MVE165/MMG631 Linear and Integer Optimization with Applications Lecture 3 Extreme points of convex polyhedra; reformulations; basic feasible solutions

Ann-Brith Strömberg

2018-03-23

1 Course evaluation

2 Convex polyhedra

- Convex polyhedra
- Extreme points of convex polyhedra
- Optimal extreme points

8 Reformulation of linear optimization models

- Notation of linear optimization
- The standard form of linear optimization
- Reformulations to the standard form of linear optimization

4 Basic feasible solution

- Basic solution
- Basic feasible solution
- Algebraic description of basic feasible solutions
- Example: basic feasible solutions

Course evaluation

- The first meeting will be held on XXXday, March XX at X.XX
- The second meeting will be held during week 17 (April, 23-27)
- Notes will be published in the course's PingPong event
- Any voluntary representative from GU is also welcome! Anyone?

Contact any student representative to present your opinion:

- Gagandeep Bhatia (MPCAS)
- Marielle Cederlund (TKTEM)
- Konstantinos Konstantinou (MPENM)
- Emma Nirvin (TKTEM)
- Samuel Sjöström (TKTEM)

Linear optimization models = linear programs (LP)

A linear optimization model: c_j , a_{ij} , b_i : constant parameters

$$\begin{array}{ll} \mbox{minimize} & z = \sum_{j=1}^n c_j x_j \\ \mbox{subject to} & \sum_{j=1}^n a_{ij} x_j \leq b_i, & i = 1, \dots, m \\ & x_j \geq 0, & j = 1, \dots, n \end{array}$$

In vector notation: $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\begin{array}{ll} \mbox{min} & z = \mathbf{c}^\top \mathbf{x} \\ s.t. & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}^n \end{array}$$

A polyhedron

The *feasible region* of a linear optimization model is defined as the intersection of halfspaces in \mathbb{R}^n defined by its constraints.

The feasible region is a polyhedron, denoted
$$X \subset \mathbb{R}^n_+$$

$$X := \left\{ \mathbf{x} \in \mathbb{R}^n_+ \mid \sum_{j=1}^n a_{ij} x_j \le b_i, i = 1, \dots, m \right\} \equiv \left\{ \mathbf{x} \ge \mathbf{0}^n \mid \mathbf{A}\mathbf{x} \le \mathbf{b} \right\}$$

Convex combinations (Ch. 4.1)

Definition (Convex combination (Def. 4.1))

A convex combination of the points $\mathbf{x}^p \in \mathbb{R}^n$, p = 1, ..., P, is any point $\mathbf{x} \in \mathbb{R}^n$ that can be expressed as

$$\mathbf{x} = \sum_{p=1}^{P} \lambda_p \mathbf{x}^p$$

where the following constraints hold:

$$\sum_{p=1}^{P} \lambda_p = 1;$$
 $\lambda_p \ge 0, \quad p = 1, \dots, P$

The variables λ_p are called *convexity weights*

Convex sets (Ch. 2.4)

Definition (Convex set (Def. 2.5))

A set $X \in \mathbb{R}^n$ is a *convex set* if, for any two points $\mathbf{x}^1 \in X$ and $\mathbf{x}^2 \in X$, and any $\lambda \in [0, 1]$, it holds that

$$\mathbf{x} := \lambda \mathbf{x}^1 + (1 - \lambda) \mathbf{x}^2 \in X$$

Linear programs and convex polyhedra (Ch. 4.1)

Let $\mathbf{x} := \lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2$, where \mathbf{x}^1 and \mathbf{x}^2 are feasible, i.e., $\mathbf{x}^1 \ge \mathbf{0}^n$, $\mathbf{x}^2 \ge \mathbf{0}^n$, $\mathbf{A}\mathbf{x}^1 \le \mathbf{b}$, and $\mathbf{A}\mathbf{x}^2 \le \mathbf{b}$.

The intersection of linear constraints forms a convex set

The feasible region of a linear program is a *convex set*, since for any two feasible points x^1 and x^2 and any $\lambda \in [0, 1]$ it holds that

$$\sum_{j=1}^{n} a_{ij} x_j = \sum_{j=1}^{n} a_{ij} \left(\lambda x_j^1 + (1-\lambda) x_j^2 \right)$$
$$= \lambda \sum_{j=1}^{n} a_{ij} x_j^1 + (1-\lambda) \sum_{j=1}^{n} a_{ij} x_j^2$$
$$\leq \lambda b_i + (1-\lambda) b_i$$
$$= b_i, \qquad i = 1, \dots, m$$

and

$$x_j = \lambda x_j^1 + (1 - \lambda) x_j^2 \geq 0,$$
 $j = 1, \dots, n$

[Draw on the board]

Extreme points (Ch. 4.1)

Definition (Extreme point (Def. 4.2))

The point \mathbf{x}^k is an *extreme point* of the polyhedron X if $\mathbf{x}^k \in X$ and it is *not* possible to express \mathbf{x}^k as a *strict convex combination* of two distinct points in X.

I.e: Given
$$\mathbf{x}^1 \in X$$
, $\mathbf{x}^2 \in X$, and $0 < \lambda < 1$, it holds that $\mathbf{x}^k = \lambda \mathbf{x}^1 + (1 - \lambda)\mathbf{x}^2$ only if $\mathbf{x}^k = \mathbf{x}^1 = \mathbf{x}^2$ hold.

Optimal solution and extreme points (Ch. 4.1)

Theorem (Optimal solution in an extreme point (Th. 4.2))

Assume that the feasible region $X = \{\mathbf{x} \ge \mathbf{0}^n \mid \mathbf{A}\mathbf{x} \le \mathbf{b}\}$ is non-empty and bounded. Then, the minimum value of the objective $\mathbf{c}^{\top}\mathbf{x}$ is attained at (at least) one extreme point \mathbf{x}^k of X.

Proof of Theorem 4.2

Assume the opposite, that there is a *non-extreme point* $\tilde{\mathbf{x}} \in X$ with a lower objective value than any of the *extreme points*, i.e.,

$$\mathbf{c}^{\top} \tilde{\mathbf{x}} < \mathbf{c}^{\top} \mathbf{x}^{k}$$
 for all extreme points \mathbf{x}^{k} of X. (1)

Since the polyhedron X is a convex set, the point $\tilde{\mathbf{x}}$ can be expressed as a convex combination of the extreme points of X, i.e.,

$$\tilde{\mathbf{x}} = \sum_{k=1}^{p} \lambda_k \mathbf{x}^k; \qquad (2)$$

$$\sum_{k=1}^{p} \lambda_k = 1 \tag{3}$$

$$\lambda_k \geq 0, \qquad k=1,\ldots,p$$
 (4)

where p is the number of extreme points. Then, it must hold that

$$\mathbf{c}^{\top}\tilde{\mathbf{x}} \underbrace{=}_{(2)} \mathbf{c}^{\top} \sum_{k=1}^{p} \lambda_{k} \mathbf{x}^{k} = \sum_{k=1}^{p} \lambda_{k} \mathbf{c}^{\top} \mathbf{x}^{k} \underbrace{>}_{(1),(4)} \sum_{k=1}^{p} \lambda_{k} \mathbf{c}^{\top} \tilde{\mathbf{x}} = \mathbf{c}^{\top} \tilde{\mathbf{x}} \sum_{k=1}^{p} \lambda_{k} \underbrace{=}_{(3)} \mathbf{c}^{\top} \tilde{\mathbf{x}}$$

which is a contradiction.

A general linear program – notation

Definition (Notation of linear programs)

minimize or maximize $c_1x_1 + \ldots + c_nx_n$

subject to
$$a_{i1}x_1 + \ldots + a_{in}x_n \quad \left\{ \begin{array}{c} \leq \\ = \\ \geq \end{array} \right\} \quad b_i, \quad i = 1, \ldots, m$$

$$x_j \quad \left\{ \begin{array}{l} \leq 0 \\ \text{unrestricted in sign} \\ \geq 0 \end{array} \right\}, \quad j = 1, \dots, n$$

The blue notation corresponds to the standard form

The standard form of the simplex method for linear programs (Ch. 4.2)

- Every linear program can be reformulated such that:
 - all constraints are expressed as *equalities* with *non-negative right hand sides*
 - all variables involved are restricted to be *non-negative*
- Referred to as the *standard form*
- These requirements streamline the calculations of the *simplex method*
- Software solvers (e.g., Cplex, GLPK, Clp, Gurobi, SCIP) handle also inequality constraints and unrestricted variables – the reformulations are made automatically

The simplex method—standard form reformulations

• Slack variables:

$$\left[\begin{array}{ccc}\sum_{j=1}^{n}a_{ij}x_{j} \leq b_{i}, \ \forall i\\x_{j} \geq 0, \ \forall j\end{array}\right] \Longleftrightarrow \left[\begin{array}{ccc}\sum_{j=1}^{n}a_{ij}x_{j} + s_{i} = b_{i}, \ \forall i\\x_{j} \geq 0, \ \forall j\\s_{i} \geq 0, \ \forall i\end{array}\right]$$

• The lego example:

$$\begin{bmatrix} 2x_1 & +x_2 \le & 6\\ 2x_1 & +2x_2 \le & 8\\ & x_1, x_2 \ge & 0 \end{bmatrix} \iff \begin{bmatrix} 2x_1 & +x_2 & +s_1 & = & 6\\ 2x_1 & +2x_2 & +s_2 & = & 8\\ & & x_1, x_2, s_1, s_2 \ge & 0 \end{bmatrix}$$

 s₁ and s₂ are called *slack variables*—they "fill out" the (positive) distances between the left and right hand sides

The simplex method—standard form reformulations

• Surplus variables:

$$\left[\begin{array}{ccc} \sum_{j=1}^n a_{ij}x_j & \geq & b_i, \quad \forall i \\ x_j & \geq & 0, \quad \forall j \end{array}\right] \iff \left[\begin{array}{ccc} \sum_{j=1}^n a_{ij}x_j & -s_i & = b_i, \quad \forall i \\ x_j & & \geq 0, \quad \forall j \\ & s_i & \geq 0, \quad \forall i \end{array}\right]$$

• Surplus variable *s*₃ (another instance):

$$\left[\begin{array}{cccc} x_1 & + & x_2 & \ge & 800 \\ & x_1, x_2 & \ge & 0 \end{array}\right] \iff \left[\begin{array}{cccc} x_1 & + & x_2 - & \mathbf{s_3} & = & 800 \\ & & x_1, x_2, \mathbf{s_3} & \ge & 0 \end{array}\right]$$

The simplex method—standard form reformulations

• Suppose that b < 0:

$$\left[egin{array}{c} \sum\limits_{j=1}^n a_j x_j \leq b \ x_j \geq 0, orall j \end{array}
ight] \Longleftrightarrow \left[egin{array}{c} \sum\limits_{j=1}^n (-a_j) x_j \geq -b \ x_j \geq 0, orall j \end{array}
ight] \Longleftrightarrow \left[egin{array}{c} -\sum\limits_{j=1}^n a_j x_j & -s & = -b \ x_j & \sum & 20, orall j \end{array}
ight] \ s & \geq 0, orall j \end{array}
ight]$$

• Non-negative right hand side:

$$\begin{bmatrix} x_1 - x_2 \leq -23 \\ x_1, x_2 \geq 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} -x_1 + x_2 \geq 23 \\ x_1, x_2 \geq 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} -x_1 + x_2 - s_4 = 23 \\ x_1, x_2, s_4 \geq 0 \end{bmatrix}$$

Evaluation Polyhedra Reformulation BFS

Notation Standard form Reformulation

The simplex method—standard form reformulations

 Suppose that some of the variables are unconstrained (here: k < n). Replace x_j with x_j¹ - x_j² for the corresponding indices:

$$\begin{bmatrix} \sum_{j=1}^{n} a_j x_j \le b \\ x_j \ge 0, j = 1, \dots, k \end{bmatrix} \iff \begin{bmatrix} \sum_{j=1}^{k} a_j x_j + \sum_{j=k+1}^{n} a_j (x_j^1 - x_j^2) + s &= b \\ x_j \ge 0, \ j = 1, \dots, k, \\ x_j^1 \ge 0, x_j^2 \ge 0, \quad j = k+1, \dots, n \\ s \ge 0 \end{bmatrix}$$

• Sign-restricted (non-negative) variables:

$$\begin{bmatrix} x_1 + x_2 \le 10 \\ x_1 \ge 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} x_1 + x_2^1 - x_2^2 \le 10 \\ x_1, x_2^1, x_2^2 \ge 0 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} x_1 + x_2^1 - x_2^2 + s_5 = 10 \\ x_1, x_2^1, x_2^2, s_5 \ge 0 \end{bmatrix}$$

Basic solutions (Ch. 4.3)

- Consider *m* equations with *n* variables, where $m \le n$
- Set n m variables to zero and solve (if possible) the remaining $(m \times m)$ system of equations
- If the solution is *unique*, it is called a *basic* solution

Definition (Def. 4.3)

A *basic* solution to the $m \times n$ system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ is obtained if n - m of the variables are set to 0 and the remaining variables get their unique values from the solution to the remaining $m \times m$ system of equations.

The variables that are set to 0 are called *nonbasic variables* and the remaining *m* variables are called *basic variables*.

Basic feasible solutions (BFS) (Ch. 4.3)

- A basic solution **x** corresponds to the *intersection* of *m* hyperplanes in \mathbb{R}^m
 - It is *feasible* if $\mathbf{x} \ge \mathbf{0}$
 - It is *infeasible* if $x \not\geq 0$
- Each extreme point of the feasible set is an intersection of m hyperplanes such that all variable values are ≥ 0
- Basic feasible solution \iff extreme point of the feasible set

$$\begin{array}{ll} a_{11}x_1 + \ldots + a_{1n}x_n = b_1 & x_1 \ge 0 \\ a_{21}x_1 + \ldots + a_{2n}x_n = b_2 & x_2 \ge 0 \\ & \ddots & & \ddots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = b_m & x_n \ge 0 \end{array}$$

Basic feasible solutions – algebraic descriptions

Assume that
$$m < n$$
 and that $b_i \ge 0$, $i = 1, ..., m$, and let
 $\mathbf{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$, $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

Consider the linear program to

 $\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & z = \mathbf{c}^{\top}\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$

• Partition **x** into *m* basic variables \mathbf{x}_B and n - m non-basic variables \mathbf{x}_N , such that $\mathbf{x} = (\mathbf{x}_B, \mathbf{x}_N)$.

• Analogously, let $\mathbf{c} = (\mathbf{c}_B, \mathbf{c}_N)$ and $\mathbf{A} = (\mathbf{A}_B, \mathbf{A}_N) \equiv (\mathbf{B}, \mathbf{N})$

• The matrix $\mathbf{B} \in \mathbb{R}^{m \times m}$ with inverse \mathbf{B}^{-1} (if it exists)

Basic feasible solutions – algebraic descriptions (Ch. 4.8)

Rewrite the linear program equivalently as

$$\begin{array}{lll} \mbox{minimize} & z = \mathbf{c}_B^\top \mathbf{x}_B + \mathbf{c}_N^\top \mathbf{x}_N & (5a) \\ \mbox{subject to} & \mathbf{B} \mathbf{x}_B + \mathbf{N} \mathbf{x}_N &= \mathbf{b} & (5b) \\ & \mathbf{x}_B \ge \mathbf{0}^m, \ \mathbf{x}_N &\ge \mathbf{0}^{n-m} & (5c) \end{array}$$

• Multiply the system of equations (5b) by ${\bf B}^{-1}$ from the left:

$$\mathbf{B}^{-1}\mathbf{B}\mathbf{x}_{B} + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_{N} = \mathbf{x}_{B} + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_{N} = \mathbf{B}^{-1}\mathbf{b}$$
$$\implies \mathbf{x}_{B} = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_{N} = \mathbf{B}^{-1}(\mathbf{b} - \mathbf{N}\mathbf{x}_{N})$$
(6)

• Replace \mathbf{x}_B in (5a) by the expression in (6):

$$\mathbf{c}_B^\top \mathbf{x}_B + \mathbf{c}_N^\top \mathbf{x}_N = \mathbf{c}_B^\top \mathbf{B}^{-1} (\mathbf{b} - \mathbf{N} \mathbf{x}_N) + \mathbf{c}_N^\top \mathbf{x}_N = \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{b} + (\mathbf{c}_N^\top - \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{N}) \mathbf{x}_N$$

$$\begin{array}{ll} \Rightarrow & \textit{minimize} \quad z = \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{b} + (\mathbf{c}_N^\top - \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{N}) \mathbf{x}_N \\ & \text{subject to} & \mathbf{B}^{-1} \mathbf{b} - \mathbf{B}^{-1} \mathbf{N} \mathbf{x}_N & \geq \mathbf{0}^m, \ \mathbf{x}_N \geq \mathbf{0}^{n-m} \end{array}$$

Basic feasible solutions – algebraic descriptions

The rewritten program

$$\begin{array}{ll} \text{minimize} \quad z = \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{b} + (\mathbf{c}_N^\top - \mathbf{c}_B^\top \mathbf{B}^{-1} \mathbf{N}) \mathbf{x}_N & (7a) \\ \text{subject to} & \mathbf{B}^{-1} \mathbf{b} - \mathbf{B}^{-1} \mathbf{N} \mathbf{x}_N \geq \mathbf{0}^m & (7b) \\ \mathbf{x}_N \geq \mathbf{0}^{n-m} & (7c) \end{array}$$

At the basic solution defined by $B \subset \{1, \ldots, n\}$:

- Each non-basic variable takes the value 0, i.e., $\mathbf{x}_N = \mathbf{0}$
- The basic variables take the values $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N = \mathbf{B}^{-1}\mathbf{b}$
- The value of the objective function is $z = \mathbf{c}_B^{\top} \mathbf{B}^{-1} \mathbf{b}$
- The basic solution is feasible if $\mathbf{B}^{-1}\mathbf{b} \ge \mathbf{0}^m$

Basic feasible solutions, example

0

• Constraints:

Add slack variables:



Basic and non-basic variables and solutions

basic	 basic solution 		non-basic	point	feasible?	
variables				variables (0,0)		
<i>s</i> ₁ , <i>s</i> ₂ , <i>s</i> ₃	23	6	85	x_1, x_2	А	yes
s_1, s_2, x_1	$-5\frac{1}{3}$	$4\frac{1}{9}$	$28\frac{1}{3}$	<i>s</i> ₃ , <i>x</i> ₂	Н	no
s_1, s_2, x_2	23	$-4\frac{5}{8}$	$10\frac{5}{8}$	<i>x</i> ₁ , <i>s</i> ₃	С	no
<i>s</i> ₁ , <i>x</i> ₁ , <i>s</i> ₃	-67	90 [°]	-185	<i>s</i> ₂ , <i>x</i> ₂	- I	no
<i>s</i> ₁ , <i>x</i> ₂ , <i>s</i> ₃	23	6	37	s_2, x_1	В	yes
<i>x</i> ₁ , <i>s</i> ₂ , <i>s</i> ₃	23	$4\frac{7}{15}$	16	s_1, x_2	G	yes
<i>x</i> ₂ , <i>s</i> ₂ , <i>s</i> ₃	-	-	-	s_1, x_1	-	-
x_1, x_2, s_1	15	5	8	<i>s</i> ₂ , <i>s</i> ₃	D	yes
x_1, x_2, s_2	23	2	$2\frac{7}{15}$	<i>s</i> ₁ , <i>s</i> ₃	F	yes
x_1, x_2, s_3	23	$4\frac{7}{15}$	$-19\frac{11}{15}$	s_1, s_2	E	no
	×2 10 5 5 1 4 1 4	(3)		D	(1) <u>E</u> (2) E G,	
	1	5	Loctur	13 20	∠⊃ togor Optimiz	ation with Applicatio

Basic *feasible* solutions correspond to solutions to the system of equations that *fulfill non-negativity*



Basic *infeasible* solutions corresp. to solutions to the system of equations with one or more variables < 0



<i>x</i> ₁	+	- <i>s</i> 1	= 23
0.067 <i>x</i> ₁	$+x_{2}$	$+s_{2}$	= 6
3 <i>x</i> ₁	$+8x_{2}$	$+s_{3}$	= 85