Lecture 8: Linear programming models

Michael Patriksson

12 February 2004

A road map

- The most important application area of optimization
- An "easy" problem to some degree: convex, linear objective and constraints
- But: Often large-scale. May not always be possible to solve directly.
- Solution: Decomposition, column generation techniques. (Generates "good" variables iteratively.)
- Example: The integer programming problems modelling staff planning at airlines.

- Problem type: $x_j \in \{0, 1\}$ is a logical variable deciding whether a particular group of staff should serve during a particular "leg" (a flight).
- Objective: Choose a cost-effective plan, one per week.
- Constraints: all legs must be covered.

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}}\ & f(\boldsymbol{x}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{subject to}\ & \boldsymbol{A}\boldsymbol{x} \geq \mathbf{1}^m, \\ & \boldsymbol{x} \geq \mathbf{0}^n, \\ & \boldsymbol{x} \text{ binary}, \end{aligned}$$

where $\boldsymbol{A}$ is a 0/1 matrix describing whether a group of staff is possible for inclusion in a leg or not.

- We call this a *set covering* problem.
- But: This is not all! How do we define x_j, that is, the column $\boldsymbol{a}_j$ of $\boldsymbol{A}$?
- The column $\boldsymbol{a}_j$ must reflect the possibility for the group to do a certain service. This depends a lot upon the timing of the leg, since the geographical location puts constraints on the staff availability, as well as union laws of working hours and conditions.

- The number of possible columns are in the millions, and cannot be generated before-hand.
- Answer: column generation. Solve subproblems that generate “feasible” columns, then solve the restricted problem to combined feasible columns into a work plan.
- This technique solves the problem of minimizing the cost over the convex hull of the feasible set; the *strong formulation* of the LP relaxation of the above integer program. Combined with effective IP techniques.
- More on this topic in course on integer programming and the Project course.

Basic method and its foundations

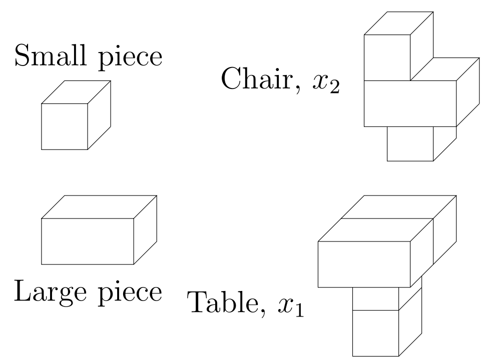
- Know that if there exists an optimal solution, one of them is an extreme point.
- Search only among extreme points.
- Extreme points can be easily described in algebraic terms.
- Find such a point.
- Generate a descent direction which leads to a better extreme point.
- Continue until convergence (finite!)

Duality and optimality

- LP problems are convex problems with CQ fulfilled (linear constraints—Abadie).
- Strong duality holds.
- KKT necessary and sufficient!
- Lagrangian dual same as LP dual.
- Simplex method: always satisfies complementarity; always primal feasibility after finding the first feasible solution; searches for a dual feasible point.

An introductory problem—A DUPLO game

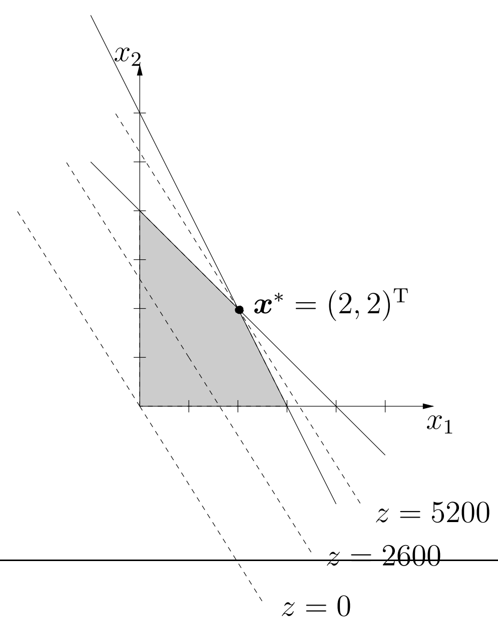
- A manufacturer produces two pieces of furniture, tables and chairs.
- The production of furniture requires two different pieces of raw-material, large and small pieces.
- One table is assembled from two pieces of each; one chair is assembled from one of the larger pieces and two of the smaller pieces.



- Data: 6 large and 8 small pieces available. Selling a table gives 1600 SEK, a chair 1000 SEK.
- Not trivial to choose an optimal production plan.

- What is the problem and how do we solve it?
- Solution by (1) the DUPLO game; (2) graphically; (3) the Simplex method.

$$
\begin{aligned}
\text{maximize} \quad z = \; & 1600x_1 + 1000x_2 \\
\text{subject to} \quad & 2x_1 + x_2 \leq 6, \\
& 2x_1 + 2x_2 \leq 8, \\
& x_1, \quad x_2 \geq 0.
\end{aligned}
$$



Further topics

- Sensitivity analysis: What happens with z^*, $\boldsymbol{x}^*$ if ...?
- A dual problem: A manufacturer (Billy) produce book shelves with same raw material. Billy wish to expand their production; interested in acquiring our resources.
- Two questions (with identical answers): (1) what is the lowest bid (price) for the total capacity at which we are willing to sell?; (2) what is the highest bid (price) that Billy are prepared to offer for the resources? The answer is a measure of the wealth of the company in terms of their resources.

A dual problem

- To study the problem, we introduce the variables

y_1 = the price which Billy offers for each large piece,

y_2 = the price which Billy offers for each small piece,

w = the total bid which Billy offers.

- Example: Net income for a table is 1600 SEK; need to get at least price bid $\boldsymbol{y}$ such that $2y_1 + 2y_2 \geq 1600$.

$$\begin{array}{lrl} \text{minimize} & w = & 6y_1 + 8y_2 \\ \text{subject to} & & 2y_1 + 2y_2 \geq 1600, \\ & & y_1 + 2y_2 \geq 1000, \\ & & y_1, \quad y_2 \geq 0. \end{array}$$

- Why the sign? $\boldsymbol{y}$ is a price!
- Optimal solution: $\boldsymbol{y}^* = (600, 200)^{\mathrm{T}}$. The bid is $w^* = 5200$ SEK.
- Remarks: (1) $z^* = w^*$! Our total income is the same as the value of our resources. (2) The price for a large piece equals its *shadow price*!

Geometric $\Longleftrightarrow$ Algebraic connections

- Must have equality constraints. Why? Inequalities cannot be manipulated while keeping the same solution set! Equalities can!
- Good to know: Every polyhedron P can be described in the form

$$P = \{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}; \quad \boldsymbol{x} \geq \boldsymbol{0}^n \}.$$

- We call this the *standard form*.

- Slack variables: ($\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$)

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} &\leq \boldsymbol{b}, \\ \boldsymbol{x} &\geq \boldsymbol{0}^n \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \boldsymbol{A}\boldsymbol{x} + \boldsymbol{I}^m\boldsymbol{s} &= \boldsymbol{b}, \\ \boldsymbol{x} &\geq \boldsymbol{0}^n, \\ \boldsymbol{s} &\geq \boldsymbol{0}^m \end{aligned}$$

- We can always assume even that $\boldsymbol{b} \geq \boldsymbol{0}^m$; otherwise, multiply necessary rows by -1.
- Idea: We describe an extreme point through this characterization of the feasible set; we then prove that moving between "adjacent" extreme points is simple.
- *Basic feasible solutions* is the buzz-word. Algebraic description of an extreme point.

- Note: $\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longrightarrow$ Linear algebra.
- $\boldsymbol{x} \geq \mathbf{0}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \Longrightarrow$ Polyhedra, convex analysis!
- Sign restrictions? If x_j is free of sign, substitute it everywhere by
$$x_j = x_j^+ - x_j^-,$$
where $x_j^+, x_j^- \geq 0$!

Basic feasible solutions (BFS)

- Consider an LP in standard form:
$$\begin{aligned} \text{minimize} \quad & z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geq \mathbf{0}^n, \end{aligned}$$
$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ with rank $\boldsymbol{A} = m$ (otherwise, delete rows), $n > m$, and $\boldsymbol{b} \in \mathbb{R}^m_+$.
- A point $\tilde{\boldsymbol{x}}$ is a *basic solution* if
1. $\boldsymbol{A}\tilde{\boldsymbol{x}} = \boldsymbol{b}$; and
2. the columns of $\boldsymbol{A}$ corresponding to the non-zero components of $\tilde{\boldsymbol{x}}$ are linearly independent.

- A basic solution that satisfies non-negativity is called a *basic feasible solution* (BFS).
- Additional terms: degenerate, non-degenerate basic solutions.
- Connection BFS–extreme points?
- Theorem 9.7: *A point $\boldsymbol{x}$ is an extreme point of the set $\{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{x} \geq \mathbf{0}^n \}$ if and only if it is a basic feasible solution.*
- Proof by the fact that the rank of $\boldsymbol{A}$ is full + Theorem 3.17. □

The Representation Theorem revisited

- Theorem 9.9: *Let $P = \{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b};\ \boldsymbol{x} \geq \mathbf{0}^n \}$ and $V = \{\boldsymbol{v}^1, \dots, \boldsymbol{v}^k\}$ its set of extreme points. If and only if P is nonempty, V is nonempty (finite). Let $C = \{ \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \mathbf{0}^m;\ \boldsymbol{x} \geq \mathbf{0}^n \}$ and $D = \{\boldsymbol{d}^1, \dots, \boldsymbol{d}^r\}$ be the set of extreme directions of C. If and only if P is unbounded D is nonempty (finite). Every $\boldsymbol{x} \in P$ is the sum of a convex combination of points in V and a non-negative linear combination of points in D:*
$$\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{v}^i + \sum_{j=1}^{r} \beta_j \boldsymbol{d}^j,$$

for some $\alpha_1, \ldots, \alpha_k \geq 0$ *such that* $\sum_{i=1}^k \alpha_i = 1$, *and* $\beta_1, \ldots, \beta_r \geq 0$.

- This is a restatement of Representation Theorem 3.22, adapted to the standard form of the LP.

Existence of optimal solutions to LP: Theorem 9.10

- *Let the sets* P, V *and* D *be defined as in Theorem 9.9 and consider the LP*

$$\begin{aligned} \text{minimize} \quad & z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x} \in P. \end{aligned}$$

  This problem has a finite optimal solution if and only if P *is nonempty and* z *is lower bounded on* P, *that is, if* $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}^j \geq 0$ *for all* $\boldsymbol{d}^j \in D$. *If the problem has a finite optimal solution, then there exists an optimal solution among the extreme points.*

- *Proof.* Let $\boldsymbol{x} \in P$. Then by the Representation

Theorem,

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \sum_{i=1}^{k} \alpha_i \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^i + \sum_{j=1}^{r} \beta_j \boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}^j. \qquad (1)$$

Now vary $\boldsymbol{x}$ over P. Then, we vary α_i and β_j only. Then, the first term above is finite, the second is finite if and only if $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}^j \geq 0$ for all $\boldsymbol{d}^j \in D$. Supposing that that is true, we choose $\beta_j = 0$ for all j.

- Now, let

$$a \in \arg \operatorname*{minimum}_{i \in \{1,\ldots,k\}} \{\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^i\}.$$

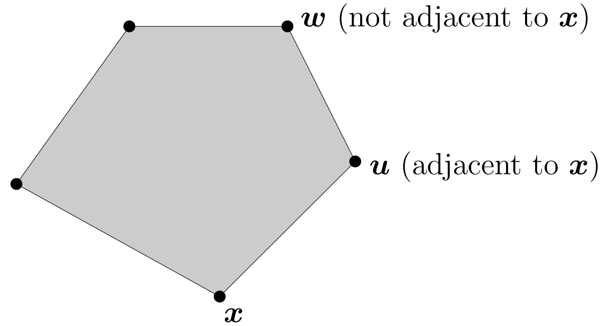
Then,

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^a = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^a \sum_{i=1}^{k} \alpha_i = \sum_{i=1}^{k} \alpha_i \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^a \leq \sum_{i=1}^{k} \alpha_i \boldsymbol{c}^{\mathrm{T}}\boldsymbol{v}^i = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x},$$

that is, the extreme point $\boldsymbol{v}^a$ is a global minimum. □

Adjacent extreme points

- Consider the following polytope.



- Every point on the line segment joining $\boldsymbol{x}$ and $\boldsymbol{u}$ cannot be written as a convex combination of any pair of points that are not on this line segment. However, this is not true for the points on the line segment between the extreme points $\boldsymbol{x}$ and $\boldsymbol{w}$. The extreme points $\boldsymbol{x}$ and $\boldsymbol{u}$ are said to be *adjacent* (while $\boldsymbol{x}$ and $\boldsymbol{w}$ are not adjacent).
- Theorem 9.13: *Two extreme points are adjacent if and only if there exist corresponding BFSs whose sets of basic variables differ in exactly* one *place.* □
- The DUPLO example!