dots, 8$, and $c = (-4, -1, -6)^T$, the corresponding master problem is

$$\begin{aligned} min\ & \sum_{i=1}^{8} \lambda_j c^T x^j \\ subject\ to\ & \sum_{j=1}^{4} \lambda_j D_1 x^j = 17 \\ & \sum_{i=1}^{8} \lambda_j = 1 \\ & \lambda_j \ge 0\ j = 1, 2, \dots, 8 \end{aligned}$$

or

$$\begin{aligned} min\ & \sum_{i=1}^{8} \lambda_j c^T x^j \\ subject\ to\ & \hat{A}\lambda = 17 \\ & \lambda \ge 0 \end{aligned}$$

with the j-th column of $\hat{A}$ be given by $\begin{bmatrix} D_1 x^j \\ 1 \end{bmatrix}$ for $j = 1, \dots, 8$. Taking, in particular, the extreme points $x^1 = (2, 2, 2)^T$ and $x^2 = (1, 1, 2)^T$, we obtain the matrix

$$B = \begin{bmatrix} 18 & 13 \\ 1 & 1 \end{bmatrix}$$

which is non-singular and, therefore, a basis for the master problem. We also have

$$B^{-1} = \begin{bmatrix} 0.2 & -2.6 \\ -0.2 & 3.6 \end{bmatrix} \text{ and } \lambda_B = B^{-1}(17, 1)^T = \begin{bmatrix} 0.2 & -2.6 \\ -0.2 & 3.6 \end{bmatrix} \begin{bmatrix} 17 \\ 1 \end{bmatrix} = (0.8, 0.2)^T \ge 0,$$

hence we have a feasible basic solution, and $c_B = (cx^1, cx^2) = (-22, -17)$. Now lets proceed with the algorithm.

Dantzig-Wolfe decomposition algorithm

1. *We calculate the dual vector*

$$p^T = (q, r_1)^T = c_B^T B^{-1} = (-22, -17) \begin{bmatrix} 0.2 & -2.6 \\ -0.2 & 3.6 \end{bmatrix} = (-1, -4)$$

2. *Form and solve the subproblem*

$$\begin{array}{rl} \min & (c^T - q^T D_1)x \\ \textit{subject to} & x \in P_1. \end{array}$$

Since $c^T - q^T D_1 = (-4, -1, -6) - (-1)(3, 2, 4) = (-1, 1, -2)$ *we obtain*

$$\begin{array}{rl} \min & -x_1 + x_2 - 2x_3 \\ \textit{subject to} & 1 \le x_1 \le 2 \\ & 1 \le x_2 \le 2 \\ & 1 \le x_3 \le 2 \end{array}$$

Clearly, the solution is $(2, 1, 2)$ *and the optimal value is* -5*. Notice that* $-5 < r_1 = -4$*. Then, we have a new extreme point* $x^3 = (2, 1, 2)$ *and we can take vector*

$$\hat{A}_3 = \begin{bmatrix} D_1 x^3 \\ 1 \end{bmatrix} = \begin{bmatrix} (3, 2, 4)(2, 1, 2)^T \\ 1 \end{bmatrix} = \begin{bmatrix} 16 \\ 1 \end{bmatrix}$$

to enter the basis. Calculating

$$u = B^{-1}\hat{A}_3 = \begin{bmatrix} 0.2 & -2.6 \\ -0.2 & 3.6 \end{bmatrix} \begin{bmatrix} 16 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix},$$

and $\min\{\lambda_1/u_1, \lambda_2/u_2\} = \min\{0.8/0.6, 0.2/0.4\} = 0.2/0.4$ *we determine that variable* λ_2 *leaves the basis. Then, the new basis is*

$$B = \begin{bmatrix} 18 & 16 \\ 1 & 1 \end{bmatrix}$$

its inverse and the new

$$B^{-1} = \begin{bmatrix} 0.5 & -8 \\ -0.5 & 9 \end{bmatrix}.$$

and we have that $c_B = (c^T x^1, c^T x^3) = (-22, -21)$ *and* $\lambda_B = (\lambda_1, \lambda_3) = B^{-1}(17, 1)^T = (0.5, 0.5)$*. Now we repeat the process*

3. *We calculate the dual vector*

$$p^T = (q, r_1)^T = c_B^T B^{-1} = (-22, -21) \begin{bmatrix} 0.5 & -8 \\ -0.5 & 9 \end{bmatrix} = (-0.5, -13)$$

4. Form and solve the subproblem

$$\begin{array}{rl} \min & (c^T - q^T D_1)x \\ \textit{subject to} & x \in P_1. \end{array}$$

Since $c^T - q^T D_1 = (-4,-1,-6) - (-0.5)(3,2,4) = (-2.5,0,-4)$ *we obtain*

$$\begin{array}{rl} \min & -2.5x_1 - 4x_3 \\ \textit{subject to} & 1 \leq x_1 \leq 2 \\ & 1 \leq x_2 \leq 2 \\ & 1 \leq x_3 \leq 2 \end{array}$$

Clearly, the solutions are $(2, s, 2)$ *for any* $s \in [1,2]$ *and the optimal value is* -13*. Notice that* $-13 = r_1$*. Then, all the reduced costs are non-negative and we have an optimal solution for the master problem. Now, it follows that*

$$x = \lambda_1 x^1 + \lambda_3 x^3 = 0.5(2,2,2)^T + 0.5(2,1,2)^T = (2,1.5,2)^T$$

is an optimal solution for the LPP (10.15)*.*

Example 10.16 *Solve the LPP*

$$\begin{array}{rl} \textit{min} & -5x_1 + x_2 \\ \textit{subject to} & x_1 \leq 8 \\ & x_1 - x_2 \leq 4 \\ & 2x_1 - x_2 \leq 10 \\ & x_1, x_2 \geq 0 \end{array}$$

We will introduce a slack variable x_3 *for the first constraint, obtaining the following reformulation*

$$\begin{array}{rl} \textit{min} & -5x_1 + x_2 \\ \textit{subject to} & x_1 + x_3 = 8 \\ & x_1 - x_2 \leq 4 \\ & 2x_1 - x_2 \leq 10 \\ & x_1,\ x_2,\ x_3 \geq 0 \end{array}$$

The first constraint is the coupling constraint and for the rest we define the sets $P_1 = \{(x_1,x_2) \,|\, x_1 - x_2 \leq 4,\ 2x_1 - x_2 \leq 10,\ x_1 \geq 0,\ x_2 \geq 0\}$ *and* $P_2 = \{x_3 \,|\, x_3 \geq 0\}$*. Hence, denoting* $x = (x_1,x_2)^T$ *the problem can be rewritten as follows*

$$\begin{array}{rl} \textit{min} & (-5,1)x \\ \textit{subject to} & (1,0)x + x_3 = 8 \\ & x \in P_1 \\ & x_3 \in P_2 \end{array} \tag{10.57}$$

Hence, in the notations of LPP (10.52), $D_1 = [1\ 0]$, $b_0 = 8$ and $D_2 = 1$, $c_1 = (-5, 1)$ and $c_2 = 0$. Drawing the set P_1 allows us to identify the extreme points $x^1 = (4, 0)$ and $x^2 = (6, 2)$. Since P_2 have the unique extreme point 0 and the unique extreme ray 1, then we can identify the variable associated with the extreme ray with x_3. Taking $\lambda_1^1, \lambda_2^1 \geq 0$ with $\lambda_1^1 + \lambda_1^2 = 1$ and setting $x = \lambda_1^1 x^1 + \lambda_1^2 x^2 = (4\lambda_1^1 + 6\lambda_1^2, 2\lambda_2^1)$, we will consider the restricted master problem

$$\begin{aligned} \min\ & (-5,1)(4\lambda_1^1 + 6\lambda_1^2, 2\lambda_1^2)^T \\ \text{subject to}\ & (1,0)(4\lambda_1^1 + 6\lambda_1^2, 2\lambda_1^2)^T + x_3 &= 8 \\ & \lambda_1^1 + \lambda_1^2 &= 1 \\ & \lambda_1^1,\ \lambda_1^2, x_3 &\geq 0, \end{aligned}$$

that is

$$\begin{aligned} \min\ & -20\lambda_1^1 - 28\lambda_1^2 \\ \text{subject to}\ & 4\lambda_1^1 + 6\lambda_1^2 + x_3 &= 8 \\ & \lambda_1^1 + \lambda_1^2 &= 1 \\ & \lambda_1^1,\ \lambda_1^2, x_3 &\geq 0. \end{aligned}$$

Now lets proceed with the algorithm. Notice that we have a trivial basic solution $x = (\lambda_1^1, \lambda_1^2, x_3) = (1, 0, 4)$ corresponding to the basis $B = \begin{bmatrix} 4 & 1 \\ 1 & 0 \end{bmatrix}$ with basic variables λ_1^1 and x_3. Let us solve this LPP by using the Simplex method. Here we have the initial tableau.

	z	λ_1^1	λ_1^2	x_3	b
x_3	0	0	2	1	4
$\leftarrow$ λ_1^1	0	1	1	0	1
z	1	0	8	0	-20
			$\uparrow$		

Then,

	z	λ_1^1	λ_1^2	x_3	b
x_3	0	-2	0	1	2
λ_1^2	0	1	1	0	1
z	1	-8	0	0	-28

and we have obtained an optimal solution $(\lambda_1^1, \lambda_1^2, x_3) = (0, 1, 2)$ associated to the basis $B = \begin{bmatrix} 6 & 1 \\ 1 & 0 \end{bmatrix}$ with $x_B = (\lambda_1^2, x^3)$. Associated to this solution we have the optimal dual solution

$$p = (q, r_1) = c_B B^{-1} = (-28, 0) \begin{bmatrix} 0 & 1 \\ 1 & -6 \end{bmatrix} = (0, -28).$$

Now we solve the subproblem

$$\begin{aligned} \min\ & (c_1^T - q^T D_1)x = c^T x = -5x_1 + x_2 \\ \text{subject to}\ & x_1 - x_2 \leq 4 \\ & 2x_1 - x_2 \leq 10 \\ & x_1, x_2 \geq 0 \end{aligned} \tag{10.58}$$

We introduce slack variables e_1*,* e_2 *and reformulate the problem above as*

$$\begin{aligned} &min\ -5x_1 + x_2 \\ &subject\ to\quad x_1 - x_2 + e_1 = 4 \\ &\qquad 2x_1 - x_2 + e_2 = 10 \\ &\qquad x_1, x_2.e_1, e_2 \geq 0 \end{aligned}$$

We will solve this problem using the Simplex method. The initial tableau is

	z	x_1	x_2	e_1	e_2	b
$\leftarrow$ e_1	0	1	-1	1	0	4
e_2	0	2	-1	0	1	10
z	1	5	-1	0	0	0
		$\uparrow$				

	z	x_1	x_2	e_1	e_2	b
x_1	0	1	-1	1	0	4
$\leftarrow$ e_2	0	0	1	-2	1	2
z	1	0	4	-5	0	-20
			$\uparrow$			

	z	x_1	x_2	e_1	e_2	b
x_1	0	1	0	-1	1	6
x_2	0	0	1	-2	1	2
z	1	0	0	3	-4	-28
				$\uparrow$		

It follows that the problem is unbounded below and we have the following extreme ray $w^1 = (-B^{-1}\hat{A}_3, 1, 0) = (1, 2, 1, 0)$ *where* $\hat{A} = \begin{bmatrix} 1 & -1 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{bmatrix}$ *is the matrix defining the constraints in the initial problem and* $B = \begin{bmatrix} 1 & -1 \\ 2 & - \end{bmatrix}$ *is the matrix of the basis corresponding to the current solution. Hence, for the subproblem (10.60) we have the following extreme ray* $w^1 = (1, 2)$ *with negative cost. We add this extreme ray to construct the new restricted master problem. Setting* $x = \lambda_1^1 x^1 + \lambda_1^2 x^2 + \theta w^1 = (4\lambda_1^1 + 6\lambda_1^2 + \theta^1, 2\lambda_1^2 + 2\theta^1)$

$$\begin{aligned} &min\ (-5, 1)(4\lambda_1^1 + 6\lambda_1^2 + \theta^1, 2\lambda_1^2 + 2\theta^1)^T \\ &\qquad subject\ to\quad (1, 0)(4\lambda_1^1 + 6\lambda_1^2 + \theta^1, 2\lambda_1^2 + 2\theta^1)^T + x_3 &= 8 \\ &\qquad\qquad \lambda_1^1 + \lambda_1^2 &= 1 \\ &\qquad\qquad \lambda_1^1,\ \lambda_1^2, \theta^1, x_3 &\geq 0, \end{aligned} \tag{10.59}$$

that is

$$\begin{aligned} &min\ -20\lambda_1^1 - 28\lambda_1^2 - 3\theta^1 \\ &\qquad subject\ to\quad 4\lambda_1^1 + 6\lambda_1^2 + \theta^1 + x_3 &= 8 \\ &\qquad\qquad \lambda_1^1 + \lambda_1^2 &= 1 \\ &\qquad\qquad \lambda_1^1,\ \lambda_1^2,\ \theta^1,\ x_3 &\geq 0. \end{aligned}$$

Let us solve this LPP by using the Simplex method. We start with the basic solution $x = (\lambda_1^1, \lambda_1^2, \theta^1, x_3) = (1, 0, 0, 4)$ *corresponding to the basis* $B = \begin{bmatrix} 4 & 1 \\ 1 & 0 \end{bmatrix}$ *with basic variables* λ_1^1 *and* x_3*. Then,*

	z	λ_1^1	λ_1^2	θ^1	x_3	b
$\leftarrow$ x_3	0	0	2	1	1	4
λ_1^1	0	1	1	0	0	1
z	1	0	8	3	0	-20
				$\uparrow$		

	z	λ_1^1	λ_1^2	θ^1	x_3	b
θ^1	0	0	2	1	1	4
$\leftarrow$ λ_1^1	0	1	1	0	0	1
z	1	0	2	0	-3	-32
			$\uparrow$			

	z	λ_1^1	λ_1^1	θ^1	x_3	b
θ^1	0	-2	0	1	1	2
λ_2^1	0	1	1	0	0	1
z	1	-2	0	0	-3	-34

Since alll the reduced costs are non-positive, we have obtained an optimal solution of the restricted master problem (10.59). The solution is $x = (\lambda_1^1, \lambda_1^2, \theta^1, x_3) = (0, 1, 2, 0)$ *associated to the basis* $B = \begin{bmatrix} 6 & 1 \\ 1 & 0 \end{bmatrix}$ *with* $x_B = (\lambda_1^2, \theta^1)$*. Associated to this solution we have the optimal dual solution*

$$p = (q, r_1) = c_B B^{-1} = (-28, -3) \begin{bmatrix} 0 & 1 \\ 1 & -6 \end{bmatrix} = (-3, -10).$$

Now we solve the subproblem

$$\begin{aligned} \mathit{min}\ (c_1^T - q^T D_1)x &= ((-5, 1) - (-3)(1, 0))\, x = -2x_1 + x_2 \\ \mathit{subject\ to}\quad x_1 - x_2 &\leq 4 \\ 2x_1 - x_2 &\leq 10 \\ x_1, x_2 &\geq 0 \end{aligned} \tag{10.60}$$

We introduce slack variables e_1, e_2 *and reformulate the problem above as*

$$\begin{aligned} \mathit{min}\ & -2x_1 + x_2 \\ \mathit{subject\ to}\quad & x_1 - x_2 + e_1 = 4 \\ & 2x_1 - x_2 + e_2 = 10 \\ & x_1, x_2 . e_1, e_2 \geq 0 \end{aligned}$$

We will solve this problem using the Simplex method. The initial tableau is

		z	x_1	x_2	e_1	e_2	b
$\leftarrow$	e_1	0	1	-1	1	0	4
	e_2	0	2	-1	0	1	10
	z	1	2	-1	0	0	0
			$\uparrow$				

		z	x_1	x_2	e_1	e_2	b
	x_1	0	1	-1	1	0	4
$\leftarrow$	e_2	0	0	1	-2	1	2
	z	1	0	1	-2	0	-8
				$\uparrow$			

	z	x_1	x_2	e_1	e_2	b
x_1	0	1	0	-1	1	6
x_2	0	0	1	-2	1	2
z	1	0	0	0	-1	-10

We have an optimal solution. Notice that the optimal value satisfies $-10 = r_1 = -10$. *Next, we solve the other subproblem.*

$$\begin{aligned} &\textit{min } (c_2^T - q^T D_2)x = (0 - (-3)1)\, x = 3x \\ &\textit{subject to } x \geq 0 \end{aligned} \tag{10.61}$$

which has the optimal solution 0. *Then, all the reduced cost for the variables of the master problem are nonpositive, and, therefore* $x = (\lambda_1^1, \lambda_1^2, \theta^1, x_3) = (0, 1, 2, 0)$ *is its optimal solution. Then,*

$$x = \lambda_1^1 x^1 = \lambda_1^2 x^2 + \theta^1 w^1 = 0(4, 0) + 1(6, 2) + 2(1, 2) = (8, 6)$$

is an optimal solution of the LPP (10.57)

Starting the algorithm
To find a basic feasible solution to the master problem we can proceed as follows:
Option 1:

1. Apply the Phase I to each one of the polyhedra P_1 and P_2 separately, and find extreme points x^1 and x^2 of P_1 and P_2 respectivel (notice that these points are not necessarily feasible for the original LPP).

2. We can assume that $D_1 x^1 + D_2 x^2 \leq b_0$ (multiplying the original constraints by -1 if necessary in the case of inequality), introduce an artificial variable $y \in \mathbb{R}_+^{m_0}$ and form

the auxiliary master problem

$$
\begin{aligned}
\min \sum_{l=1}^{m_0} y_l & \\
s.t. \quad \sum_{i=1,2} \left(\sum_{j \in J_i} \lambda_i^j D_i x_i^j + \sum_{k \in K_i} \theta_i^k D_i w_i^k \right) + y &= b_0 \\
\sum_{i \in J_1} \lambda_1^j &= 1 \\
\sum_{i \in J_i} \lambda_2^j &= 1 \\
\lambda_i^j,\ \theta_i^k,\ y_l \geq 0 \quad & \forall\, i, j, k, l.
\end{aligned}
\tag{10.62}
$$

Notice that we have the feasible basic solution $\lambda_1^1 = \lambda_1^2 = 1$, $\lambda_j^i = 0$ for $j \in J_i - \{1\}$ and $i = 1, 2$, and $y = b_0 - D_1 x^1 - D_2 x2$, so we can apply the D-W decomposition algorithm to solve this problem. If the optimal cost is positive then the master problem is infeasible. On the other hand, if the optimal cost is zero then we can obtain a basic fesible solution for the master problem.

Option 2:

1. Apply the Phase I to each one of the polyhedra P_1 and P_2 separately, and find extreme points x^1 and x^2 of P_1 and P_2 respectively.

2. Set $J_1^1 = K_i^1 = \{1\}$ and form the restricted master problem

$$
\begin{aligned}
\min \sum_{j \in J_1^l} \lambda_1^j c_1^T x_1^j + \sum_{k \in K_1^l} \theta_1^j c_1^T w_1^k + \sum_{j \in J_2^l} \lambda_2^j c_2^T x_2^j + \sum_{k \in K_2^l} \theta_2^k c_2^T x_2^k & \\
s.t. \quad \sum_{i=1,2} \left(\sum_{j \in J_i^1} \lambda_j^i D_i x_j^i + \sum_{k \in K_i^1} \theta_i^k D_i w_k^i \right) &= b_0 \\
\sum_{i \in J_1^1} \lambda_1^j &= 1 \\
\sum_{i \in J_2^1} \lambda_2^j &= 1 \\
\lambda_i^j,\ \theta_i^k,\ y_l \geq 0 \quad & \forall\, i, j, k, l.
\end{aligned}
\tag{10.63}
$$

and apply the D-W procedure presented in Remark 10.14.

Bounds on the optimal cost

Theorem 10.17 *Suppose that the master problem is feasible and its optimal cost z^* is finite. Let z be the cost of the feasible solution obtained at some intermediate stage of the decomposition algorithm. Also, let r_i be the value of the dual variable associated with the* *convexity constraint for the i-th subproblem. Finally, let z_i, be the optimal cost in the ith subproblem, assumed finite. Then,*

$$z + \sum_i (z_i - r_i) \leq z^* \leq z$$

The inequality $z^* \leq z$ is trivial, since any intermediate solution is feasible for the original problem. To prove the other inequality, first notice that the dual of the Master Problem (10.56) is the LPP

$$\begin{array}{llcll} \max\ q^T b_0 + r_1 + r_2 & & & & \\ \quad \text{subject to} \quad q^T D_1 x_j^1 & + r_1 & \leq & c_1^T x_i^i & \forall j \in J_1, \\ \quad\quad q^T D_1 w_k^1 & & \leq & c_2^T x_2^i & \forall k \in K_1, \\ \quad\quad q^T D_2 x_j^2 & + r_2 & \leq & c_2^T x_2^i & \forall j \in J_2, \\ \quad\quad q^T D_2 w_2^k & & \leq & c_2^T w_2^k & \forall k \in K_2. \end{array} \tag{10.64}$$

Let x be a feasible solution for the master problem at some stage of the algorithm with value of the objective function equal to z and let $p = c_B^T B^{-1} = (q, r_1, r_2)$ be the corresponding vector of dual multipliers. Simple calculations show that

$$, z = c_B^T B^{-1} b = b^T p = (b_0^T, 1, 1) p = b_0^T q + r_1 + r_2, \tag{10.65}$$

that is q is a (possibly infeasible) solution for the dual problem with the same objective function. For $i = 1, 2$, let $z_i = \min_{x \in P_i} \left(c_i^T - q^T D_i\right) x$. Since by the assumptions z_1 and z_2 are finite, it holds, for $i = 1, 2$,

1. $\min_{i \in J_i} \left(c_i^T - q^T D_i\right) x_i^j = z_i$, since at least one optimal solution of the subproblem is an extreme point of P_i
2. $\min_{k \in K_i} \left(c_i^T - q^T D_i\right) w_i^k \geq 0$, since $t w_i^k \in P_i\ \forall t > 0$ and $k \in K_i$.

Hence (q, z_1, z_2) is a feasible solution for the dual problem (10.64). Now, by weak duality and relation (10.65), it follows that

$$z^* \geq b_0^T q + z_1 + z_2 = b_0^T q + r_1 + r_2 + (z_1 - r_1) + (z_2 - r_2) = z + (z_1 - r_1) + (z_2 - r_2)$$

which ends the proof.

10.3 Benders decomposition

Consider the following LPP

$$\begin{array}{rl} \max & b_1^T \pi_1 + b_2^T \pi_2 \\ \text{subject to} & A_1 \pi_1 + A_2 \pi_2 \leq c. \end{array} \tag{10.66}$$

Let us rewrite this problem as

$$\begin{array}{rl} \max & b_1^T\pi_1 + b_2^T\pi_2 \\ \text{subject to} & A_2\pi_2 \le c - A_1\pi_1. \end{array} \tag{10.67}$$

For ease of exposition, we will consider the following two assumptions:

A.1) the LPP (10.66) is feasible and has a finite optimum;

A.2) given π_1 is always possible to find π_2 such that $A_1\pi_1 + A_2\pi_2 \le c$.

It is easy to see that solving (10.67) is equivalent to solving

$$\max_{\pi_1} b_1^T\pi_1 \ + \phi(\pi_1) \tag{10.68}$$

where

$$\begin{aligned} \phi(\pi_1) = &\max \ b_2^T\pi_2 \\ &\text{subject to } A_2\pi_2 \le c - A_1\pi_1. \end{aligned} \tag{10.69}$$

The dual of the LPP (10.69) is

$$\begin{array}{rl} \min\limits_{x} & (c - A_1\pi_1)^T x \\ \text{subject to} & A_2^T x = b_2 \\ & x \ge 0. \end{array} \tag{10.70}$$

Let $\{x_i,\ i \in I\}$ be the set of extreme points of $P = \{x \in \mathbb{R}^n \,|\, A_2^T x = b_2,\ x \ge 0\}$, and let $\{w_k,\ k \in K\}$ be a complete set of extreme rays of P. Then, any $x \in P$ can be written as

$$x = \sum_{i \in I}\lambda_i x_i + \sum_{k \in K}\theta_k w_k$$

where $\lambda_i \ge 0$, $\theta_k \ge 0$ for all $i \in I$ and $k \in K$ and $\sum\limits_{i \in I}\lambda_i = 1$. Hence, we can reformulate the LPP (10.70) as follows

$$\begin{array}{rl} \min\limits_{\lambda,\theta} & \sum\limits_{i \in J}\left[(c - A_1\pi_1)^T x_i\right]\lambda_i \ + \ \sum\limits_{k \in K}\left[(c - A_1\pi_1)^T w_k\right]\theta_k \\ \text{subject to} & \sum\limits_{j \in J}\lambda_i \ = 1 \\ & \lambda_i,\ \theta_k \ge 0 \ \forall i \in I,\ k \in K. \end{array} \tag{10.71}$$

Regarding the LPP (10.71) we have two important facts:

- Notice that view of Assumptions A1 and A2, it follows that problem (10.70), and, therefore, problem (10.71), must have a finite optimum, which implies that

$$(c - A_1\pi_1)^T w_k \ge 0 \ \forall\, k \in K$$

- The optimum of problem (10.70) must occur at an extreme point.

Hence we can reformulate the LPP (10.71), and, therefore, problem (10.70), as follows

$$\begin{array}{rl} \min\limits_{i \in I} & (c - A_1\pi_1)^T x_i \\ \text{subject to} & (c - A_1\pi_1)^T w_k \geq 0 \quad \forall\, k \in K. \end{array}$$

which is equivalent to (in terms of optimal value) to the LPP below

$$\begin{array}{rl} \max\limits_{\pi_1,\,\gamma} & \gamma \\ \text{subject to} & (c - A_1\pi_1)^T x_i \geq \gamma \quad \forall\, i \in I, \\ & (c - A_1\pi_1)^T w_k \geq 0 \quad \forall\, k \in K. \end{array} \tag{10.72}$$

Now it follows that we can reformulate the LPP (10.68), and, therefore, the LPP (10.67), as follows

$$\begin{array}{rl} \max\limits_{\pi_1,\,\gamma} & b_1^T\pi_1 + \gamma \\ \text{subject to} & (c - A_1\pi_1)^T x_i - \gamma \geq 0 \quad \forall i \in I, \\ & (c - A_1\pi_1)^T w_k \geq 0 \quad \forall\, k \in K. \end{array} \tag{10.73}$$

Introducing the notations

$$G_i = A_1 x_i, \quad g_i = c^T x_i, \quad H_k = A_1 w_k, \quad h_k = c^T w_k \quad i \in I,\ k \in K$$

we obtain *Full Benders Master Problem*

$$\begin{array}{rl} \max\limits_{\pi_1,\psi,\,\gamma} & \psi \\ \textit{subject to} & b_1^T\pi_1 + \gamma = \psi \\ & G_i^T\pi_1 + \gamma \leq g_i \quad \forall i \in I, \\ & H_k^T\pi_1 \leq h_k \quad \forall\, k \in K. \end{array} \tag{10.74}$$

We summarize the previous discussion in the following theorem.

Theorem 10.18 *(Benders Transformation Into an Equivalent LP) Benders Full Master Program (10.74) transforms the original $n \times (m_1 + m_2)$ linear program (10.66) into an equivalent linear program with fewer columns and possibly many more inequalities, namely, an $(L+M+1) \times (m_1+2)$, where m_1 is the number of columns of A_1, L is the number of extreme solutions, and M is the number of extreme rays of the set $\left\{x \mid A_2^T x = b_2,\ x \geq 0\right\}$.*

Since for any (π_1, π_2) feasible for problem (10.66) it holds

$$b_1^T\pi_1 + b_2^T\pi_2 \leq b_1^T\pi_1 + \phi(\pi_1)$$

with ϕ defined as in (10.69), it follows that the optimal value of the LP (10.68) is an upper bound for the optimal value of the LP (10.66). On the other hand, if $z^* < +\infty$ is the optimal value for the LP (10.66), then, for any (π_1, π_2) with $A_2\pi_2 \le c - A_1\pi_1$, it holds

$$b_1^T\pi_1 + b_2^T\pi_2 \le z^*$$

which, combined with Assumption A.2, implies that, for any π_1, it holds

$$b_1^T\pi_1 \; + \; \begin{pmatrix} \max\limits_{\pi_2} & b_2^T\pi_2 \\ \textit{subject to} & A_2\pi_2 \le c - A_1\pi_1. \end{pmatrix} \le z^*$$

which means that the optimal value of the LP (10.66) is an upper bound for the optimal value of the LP (10.68). We conclude then that the LPPs (10.66) and (10.68) have the same optimal value. Hence, these problems are equivalent in the sense that the component π_1 of the optimal solutions of either one of the problems can be found by solving the other problem. Since by the assumptions and the above relation, the LPP (10.69) is feasible and is upper bounded, it follows that its dual, the LPP (10.70), has an optimal solution and strong duality holds. Then, we can reformulate the LPP (10.68) as the following equivalent problem

$$\max_{\pi_1} b_1^T\pi_1 \; + \; \begin{array}{rl} \min\limits_{x} & (c - A_1\pi_1)^T x \\ \textit{subject to} & A_2^T x = b_2 \\ & x \ge 0. \end{array}$$

Applying the Dantzig-Wolfe reformulation to the LPP in the second term above (see Remark 10.13), we obtain problem (10.71), which, as shown in the arguments previous to the statement of the theorem, can be reformulated as(10.72). This fact, combined with the formulation above, finally leads the equivalent reformulations (10.73) and (10.74), and ends the proof

The Benders algorithm. *The algorithm consists in applying a constraint generation technique to the reformulation (10.74) of the original problem.*

1. *To initiate an iterative step, assume that we have already inherited from earlier iterations a set of inequalities, that is, a subset of $\bar{L}$ inequalities of the first set of L inequalities and a subset of $\bar{M}$ inequalities of the second set of M inequalities in (10.74), corresponding to certain subsets $\{x_i \,|\, i \in I_{\bar{L}} \subset I\}$ and $\{w_k \,|\, k \in K_{\bar{M}} \subset K\}$ of extreme points and extreme rays of P, respectively. The linear program with $\bar{L} + \bar{M}$ inequalities is called the Benders Restricted Master Program.*

2. *Assuming that the $\bar{L} + \bar{M}$ inequality Benders Restricted Master has an optimal solution $\bar{\pi}_1$ and $\bar{\gamma}$, we generate a new inequality by letting the adjusted cost be*

$$\bar{p} = c - A_1^T\bar{\pi}_1$$

and solving

$$\begin{array}{rl} \min & \bar{p}^T x \\ \textit{subject to} & A_2^T x = b_2 \\ & x \geq 0, \end{array} \tag{10.75}$$

or equivalently

$$\begin{array}{rl} \max & b_2^T \pi_2 \\ \textit{subject to} & A_2 \pi_2 \leq c - A_1 \pi_1. \end{array}$$

3. *The solution of* (10.75) *gives rise to two options for increasing the number of inequalities in the Benders Restricted Master Program.*

   (a) **Optimality Cut**. *If a finite optimal basic solution* $x = x^* \notin \{x_i \,|\, i \in I_{\bar{L}}\}$ *is obtained for* (10.75), *then we can generate a new inequality*

   $$G_{\bar{L}+1}^T \pi_1 + \gamma \leq g_{\bar{L}} \tag{10.76}$$

   where $G_{\bar{L}+1} = A_1 x^*$ *and* $g_{\bar{L}+1} = c^T x^*$. *This inequality is called an optimality cut.*

   (b) **Feasibility Cut**. *On the other hand, if the problem is unbounded and an extreme ray solution* $x = w^* \notin \{w_k \,|\, k \in K_{\bar{M}}\}$ *is obtained for* (10.75), *then we generate the new inequality*

   $$H_{\bar{M}+1}^T \pi_1 \leq h_{\bar{M}+1}$$

   where $H_{\bar{M}+1} = A_1 w^*$ *and* $h_{\bar{M}+1} = c^T w^*$. *This inequality is called a feasibility cut.*

4. *After augmenting the Benders Restricted Master Program by either the new optimality cut or feasibility cut indexed by* $i = \bar{L} + 1$ *or* $k = \bar{M} + 1$, *it is re-solved (see Lemma 10.19 and Theorem 10.20 below).*

Regarding step 4 above, we should take in to consideration the following results.

Lemma 10.19 *If, after augmenting the Benders Restricted Master Program by a new feasibility cut and resolving, we find that it is infeasible, then the original linear program is infeasible.*

The claim follows trivially since all the restricted master problems are relaxations of the Full Master Problem (10.74).

Theorem 10.20 *If the new optimality cut is feasible for* $\pi_1 = \bar{\pi}_1$ *and* $\gamma = \bar{\gamma}$, *then* $(\pi_1, \pi_2) = (\bar{\pi}_1, \bar{\pi}_2)$ *is an optimal solution to the original LPP 10.66, where* $\pi_2 = \bar{\pi}_2$ *is an optimal dual solution to* (10.75).

In this case we are assuming that the LPP (10.75) *has an optimal solution* x^*. *Then, since the problem is bounded, it follows that for any extreme ray of the set* P, $w \in K$, *it holds*

$$\bar{p}^T w = c^T w - (A_1 \bar{\pi}_1)^T w \geq 0.$$

Now, let us assume that $(\bar{\pi}_1, \bar{\gamma})$ satisfies the inequality (10.76). Then, we have that

$$\bar{\gamma} \leq g_{\bar{L}+1} - G_{\bar{L}+1}^T \bar{\pi}_1 = c^T x^* - (A_1 x^*)^T \bar{\pi}_1 = \left(c - A_1^T \bar{\pi}_1\right)^T x^* \leq \left(c - A_1^T \bar{\pi}_1\right)^T x \quad \forall\, x \in P.$$

The above two relations show that $(\bar{\pi}_1, \bar{\gamma})$ is a feasible solution for the LPP (10.73), and, since it is also an optimal solution for a relaxation of this problem, it follows that it is an optimal solution. Let π_2 be an optimal dual solution to (10.75). Since the dual of this problem is the LPP (10.69), from the optimality of π_2 for this problem it follows the optimality (π_1, π_2) for the LPP (10.66).

Remark 10.21 *Since the total number of optimality and feasibility cuts is finite ($L+M$ with L and M as in Theorem 10.18), the algorithm terminates in a finite number of iterations.*

Example 10.22 *Solve the LPP*

$$\begin{array}{rrcl}
\textit{min} & x_1 + x_2 + x_3 + 3x_4 + 2x_5 + x_6 & & \\
\textit{subject to} & x_1 + 2x_3 + 3x_3 & = & 6 \\
 & 3x_1 + 2x_2 + x_3 & = & 6 \\
 & x_1 + x_2 + x_3 + 4x_4 - x_5 + x_6 & = & 9 \\
 & 3x_1 + 2x_2 + x_3 + 3x_4 + 2x_5 + x_6 & = & 15 \\
 & 4x_1 - x_2 + x_3 + x_4 + x_5 + x_6 & = & 9 \\
 & x_i & \geq 0 & i = 1, \ldots, 6
\end{array} \tag{10.77}$$

Let us rewrite the LPP above as follows

$$\begin{array}{rrcl}
\textit{min} & b_1 \pi_1 + b_2 \pi_2 & & \\
\textit{subject to} & D\pi_1 & = & b_0 \\
 & A_1 \pi_1 + A_2 \pi_2 & = & c \\
 & \pi_1,\ \pi_2 & \geq & 0 \quad i = 1, \ldots, 6
\end{array} \tag{10.78}$$

where $\pi_1 = (x_1, x_2, x_3)$, $\pi_2 = (x_4, x_5, x_6)$, $b_0 = (6,6,6)^T$, $b_1 = (1,1,1)^T$, $b_2 = (3,2,1)^T$, $c = (9,15,9)^T$, $D = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$, $A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 4 & -1 & 1 \end{pmatrix}$ and $A_2 = \begin{pmatrix} 4 & -1 & 1 \\ 3 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Notice that, in order to obtain the Full Benders Master Problem, we must dualize the LPP

$$\begin{array}{rcl}
\phi(\pi_1) = & \min & b_2^T \pi_2 \\
 & \textit{subject to} & A_2 \pi_2 = c - A_1 \pi_1 \\
 & & \pi_2 \geq 0,
\end{array}$$

which results in the LPP

$$\begin{array}{rcl}
\phi(\pi_1) = & \max & (c - A_1 \pi_1)^T x \\
 & \textit{subject to} & A_2^T x \leq b_2.
\end{array}$$

The Full Benders Master Problem, with the notations of (10.73)*, and after some simple adaptations, since* (10.78) *is slightly different than* (10.66)*, is*

$$\begin{aligned} \min_{\pi_1,\gamma}\ & b_1^T\pi_1+\gamma \\ \textit{subject to}\ \ & D\pi_1 = \ b_0 \\ & (c-A_1\pi_1)^T x_i-\gamma\le 0 \quad \forall i\in I, \\ & (c-A_1\pi_1)^T w_k\le 0\ \forall\, k\in K, \\ & \pi_1\ge 0 \end{aligned} \tag{10.79}$$

where $\{x^i,\ i\in I\}$ *and* $\{w^k,\ k\in K\}$ *are, respectively, the set of extreme points and a complete set of extreme rays of the set* P*, defined as* $P=\{x\in\mathbb{R}^n\,|\,A_2^Tx\le b_2\}$*. Since currently we do not have any extreme points or rays, we will consider the following Restricted Benders Master Problem*

$$\begin{aligned} \min_{\pi_1}\ & b_1^T\pi_1 \\ \textit{subject to}\ \ & D\pi_1 = \ b_0 \\ & \pi_1\ge 0 \end{aligned}$$

That is

$$\begin{array}{rrcl} \textit{min} & x_1+x_2+x_3 & & \\ \textit{subject to} & x_1+2x_3+3x_3 & = & 6 \\ & 3x_1+2x_2+x_3 & = & 6 \\ & x_i & \ge 0 & i=1,\dots,3. \end{array} \tag{10.80}$$

Since the number of variables 3 and we have 2 equations, all basic solutions will have 2 basic variables and 1 nonbasic variable. So, in this particular case, we can look for basic solutions directly by setting one of the variable (a possible non-basic one in some basic solution) equal to 0*. Setting* $x_1=0$*, it follows that* $x_2=3$ *and* $x_3=0$*, then we have a (degenerated) basic solution* $\pi_1=(0,3,0)^T$ *associated to the basis* $B=\begin{pmatrix}2&2\\3&1\end{pmatrix}$*. Notice that this is the same basic solution obtained considering* $x_3=0$ *and the basis* $B=\begin{pmatrix}1&3\\2&2\end{pmatrix}$*. Setting* $x_2=0$*, it follows that* $x_1=x_3=1,5$*, then we have a basic solution* $\pi_1=(1.5,0,1.5)^T$ *associated to the basis* $B=\begin{pmatrix}1&3\\3&1\end{pmatrix}$*. All three solutions are optimal, since they all generate the same value of the objective function. We will use the third solution. To construct the first subproblem we generate the reduced cost*

$$\bar p=c-A_1\pi_1=(9,15,9)^T-\begin{pmatrix}1&1&1\\3&2&1\\4&-1&1\end{pmatrix}\begin{pmatrix}1.5\\0\\1.5\end{pmatrix}=(6,9,1.5)^T$$

and obtain

$$\begin{array}{rl} \textit{max} & \bar p^Ty \\ \textit{subject to} & A_2^Ty\ \le b_2 \end{array}$$

that is

$$\begin{array}{rrcl} \textit{max} & 6y_1 + 9y_2 + 1.5y_3 & & \\ \textit{subject to} & 4y_1 + 3y_2 + y_3 & \leq & 3 \\ & -y_1 + 2y_2 + y_3 & \leq & 2 \\ & y_1 + y_2 + y_3 & \leq & 1, \end{array} \tag{10.81}$$

This problem is unbounded and $\hat{w} = (-0.2, 1, -2.2)^T$ *is an extreme ray solution. Hence, we can generate the feasibility cut*

$$(c - A_1\pi_1)^T \hat{w} = (9, 15, 9)(-0.2, 1, -2.2)^T - (-0.2, 1, -2.2)\begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 4 & -1 & 1 \end{pmatrix}^T \pi_1 = -6.6 + 6x_1 - 4x_2 + 1.4x_3 \leq 0$$

and add this constraint to define our new restricted master problem, which is

$$\begin{array}{rrcl} \textit{min} & x_1 + x_2 + x_3 & & \\ \textit{subject to} & x_1 + 2x_3 + 3x_3 & = & 6, \\ & 3x_1 + 2x_2 + x_3 & = & 6, \\ & -6x_1 + 4x_2 - 1.4x_3 & \geq & -6.6 \quad 0 \\ & x_i & \geq & 0 \quad i = 1, \ldots, 3 \end{array} \tag{10.82}$$

The optimal solution to this problem is $\pi_1 = (1.207792, 0.584416, 1.207792)$ *with optimal value* $z^* = 3$. *To construct the next subproblem we generate the reduced cost*

$$\bar{p} = c - A_1\pi_1 = (6, 9, 3.545456)^T$$

and obtain

$$\begin{array}{rl} \textit{max} & \bar{p}^T y \\ \textit{subject to} & A_2^T y \leq b_2 \end{array}$$

that is

$$\begin{array}{rrcl} \textit{max} & 6y_1 + 9y_2 + 3.545456y_3 & & \\ \textit{subject to} & 4y_1 + 3y_2 + y_3 & \leq & 3 \\ & -y_1 + 2y_2 + y_3 & \leq & 2 \\ & y_1 + y_2 + y_3 & \leq & 1, \end{array} \tag{10.83}$$

The optimal solution of this problem is $x^* = (0, 1, 0)$ *with optimal value* $w^* = 9$. *Next, we construct an optimality cut*

$$(c - A_1\pi_1)^T x^* - \gamma = (9, 15, 9)(0, 1, 0)^T - (0, 1, 0)\begin{pmatrix} 1 & 1 & 1 \\ 3 & 2 & 1 \\ 4 & -1 & 1 \end{pmatrix}^T \pi_1 - \gamma = 15 - 3x_1 - 2x_2 - x_3 - \gamma \leq 0$$

and add this constraint to define our new restricted master problem, which is

$$\begin{array}{rrcl} \textit{min} & x_1 + x_2 + x_3 + \gamma & & \\ \textit{subject to} & x_1 + 2x_3 + 3x_3 & = & 6, \\ & 3x_1 + 2x_2 + x_3 & = & 6, \\ & -6x_1 + 4x_2 - 1.4x_3 & \geq & -6.6 \quad 0 \\ & 3x_1 + 2x_2 + x_3 & \geq & 15 \quad 0 \\ & x_i & \geq & 0 \quad i = 1, \ldots, 3 \end{array} \tag{10.84}$$

The optimal solution of this problem is $\pi_1 = (1.207792, 0.584416, 1.207792)$, $\gamma = 9$, *with optimal value* $z^* = 12$. *Since* $w^* = \gamma = 9$, *it follows that* (π_1, γ) *is feasible for the the Full Benders Master Problem* (10.79). *Since it is also optimal for the relaxation* (10.84) *of this problem, we conclude that its is an optimal solution. Recovering the dual optimal solution in problem* (10.83) *allows to compute an optimal solution to the LPP* (10.77), $(\pi_1, \pi_2) = ()$.

10.4 Branch and Price

The general ILPP we consider is of the form:

$$\max c^T x \tag{10.85}$$
$$\text{subject to: } Ax \le b, \quad x \in S, \quad x \text{ integer}$$

In Branch and Price, columns are excluded to form the LP relaxation because there are too many to handle efficiently, and most would have their associated variables equal to zero in an optimal solution. Then to check the optimality of an LP solution of the relaxed LP, a subproblem called the pricing problem (a separation problem for the dual of the original LP) is solved to identify columns that should enter the basis. If such columns are found, a new relaxed LP is formed and then reoptimized. Branching occurs when no columns price out to enter the basis and the LP solution does not meet the integrality conditions. Branch and Price, a generalization of branch and bound that incorporates LP relaxations, enables the continuous use of column generation within the branch-and-bound tree structure. Some difficulties arise when devising this process. For example,

- *Conventional integer programming branching on variables may not effective because fixing variables can destroy the structure of the pricing problem.*
- *Solving these LPs and the subproblems to optimality may not be effcient, in which case different rules will apply for managing the branch-and-price tree.*
- *The LP relaxation of the relaxed master problem with the existing columns may not have an integral optimal solution. Hence, applying standard branch-and-bound techniques to this relaxation is unlikely to yield an optimal, feasible solution for the original problem. Therefore, it may be necessary to generate additional columns in order to solve the linear programming relaxations of the master problem at non root nodes of the search tree.*

The fundamental idea of column generation inthe context of the ILPP (10.85) *is that the set*

$$S^* = \{x \in S \mid x \text{ integer}\},$$

if bounded, it is just a finite set of points, say $S^* = \{y_1, \ldots, y_p\}$, *which coincides with the extreme points of its convex hull, denoted by* $conv(S)$, *in the case when* x *is binary. For the results presented in this section we will assume that the set* S *is bounded.*

Remark 10.23 *When S is unbounded, a classical results of Minkowski and Weyl states that conv(S) is a polyhedron and is represented by a convex combination of a finite set of points and a linear combination of a finite set of rays. Thus, column generation for integer programming is still possible when S is unbounded.*

Given $S^* = \{y_1, \dots, y_p\}$, any point $y \in S^*$ can be represented as:

$$y = \sum_{k=1}^{p} \lambda_k y_k$$

subject to the convexity constraint:

$$\sum_{k=1}^{p} \lambda_k = 1, \quad \lambda_k \in \{0, 1\} \quad k = 1, \dots, p$$

Let $c_k = c^T y_k$ and let $a_k = Ay_k$. The master problem (MP) of P is the following integer linear programming problem

$$\begin{aligned}
\max \ & \sum_{k=1}^{p} c_k \lambda_k \\
\textit{subject to:} \ & \sum_{k=1}^{p} a_k \lambda_k \leq b, \\
& \sum_{k=1}^{p} \lambda_k = 1, \quad \lambda_k \in \{0, 1\} \quad k = 1, \dots, p
\end{aligned}$$

If S can be decomposed, i.e., $S = \cup_{j=1}^{n} S_j$, we can represent each set:

$$S_j^* = \{x_j \in S_j \mid x_j \textit{ integer}\}$$

as:

$$S_j' = \{y_{j1}, \dots, y_{jp_j}\}$$

Now let $c(y_{jk}) = c_{jk}$ and $Ay_{jk} = a_{jk}$. This yields the master problem (or column generation form) of P with separate convexity constraints for each S_j given by:

$$\begin{aligned}
\max \ & \sum_{j=1}^{n} \sum_{k=1}^{p_j} c_{jk} \lambda_{jk} \\
\textit{subject to:} \ & \sum_{j=1}^{n} \sum_{k=1}^{p_j} \lambda_{jk} a_{jk} \leq b, \\
& \sum_{k=1}^{p_j} \lambda_{jk} = 1 \quad j = 1, \dots, n, \\
& \lambda_{jk} \in \{0, 1\} \quad j = 1, \dots, n, \quad k = 1, \dots, p_j
\end{aligned}$$

If the subsets in the decomposition are identical, i.e., $S_j = S = \{y_1, \ldots, y_p\}$ for $j = 1, \ldots, n$, then they can be represented by one subset S with $\lambda_k = \sum_j \lambda_{jk}$ and the convexity constraints replaced by an aggregated convexity constraint:

$$\sum_{k=1}^{p} \lambda_k = n$$

where $\lambda_k \geq 0$ and integer. This results in the column generation form:

$$\begin{aligned} \max \ & \sum_{k=1}^{p} c_k \lambda_k \\ \textit{subject to:} \ & \sum_{k=1}^{p} \lambda_k a_k \leq b, \\ & \sum_{k=1}^{p} \lambda_k = n, \\ & \lambda_k \geq 0 \textit{ and integer} \quad k = 1, \ldots, p \end{aligned}$$

Remark 10.24 *The essential difference between P and its column generation form is that S^* is represented by a finite set of points. We see that any fractional solution to the linear programming relaxation of P is a feasible solution to the linear programming relaxation of its column generation form if and only if it can be represented by a convex combination of extreme points of conv(S). In particular, Geoffrion (1974) has shown that if the polyhedron conv(S) does not have all integral extreme points, then the linear programming relaxation of the column generation form of P will be tighter than that of P for some objective functions.*

Example 10.25

1. **Generalized Assignment Problem (GPA)** *The objective is to find a maximum profit assignment of m tasks to n machines such that each task is assigned to precisely one machine subject to capacity restrictions on the machines.*

   - *Unrelated Machines: The standard ILP formulation of GAP is*

$$\begin{aligned} \max & \sum_{i=1}^{m} \sum_{j=1}^{n} p_{ij} z_{ij} \\ s.t. & \sum_{j=1}^{n} z_{ij} = 1, \quad i = 1, \ldots, m \\ & \sum_{i=1}^{m} w_{ij} z_{ij} \leq d_j, \quad j = 1, \ldots, n \\ & z_{ij} \in \{0, 1\}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n \end{aligned}$$

where p_{ij} *is the profit associated with assigning task* i *to machine* j*,* w_{ij} *is the amount of the capacity of machine* j *used by task* i*,* d_j *is the capacity of machine* j*, and* z_{ij} *is a binary variable indicating whether task* i *is assigned to machine* j*. Setting* $S_j = \{z = (z_1, \ldots, z_m) \in \mathbb{R}^m \mid \sum_{i=1}^m w_{ij} z_i \leq d_j\}$ *for* $j = 1, \ldots, n$*. The corresponding master problem is*

$$\begin{aligned} \max \sum_{j=1}^{n} \sum_{k=1}^{K_j} \left(\sum_{i=1}^{m} p_{ij} y_{ik}^{j} \right) \lambda_k^j & \\ s.t. \sum_{j=1}^{n} \sum_{k=1}^{K_j} \lambda_k^j y_{jk} = 1, \quad & i = 1, \ldots, m \\ \sum_{k=1}^{K_j} \lambda_k^j = 1, \quad & j = 1, \ldots, n \\ \lambda_k^j \in \{0, 1\}, \quad & j = 1, \ldots, n, \quad k \in K_j \end{aligned} \tag{10.86}$$

where the first m *entries of a column, given by* $y_{jk} = (y_{jk}^1, y_{jk}^2, \ldots, y_{jk}^m)$*, satisfy the knapsack constraint*

$$\sum_{i=1}^{m} w_{ij} x_i \leq d_j, \quad x_i \in \{0, 1\}, \quad i = 1, \ldots, m$$

and where K_j *denotes the number of feasible solutions to the above knapsack constraint. In other words, a column represents a feasible assignment of tasks to a machine. Its LP relaxation is*

$$\begin{aligned} \max \sum_{j=1}^{n} \sum_{k=1}^{K_j} \left(\sum_{i=1}^{m} p_{ij} y_{ik}^{j} \right) \lambda_k^j & \\ s.t. \sum_{j=1}^{n} \sum_{k=1}^{K_j} \lambda_k^j y_{ik}^j = 1, \quad & i = 1, \ldots, m \\ \sum_{k=1}^{K_j} \lambda_k^j = 1, \quad & j = 1, \ldots, n \\ \lambda_k^j \geq 0, \quad & j = 1, \ldots, n, \quad k \in K_j \end{aligned}$$

Remark 10.26

- *The LP relaxation of the master problem solved by column generation may not have an integral optimal solution and applying a standard branch-and-bound procedure to the master problem over the existing columns is unlikely to find an optimal, or even good, or even feasible, solution to the original problem. Therefore it may be necessary to generate additional columns in order to solve the linear programming relaxations of the master problem at non-root nodes of the search tree.*

– *Standard branching on the z_{ij} variables creates a problem along a branch where a variable has been set to zero. Recall that y_k^j represents a particular solution to the jth knapsack problem. Thus $\lambda_k^j = 0$ means that this solution is excluded. However, it is possible (and quite likely) that the next time the jth knapsack problem is solved the optimal knapsack solution is precisely the one represented by y_k^j. In that case, it would be necessary to find the second best solution to the knapsack problem. At depth l in the branch-and-bound tree we may need to find the lth best solution. Fortunately, there is a simple remedy to this difficulty. Instead of branching on the αs in the master problem, we use a branching rule that corresponds to branching on the original variables z_{ij}. When $z_{ij} = 1$, all existing columns in the master that don't assign task i to machine j are deleted and task i is permanently assigned to machine j, i.e., variable x_i is fixed to 1 in the jth knapsack. When $z_{ij} = 0$, all existing columns in the master that assign job i to machine j are deleted and task i cannot be assigned to machine j, i.e., variable x_i is removed from the jth knapsack. Note that each of the knapsack problems contains one fewer variable after the branching has been done. Observe that since we know how to fix a single original variable z_{ij} to 1, we can also use branching rules based on fixing sets of original variables to 1.*

- *Identical Machines This is a special case of the problem with unrelated machines and therefore the methodology described above applies. However, we need only one subproblem since all of the machines are identical, which implies that the λ_k^j can be aggregated by defining $\lambda_k = \sum_j \lambda_k^j$ and that the convexity constraints can be combined into a single constraint $\sum_{k \in K} \lambda_k^j = n$ where λ_k is restricted to be integer. In some cases the aggregated constraint will become redundant and can be deleted altogether. An important issue here concerns symmetry, which causes branching on the original variables to perform very poorly. With identical machines, there are an exponential number of solutions that differ only by the names of the machines, i.e., by swapping the assignments of the machines we get solutions that are the same but have different values for the variables. This statement is true for fractional as well as integer solutions. The implication is that when a fractional solution is excluded at some node of the tree, it pops up again with different variable values somewhere else in the tree. In addition, the large number of alternate optima dispersed throughout the tree renders pruning by bounds nearly useless.*

10.5 Algorithmic schemes and branching strategies

General column generation cost scheme:

1. *Determine an initial feasible restricted master problem*

2. Initialize the column pool to be empty

3. Solve the current restricted master problem

4. Delete nonbasic columns with high negative reduced costs from the restricted master problem

5. If the column pool still contains columns with positive reduced costs select a subset of them add them to the restricted master and go to 3

6. Empty the column pool

7. Invoke an approximation algorithm for the pricing problem to generate one or more columns with positive reduced cost. If columns are generated add them to the column pool and go to 5.

8. Stop

Generic Branch-and-Price scheme:

1. Select an active node in the branch-and-bound tree.

2. Determine an initial feasible restricted master problem

3. Initialize the column pool to be empty

4. Solve the current restricted master problem

5. Delete nonbasic columns with high negative reduced costs from the restricted master problem

6. If the column pool still contains columns with positive reduced costs select a subset of them add them to the restricted master and go to 4

7. Empty the column pool

8. Invoke an approximation algorithm for the pricing problem to generate one or more columns with positive reduced cost. If columns are generated add them to the column pool and go to 6.

9. If the solution is integer, update the lower bound and set the node incactive. Otherwise, use a branching procedure to generate new active non-root nodes and go to step 1.

Branching strategies:

1. Branching strategy for the Generalized Assignment Problem (GPA) (and in general for set partitioning problems) based on the following proposition.

Proposition 10.27 ((Ryan and Foster(1981))) *If Y is a (matrix size) matrix and a basic solution to Y is fractional (i.e., at least one of the components of the solution is fractional), then there exist two rows r and s of the master problem such that:*

$$\sum_{k:y_{rk}=1,y_{sk}=1} \lambda_k < 1.$$

We can use the pair r, s to define the pair of branching constraints:

$$\sum_{k:y_{rk}=1,y_{sk}=1} \lambda_k = 1 \quad \text{and} \quad \sum_{k:y_{rk}=1,y_{sk}=1} \lambda_k = 0$$

i.e., the rows r and s have to be covered by the same column on the first (left) branch and by different columns on the second (right) branch. The proposition above implies that if no branching pair can be identified, then the solution to the master problem must be integer. Notice that the branch and bound algorithm must terminate after a finite number of branches since there are only a finite number of pairs of rows. This branching scheme requires that elements r and s belong to the same subset on the left branch and to different subsets on the right branch. Thus, on the left branch, all feasible columns must have $y_{rk} = y_{sk} = 0$ or $y_{rk} = y_{sk} = 1$ while on the right branch all feasible columns must have $y_{rk} = y_{sk} = 0$ or $y_{rk} = 0$, $y_{sk} = 1$ or $y_{rk} = 1$ or $y_{sk} = 0$. Instead of adding the branching constraints to the master problem explicitly, the infeasible columns in the master problem can be eliminated. This has the advantage of not introducing new dual variables that have to be dealt with in the pricing problem

2. *Branching strategy for general mixed integer master problems*

   - *Different restrictions on subsets: The optimal solution to the linear programming relaxation is infeasible if and only if*

     $$x_j = \sum_{k=1}^{p_j} \lambda_k^j y_k^j$$

     has a fractional component r for some j, say with value α. This suggests the following branching rule: on one branch we require

     $$\sum_{k=1}^{p_j} \lambda_k^j y_k^j \leq \lfloor \alpha \rfloor$$

     and on the other branch we require

     $$\sum_{k=1}^{p_j} \lambda_k^j y_k^j \geq \lceil \alpha \rceil.$$

     This branching rule amounts to branching on the original variable x_j.

- *Identical Restrictions on Subsets. If the solution to (reference) is fractional, we may be able to identify a single row r and an integer α_r such that*

$$\sum_{k:(a_k)_r \geq \alpha_r} \lambda_k = \beta_r$$

and β_r is fractional. We can then branch on the constraints

$$\sum_{k:(a_k)_r \geq \alpha_r} \lambda_k \leq \lfloor \alpha_r \rfloor$$

and

$$\sum_{k:(a_k)_r \geq \alpha_r} \lambda_k \geq \lceil \alpha_r \rceil.$$

These constraints place upper and lower bounds on the number of columns with $(a_k)_r \geq \alpha_r$ that can be present in the solution. In general, these constraints cannot be used to eliminate variables and have to be added to the formulation explicitly. Each branching constraint will contribute an additional dual variable to the reduced cost of any new column with $akr = \alpha_r$. This may complicate the pricing problem. Finally, it is easy to see that a single row may not be sufficient to define a branching rule.

10.6 Langrangean relaxation

Consider the LPP

$$\begin{array}{ll} \min & c^T x \\ s.t. & Ax \leq b \\ & B_i x_i = d_i \quad i = 1, \ldots, k. \\ & x \geq 0. \end{array} \tag{10.87}$$

where $x = (x_1, \ldots, x_k) \in \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_k} = \mathbb{R}^n$. We will call the first block of constraints "coupling contraints" since they establish relations that should hold for the whole set of variables. The second block of constraints have a separable structure which as we shall see is very convenient.

We will consider the following problem that we call a lagrangian relaxation of (**??**)

$$\begin{array}{l} \min L(x;p) = c^T x + p^T(b - Ax) \\ x \quad s.t.\ B_i x_i = d_i \ i = 1, \ldots, k. \\ \qquad\qquad x \geq 0. \end{array} \tag{10.88}$$

where $p \leq 0$ is a vector of Lagrange multipliers. Taking in to account to the particular structure of the constraints (separability) in the LPP (10.88), and denoting $A = [A_1, \ldots, A_k]$

the partition of the matrix A in blocks corresponding to the partition of the variable x, we can rewrite this problem as follows

$$\begin{array}{c} \min_x\ p^T b + \sum_{i=1}^{k} (c_i^T - A_i^T p)^T x_i \\ s.t.\ B_i x_i = d_i\ i = 1, \ldots, k. \\ x \geq 0. \end{array} \Leftrightarrow \begin{array}{c} p^T b + \min_x \sum_{i=1}^{k} (c_i^T - A_i^T p)^T x_i \\ s.t.\quad B_i x_i = d_i\ i = 1, \ldots, k. \\ x \geq 0. \end{array} \Leftrightarrow \begin{array}{c} p^T b\ +\ \sum_{i=1}^{k} \min_{x_i} (c_i^T - A_i^T p)^T x_i \\ s.t.\quad B_i x_i = d_i \\ x_i \geq 0. \end{array} \tag{10.89}$$

Assuming that all the subproblems in the last expression in (10.89) have solution, we will consider the lagrangian dual function $g : \mathbb{R}^m_- \to \mathbb{R}$ which will assign to each $p \leq 0$ the result of the above problem, i.e. $g(p) = \inf_{x \in \mathcal{X}} L(x; p)$ where $\mathcal{X} = \{x = (x_1, x_2, \ldots, x_k) \in \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_k} = \mathbb{R}^n \,|\, B_i x_i = d_i\ i = 1, \ldots, k,\ x \geq 0\}$. Notice that g is a closed concave function, since it is the supremum of family of linear functions.

Proposition 10.28 (Weak duality) *Let f^* be the optimal value of the LPP (10.87) and $g^* = \sup_{p \leq 0} g(p)$. Then $f^* \geq g^*$*

Simple calculations show that for any x feasible for the LPP (10.87) and any $p \leq 0$ it holds

$$c^T x + p^T (b - Ax) \leq c^T x$$

Denoting by C the set of feasible points for problem (10.87), we have that $C \subset \mathcal{X}$ and therefore it holds

$$f^* = \inf\{c^T x \,|\, x \in C\} \geq \inf\{c^T x + p^T (b - Ax) \,|\, x \in C\} \geq \inf\{c^T x + p^T (b - Ax) \,|\, x \in \mathcal{X}\} = g^*.$$

which ends the proof;

Proposition 10.29 (Strong duality) *(??) Let f^* and g^* be defined as in the previous proposition, and assume that f^* is finite. Then $f^* = g^*$*

Let x^* be an optimal solution of the LPP (10.87). Then, by Theorems 7.19 and 7.18, and after some calculations, it follows that there exists an optimal solution of the corresponding dual problem, $(\lambda_1^*, \lambda_2^*)$, which satisfies

$$\lambda_1^* \leq 0, \quad c^T f^* = b^t \lambda_1^* + d^T \lambda_2^*, \quad \lambda_1^*(Ax^* - b) = 0$$

11 Listas de exercícios

Lista 1

Convexity

Exercice 1. *Let A be an $m \times n$ matrix, let $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$. Solve the optimization problem: minimize $c^T x$ subject to the constraints $Ax = b$.*

Exercice 2. *Show that the following sets are convex:*

a) $\{\sum_{i=1}^n \lambda_i x_i : \lambda \geq 0, \sum_{i=1}^n \lambda_i = 1\}$ *where $x_1, \ldots, x_n$ are given in $\mathbb{R}^n$.*

b) $\{x \in \mathbb{R}^n : \|x - x_0\| \leq r\}$ *where $r > 0$, $x_0 \in \mathbb{R}^n$, and $\|\cdot\|$ is a norm on $\mathbb{R}^n$.*

c) $\{x \in \mathbb{R}^n : a^T x = b\}$ *and* $\{x \in \mathbb{R}^n : a^T x \leq b\}$ *where $a \in \mathbb{R}^n, b \in \mathbb{R}$.*

d) $\{x \in \mathbb{R}^n : Ax = b, Cx \leq d\}$.

Exercice 3. *Show that if a nonempty set is an intersection of convex sets then it is convex. Deduce that the set of semidefinite positive matrices is convex.*

Exercice 4. *Show that the set of optimal solutions of a convex optimization problem is convex.*

Exercice 5. *a) Show that if X is convex its interior is convex too.*
b) We define the ε-enlargement of set X by

$$X^\varepsilon = \{y : dist(y, X) \leq \varepsilon\}$$

for $\varepsilon > 0$ where dist is the distance function:

$$dist(y, X) = \left\{ \begin{array}{l} \inf\ \|y - x\| \\ x \in X. \end{array} \right.$$

Show that if X is convex then the set X^ε is convex too.
c) Let $B_F(0, \varepsilon) = \{x : \|x\| \leq \varepsilon\}$. Show that if X is convex and closed then $X^\varepsilon = X + \mathbb{B}_F(0, \varepsilon)$. If X is closed and convex, deduce another proof of the convexity of X^ε.

Exercice 6. *Let C be a convex set and let $\alpha_1, \alpha_2 > 0$. Show that*

$$(\alpha_1 + \alpha_2)C = \alpha_1 C + \alpha_2 C.$$

Find a set C for which the above relation does not hold.

Exercice 7. *Let S, T be two nonempty convex sets in $\mathbb{R}^n$. Show that a separates S, T if and only if*

$$\sup_{x \in S} a^T x \leq \inf_{y \in T} a^T y$$

and

$$\inf_{x \in S} a^T x < \sup_{y \in T} a^T y.$$

Show that the separation is strong if and only if

$$\sup_{x \in S} a^T x < \inf_{y \in T} a^T y.$$

Lista 2

Theorem on Alternative, cones, and conic programs

Exercice 1. *Let K be a regular cone and let $\bar{x} >_K 0$. Prove that $x >_K 0$ if and only if there exists a positive real t such that $x \geq_K t\bar{x}$.*

Exercice 2. *Let K be a regular cone. 1) Prove that if $x \neq 0 \geq_K 0$ and $\lambda >_{K_*} 0$ then $\langle \lambda, x \rangle > 0$.*

2) Assume that $\lambda \in K_*$. Prove that $\lambda >_{K_*} 0$ if and only if for every $x \neq 0$ with $x \geq_K 0$ we have $\langle \lambda, x \rangle > 0$.

3) Prove that $\lambda >_{K_*}$ if and only if the set

$$\{x \geq_K 0 \ : \ \langle \lambda, x \rangle \leq a\}$$

where a is a fixed postive real is compact.

Exercice 3. *Derive the General Theorem on Alternative from Homogeneous Farkas Lemma.*

Hint: Verify that the system

$$(\mathcal{S}) : \left\{ \begin{array}{l} a_i^T x > b_i, \ i = 1, \dots, m_s, \\ a_i^T x \geq b_i, \ i = m_s + 1, \dots, m, \end{array} \right.$$

in variable x has no solution if and only if the homogeneous inequality

$$\varepsilon \leq 0$$

in variables x, ε, t, is a consequence of the system of homogeneous inequalities

$$(\mathcal{T}) : \left\{ \begin{array}{rcl} a_i^T x - b_i t - \varepsilon & \geq & 0, \ i = 1, \dots, m_s, \\ a_i^T x - b_i t & \geq & 0, \ i = m_s + 1, \dots, m, \\ t & \geq & \varepsilon \end{array} \right.$$

in these variables.

Exercice 4. *Prove the following corollaries of General Theorem on Alternative:*

1) Gordan's Theorem on Alternative. One of the following systems of inequalities

$$(I) \ \ Ax < 0, \ x \in \mathbb{R}^n,$$

in variable x and

$$(II)\ \ A^T y = 0,\ y \neq 0,\ y \geq 0, y \in \mathbb{R}^m,$$

in variable y, with A an $m{\times}n$ matrix has a solution if and only if the other has no solution.

2) Inhomogeneous Farkas Lemma. A linear inequality $a^T x \leq p$ in variable x is a consequence of a solvable system of linear inequalities $Nx \leq q$ iff there exists $\nu \geq 0$ such that $a = N^T \nu$ and $\nu^T q \leq p$.

3) Motzkin's Theorem on Alternative. The system

$$Sx < 0,\ Nx \leq 0$$

in variables x has no solution if and only if the system

$$S^T \sigma + N^T \nu = 0,\ \sigma \geq 0,\ \nu \geq 0,\ \sigma \neq 0,$$

in variables σ, ν has a solution.

Exercice 5. *Show that if P is an orthogonal matrix then*

(i) $\|PA\|_2 = \|A\|_2$,

(ii) $\|PA\|_F = \|AP\|_F = \|A\|_F$.

Exercice 6. *For matrices A of size $m{\times}n$ and B of size $n{\times}q$ show that*

$$\|AB\|_F \leq \|A\|_2 \|B\|_F.$$

Exercice 7. *Let A be a semidefinite positive matrix and let B be definite positive. If $A \succeq B$ show that $B^{-1/2} A B^{-1/2} \succeq I \succeq A^{-1/2} B A^{-1/2}$.*

Exercice 8. *Let A, B be two positive definite matrices. If $A \succ B$ show that $B^{-1} \succ A^{-1}$. If $A \succeq B$ show that $B^{-1} \succeq A^{-1}$.*

Lista 3

Dualidade

- ***Exercise 1.*** *Considere o problema de programação linear (PPL)*

$$
\begin{aligned}
P) \qquad & \max \; x_1 - x_2 + x_3 + x_4 \\
& s.t \quad 2x_1 + 2x_2 + 3x_3 + 7x_4 = 16 \\
& \qquad 10x_1 + x_2 + x_3 + x_4 \geq 10 \\
& \qquad -2x_1 + x_2 + x_3 + x_4 \leq 1 \\
& \quad x_1 \geq 0, \; x_2 \geq 0, \; x_3 \leq 0.
\end{aligned}
$$

  a) Escreva o problema dual do problema $P)$.

  b) Verifique que o dual do dual é o primal.

- ***Exercise 2.*** *Considere o problema de programação linear (PPL)*

$$
\begin{aligned}
P) \qquad & \max \; 2x_1 + x_2 + x_3 \\
& s.t \quad x_1 + x_3 \leq 1 \\
& \qquad x_2 - x_3 \leq 1 \\
& \quad x_1, \; x_2, \; x_3 \geq 0
\end{aligned}
$$

  a) Escreva o problema dual do problema $P)$.

  b) Use o problema dual para encontrar o valor ótimo de $P)$. Justifique sua resposta.

- ***Exercise 3.*** *Considere o problema de programação linear (PPL)*

$$
\begin{aligned}
P) \qquad & \max \; x_1 + x_2 \\
& \quad x_1 - x_2 \leq -1 \\
& \quad x_2 - x_1 \leq -2 \\
& \quad x_1, \; x_2 \geq 0
\end{aligned}
$$

  a) Escreva o problema dual do problema $P)$.

  b) Mostre que nenhum dos problemas possui solução viável.

- ***Exercise 4.*** *Considere o problema de programação linear (PPL)*

$$\begin{aligned} \min \ & q^T z \\ s.t \ & Mz \geq -q \\ & z \geq 0 \end{aligned}$$

onde M é uma matriz "skew symmetric" , isto é, $M^T = -M$.

a) Mostre que o problema (P) e seu dual são iguais (isto é chamado de PPL auto-dual (self-dual)).

b) Mostre que um PPL auto-dual tem uma solução ótima se e somente se é viável.

Lista 4

SIMPLEX METHOD

Exercise 1. *Consider the problem*

$$\begin{aligned} \textit{minimize} \ & -2x_1 - x_2 \\ \textit{subject to} \ & x_1 - x_2 < 2 \\ & x_1 + x_2 < 6 \\ & x_1,\ x_2 \geq 0. \end{aligned}$$

(a) *Convert the problem into standard form and construct a basic feasible solution at which* $(x_1, x_2) = (0, 0)$.

(b) *Carry out the full tableau implementation of the simplex method, starting with the basic feasible solution of part (a).*

(c) *Draw a graphical representation of the problem in terms of the original variables* x_1, x_2, *and indicate the path taken by the simplex algorithm.*

Exercise 2. *Using the simplex procedure, solve minimize*

$$\begin{aligned} \textit{minimize}\ & 2x_1 + 4x_2 + x_3 + x_4 \\ \textit{subject to}\ & x_1 + 3x_2 + x_4 \leq 4 \\ & 2x_1 + x_2 \leq 3 \\ & x_2 + 4x_3 + x_4 \leq 3 \\ & x_i \geq 0 \ i = 1, 2, 3, 4. \end{aligned}$$

(a) *How much can the elements of* $b = (4, 3, 3)$ *be changed without changing the optimal basis?*

(b) *How much ca nth eelements of* $c = (2, 4, 1, 1)$ *be changed without changing the optimal basis?*

(c) *What happens to the optimal cost for small changes in* b*?*

(d) *What happens to the optimal cost for small changes in* c*?*

Exercise 3. *Consider the simplex method applied to a standard form problem and assume that the rows of the matrix* A *are linearly independent. For each of the statements that follow, give either a proof or a counterexample.*

(a) *An iteration of the simplex method may move the feasible solution by a positive distance while leaving the cost unchanged.*

(b) *A variable that has just left the basis cannot renter in the very next iteration.*

(c) *A variable that has just entered the basis cannot leave in the very next iteration.*

(d) *If there is a nondegenerate optimal basis, then there exists a unique optimal basis.*

(e) *If x is an optimal solution found by the simplex method, no more than m of its components can be positive, where m is the number of equality constraints.*

Exercise 4. *Using the two-phase simplex procedure solve*

$$\begin{aligned} \text{minimize} \quad & -3x_1 + x_2 + 3x_3 - x_4 \\ \text{subject to} \quad & x_1 + 2x_2 - x_3 + x_4 = 0 \\ & 2x_1 - 2x_2 + 3x_3 + 3x_4 = 9 \\ & x_1 - x_2 + 2x_3 - x_4 = 6 \\ & x_i \geq 0 \ i = 1, 2, 3, 4. \end{aligned}$$

Exercise 5. *The following tableau is an intermediate stage in the solution of a minimization problem:*

	y_1	y_2	y_3	y_4	y_5	y_6	y_0
	1	$2/3$	0	0	$4/3$	0	4
	0	$-7/3$	3	1	$-2/3$	0	2
	0	$-2/3$	-2	0	$2/3$	1	2
r^T	0	$8/3$	-11	0	$4/3$	0	-8

(a) *Determine the next pivot element.*

(b) *Given that the inverse of the current basis is*

$$B^{-1} = \frac{1}{3}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 1 \end{pmatrix}$$

and the corresponding cost coefficients are

$$c_B^T = (c_1, c_4, c_6) = (-1, -3, 1),$$

find the original problem.

Exercise 6. *Consider the following LPP*

$$\begin{aligned} &\text{minimize } -10x_1 - 12x_2 - 12x_3 \\ &\text{subject to } x_1 + 2x_2 + 2x_3 \leq 20 \\ &\qquad 2x_1 + x_2 + 2x_3 \leq 20 \\ &\qquad 2x_1 + 2x_2 + x_3 \leq 20 \\ &\qquad x_i \geq 0 \;\; i = 1, 2, 3 \end{aligned}$$

(a) *Verify that that the inverse of the matrix* $B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ *is given by* $B^{-1} = \begin{pmatrix} -0.6 & 0.4 & 0.4 \\ 0.4 & -0.6 & 0.4 \\ 0.4 & 0.4 & -0.6 \end{pmatrix}$

(b) *Prove that the basic solution that corresponds to the basis* B *is optimal.*

(c) *Compute the optimal solution and the optimal objective value.*

Exercise 7 *(Sensitivity with respect to changes in a basic column o f A) In this problem we study the change in the value of the optimal cost when an entry of the matrix* A *is perturbed by a small amount. We consider a linear programming problem in standard form, under the usual assumption that* A *has linearly independent rows. Suppose that we have an optimal basis* B *that leads to a nondegenerate optimal solution* x^**, and a nondegenerate dual optimal solution* p*. We assume that the first column is basic. We will now change the first entry of* A_1 *from a* a_{11} *to* $a_{11} + \delta$*, where* δ *is a small scalar. Let* E *be a matrix of dimensions* $m \times m$ *(where* m *is the number of rows of* A*), whose entries are all zero except for the top left entry* e_{11} *which is equal to* 1.

(a) *Show that if* δ *is small enough,* $B + \delta E$ *is a basis matrix for the new problem.*

(b) *Show that under the basis* $B + \delta E$*, the vector* x_B *of basic variables in the new problem is equal to* $(I + \delta B^{-1}E)B^{-1}b$.

(c) *Show that if* δ *is sufficiently small,* $B + \delta E$ *is an optimal basis for the new problem.*

Exercise 8 *Consider the following linear programming problem:*

$$\begin{aligned} &\text{minimize } 4x_1 + 5x_3 \\ &\text{subject to } 2x_1 + x_2 - 5x_3 = 1 \\ &\qquad -3x_1 + x_3 + x_4 = 2 \\ &\qquad x_1,\, x_2,\, x_3,\, x_4 \geq 0 \end{aligned}$$

(a) *Write down a simplex tableau and find an optimal solution. Is it unique?*

(b) *Write down the dual problem and find an optimal solution. Is it unique?*

(c) *Suppose now that we change the vector* b *from* $b = (1, 2)$ *to* $b = (1 - 2\theta, 2 - 3\theta)$*, where* θ *is a scalar parameter. Find an optimal solution and the value of the optimal cost, as a function of* θ*. (For all* θ*, both positive and negative.)*

Lista 5

Modeling 1

Exercício 1 *Uma fábrica possui duas usinas U1 e U2. A usina U1 dispõe de 400 unidades de um produto e a U2 de 300 unidades do mesmo produto. A fábrica tem três clientes E1, E2 e E3 cujas demandas respectivas para o produto são 100 unidades para E1, 200 unidades para E2 e 300 unidades para E3. Os custos de transportes são resumidos na tabela seguinte:*

	$E1$	$E2$	$E3$
$U1$	1	1.5	3.5
$U2$	2	1	2

Por exemplo, cada unidade fornecida a E1 a partir de U2 custa 2 reais. Como obter um sistema de distribuição ótimo?

Exercício 2 *Numa usina química, produzimos dois tipos de produtos a partir de três fertilizantes. O produto I, composto por um quilograma de nitratos e dois quilogramas de sal de potássio é vendido por R\$7, e o produto II, composto de um quilograma de nitratos, um quilograma de fosfatos e três quilogramas de sal de potássio é vendido por R\$9. Sobram no estoque 8kg de nitratos, 4kg de fosfatos e 19kg de sal de potássio.*

1. *Qual quantidade de cada produto a empresa tem que produzir para maximizar o lucro?*
2. *Uma cooperativa agrícola quer negociar (i.e., minimizar) o preço de um quilograma de cada componente para comprar a granel todos os fertilizantes do estoque. Como determinar os preços para que a venda a granel seja pelo menos tão lucrativa como a venda dos produtos?*

Exercício 3 *Um modelo de carro é montado em 3 usinas situadas em cidades V1, V2 e V3. O motor destes modelos é fornecido por duas outras usinas situadas nas cidades U1 e U2. As usinas de montagem precisam de pelo menos 5, 4 e 3 motores. Cada usina pode fornecer no máximo 6 motores. A direç ao da empresa quer minimizar o custo de transporte dos motores entre os dois sítios de fábrica e os três sítios de montagem. Os custos unitários (por motor transportado) para todos os itinerários possíveis são:*

	$V1$	$V2$	$V3$
$U1$	38	27	48
$U2$	37	58	45

Como minimizar o custo total de transporte respeitando a oferta e a demanda?

Como minimizar o custo total de transporte respeitando a oferta e a demanda?

Exercício 4 *Um sapateiro faz 6 sapatos por hora, se fizer somente sapatos e 5 cintos por hora, se fizer somente cintos. Ele trabalha 10 horas por dia e gasta 2 unidades de couro para fabricar 1 unidade de sapato e 1 unidade de couro para fabricar 1 unidade de cinto. Sabendo-se que o total disponível de couro é de 78 unidades por dia e que o lucro unitário por sapato é de 5 reais e o de cinto é de 4 reais, pede-se: o modelo do sistema de produç ao diário do sapateiro, se o objetivo é maximizar seu lucro diário.*

Exercício 5 *Certo fabricante de combustível para avião vende 2 tipos de combustível, A e B. O combustível de tipo A possui 25% de gasolina 1, 25% de gasolina 2 e 50% de gasolina 3. O combustível B tem 50% de gasolina 2 e 50% de gasolina 3. Há disponível para produção 500 galões de gasolina 1 e 200 galões de cada gasolina 2 e 3. Os lucros pela venda dos combustíveis A e B são, respectivamente, 20 e 30 dólares. Quanto se deve fazer de cada combustível para se obter um lucro máximo? Formule e resolva o*

Exercício 6 *Uma fábrica de petróleo deseja utilizar quatro tipos de petróleos para produzir três tipos de diesel: A, B, e C. A respeito dos tipos de petróleo, temos as seguintes informações:*

Tipo de petróleo	*Quantidade máxima disponível por dia*	*Custo (reais por barril)*
1	3000	3
2	2000	6
3	4000	4
4	1000	5

O diesel A não pode conter mais de 30% do petróleo do tipo 1, nem mais de 50% do tipo 3, mas deve conter no mínimo 40% do tipo 2. O preço de venda deste diesel é de 5.5 reais por barril.

O diesel B, cujo preço de venda é 4.5 reais por barril, deverá ser composto de pelo menos 10% do tipo 2 mas no máximo de 50% do tipo 1.

O diesel C não poderá conter mais de 70% do petróleo do tipo 1 e o seu preço de venda é de 3.5 reais por barril.

A fábrica gostaria de saber a quantidade de barris de cada tipo de petróleo que deveria ser utilizada na fabricação de cada um dos tipos de diesel para poder maximizar seu lucro.

qtd de energia gerada por cada usina
↪ var. dec.

Lista 6

Modeling 2

Exercício 1 *Gestão da produção de eletricidade. As usinas térmicas e hidrelétricas brasileiras podem ser agrupadas em 4 subsistemas intercambiando energia entre eles. Supomos que cada região contém uma usina hidrelétrica (agregação das usinas hidrelétricas deste subsistema) e um número arbitrário de usinas térmicas. Cada usina hidrelétrica funciona com um reservatório. Somente 80% das afluências de uma região dada é armazenado nos reservatórios. O restante é diretamente convertido em eletricidade por usinas ao fil d'água. O custo de produção da eletricidade com as usinas térmicas é uma função linear da produção enquanto ele é considerado nulo com usinas hidrelétricas. Cada usina tem uma capacidade de produção conhecida e os níveis dos reservatórios devem ficar entre determinados valores mínimos e máximos. Cada dia, a demanda dos clientes deve ser atendida, eventualmente comprando energia no mercado spot a um custo unitário mais alto que o maior custo unitário das térmicas.*

Explicar como determinar as produções diárias das usinas térmicas e hidrelétricas para o mês seguinte, de maneira a minimizar o custo e satisfazendo a demanda e as restrições de funcionamento das usinas.

Exercício 2 *Gestão de carteiras. Queremos investir M reais em n ativos financeiros. O retorno do ativo i no período de investimento é r_i. Escrever um problema de otimização linear que permita determinar a quantidade de dinheiro a investir em cada ativo para maximizar o lucro. Qual é a solução deste problema?*

Exercício 3 *Planificação da expansão da produção. Consideramos o problema de expansão da capacidade de produção de uma usina produzindo m produtos. Cada uma das n máquinas é flexível e cada produto pode ser produzido por qualquer máquina. A máquina j está agora disponível para h_j horas de funcionamento por semana e horas adicionais podem ser adquiridas num custo atualizado de c_j por hora. O uso da máquina j é limitado por uma cota superior de u_j horas, por outra parte, uma revisão de t_j horas da máquina j é necessária para cada hora de funcionamento. O tempo total gasto em revisão não pode ultrapassar T horas. A taxa de produção do produto i na máquina j é a_{ij}, com um custo associado de g_{ij} por hora.*

Cada semana, a empresa deve satisfazer a demanda em cada um dos m produtos. Cada unidade de produto i não vendida acarreta um custo p_i. A empresa quer decidir quantas horas adicionais são necessárias para cada máquina com os dados a seguir:

– $n = 4$, $m = 3$, $T = 100$, $p_i = (400, 400, 400)$;

– $c_j = (2.5, 3.75, 5.0, 3.0)$, $t_j = (0.08, 0.04, 0.03, 0.01)$;

– $h_j = (500, 500, 500, 500)$, $u_j = (2000, 2000, 3000, 3000)$;

– $[a_{ij}] = \begin{pmatrix} 0.6 & 0.6 & 0.9 & 0.8 \\ 0.1 & 0.9 & 0.6 & 0.8 \\ 0.05 & 0.2 & 0.5 & 0.8 \end{pmatrix}$

– $[g_{ij}] = \begin{pmatrix} 2.6 & 3.4 & 3.4 & 2.5 \\ 1.5 & 2.4 & 2.0 & 3.6 \\ 4.0 & 3.8 & 3.5 & 3.2 \end{pmatrix}$

– *as demandas nos diferentes produtos (numa semana dada) são dadas por* {*1800, 600, 3000*}.

Exercício 4 *Gestão de contratos com opção de cancelamento. Uma empresa deve entregar cada mês, via um gasoduto (visto como um armazém de gás com capacidades mínimas e máximas), uma determinada quantidade de gás a seus clientes. Para isto, ela dispõe de um armazém de gás e de um contrato (já pago) com um país produtor de gás que garante uma determinada chegada de gás cada mês. Este gás pode ser enviado no armazém ou no gasoduto diretamente. Além disto, a empresa passou contratos de Gás Natural Liquefeito (GNL) com opção de cancelamento. Cada um destes contratos permite a entrega de uma determinada quantidade de gás, para uma data fixada no contrato. Porém, esta carga pode ser cancelada até um mês antes do dia da entrega, pagando uma multa, dependendo do momento em que é feito o cancelamento (quanto mais tarde o cancelamento, mais elevada a multa). O preço pago pelo GNL é o preço spot do gás natural no dia da entrega. O GNL é entregue por navios que ficam no porto até serem esvaziados. Para evitar que os navios permaneçam muito tempo no porto, um custo (função linear do armazém) é pago cada mês para o gás que sobra nos navios. O gás é vendido aos clientes 30 acima do preço spot.*

Escrever um programa de otimização linear que permite determinar que contratos têm que cancelar assim como os fluxos de gás na rede de modo a maximizar o lucro satisfazendo as restrições do sistema.

Exercício 5 *Problema da mochila. Mickey está preparando sua mochila para um trekking na Cordilheira dos Andes. Cada objeto que ele pode levar tem uma certa utilidade (expressa por um número positivo). Cada objeto tem um peso conhecido e Mickey não quer carregar mais de P kg. Escrever um problema de otimização que permita determinar os objetos a serem colocados na mochila de modo a maximizar a utilidade. Como modificar este problema se tomarmos em consideração o volume de cada objeto; o volume da mochila sendo V?*

Exercício 6 *Problema de "Unit commitment". Uma empresa dispõe de 10 usinas térmicas que estão por enquanto desligadas. As usinas devem ser usadas para satisfazer, para cada um dos meses do ano seguinte, as demandas em eletricidade dos clientes da empresa. Para cada usina, o custo de produção é uma função linear da produção. Por outra parte, ligar uma usina acarreta um custo fixo que depende da usina. Uma vez ligada, a usina fica funcionando até o final do ano. Escrever um problema de otimização que permita saber que usinas ligar e quando, assim como as produções das usinas ligadas de modo a satisfazer as demandas minimizando o custo de produção.*

Lista 7

Branch and Bound

Exercise 1. *Solve the following ILPP using Branch and Bound*

$$
\begin{array}{llrl}
\textit{Problem N:} & \min & x_1 - 2x_2 & \\
 & \textit{subject to} & -4x_1 + 6x_2 & \leq 9 \\
 & & x_1 + x_2 & \leq 4 \\
 & & x_1,\ x_2 & \geq 0 \\
 & & x_1,\ x_2 & \textit{integer.}
\end{array}
\tag{11.90}
$$

12 Provas

EMAp-FGV ***2024***

Primeira prova de Programação Linear - Duração 3h
Abril 2024

Todas as respostas devem ser justificadas.

Exercício 1. 2 pontos. *Uma empresa pode escolher quatro tipos de líquidos : 8 000 litros do líquido A ao custo unitario 5,50 \$, 4 250 litros de B ao custo unitario 4,50 \$, 16 000 litros de C ao custo unitario 7,50 \$, e 2 000 litros de D ao custo unitario 11,25 \$.*

A empresa pode revender estes líquidos diretamente, sem transforma-los, e vende-los por 6 \$ por litro.

Ela pode também elaborar as misturas E, F e G. As misturas devem apresentar as características dadas na tabela 3.

Mistura	Líquido A	Líquido B	Líquido C	Líquido D
E	30%	Pelo menos 10%	40%	No máximo 5%
F	Pelo menos 25%	No máximo 20%	20%	Pelo menos 10%
G	20%	Pelo menos 15%	40%	No máximo 20%

Table 3: Proporção de cada líquido num litro de mistura

As misturas se vendem respectivamente 11 \$, 15 \$ e 14 \$ por litro, e o mercado pode comprar todas as misturas produzidas.

A empresa é obrigada a produzir pelo menos 400 litros de E, pelo menos 800 litros de F e pelo menos 200 litros de G.

Enfim, misturando 2 partes de G com uma parte de E, podemos obter um produto P vendido 22 \$ por litro e cuja demande é suficientemente grande para ser considerada ilimitada.

Modelar este problema por um problema de otimização linear dado que a empresa quer maximizar seu lucro.

Exercício 2. 1.75 pontos. *A companhia energética Dark necessita realizar o planejamento energético para um novo prédio. A energia necessária é classificada em 3 categorias: (a) iluminação; (b) aquecimento ambiente; (c) aquecimento água. As necessidades mensais de cada categorias são:*

Iluminação	*20MW*
Aquecimento ambiente	*10MW*
Aquecimento água	*30MW*

3 fontes de energia podem ser instaladas para suprir a energia necessária: (a) eletricidade; (b) painéis solares; (c) gás natural. O suprimento máximo mensal de energia de cada fonte é:

Eletricidade	*50MW*
Painéis solares	*50MW*
Gás natural	*20MW*

A iluminação só pode ser suprida pela energia "eletricidade", a um custo de R\$50 por MW. As demais categorias ("aquecimento ambiente" e "aquecimento água") podem ser supridas por qualquer fonte. Os custos unitários de suprimento para estas categorias (em R\$/MW) são os seguintes:

	Eletricidade	*Gás natural*	*Painéis solares*
Aquecimento ambiente	*90*	*60*	*30*
Aquecimento água	*80*	*50*	*40*

O objetivo é minimizar o custo total de energia mensal do prédio. Escrever um programa linear para modelar este problema.

Exercício 3. Cones.

Exercício 4. 2 pontos. *O epígrafo de uma função $f : \mathbb{R}^n \to \mathbb{R}$ é o conjunto*

$$epi(f) = \left\{(x, y) \in \mathbb{R}^{n+1} \,|\, y \geq f(x)\right\}.$$

Uma função $f : \mathbb{R}^n \to \mathbb{R}$ é convexa se

$$f\left(\alpha x + (1-\alpha)y\right) \leq \alpha f(x) + (1-\alpha) f(y) \;\forall\, x, y \in \mathbb{R}^n, \;\forall\, \alpha \in [0,1].$$

a) *Prove que se $f : \mathbb{R}^n \to \mathbb{R}$ é uma função convexa e contínua então $epi(f)$ é igual a seu fecho convexo.*

b) *Mostre que o interior relativo do conjunto $epi(f)$ é o conjunto*

$$ri\,(epi(f)) = \left\{(x, y) \in \mathbb{R}^{n+1} \,|\, y > f(x)\right\}.$$

b) *Suponha que f é diferenciável num ponto x_0 e que $H = \{(x, y) \,|\, \langle (d, 1), (x, y)\rangle = 0\}$ é um hiperplano suporte do $epi(f)$ no ponto $(x_0, f(x_0))$. Prove que o vetor d é paralelo ao vetor $\nabla f(x_0)$*

Exercício 5. 2 pontos. *Considere o PPL*

$$\begin{array}{lll} P) & \max & 4.2976x_1 + 2.7x_2 + 2.5x_3 + 2.1976x_4 + 4.4976x_5 \\ & s.a. & x_1 + x_2 + x_3 + x_4 + x_5 \leq 10, \\ & & -x_2 - x_3 - x_4 \leq -4, \\ & & 6x_1 + 6x_2 - 4x_3 - 4x_4 + 36x_5 \leq 0, \\ & & 4x_1 + 10x_2 - x_3 - 2x_4 - 3x_5 \leq 0, \\ & & x_i \geq 0 \; i = 1,2,3,4,5. \end{array}$$

a) *Escreva o problema dual de P).*

b) *Sabendo que a solução do problema dual de P) é* $y^* = (2.94, 0, 0.0636, 0.244)$, *determine o valor ótimo e uma solução óptima de P).*

c) *Seja* P_1) *o novo PPL obtido ao multiplicar a primeira inequação do problema* P) *por* 10. *Encontre uma solução ótima do dual do problema* P_1).

Exercício 6. 1 ponto. *Mostre que um conjunto* $H \subset \mathbb{R}^n$ *é um hiperplano se e somente se ele é da forma* $H = \varphi^{-1}(\alpha)$ *onde* $\varphi : X \to \mathbb{R}$ *é linear não nula e* $\alpha \in \varphi(\mathbb{R}^n)$

13 References

1. *A. Ben-Tal and A. Nemirovski. Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, 2001. isbn: 9780898718829. url: https://books.google.com.br/books?id=CENjbXz2SDQC.*

2. *Dimitris Bertsimas and John Tsitsiklis. Introduction to Linear Optimization. 1st. Athena Scientific, 1997. isbn: 1886529191.*

3. *S.P. Bradley, A.C. Hax, and T.L. Magnanti. Applied Mathematical Programming. Addison-Wesley Publishing Company, 1977. isbn: 9780201004649. url: https://books.google.com.br/books?id=MSWdWv3Gn5cC.*

4. *D.G. Luenberger and Y. Ye. Linear and Nonlinear Programming. International Series in Operations Research & Management Science. Springer US, 2008. isbn: 9780387745022. url: https://books.google.com.br/books?id=-pD62uvi9lgC.*

5. *George B. Dantzig, Mukund N. Thapa. Linear Programming 1. Springer Series in Operations Research and Financial Engineering. Springer New York, NY. ISBN 978-0-387-94833-1. https://doi.org/10.1007/b97672*

6. *Branch-and-Price: Column Generation for Solving Huge Integer Programs Cynthia Barnhart, Ellis L. Johnson, George L. Nemhauser, Martin W. P. Savelsbergh and Pamela H. Vance Operations Research Vol. 46, No. 3 (May - Jun., 1998), pp. 316-329 (14 pages)*