

Optimization Handbook

Henri Lefebvre

May 26, 2019

Contents

I	Linear Optimization	2
1	Introduction	3
2	Lagrangian duality	4
3	The Simplex algorithm	5
4	Relaxation techniques	6
5	The branch-and-bound algorithm	7
6	The branch-and-cut algorithm	8
7	Column generation and the branch-and-price algorithm	9
8	Benders decomposition	10
1	Introduction	10
2	Formal derivation	10
9	Dantzig-Wolfe decomposition	12
II	Non-linear Optimization	13
	Appendices	14
A	Convex sets	15
1	Polyhedra	15
1.1	Definitions	15
1.2	Theorems	15

Part I

Linear Optimization

Chapter 1

Introduction

Chapter 2

Lagrangian duality

Chapter 3

The Simplex algorithm

Chapter 4

Relaxation techniques

Chapter 5

The branch-and-bound algorithm

Chapter 6

The branch-and-cut algorithm

Chapter 7

Column generation and the branch-and-price algorithm

Chapter 8

Benders decomposition

1 Introduction

Benders decomposition is a solution method for solving certain large-scale optimization problems. It is particularly suited for problems in which a set of variables are said to be *complicating* in the sense that fixing them to a given value makes the problem easy. Briefly, the Benders decomposition approach separates an original problem into several decision stages. A first-stage *master* problem is solved using only a subset of variables, then, the values of the remaining variables are determined by a so-called *subproblem* depending on the first-stage variables. If the master problem's optimal solution yields an infeasible subproblem, a *feasibility cut* is added to the master problem, which is then re-solved. Due to the structure of the reformulation, the Benders algorithm starts with a *restricted master problem* where only a subset of constraints are considered while the others are iteratively added.

This technique was first introduced in [1] and has since been generalized to non-linear mixed integer problems.

2 Formal derivation

Consider the following problem :

$$\text{minimize } c^T x + f(y) \quad (8.1)$$

$$\text{s.t. } Ax + g(y) = b \quad (8.2)$$

$$y \in Y, x \geq 0 \quad (8.3)$$

where variable y is a *complicating constraint*. Note that it may be complicating due to the form of f or g but also by our ability to enforce the constraint $y \in Y$. We assume that fixing y to a given value \hat{y} turns our problem into an easy-to-solve problem.

We can notice that our problem is equivalent to the following one :

$$\min \{ f(y) + \min \{ c^T x : Ax = b - g(y) \} : y \in Y \}$$

Let us denote by $q(y)$ the value of the minimization problem over x : $q(y) = \min \{ c^T x : Ax = b - g(y) \}$. By duality, the following holds

$$q(y) = \max \{ (b - g(y))^T \pi : A^T \pi \leq c \}$$

Note that the feasibility space of the dual does not depend on the values of y and let us apply the decomposition theorem for polyhedra 3 on it :

$$\{ A^T \pi \leq c \} = \left\{ \sum_i u^i \alpha_i + \sum_j v^j \beta_j \right\}$$

$$\sum_i \alpha_i = 1 \quad \text{and} \quad \alpha, \beta \geq 0$$

where $\{u^i\}_i$ denotes the extreme points of $\{x | Ax \leq c\}$ and $\{v^j\}_j$ denotes the extreme rays of the polyhedral cone $\{x | Ax \leq 0\}$. Intuitively, the convex combination of the extreme points of $\{x | Ax \leq c\}$ defines the optimal solutions (since we know that there exists at least one optimal solution corresponding to an extreme point of the considered polyhedron) while the conical combination of extreme rays of $\{x | Ax \leq 0\}$

defines the feasibility region. The following theorem will allow us to reformulate our problem :

Theorem 1. *Let (\mathcal{P}) be the following problem :*

$$\max\{c^T x : Ax \leq b, x \geq 0\} \quad (\mathcal{P})$$

Then, (\mathcal{P}) is upper bounded if and only if

$$c^T v^j \leq 0 \quad \forall j = 1, \dots, J$$

where $\{v^j\}_{j=1, \dots, J}$ denotes the set of extreme rays of $\{x | Ax \leq 0\}$.

Proof. \Rightarrow : By contradiction, let us suppose that (\mathcal{P}) is upper bounded and that there exists k such that $c^T v^k > 0$. Let us consider a feasible solution to (\mathcal{P}) denoted by $u = z + t$ where z is a convex combination of the extreme points of $\{x | Ax \leq b\}$ and t an element of the conical combinations of the extreme rays of $\{x | Ax \leq 0\}$ (see theorem 3). Let $\lambda \in \mathbb{R}_+$. Since v^k is in the conical polyhedra $\{x | Ax \leq 0\}$, it holds that $\lambda A v^k \leq 0$. Moreover, we have $A t \leq 0$. Hence, $A(t + \lambda v^k) \leq 0$. Now, since $v^k \geq 0$, we have $\lambda v^k \geq 0$ and since $t \geq 0$ it holds that $t + \lambda v^k \geq 0$. This shows that, for any λ , $z + \lambda v^k$ is a feasible solution for problem (\mathcal{P}) . However, its associated objective value is given by $c^T u + \lambda c^T v^k \rightarrow +\infty$ when $\lambda \rightarrow +\infty$ since, by assumption, $c^T v^k > 0$. This contradicts the fact that (\mathcal{P}) is upper bounded.

\Leftarrow : Let us consider a solution $u = z + t$ then $c^T u = c^T z + \sum_j c^T v^j \beta_j \leq c^T z$ since $c^T v^j \leq 0, \forall j = 1, \dots, J$. Problem (\mathcal{P}) is therefore upper bounded by $\sum_i \alpha_i \max_i \{c^T u^i\}$. \square

Theorem 1 can be intuitively understood by interpreting $c^T v^j$ as $\langle c, v^j \rangle$ (scalar product). The necessary and sufficient condition that $\langle c, v^j \rangle \leq 0$ simply expresses that the hyperplane defining the objective function must be *oriented* (i.e., increasing) in the direction of the origin of the cone formed by the extreme rays of the polyhedron. It is depicted in figure 8.1.

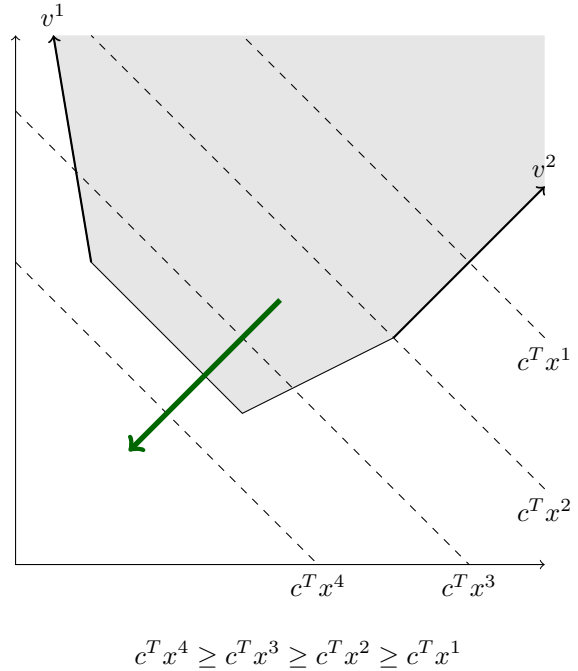


Figure 8.1: Illustration of theorem 1

Chapter 9

Dantzig-Wolfe decomposition

Part II

Non-linear Optimization

Appendices

Appendix A

Convex sets

1 Polyhedra

1.1 Definitions

Combinations and hulls

First, the following definitions will allow us to introduce some geometrical notions :

Definition 1 (Convex combination). *Let x_1, \dots, x_n be a finite set of vectors in a real vector space, a convex combination of these vectors is a vector of the form*

$$\sum_{i=1}^n \alpha_i x_i \quad \text{with } \sum_{i=1}^k \alpha_i = 1, \alpha \geq 0$$

Definition 2 (Convex hull).

$$\text{conv}(X) = \left\{ \sum_{i=1}^n \alpha_i x_i \mid x_i \in X, \sum_{i=1}^n \alpha_i = 1, \alpha \geq 0 \right\}$$

Definition 3 (Conical combination). *Let x_1, \dots, x_n be a finite set of vectors in a real vector space, a conical combination of these vectors is a vector of the form*

$$\sum_{i=1}^k \alpha_i x_i \quad \text{with } \alpha \geq 0$$

Definition 4 (Conical hull).

$$\text{cone}(X) = \left\{ \sum_{i=1}^n \alpha_i x_i \mid x_i \in X, \alpha_i \geq 0 \right\}$$

Polyhedra, polytopes and polyhedral cones

Definition 5 (Polyhedron). *(plural : polyhedra)*

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b\}$$

Definition 6 (Polytope). *A polytope is the convex set a finite number of points. Alternatively,*

$$P = \{x \in \mathbb{R}^n \mid Ax \leq b, u^- \leq x \leq u^+\}$$

Definition 7 (Polyhedral cone).

$$P = \{x \in \mathbb{R}^n \mid Ax \leq 0, x \geq 0\}$$

Observation 1 (Important). *Note that no real consensus has been reached on the definitions of polyhedra and polytopes. Sometimes, polyhedra corresponds to three dimensional solids with polygonal faces while polytopes denote the extension of a polyhedra to higher dimensions.*

It comes from these definitions that every polytopes and every polyhedral cones are polyhedra. These geometrical objects are depicted in figure A.1.

1.2 Theorems

We first recall the following definition in order to properly enounce some important theorems.

Definition 8 (Minkowski sum). *Let A and B be two vector spaces, the Minkowski sum is defined as*

$$A + B = \{a + b \mid a \in A, b \in B\}$$

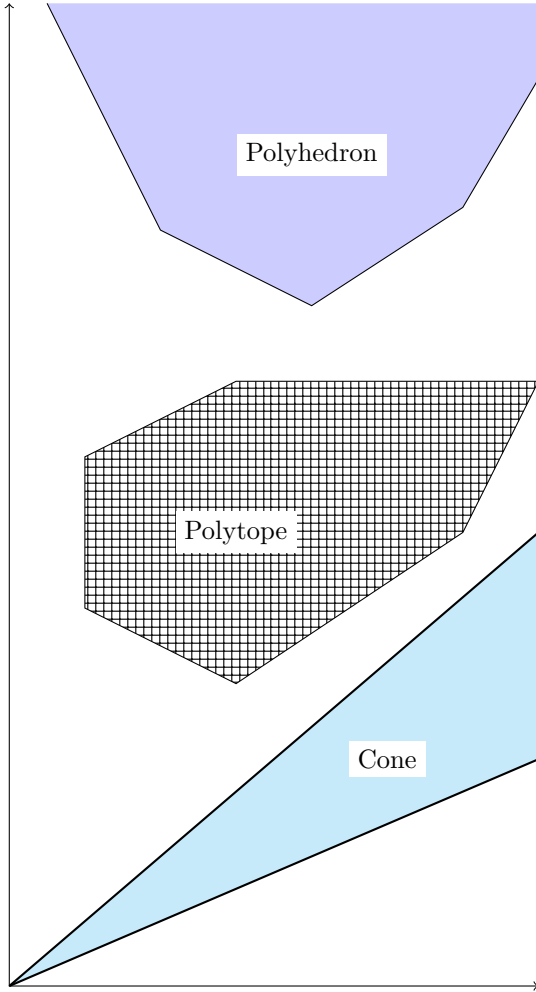


Figure A.1: A polyhedron, a polytope and a polyhedral cone depicted in 2D

Theorem 2 (Affine Minkowski-Weyl). *Let there be a polyhedron defined by a set of inequalities, $P = \{x \in \mathbb{R}^n | Ax \leq b\}$. There exists vectors $x_1, \dots, x_q \in \mathbb{R}^n$ and $y_1, \dots, y_r \in \mathbb{R}^n$ such that*

$$P = \text{cone}(x_1, \dots, x_q) + \text{conv}(y_1, \dots, y_r)$$

Theorem 3 (Decomposition theorem for polyhedra). *A set P of vectors in a Euclidean space is a polyhedron if and only if it is the Minkowski sum of a poly-*

tope Q and a polyhedral cone C .

$$P = Q + C$$

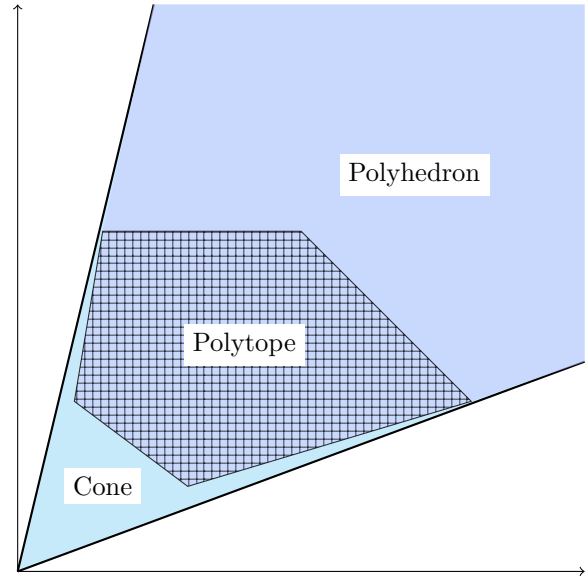


Figure A.2: Illustration of theorem 3

This decomposition theorem is illustrated in figure A.2. Intuitively, the polytope defines the *lower* shape of the polyhedron and the polyhedral cone defines the *queue* of the polyhedron.

Bibliography

- [1] J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, Dec 1962.