Advanced Linear Programming

Organisation

Lecturers:
Leen Stougie,
Vrije Universiteit and CWI in Amsterdam
E-mail: l.stougie@vu.nl or stougie@cwi.nl

Marjan van den Akker
Universiteit Utrecht
marjan@cs.uu.nl

Advanced Linear Programming suggests that there is also something like Basic Linear Programming. And indeed, I assume everyone has had a course or part of a course on Basic Linear Programming. Usually LP is taught in the bachelor's as the first subject in a course on Optimization. Most importantly I assume that everyone has had a course in Linear Algebra, at least covering linear (sub)spaces, linear independence, bases of linear spaces, linear transformations, matrix theory including ranks, inverse, solutions of linear systems (the Gaussian elimination method), etc.

In this course I will refresh very briefly your memory on LP. I expect everyone to have access to the book

D. Bertsimas and J.N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont, Massachusetts, 1997. (ISBN 1-886529-19-1). The first five chapters are on my website (see below).

The book also allows you to refresh your memory, since it is a complete course on LP, and I will skip or go very briefly over parts that I suppose you to know already. I'll try to indicate as clearly as possible which parts of the book I assume to be known at the end of the course. Our ambition is to cover almost the whole book, indeed relying on your knowledge of basic LP. Please feel free to ask me questions during the lectures if anything remains unclear.

Detailed week-to-week information can be found on the website http://personal.vu.nl/l.stougie
The website also includes the lecture notes I made for myself for preparing the classes. In the lecture notes you also find the Exercises for the week. Answers to the exercises will become available on the website. If an answer is missing, please hand in the exercise, which I will check and correct and make it available at the website.

Last organisational detail: The exam is a written exam. There will be a re-exam, but there will NOT be a third exam. So plan your agendas very well and make sure you are present at the exam. Time and place will be announced through my own website and (hopefully) through the website of Mastermath. Examples of exams are at the website.

1 Introduction to LP

In general an optimization problem is

$$\begin{array}{ll} \min & f(x) \\ s.t. & x \in S, \end{array}$$

with $f : S \to \mathbb{R}$ and S some set of *feasible* solutions. Weierstrass proved that the optimum exists if f is a continuous function and S is a compact set.

In linear programming (LP) f is a linear function and the set S is defined by a set of linear inequalities. Read Chapter 1 for the modelling power of LP and Integer LP (ILP). The general LP is given by

$$\begin{array}{lll} \min & c^T x & \\ \text{s.t.} & a_i^T x \geq b_i, & i \in M_1 \\ & a_i^T x \leq b_i, & i \in M_2 \\ & a_i^T x = b_i, & i \in M_3 \\ & x_j \geq 0, & j \in N_1 \\ & x_j \leq 0, & j \in N_2 \\ & x_j \text{ unrestricted}, & j \in N_3, \end{array} \tag{1}$$

or in matrix notation, after appropriately rewriting some of the inequalities

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax \geq b, \end{array} \tag{2}$$

or as the standard form to be used in the simplex method for solving LPs

$$\begin{array}{ll} \min & c^T x \\ \text{s.t.} & Ax = b, x \geq \underline{0}. \end{array} \tag{3}$$

For reasons that are obscure to me, the authors of the book have chosen for minimization as the standard optimization criterion. I will follow this unless it bothers me too much.

The first system is usually the one that is obtained after modelling the verbal description of the problem. It is easy to see that any such system can be transformed into an equivalent system with only $\geq$ (or only $\leq$) constraints and only non-negative *decision variables*, or, by using slack- or surplus variables, into a system with only $=$ constraints and only non-negative *decision variables*, obtaining the standard LP formulation.

Chapter 1 of the book gives an abundance of examples, some of which really show how LP can capture a large variety of optimization problems. Specifically, the example in Section 1.3 shows how minimizing a piecewise linear convex function can be solved using LP, with problems concerning absolute values as a specific example. Some of the problems in Chapter 1 model as integer LP's, a subject Marjan will cover in her part of this course. Please read these modelling parts of the chapter.

A basic LP-course is heavily based on Linear Algebra. However, the insights that lead to the solution of LP-problems and to a deeper understanding of duality are based on geometry. The book puts a lot of emphasis on the geometrical interpretation of LP, and I will use it in this course to show how geometric and algebraic interpretations coincide, re-deriving known theorems about LP from another theory.

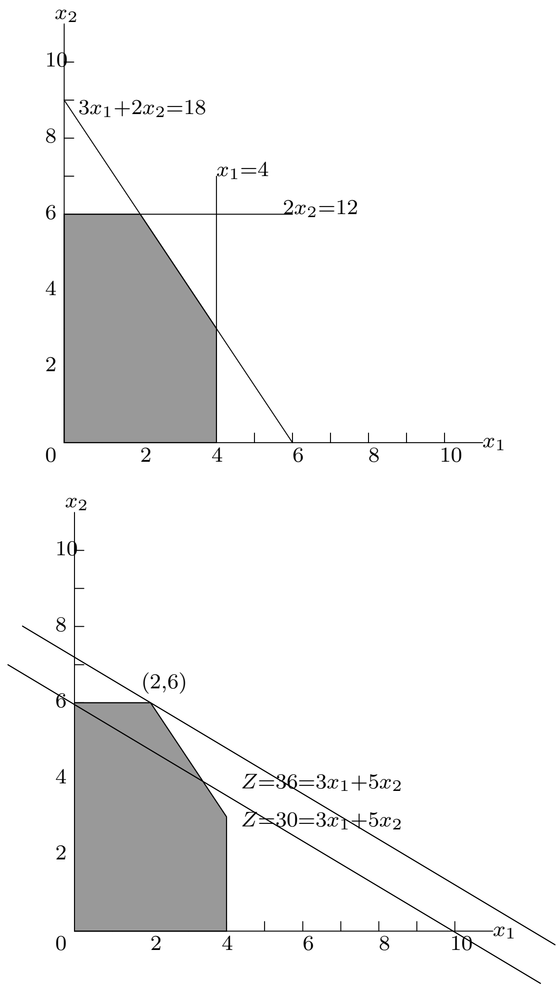

Figure 1: Feasible region for WGC

1.1 Graphical solution

As a step-up let us first have a look at a simple example that allows a graphical solution. The following problem is extracted from Hillier and Lieberman, Introduction to Operations Research, McGraw Hill, New York. I skip the verbal description and the modelling and switch immediately to the LP-formulation.

$$
\begin{array}{rrcrcr}
\text{maximize} & 3x_1 & + & 5x_2 & & \\
s.t. & x_1 & & & \leq & 4 \\
 & & & 2x_2 & \leq & 12 \\
 & 3x_1 & + & 2x_2 & \leq & 18
\end{array}
\tag{4}
$$

$$x_1 \geq 0, \qquad x_2 \geq 0.$$

So for once we use a maximization problem here. We have two *decision variables*, which allows a graphical description of the problem:

The feasible set of an LP in the $\mathbb{R}^n$, i.e., an LP with n decision variables, is bounded by $(n-1)$-dimensional hyperplanes. A vertex of the feasible set is the intersection of n such bounding

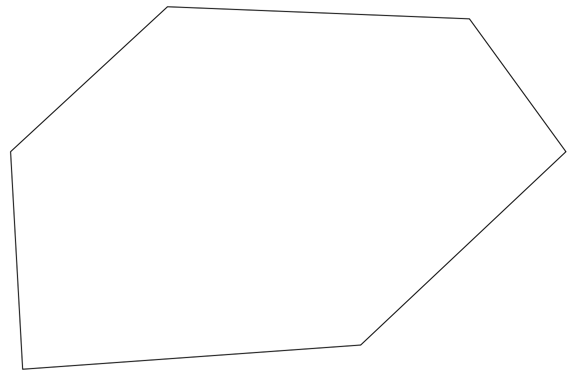

Figure 2: Bounded polyhedron

hyperplanes of dimension $(n-1)$. Moreover, the optimum is attained in one of the vertices of the feasible region. The optimum in the above 2-dimensional problem is found by pressing a line parallel to $3x_1 + 2x_2 = v$ through the feasible set until the line when pressed a bit higher would miss any point of the feasible set.

Similarly, think of optimizing a linear function of 3 variables as pushing a two-dimensional plane through a diamond.

This geometric insight is generalised in the following part.

1.2 Geometric interpretation

The feasible set of an LP has the property of being *convex*.

Definition 1.1 A set $S \subset \mathbb{R}^n$ is **convex** if for any two elements $x, y \in S$ and any constant $0 \leq \lambda \leq 1$ the vector $\lambda x + (1-\lambda)y \in S$.

Using the definition it is obvious that the intersection of convex sets is a convex set. It also easy to show that any set $\mathcal{H} =: \{x \in \mathbb{R}^n | a^T x \geq b\}$ is convex. Such a set is called a *halfspace*. Such a halfspace contains all points on or at one side of the $(n-1)$-dimensional *hyperplane* $H =: \{x \in \mathbb{R}^n | a^T x = b\}$. Since the feasible set $P = \{x \in \mathbb{R}^n | Ax \geq b\}$ of an LP instance is an intersection of m halfspaces it must be convex.

In fact, there is a theorem, that states the equivalence of convex sets and intersections of halfspaces, saying that every convex set is the intersection of a set of halfspaces. It is not mentioned in the book, probably because its proof is not really easy combined with the fact that it is not relevant for LP. We will be interested in convex sets that are intersections of a finite number of halfspaces:

Definition 1.2 A *polyhedron* is the intersection of a finite set of halfspaces.

We will always work with closed sets in this course. An example of a polyhedron is given in Figure 2.

Not every polyhedron needs to be bounded (see Figure 3).

A bounded polyhedron is called a polytope.

Definition 1.3 Let $x^1, \ldots, x^k \in \mathbb{R}^n$ and $\lambda_1, \ldots, \lambda_k \in \mathbb{R}_+ \cup \{0\}$ such that $\sum_{i=1}^k \lambda_i = 1$ then the vector $\sum_{i=1}^k \lambda_i x^i$ is called a *convex combination* of the vectors $x^1, \ldots, x^k$. The set of all convex combinations of $x^1, \ldots, x^k$ is called the *convex hull* of $x^1, \ldots, x^k$.

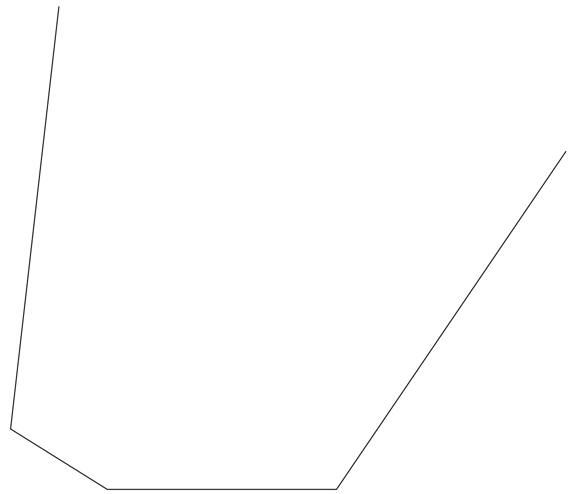

Figure 3: Unbounded polyhedon

It is easy to see that the convex hull of a finite set of vectors is a convex set. It is also intuitively clear that the convex hull of a finite set of vectors is a polytope. But the proof of this is not trivial and we postpone it until treating Chapter 4.

It is intuitively clear that, if the LP has a finite optimal value then there is a vertex or extreme point of the feasible polyhedron which is an optimum.

Definition 1.4 An *extreme point* is a point in the polyhedron which cannot be written as a convex combination of two other points in the polyhedron.

Another definition is that of a vertex

Definition 1.5 A *vertex* of a polyhedron is the only intersection point of the polyhedron with some hyperplane.

However, for solution methods an algebraic definition is much more useful, which views an extreme point as a point on the intersection of a number of hyperplanes that bound the polyhedron.

As a preliminary, I state without proof the following theorem from [B& T] which I will use later. The proof is straightforward from your linear algebra course.

Theorem 1.1 (2.2 in [B& T]) *Let x^* be an element of $\mathbb{R}^n$ and let $I = \{i \mid a_i^T x^* = b_i\}$ be the set of indices of constraints that hold with equality at x^*. Then, the following are equivalent:*

(a) There exist n vectors in the set $\{a_i \mid i \in I\}$, which are linearly independent.

(b) The span of the vectors a_i, $i \in I$, is all of $\mathbb{R}^n$, that is, every element of $\mathbb{R}^n$ can be expressed as a linear combination of the vectors a_i, $i \in I$.

(c) The system of equations $a_i^T x = b_i$, $i \in I$, has a unique solution.

Allow me to abuse mathematical language slightly and say that hyperplanes or linear constraints are linearly independent if the a_i's in their description $\{x \in \mathbb{R}^n | a_i^T x = b\}$ are linearly independent. In the $\mathbb{R}^n$ any point is determined by the intersection of n linearly independent $n-1$-dimensional hyperplanes. In the book it is shown and it is indeed intuitively clear, that any extreme point of a polyhedron P in $\mathbb{R}^n$ is defined as the intersection of n linearly independent hyperplanes that are bounding P, among them being all on which P lies entirely. We say that a constraint is *tight* or *active* or *supporting* in a point of P if it is satisfied with equality in that point. In LP a point that represents a feasible solution in which n linearly independent constraints are tight (active or supporting) is called a *basic feasible solution*.

Theorem 1.2 (2.3 in [B& T]) *For a non-empty polyhedron P, the following three statements are equivalent*

(a) $x^* \in P$ is an extreme point of P;

(b) x^* is a basic feasible solution;

(c) x^* is a vertex.

Proof. (a)⇒(b) Suppose x^* is not bfs. Let $I = \{i | a_i^T x = b_i\}$ be the set of constraints active in x^*. There are no n linearly independent vectors a_i among the ones in I. Thus, the vectors a_i, $i \in I$ lie in a subspace of $\mathbb{R}^n$, i.e., there is at least one vector d orthogonal to all a_i, $a_i^T d = 0$, $i \in I$. Consider vectors $y = x^* + \epsilon d$ and $z = x^* - \epsilon d$, for $\epsilon > 0$ sufficiently small. $y \in P$ because

$$a_i^T y = a_i^T x^* + \epsilon a_i^T d = b_i + 0 = b_i, \ \forall i \in I,$$

and since for $i \notin I$, $a_i^T x^* > b_i$, whence for ϵ small enough $a_i^T y \geq b_i$. Similarly $z \in P$. But $x^* = \frac{1}{2}(y + z)$ hence not an extreme point.

(b)⇒(c) One shows that the choice $c = \sum_{i \in I} a_i$ certifies that x^* is a vertex of P.

(c)⇒(a) Take any c for which x^* is a unique optimal point. For any two other points $y, z \in P$ we therefore have $c^T y > c^T x^*$ and $c^T z > c^T x^*$. Hence, we necessarily have $c^T(\lambda y + (1 - \lambda) z) > c^T x^*$, concluding that x^* cannot be written as a convex combination of any two other points in P. □

Two remarks: Firstly, two different basic feasible solutions are said to be *adjacent* on the polyhedron P if they share $n-1$ out of the n hyperplanes or active constraints that define each of them.

Secondly, the last theorem implies that there are a finite number, at most $O(m^n)$, of extreme points in any polyhedron P, since P is the intersection of a finite number m of halfspaces.

One of the main results of Chapter 2 concerns the existence of an extreme point.

Theorem 1.3 (2.6 in [B& T]) *Given non-empty polyhedron $P = \{x \in \mathbb{R}^n | a_i^T x \geq b_i, \ i = 1, \ldots, m\}$, the following statements are equivalent:*
a) P has at least one extreme point;
b) P does not contain a line; i.e., $\nexists\, x \in P$, $d \neq \underline{0} \in \mathbb{R}^n$, $\forall\, \lambda \in \mathbb{R}$: $x + \lambda d \in P$;
c) There exist n vectors out of the family of $a_1, \ldots, a_m$ which are linearly independent.

Proof. a) ⇒ c) is a direct consequence of the last theorem. I will prove b) ⇒ a) and leave the proof of c) ⇒ b) to be read by yourself.

Let $x \in P$ and $I(x) = \{i | a_i^T x = b_i\}$. If $\{a_i | i \in I(x)\}$ contains n linearly independent a_i's, then x is a bfs and hence an extreme point. Otherwise, as before, they span a subspace of the $\mathbb{R}^n$ and therefore there exists a vector d such that $a_i^T d = 0$, $\forall\, i \in I(x)$. Consider $y = x + \lambda d$. Clearly $a_i^T y = a_i^T x = b_i \ \forall\, i \in I(x)$. Thus $I(x) \subseteq I(y)$. Since P does not contain any line it must be that if we decrease or increase λ long enough, then some other constraint than those in $I(x)$ will become active, say this happens at λ^* and $a_j^T(x + \lambda^* d) = b_j$, for some $j \notin I(x)$. Since $a_j^T x > b_j$, $\lambda^* a_j^T d = b_j - a_j^T x \neq 0$. Hence, because $a_j^T d \neq 0$ and $a_i^T d = 0$, $\forall\, i \in I(x)$, we have that a_j is linearly independent of $\{a_i | i \in I(x)\}$. This procedure is repeated until we arrive at a set of n linearly independent active constraints, corresponding to a bfs, i.e., an extreme point of P. □

Another result that most of you have always used but probably never seen proven rigorously.

Theorem 1.4 (2.7 in [B& T]) *Given an LP-problem, the feasible polyhedron of which contains at least one extreme point, and for which an optimal solution exists, there is always an extreme point that is an optimal solution.*

PROOF. Let $Q = \{x \in \mathbb{R}^n | Ax \geq b, c^T x = v\}$ be the set of optimal solutions, having objective value v. Q contains no line, since it is inside P and P has an extreme point, hence contains no line. Thus Q also has an extreme point, say x^*. We show by contradiction that x^* is also extreme point of P. Suppose not, then $\exists\, y, z \in P,\ \lambda \in [0,1]$: $x^* = \lambda y + (1-\lambda) z$. $c^T y \geq v$ and $c^T z \geq v$ and $c^T x^* = v$ implies that $c^T y = c^T z = v \Rightarrow y, z \in Q$ and therefore x^* is not extreme point of Q. $\square$

A slightly stronger result is proved along the same lines as the proof of Theorem 2.6 in the book.

Theorem 1.5 (2.8 in [B& T]) *Given an LP-problem, the feasible polyhedron of which contains at least one extreme point, then either the solution value is unbounded or there is an extreme point that is optimal.*

In my discussions of Chapters 1 and 2 I made a selection of what to treat here in class and what to leave to yourself for study. For those who did not have a course on complexity theory I advise to read Section 1.6 of [B& T], that shows you how to measure the "running time" of an algorithm and that making a distinction between so-called polynomial running time and exponential running time is relevant. Complexity theory is a beautiful theory, essential in the study of combinatorial problems. In this book it only plays an essential role in Chapters 8 and 9.

There are intriguing open problems in the study of polyhedra in relation to complexity theory.

You may skip Sections 2.7 and 2.8 in the book, since the material is covered again in Chapter 4. I do advise you to study 2.4 on degeneracy.

Material of Week 1 from [B& T]

Chapter 1 and Chapter 2, apart from 2.7 and 2.8.

Exercises of Week 1

1.5, 1.15, 2.6, 2.15.

Next week

Chapter 3