Operations Research - LINMA 2491

Course 3 - Insights on Extended Formulations

Mehdi Madani

UCLouvain & N-SIDE

February 2023

Table of contents

Polyhedra, polytopes and cones

Definition (Polyhedron)

A polyhedron $P \subseteq \mathbb{R}^n$ is a set that can be described by a finite set of linear inequalities:

$$P = \{x \mid A_i x \leq b_i, i = 1, ..., m\} \tag{1}$$

N.B. Alternative definitions can be given, see the Minkowski-Weyl theorem below.

Definition

A polytope is a bounded polyhedron

Definition

The finitely generated cone $C \subseteq \mathbb{R}^n$ by a finite set of vectors $r_1, ..., r_k$ is the cone $C = cone(\{r_1, ..., r_k\}) = \{x \in \mathbb{R}^n \mid x = \mu_1 r_1 + ... + \mu_k r_k, \mu \geq 0\}$

The Minkowski-Weyl theorem

Theorem (The Minkowski-Weyl theorem)

A set P is a polyhedron if and only if P can be described as

$$P = Q + C \tag{2}$$

with $Q = conv(\{v_1, ..., v_K\})$ and $C = cone(\{r_1, ..., r_H\})$

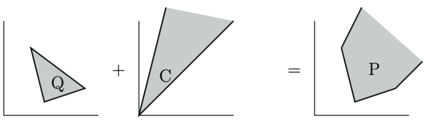


Figure: Illustration of the Minkowski-Weyl theorem taken from [1]

A polyhedron is a polytope (i.e. is bounded) if and only if it is the convex hull of a finite set of points

$\mathcal{H}$-representations and $\mathcal{V}$-representations

In view of the Minkowski-Weyl theorem, a polyhedron can be described in two different ways:

1. $\mathcal{H}$-representations: as the intersection of a finite number of half-spaces

2. $\mathcal{V}$-representations: as the convex hull of a finite set of points (only vertices = extreme points are necessary) + the conic hull of a finite set of vectors (only the "extreme rays" are necessary)

Moving from a $\mathcal{H}$-representation to a $\mathcal{V}$-representation is called a vertex enumeration problem.

Moving from a $\mathcal{V}$-representation to a $\mathcal{H}$-representation is called a convex hull problem (standard problem in computational geometry).

Perfect formulations of Mixer Integer Linear Sets

A perfect formulation for a set $X \subseteq \mathbb{R}^n$ is a description $Ax \leq b$ such that $conv(X) = \{x \mid Ax \leq b\}$

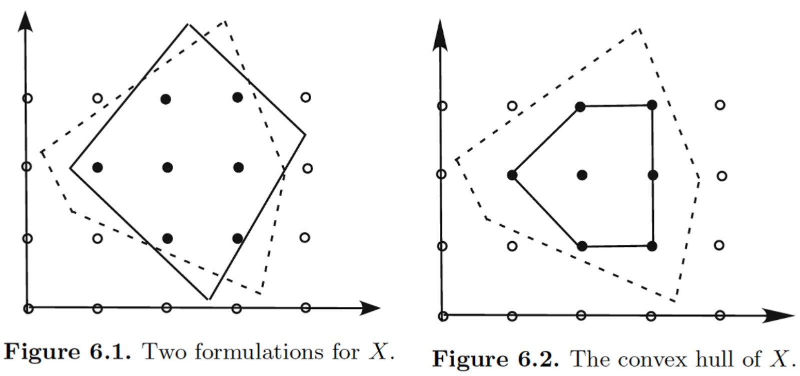


Figure: A formulation and a perfect formulation for X, taken from [3].

A perfect formulation might require a large number of constraints and fully characterizing such constraints might be difficult (there are problems for which a tight explicit formulation of X in the "original space" of the variables x is not known), and it may be easier to describe $conv(X)$ as the projection of a higher-dimensional, but easier to describe, polyhedron.

Projections

Consider a polyhedron $P = \{(x, y) \mid Ax + Gy \leq b\}$.

Its orthogonal projection on the x space (i.e. on the subspace $\{(x, y) \mid y = 0\}$) is $Proj_x(P) = \{x | \exists y, Ax + Gy \leq b\}$

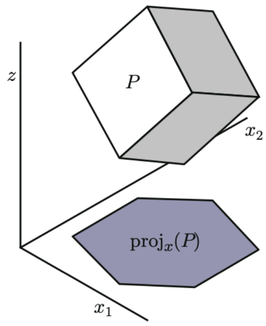


Figure: Illustration of an orthogonal projection, taken from [1]

Observation

$$\max\{c^T x \mid x \in Proj_x(P)\} = \max\{c^T x + 0y \mid (x, y) \in P\} \tag{3}$$

Extended Formulations

Definition (Extended formulation of a polyhedron)

An extended formulation $P = \{(x, y) \mid Ax + Gy \leq b\}$ of a polyhedron Q is a (higher-dimensional) polyhedron such that $Proj_x(P) = Q$.

Definition (Extended formulation for mixed integer linear sets)

A polyhedron $P = \{(x, y) \mid Ax + Gy \leq b\}$ is an extended formulation of a mixed integer linear set $X = \{x \mid Cx \leq d, x_1, ..., x_p \in \mathbb{Z}\}$ if $Proj_x(P) \cap \mathbb{Z}^p = X$, i.e. $Proj_x(P)$ is a *formulation* of X (in the sense that $Proj_x(P)$ combined with the integer requirements describe X)

Definition (Tight Extended formulations)

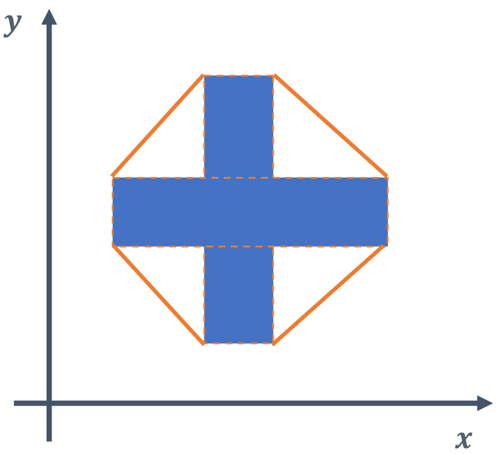
A polyhedron $P = \{(x, y) \mid Ax + Gy \leq b\}$ is a *tight* extended formulation of a mixed integer linear set $X = \{x \mid Cx \leq d, x_1, ..., x_p \in \mathbb{Z}\}$ if $Proj_x(P) = conv(X)$.

Extended Formulations are interesting when they are given by a "small" number of constraints and variables, compared to their projection of interest which would require much more constraints/variables.

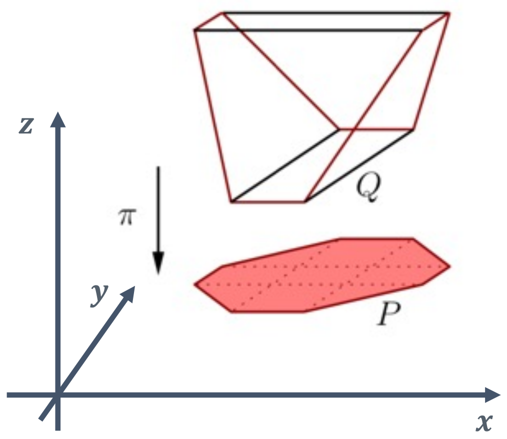
Illustration

Convex hull of the union of 2 polytopes

Same convex hull described as the projection of an Extended Formulation



Convex hull described by 8 constraints and 2 variables (x,y)



Same convex hull can be described by an Extended Formulation with 6 constraints and 3 vars (x,y,z)

Illustration on the right is from: https://arxiv.org/pdf/1107.0371.pdf

Production Planning example

A manufacturer produces bicycles[1]. Forecasts of the demand are given, corresponding to a non-constant demand over the coming months. The objective is to set a production plan given the following information:

- 1 batch per month at most: due to special equipment and tools to install at the beginning of each batch, at most one batch of bicycles is produced per month
- Setup cost of 5000 €: installation of equipment and tools
- Marginal cost of 100 €: in a batch, each bicycle costs 100€
- Forecasted demand:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
400	400	800	800	1200	1200	1200	1200

- Inventory costs: 5€ per bicycle (capital and storage costs)
- Initial stock contains 200 bicycles
- Assumption: inventory evolves linearly from one period to the next, so that inventory costs depend on the average $\dfrac{INV_{t-1} + INV_t}{2}$

[1]Small example taken from [3]

A MIP for our tiny example

MC marginal cost, *FC* the setup cost, D_t is the demand given above for each period and *h* the inventory cost.

- Decision variables?
  - x_t number of bicycles produced at period t
  - y_t binary variable *indicating* if a batch is produced at period t (indicator variable)
  - s_t, inventory level at period t
- Objective? Minimize total costs :=
  - $\sum_{t=1}^{N}(MC\ x_t + FC\ y_t) + \sum_{t=1}^{N} hs_t$
- Constraints
  - $s_{t-1} + x_t = D_t + s_t$, for all $t = 1, ..., N$ (demand satisfaction)
  - $x_t \leq\ y_t(\sum_{k=t}^{N} D_t)$, for all $t = 1, ..., N$ (variable upper bound)
  - $s_0 = 200, s_N = 0$ (zero initial and final inventories)
  - $x_t, s_t \geq 0, y_t \in \{0, 1\}$

Uncapacitated Lot Sizing problem (LS-U)

$$\min_{x,y,s} \sum_{t=1}^{N}(p_t\ x_t + q_t\ y_t + h_t s_t) \tag{4}$$

subject to:

$$s_{t-1} + x_t = D_t + s_t \qquad \forall t \in PERIODS \tag{5}$$

$$x_t \leq (\sum_{k=t}^{N} D_t)y_t \qquad \forall t \in PERIODS \tag{6}$$

$$s_0 = s_{ini}, s_N = 0 \tag{7}$$

$$x_t, s_t \geq 0 \qquad \forall t \in PERIODS \tag{8}$$

$$y_t \in \{0, 1\} \qquad \forall t \in PERIODS \tag{9}$$

In the small example: $N = 8, s_{init} = 200, p = 100, q = 5000, h = 5$, $D = [400, 400, 800, 800, 1200, 1200, 1200, 1200]$

"Facility Location" extended formulation for LS-U

It can be shown that the following extended formulation for LS-U is tight:

$$s_{u-1} + x_u = D_u + s_u \qquad \forall u \in PERIODS \tag{10}$$

$$\sum_{u=1}^{t} w_{ut} = D_t \qquad \forall u \in PERIODS \tag{11}$$

$$w_{ut} \leq D_t y_u \qquad \forall 1 \leq u \leq t \leq NT \tag{12}$$

$$x_u = \sum_{t=u}^{NT} w_{ut} \qquad \forall u \in PERIODS \tag{13}$$

$$s_0 = s_{ini}, s_N = 0 \tag{14}$$

$$w_{ut}, s_t \geq 0 \qquad \forall t \in PERIODS \tag{15}$$

$$y_t \in \{0,1\} \qquad \forall t \in PERIODS \tag{16}$$

- For any objective function, we can optimize the continuous relaxation and obtain an optimal solution with y_t binary, solving the original problem.
- Let us again verify this on the small example, using a computer

Retrieving an integer feasible optimal solution

Consider a mixed integer linear set $Q = \{x \mid Ax \leq b, x_1, ..., x_p \in \mathbb{Z}\}$ and a tight extended formulation P for Q. One has:

$$\max\{c^T x \mid x \in Q\} = \max\{c^T x \mid x \in conv(Q)\} = \max\{c^T x \mid x \in Proj_x(P)\} \tag{17}$$

With $conv(Q)$, we know that vertices are "integral" (the variable values $x_1, ..., x_p$ are integer values). However, even though $Proj_x(P) = conv(Q)$, the projection of a vertex of $Proj_x(P)$ may not be a vertex of $conv(Q)$.

Hence, we may need some extra work to be able to retrieve an optimal solution which is a vertex $conv(Q)$, see Figure 5.

Vertices of an EF versus vertices of the projection

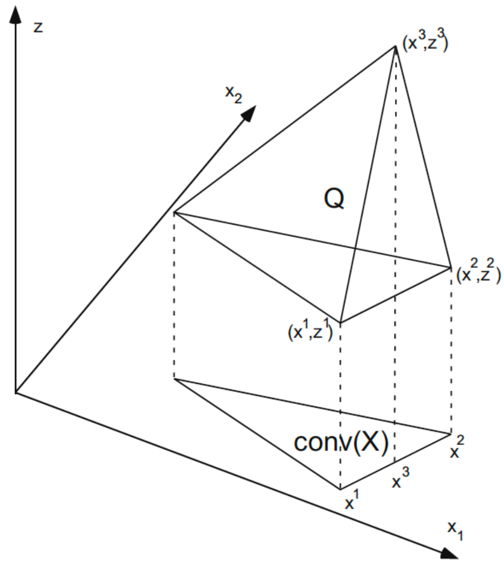


Figure: Illustration of vertices of an extended formulation not necessarily vertices of the projection, taken from [3].

Union of polyhedra

The notes focus on polytopes (i.e. bounded polyhedra), but also point how to consider more generally polyhedra.
Consider N *non-empty* polytopes $P_i = \{x \mid A^i x^i \leq b^i\}, i = 1, ..., N$. One can easily express that a point x belongs to the union of these polytopes as follows:

$$\cup_i P_i = Proj_x(X), where\ X\ is\ given\ by : \tag{18}$$

$$x = \sum_i x^i \tag{19}$$

$$A^i x^i \leq b^i\ y_i \tag{20}$$

$$\sum_i y_i = 1 \tag{21}$$

$$y \in \{0, 1\}^N \tag{22}$$

N.B. Since we consider non-empty polytopes, if $y_i = 0$, (20) will imply $x_i = 0$ (why?).

N.B. A union of polytopes/polyhedra correspond to a logical disjunction ($\wedge$ means 'OR') $p_1 \wedge ... \wedge p_N$ where p_i is the statement $x_i \in P_i$

Balas formulation for the convex hull of an union of polyhedra

It turns out that the continuous relaxation below (dropping the binary requirements and letting $y_i \in [0,1]$) describes $\overline{conv(\cup_i P_i)}$:

$$\overline{conv(\cup_i P_i)} = Proj_x(X), \textit{where } X \textit{ is given by :} \quad (23)$$

$$x = \sum_i x^i \quad (24)$$

$$A^i x^i \leq b^i \, y_i \qquad \forall i \quad (25)$$

$$\sum_i y_i = 1 \quad (26)$$

$$y_i \leq 1 \qquad \forall i \quad (27)$$

$$y \geq 0 \quad (28)$$

N.B. The above result holds true for polyhedra as well, in which case considering $\overline{conv(\cup_i P_i)}$ instead of $conv(\cup_i P_i)$ is important (see next slides). For polytopes however, $\overline{conv(\cup_i P_i)} = conv(\cup_i P_i)$.

Proof: Exercise. (Hint. Start for example with 2 polytopes and express that a point is a convex combination of points in both polytopes.) □

Importance of considering $\overline{conv(\cup_i P_i)}$ instead of $conv(\cup_i P_i)$ for general polyhedra in Balas formulation

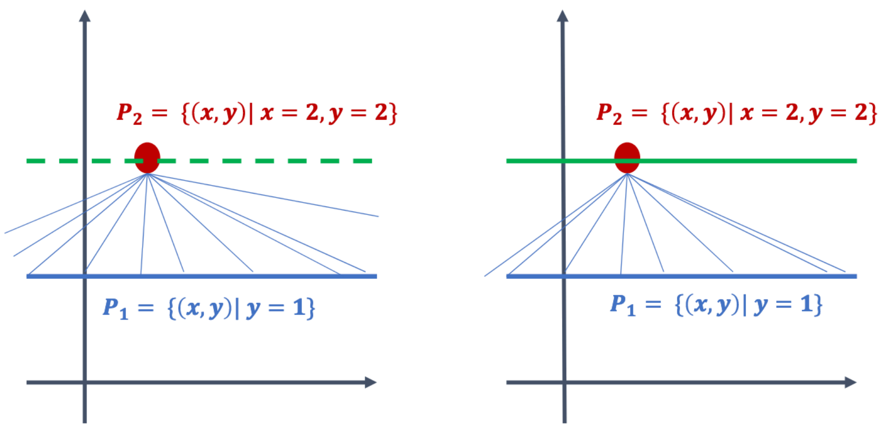


Figure: Illustration of a difference between $\overline{conv(\cup_i P_i)}$ and $conv(\cup_i P_i)$

$$conv(P_1 \cup P_2) = \{(x,y) \mid y \geq 1, y < 2\} \cup \{(2,2)\}$$

$$\overline{conv(P_1 \cup P_2)} = \{(x,y) \mid 1 \leq y \leq 2\}$$

Application to Financial (electricity) Transmission Rights (FTRs)

Separate deck of slides.

Generalization of Balas formulation for the union of polyhedra

The Balas formulation above for the convex hull of the union of polytopes (restated below) essentially describes $y_1 P_1 + ... + y_N P_N$ for

$$y \in \{y_i \mid y_i \in [0,1], \sum_i y_i = 1\} = conv(\{y_i \mid y_i \in \{0,1\}, \sum_i y_i = 1\}) \quad (29)$$

.

Balas formulation:

$$x = \sum_i x^i \quad (30)$$

$$A^i x^i \leq b^i \; y_i \qquad \forall i \quad (31)$$

$$\sum_i y_i = 1 \quad (32)$$

$$y_i \leq 1 \qquad \forall i \quad (33)$$

$$y \geq 0 \quad (34)$$

N.B. $y_1 P_1 + ... + y_N P_N$ denotes $\{y_1 x_1 + ... + y_N x_N \mid x_i \in P_i, \; i = 1, ..., N\}$
What can we say if we replace (29) by $y \in Y$ with Y some other integral polyhedron Y (i.e. whose vertices are integral) ?

Generalization of Balas formulation for the union of polyhedra

Consider non-empty polytopes $P_i = \{x \mid A^i x \leq b^i\}$, $i = 1, \ldots, m$ and $Y \subseteq \mathbb{R}^m_+$ a non-empty polytope (in the non-negative orthant) with integer vertices and which contains at least one $y > 0$ (i.e., $y_i > 0 \forall i = 1, \ldots, m$).

Consider Q defined by

$$A^i x^i \leq b^i y_i \qquad i = 1, \ldots, m \tag{35}$$

$$x = \sum_{i=1}^{m} x^i \tag{36}$$

$$(y_1, \ldots, y_m) \in Y \tag{37}$$

Then

1. $Proj(Q) = P := \bigcup_{y \in Y} (\sum_{i=1}^{m} y_i P_i)$
2. all the vertices of Q have integer $(y_1, \ldots, y_m)$

See [2] and references therein.

References I

[1] M. Conforti, G. Cornuéjols, G. Zambelli, et al.
Integer programming, volume 271.
Springer, 2014.

[2] B. Knueven, J. Ostrowski, and J. Wang.
The ramping polytope and cut generation for the unit commitment problem.
INFORMS Journal on Computing, 30(4):739–749, 2018.

[3] Y. Pochet and L. A. Wolsey.
Production Planning by Mixed Integer Programming.
Springer, 2006.