Prelims
Optimization

Colin McDiarmid

TT 2015

1 Introduction

A general optimization problem is of the form: choose $\mathbf{x}$ to

$$\begin{aligned} &\text{maximise} \quad f(\mathbf{x}) \\ &\text{subject to} \quad \mathbf{x} \in S \end{aligned}$$

where

$\mathbf{x} = (x_1, \ldots, x_n)^T$

$f : \mathbb{R}^n \to \mathbb{R}$ is the *objective function*

$S \subseteq \mathbb{R}^n$ is the *feasible set*.

We might write this problem:

$$\max_{\mathbf{x}} f(\mathbf{x}) \quad \text{subject to } \mathbf{x} \in S.$$

For example

- $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$ for some vector $\mathbf{c} \in \mathbb{R}^n$
- $S = \{\mathbf{x} : A\mathbf{x} \leqslant \mathbf{b}\}$ for some $m \times n$ matrix A and some vector $\mathbf{b} \in \mathbb{R}^m$.

If f is linear and $S \subseteq \mathbb{R}^n$ can be described by linear equalities/inequalities then we have a *linear programming* (LP) problem.

If $\mathbf{x} \in S$ then $\mathbf{x}$ is called a *feasible solution*.

If the maximum of $f(\mathbf{x})$ over $\mathbf{x} \in S$ occurs at $\mathbf{x} = \mathbf{x}^*$ then

- $\mathbf{x}^*$ is an *optimal solution*
- $f(\mathbf{x}^*)$ is the *optimal value*.

Questions

In general:

- does a feasible solution $\mathbf{x} \in S$ exist?
- if so, does an optimal solution exist?
- if so, is it unique?
- how can we find such solutions?

Example: activity analysis

A company produces drugs A and B using machines M_1 and M_2.

- 1 ton of drug A requires 1 hour of processing on M_1 and 2 hours on M_2
- 1 ton of drug B requires 3 hours of processing on M_1 and 1 hour on M_2
- 9 hours of processing on M_1 and 8 hours on M_2 are available each day
- Each ton of drug produced (of either type) yields £1 million profit

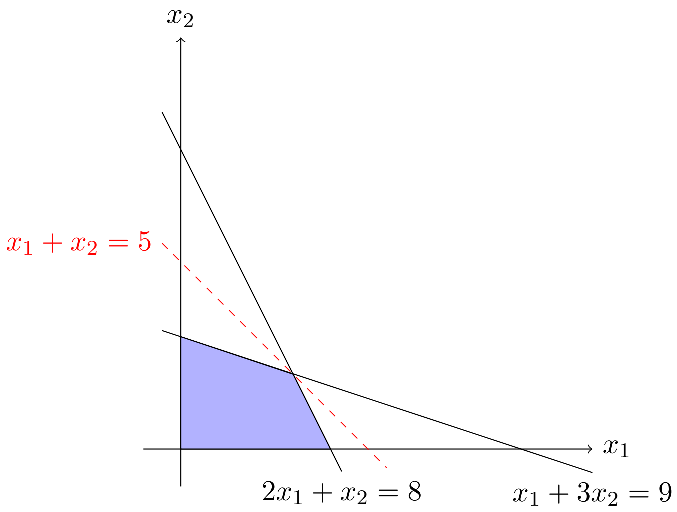
To maximise its profit, how much of each drug should the company make per day?

Solution

Let

- x_1 = number of tons of A produced
- x_2 = number of tons of B produced

$$
\begin{array}{rrcl}
P1: \text{maximise} & x_1 + \ \ x_2 & & \text{(profit in £ million)} \\
\text{subject to} & x_1 + 3x_2 & \leqslant 9 & (M_1 \text{ processing}) \\
 & 2x_1 + \ \ x_2 & \leqslant 8 & (M_2 \text{ processing}) \\
 & x_1, x_2 & \geqslant 0 &
\end{array}
$$



The shaded region is the feasible set for $P1$. The maximum occurs at $\mathbf{x}^* = (3, 2)^T$ with value 5.

Diet problem

A pig-farmer can choose between four different varieties of food, providing different quantities of various nutrients.

		food 1	2	3	4	required amount/wk
	A	1.5	2.0	1.0	4.1	4.0
nutrient	B	1.0	3.1	0	2.0	8.0
	C	4.2	1.5	5.6	1.1	9.5
cost/kg		£5	£7	£7	£9	

The (i, j) entry is the amount of nutrient i per kg of food j.

Let x_j = number of units of food F_j in the diet.

Problem $P2$:

$$\begin{array}{lrl} \text{minimise} & 5x_1 + 7x_2 + 7x_3 + 9x_4 & \\ \text{subject to} & 1.5x_1 + 2x_2 + x_3 + 4.1x_4 & \geqslant 4 \\ & x_1 + 3.1x_2 \qquad\quad + 2x_4 & \geqslant 8 \\ & 4.2x_1 + 1.5x_2 + 5.6x_3 + 1.1x_4 & \geqslant 9.5 \\ & x_1,\, x_2,\, x_3,\, x_4 & \geqslant 0 \end{array}$$

General form of the diet problem

Foods F_j for $j = 1, \ldots, n$, nutrients N_i for $i = 1, \ldots, m$.

Data:

- a_{ij} = amount of nutrient N_i in one unit of food F_j
- b_i = required amount of nutrient N_i
- c_j = cost per unit of food F_j

Let x_j = number of units of food F_j in the diet.

The diet problem is

$$\begin{array}{lll} \text{minimise} & c_1x_1 + \cdots + c_nx_n & \\ \text{subject to} & a_{i1}x_1 + \cdots + a_{in}x_n \geqslant b_i & \text{for } i = 1, \ldots, m \\ & x_1, \ldots, x_n \geqslant 0. & \end{array}$$

In matrix notation the diet problem is

$$\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to } A\mathbf{x} \geqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}.$$

Note that our vectors are always column vectors.

We write $\mathbf{x} \geqslant \mathbf{0}$ to mean $x_i \geqslant 0$ for all i. ($\mathbf{0}$ is a vector of zeros.)

Similarly $A\mathbf{x} \geqslant \mathbf{b}$ means $(A\mathbf{x})_i \geqslant b_i$ for all i.

General form of activity analysis

Goods or activities G_j for $j = 1, \ldots, n$.
Scarce resources R_i for $i = 1, \ldots, m$.

Data:

- a_{ij} = amount of R_i required to make one unit of G_j
- b_i = amount of R_i available
- c_j = profit contribution per unit of G_j

We want to maximise profit.

Let x_j = number of units of good G_j made.

The activity analysis LP is

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to } A\mathbf{x} \leqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}.$$

Real applications

"Programming" = "planning"

May be many thousands of variables or constraints

- Production management: activity analysis, large manufacturing plants, farms, etc
- Scheduling, e.g. airline crews:
  - need all flights covered
  - restrictions on working hours and patterns
  - minimise costs: wages, accommodation, use of seats by non-working staff

  shift workers (call centres, factories, etc)
- Yield management (airline ticket pricing: multihops, business/economy mix, discounts, etc)
- Network problems: transportation capacity planning in telecoms networks
- Game theory: economics, evolution, animal behaviour

Free variables

In an LP model, some variables may be positive or negative, e.g. there may not be a constraint $x_1 \geqslant 0$.

Such a *free variable* can be replaced by

$$x_1 = u_1 - v_1$$

where $u_1, v_1 \geqslant 0$.

Slack variables

In $P1$ we had

$$\begin{aligned} \text{maximise} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 + 3x_2 \leqslant 9 \\ & 2x_1 + x_2 \leqslant 8 \\ & x_1, x_2 \geqslant 0. \end{aligned}$$

We can rewrite as

$$\begin{aligned} \text{maximise} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 + 3x_2 + x_3 = 9 \\ & 2x_1 + x_2 + x_4 = 8 \\ & x_1, \ldots, x_4 \geqslant 0. \end{aligned}$$

- x_3 = unused time on machine M_1
- x_4 = unused time on machine M_2

x_3 and x_4 are called *slack variables*.

With the slack variables included, the problem has the form

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to} \quad A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geqslant \mathbf{0}.$$

Two standard forms

In fact any LP (with equality constraints, weak inequality constraints, or a mixture) can be converted to the form

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to} \quad \begin{aligned} A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geqslant \mathbf{0} \end{aligned}$$

since:

- minimising $\mathbf{c}^T\mathbf{x}$ is equivalent to maximising $-\mathbf{c}^T\mathbf{x}$
- inequalities can be converted to equalities by adding slack variables
- free variables can be replaced as above.

Similarly, any LP can be put into the form

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to} \quad \begin{aligned} A\mathbf{x} &\leqslant \mathbf{b} \\ \mathbf{x} &\geqslant \mathbf{0} \end{aligned}$$

since e.g.

$$A\mathbf{x} = \mathbf{b} \iff \begin{cases} A\mathbf{x} \leqslant \mathbf{b} \\ -A\mathbf{x} \leqslant -\mathbf{b} \end{cases}$$

(more efficient rewriting may be possible!).

So it is OK for us to concentrate on LPs in these forms.

Remark

We always assume that the underlying space is $\mathbb{R}^n$.

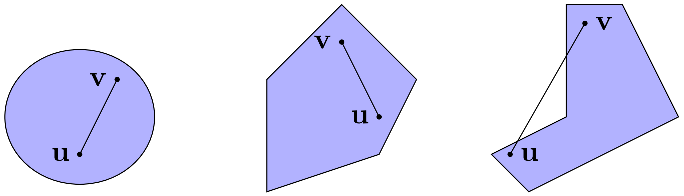
In particular $x_1, \ldots, x_n$ need not be integers. If we restrict to $\mathbf{x} \in \mathbb{Z}^n$ we have an *integer linear program* (ILP).

ILPs are in a sense *harder* than LPs. Note that the optimal value of an LP gives a *bound* on the optimal value of the associated ILP.

2 Geometry of linear programming

Definition 2.1

A set $S \subseteq \mathbb{R}^n$ is called *convex* if for all $\mathbf{u}, \mathbf{v} \in S$ and all $\lambda \in (0, 1)$, we have $\lambda\mathbf{u} + (1 - \lambda)\mathbf{v} \in S$.



convex convex not convex

Thus S is convex when each line segment joining points in S stays in S.

For now we will consider LPs in the form

$$\max \mathbf{c}^T\mathbf{x} \quad \text{subject to} \quad A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geqslant \mathbf{0}.$$

Theorem 2.2

The feasible set

$$S = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\}$$

is convex.

Proof.

Suppose $\mathbf{u}, \mathbf{v} \in S$, $\lambda \in (0, 1)$. Let $\mathbf{w} = \lambda\mathbf{u} + (1 - \lambda)\mathbf{v}$. Then

$$\begin{aligned} A\mathbf{w} &= A[\lambda\mathbf{u} + (1 - \lambda)\mathbf{v}] \\ &= \lambda A\mathbf{u} + (1 - \lambda)A\mathbf{v} \\ &= [\lambda + (1 - \lambda)]\mathbf{b} \\ &= \mathbf{b} \end{aligned}$$

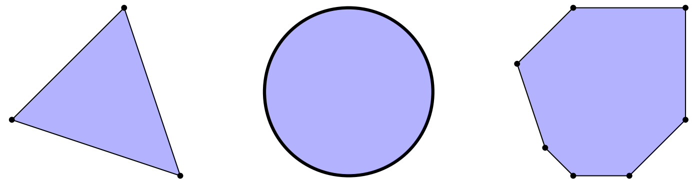
and $\mathbf{w} \geqslant \lambda\mathbf{0} + (1 - \lambda)\mathbf{0} = \mathbf{0}$. So $\mathbf{w} \in S$. $\square$

Extreme points

Definition 2.3

A point $\mathbf{x}$ in a convex set S is called an *extreme point* of S if there are no two distinct points $\mathbf{u}, \mathbf{v} \in S$, and $\lambda \in (0, 1)$, such that $\mathbf{x} = \lambda\mathbf{u} + (1 - \lambda)\mathbf{v}$.

Thus $\mathbf{x}$ is an extreme point when it is not in the interior of any line segment lying in S.



Theorem 2.4

If an LP has an optimal solution, then it has an optimal solution at an extreme point of the feasible set.

Proof.

Idea: If a given optimal point is not extremal, it's on some line segment within S all of which is optimal: move along the line until we find an optimal point with more zero co-ordinates.

Since there exists an optimal solution, there exists an optimal solution $\mathbf{x}^*$ with a minimal number of non-zero components.

Suppose $\mathbf{x}^*$ is not extremal, so that

$$\mathbf{x}^* = \lambda \mathbf{u} + (1-\lambda)\mathbf{v}$$

for some $\mathbf{u} \neq \mathbf{v} \in S$ and $\lambda \in (0,1)$.

Since $\mathbf{x}^*$ is optimal, $\mathbf{c}^T\mathbf{u} \leqslant \mathbf{c}^T\mathbf{x}^*$ and $\mathbf{c}^T\mathbf{v} \leqslant \mathbf{c}^T\mathbf{x}^*$.
But also $\mathbf{c}^T\mathbf{x}^* = \lambda\mathbf{c}^T\mathbf{u} + (1-\lambda)\mathbf{c}^T\mathbf{v}$ so in fact $\mathbf{c}^T\mathbf{u} = \mathbf{c}^T\mathbf{v} = \mathbf{c}^T\mathbf{x}^*$.

Consider the line defined by

$$\mathbf{x}(\varepsilon) = \mathbf{x}^* + \varepsilon(\mathbf{u} - \mathbf{v}) \quad \text{for } \varepsilon \in \mathbb{R}.$$

Then

(a) $A\mathbf{x}^* = A\mathbf{u} = A\mathbf{v} = \mathbf{b}$ so $A\mathbf{x}(\varepsilon) = \mathbf{b}$ for all ε

(b) $\mathbf{c}^T\mathbf{x}(\varepsilon) = \mathbf{c}^T\mathbf{x}^*$ for all ε

(c) if $x_j^* = 0$ then $u_j = v_j = 0$, so $\mathbf{x}(\varepsilon)_j = 0$ for all ε

(d) if $x_j^* > 0$ then $\mathbf{x}(0)_j > 0$, so $\mathbf{x}(\varepsilon)_j \geqslant 0$ if $|\varepsilon|$ is sufficiently small

(e) for some j, $u_j \neq v_j$ and $x_j^* > 0$.

It follows that we can move ε from zero, in a positive direction (if some $u_j < v_j$) or a negative direction (otherwise), keeping $\mathbf{x}(\varepsilon) \geqslant \mathbf{0}$, until at least one extra co-ordinate of $\mathbf{x}(\varepsilon)$ becomes zero.

This gives an optimal solution with strictly fewer non-zero co-ordinates than $\mathbf{x}^*$, contradicting the choice of $\mathbf{x}^*$.

So $\mathbf{x}^*$ must be extreme. □

'Extreme point' has a geometric flavour - algebraic next.

Basic solutions

Let $\mathbf{a}_j$ be the jth column of the $m \times n$ matrix A, so that

$$A\mathbf{x} = \mathbf{b} \iff \sum_{j=1}^{n} x_j \mathbf{a}_j = \mathbf{b}.$$

Definition 2.5

(1) A solution $\tilde{\mathbf{x}}$ of $A\mathbf{x} = \mathbf{b}$ is called a *basic solution* if the family of vectors $(\mathbf{a}_j : \tilde{x}_j \neq 0)$ is linearly independent.

(2) A basic solution satisfying $\mathbf{x} \geqslant \mathbf{0}$ is called a *basic feasible solution* (BFS).

Note. Since A has m rows, at most m columns can be linearly independent. So any basic solution has at least $n-m$ zero co-ordinates. More later.

Theorem 2.6

$\tilde{\mathbf{x}}$ is an extreme point of

$$S = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}, \mathbf{x} \geqslant \mathbf{0}\}$$

if and only if $\tilde{\mathbf{x}}$ is a BFS.

Proof.

(1) Let $\tilde{\mathbf{x}}$ be a BFS. Suppose $\tilde{\mathbf{x}} = \lambda\mathbf{u} + (1-\lambda)\mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in S$ and $\lambda \in (0,1)$. To show $\tilde{\mathbf{x}}$ is extreme we want to show $\mathbf{u} = \mathbf{v}$.

Let $J = \{j : \tilde{x}_j > 0\}$.

(a) If $j \notin J$ then $\tilde{x}_j = 0$, which implies $u_j = v_j = 0$.

(b) $A\mathbf{u} = A\mathbf{v} = \mathbf{b}$, so $A(\mathbf{u} - \mathbf{v}) = \mathbf{0}$. Thus

$$\mathbf{0} = \sum_{j=1}^{n}(u_j - v_j)\mathbf{a}_j = \sum_{j \in J}(u_j - v_j)\mathbf{a}_j$$

since $u_j = v_j = 0$ for $j \notin J$.

This implies that $u_j = v_j$ for $j \in J$ since $(\mathbf{a}_j : j \in J)$ is linearly independent.

Hence $\mathbf{u} = \mathbf{v}$, and so $\tilde{\mathbf{x}}$ is an extreme point.

(2) Suppose $\tilde{\mathbf{x}}$ is *not* a BFS, i.e. $(\mathbf{a}_j : j \in J)$ is linearly dependent, where as before $J = \{j : \tilde{x}_j > 0\}$.

Then there exists $\mathbf{u} \neq \mathbf{0}$ with $u_j = 0$ for $j \notin J$ such that $A\mathbf{u} = \mathbf{0}$.
For small enough $\varepsilon > 0$, $\tilde{\mathbf{x}} \pm \varepsilon\mathbf{u}$ are both feasible, and

$$\tilde{\mathbf{x}} = \frac{1}{2}(\tilde{\mathbf{x}} + \varepsilon\mathbf{u}) + \frac{1}{2}(\tilde{\mathbf{x}} - \varepsilon\mathbf{u})$$

so $\tilde{\mathbf{x}}$ is not extreme. □

Corollary 2.7

If there is an optimal solution, then there is an optimal BFS (that is, an optimal solution which is also a BFS).

Proof.

This is immediate from Theorems 2.4 and 2.6. □

Discussion

Recall: our constraints are $A\mathbf{x} = \mathbf{b}$, where A is $m \times n$.

Typically we may *assume* A has rank m (its rows are linearly independent): for if not, either we have a contradiction, or redundancy which we can remove).

Then ($n \geqslant m$ and) $A\mathbf{x} = \mathbf{b}$ always has a solution.

Indeed we may assume $n > m$ (more variables than constraints): for if $n = m$ there is a unique solution, easily found.

Also: $\tilde{\mathbf{x}}$ is a basic solution $\iff$ there is a set $B \subseteq \{1, \ldots, n\}$ of size m such that

- $\tilde{x}_j = 0$ if $j \notin B$,
- $(\mathbf{a}_j : j \in B)$ is linearly independent.

Proof.

Simple exercise. Augment $\{\mathbf{a}_j : \tilde{x}_j \neq 0\}$ to a larger linearly independent set if necessary. □

To look for basic solutions:

- choose $B \subseteq \{1, \ldots, n\}$ of size m.
- set $x_j = 0$ for $j \notin B$,
- look at the m columns $(\mathbf{a}_j : j \in B)$.
  Are they linearly independent? If so we have an invertible $m \times m$ matrix.
  Solve for $\{x_j : j \in B\}$ to give $\sum_{j \in B} x_j \mathbf{a}_j = \mathbf{b}$.
  Then
  $$A\mathbf{x} = \sum_{j=1}^{n} x_j \mathbf{a}_j = \sum_{j \in B} x_j \mathbf{a}_j = \mathbf{b}$$
  as required.

In this way we obtain all basic solutions (at most $\binom{n}{m}$ of them).

Bad algorithm:

- look through all basic solutions
- which are feasible?
- what is the value of the objective function?

We can do much better!

Simplex algorithm:

- move from one BFS to another, improving the value of the objective function at each step.

3 The simplex algorithm (1)

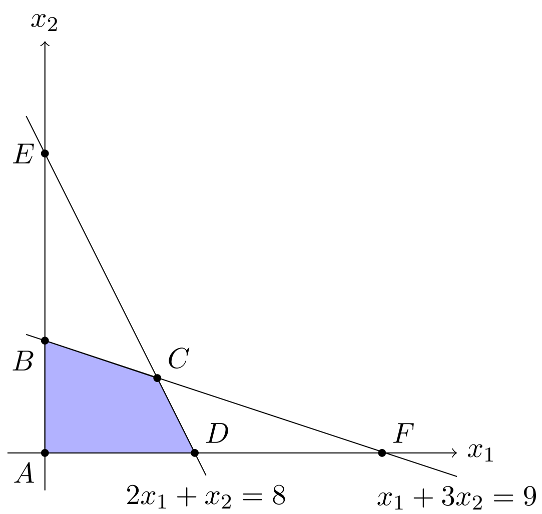
The simplex algorithm works as follows.

1. Start with an initial BFS.
2. Is the current BFS optimal?
3. If YES, stop.
   If NO, move to a new and improved BFS, then return to 2.

From Corollary 2.7, it is sufficient to consider only BFSs when searching for an optimal solution (though this will emerge anyway).

Recall the first activity analysis $P1$, expressed without slack variables:

$$\begin{array}{ll} \text{maximise} & x_1 + x_2 \\ \text{subject to} & x_1 + 3x_2 \leqslant 9 \\ & 2x_1 + x_2 \leqslant 8 \\ & x_1, x_2 \geqslant 0 \end{array}$$



Rewrite:

$$\begin{array}{rrrrcrl} x_1 & +3x_2 & +x_3 & & = & 9 & (1) \\ 2x_1 & + \ x_2 & & +x_4 & = & 8 & (2) \\ x_1 & + \ x_2 & & & = & f(\mathbf{x}) & (3) \end{array}$$

Put $x_1, x_2 = 0$, giving $x_3 = 9$, $x_4 = 8$, $f = 0$ (we're at the BFS $\mathbf{x} = (0, 0, 9, 8)^T$).

Note: In writing the three equations as (1)–(3) we are effectively expressing x_3, x_4, f in terms of x_1, x_2.

We call x_3, x_4 the *basic variables*, and x_1, x_2 the *non-basic variables*.

1. Start at the initial BFS $\mathbf{x} = (0, 0, 9, 8)^T$, vertex A, where $f = 0$.
2. From (3), increasing x_1 or x_2 will increase $f(\mathbf{x})$. Let's increase x_1.

   From (1): we can increase x_1 to 9, when x_3 decreases to 0.
   From (2): we can increase x_1 to 4, when x_4 decreases to 0.

   The *stricter* restriction on x_1 is from (2), the *pivot row*.
3. So (keeping $x_2 = 0$),
   (a) increase x_1 to 4, decrease x_4 to 0 – using (2), this maintains equality in (2),
   (b) and, using (1), decreasing x_3 to 5 maintains equality in (1).

   With these changes we move to a new and improved BFS $\mathbf{x} = (4, 0, 5, 0)^T$, $f(\mathbf{x}) = 4$, vertex D. The new basic variables are x_1, x_3.

To see if this new BFS is optimal, rewrite (1)–(3) so that

- each basic (non-zero) variable appears in exactly one constraint,
- f is in terms of the non-basic variables (which are zero at vertex D).

Alternatively, we want to express x_1, x_3, f in terms of the non-basic variables x_2, x_4.

How? Add multiples of the pivot equation (2) to the other equations.

$$
\begin{array}{rccccl}
(1) - \frac{1}{2} \times (2): & & \frac{5}{2}x_2 & +x_3 & -\frac{1}{2}x_4 & = 5 \qquad (4) \\
\frac{1}{2} \times (2): & x_1 & +\frac{1}{2}x_2 & & +\frac{1}{2}x_4 & = 4 \qquad (5) \\
(3) - \frac{1}{2} \times (2): & & \frac{1}{2}x_2 & & -\frac{1}{2}x_4 & = f - 4 \qquad (6)
\end{array}
$$

Now $f = 4 + \frac{1}{2}x_2 - \frac{1}{2}x_4$ for each feasible $\mathbf{x}$.

So we should increase x_2 to increase f.

1′. We are at vertex D, $\mathbf{x} = (4, 0, 5, 0)^T$ and $f = 4$.

2′. From (6), increasing x_2 will increase f (increasing x_4 would decrease f).

   From (4): we can increase x_2 to 2, if we decrease x_3 to 0.
   From (5): we can increase x_2 to 8, if we decrease x_1 to 0.

   The *stricter* restriction on x_2 is from (4), the new pivot row.

3′. So increase x_2 to 2, decrease x_3 to 0 (x_4 stays at 0, and from (5) x_1 decreases to 3).
   With these changes we move to the BFS $\mathbf{x} = (3, 2, 0, 0)^T$, vertex C.

Rewrite (4)–(6) so that they correspond to vertex C:

$$\begin{array}{rccccl}
\frac{2}{5}\times(4): & & x_2 & +\frac{2}{5}x_3 & -\frac{1}{5}x_4 & = \; 2 & (7)\\
(5)-\frac{1}{5}\times(4): & x_1 & & -\frac{1}{5}x_3 & +\frac{3}{5}x_4 & = \; 3 & (8)\\
(6)-\frac{1}{5}\times(4): & & & -\frac{1}{5}x_3 & -\frac{2}{5}x_4 & = \; f-5 & (9)
\end{array}$$

$1''$. We are at vertex C, $\mathbf{x} = (3,2,0,0)^T$ and $f = 5$.

$2''$. We have deduced that $f = 5 - \frac{1}{5}x_3 - \frac{2}{5}x_4 \leqslant 5$ for each feasible $\mathbf{x}$.

So $x_3 = x_4 = 0$ is the best we can do!

In that case we can read off $x_1 = 3$ and $x_2 = 2$.

So $\mathbf{x} = (3,2,0,0)^T$, which has $f = 5$, is optimal.

Summary

At each stage:

- let $B = \{j : x_j \text{ is basic}\}$
- we express x_j, $j \in B$ and f in terms of x_j, $j \notin B$
- setting $x_j = 0$, $j \notin B$, we can read off f and x_j, $j \in B$ (gives a BFS!).

At each update:

- look at f as expressed in terms of x_j, $j \notin B$
- which x_j, $j \notin B$, would we like to increase?
- if none, STOP!
- otherwise, choose one and increase it as much as possible, i.e. until a variable x_j, $j \in B$, becomes 0.

Summary continued

So at each update

- one new variable *enters* B (becomes basic, typically becomes non-zero)
- another one *leaves* B (becomes non-basic, becomes 0).

This gives a new BFS.

We update our expressions to correspond to the new B.

Simplex algorithm

We can write equations

$$\begin{array}{rrrrcrl}
x_1 & +3x_2 & +x_3 & & = & 9 & (1)\\
2x_1 & +\,x_2 & & +x_4 & = & 8 & (2)\\
x_1 & +\,x_2 & & & = & f-0 & (3)
\end{array}$$

as a 'tableau'

$$\begin{array}{c|cccc|c|c}
 & x_1 & x_2 & x_3 & x_4 & & \\
x_3 & 1 & 3 & 1 & 0 & 9 & \rho_1\\
x_4 & 2 & 1 & 0 & 1 & 8 & \rho_2\\
\hline
 & 1 & 1 & 0 & 0 & 0 & \rho_3
\end{array}$$

This initial tableau represents the BFS $\mathbf{x} = (0,0,9,8)^T$ at which $f = 0$. The basic variables are specified on the left.

Note the identity matrix in the x_3, x_4 columns (first two rows), and the zeros in the bottom row below it.

The tableau

At a given stage, the tableau has the form

$$\begin{array}{c|c} (\overline{a}_{ij}) & \overline{\mathbf{b}} \\ \hline \overline{\mathbf{c}}^T & \overline{f} \end{array}$$

which means:

$\overline{A}\mathbf{x} = \overline{\mathbf{b}}$ which has the same solutions as $A\mathbf{x} = \mathbf{b}$

and

$f(\mathbf{x}) = \overline{\mathbf{c}}^T\mathbf{x} - \overline{f}$ for each $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{b}$.

We start from $\overline{A} = A$, $\overline{\mathbf{b}} = \mathbf{b}$, $\overline{\mathbf{c}} = \mathbf{c}$ and $\overline{f} = 0$.

Updating the tableau is called *pivoting*.

To update ('pivot')

1. *Choose a pivot column*
   Choose a j such that $\overline{c}_j > 0$ (corresponds to the non-basic variable x_j that we want to increase from 0).
   Here we can take $j = 1$.

2. *Choose a pivot row*
   Among the i's with $\overline{a}_{ij} > 0$, choose i to minimize $\overline{b}_i/\overline{a}_{ij}$ (strictest limit on how much we can increase x_j).
   Here $i = 2$ since $8/2 < 9/1$.

3. Do row operations so that column j gets a 1 in row i and 0s elsewhere:
   - multiply row i by $\dfrac{1}{\overline{a}_{ij}}$
   - for $i' \neq i$
     add $-\dfrac{\overline{a}_{i'j}}{\overline{a}_{ij}} \times (\text{row } i)$ to row i'
   - add $-\dfrac{\overline{c}_j}{\overline{a}_{ij}} \times (\text{row } i)$ to objective function row.

(These row operations do not change which sets of constraint columns are linearly independent.)

In our example we pivot on $j = 1$, $i = 2$. The updated tableau is

$$\begin{array}{c|cccc|c|l} & x_1 & x_2 & x_3 & x_4 & & \\ x_3 & 0 & \frac{5}{2} & 1 & -\frac{1}{2} & 5 & \rho_1' = \rho_1 - \rho_2' \\ x_1 & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 4 & \rho_2' = \frac{1}{2}\rho_2 \\ \hline & 0 & \frac{1}{2} & 0 & -\frac{1}{2} & -4 & \rho_3' = \rho_3 - \rho_2' \end{array}$$

which means

$$\begin{array}{rrrrcl} & \frac{5}{2}x_2 & +x_3 & -\frac{1}{2}x_4 & = & 5 \\ x_1 & +\frac{1}{2}x_2 & & +\frac{1}{2}x_4 & = & 4 \\ & \frac{1}{2}x_2 & & -\frac{1}{2}x_4 & = & f - 4 \end{array}$$

non-basic variables x_2, x_4 are 0, $\mathbf{x} = (4, 0, 5, 0)^T$.

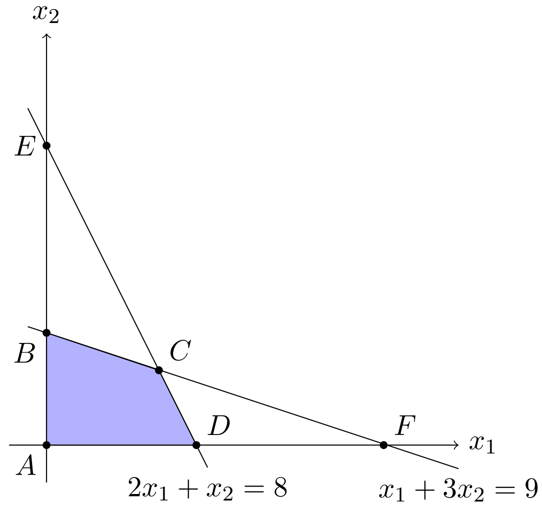
Note the identity matrix inside $(\overline{a}_{ij})$ telling us this.

Next pivot: column 2, row 1 since $\dfrac{5}{5/2} < \dfrac{4}{1/2}$.

$$
\begin{array}{c|cccc|c|l}
 & x_1 & x_2 & x_3 & x_4 & & \\
x_2 & 0 & 1 & \frac{2}{5} & -\frac{1}{5} & 2 & \rho_1' = \frac{2}{5}\rho_1 \\
x_1 & 1 & 0 & -\frac{1}{5} & \frac{3}{5} & 3 & \rho_2' = \rho_2 - \frac{1}{2}\rho_1' \\
\hline
 & 0 & 0 & -\frac{1}{5} & -\frac{2}{5} & -5 & \rho_3' = \rho_3 - \frac{1}{2}\rho_1'
\end{array}
$$

Now we have only non-positive entries in the bottom row: STOP.
$\mathbf{x} = (3, 2, 0, 0)^T$, $f(\mathbf{x}) = 5$ optimal.

Geometric picture for $P1$



	x_1	x_2	x_3	x_4
A	0	0	9	8
B	4	0	5	0
C	3	2	0	0
D	0	3	0	5

A, B, C, D are BFSs. E, F are basic solutions but not feasible.

Simplex algorithm: $A \to D \to C$ (or $A \to B \to C$ is we choose a different column for the first pivot).

Higher-dimensional problems are less trivial!

Comments on simplex tableaux

- We always find an $m \times m$ identity matrix embedded in $(\overline{a}_{ij})$, in the columns corresponding to the basic variables x_j, $j \in B$. (We are assuming A has rank m.)
- In the objective function row (bottom row) we find *zeros* in these columns.

Hence f and x_j, $j \in B$, are all written in terms of x_j, $j \notin B$. Since we set $x_j = 0$ for $j \notin B$, it's then trivial to read off the values of f and x_j, $j \in B$.

Comments on simplex algorithm

- *Choosing a pivot column*
  We may choose any j such that the *reduced profit* $\bar{c}_j > 0$.
  In general, there is no easy way to tell which such j will result in fewest pivot steps.

- *Choosing a pivot row*
  Having chosen pivot column j (which variable x_j to increase), we look for rows with $\bar{a}_{ij} > 0$.

  If $\bar{a}_{ij} \leqslant 0$, constraint i places no restriction on the increase of x_j.

  If $\bar{a}_{ij} \leqslant 0$ for all i, x_j can be increased *without limit*: the objective function is *unbounded*.

  Otherwise, the most stringent limit comes from an i that minimises $\bar{b}_i/\bar{a}_{ij}$.

4 The simplex algorithm (2)

Two issues to consider:

- can we always find a BFS from which to *start* the simplex algorithm?
- does the simplex algorithm always *terminate*, i.e. find an optimal BFS or show the objective function is unbounded?

Initialisation

To start the simplex algorithm, we need to start from a BFS, with basic variables x_j, $j \in B$, written in terms of non-basic variables x_j, $j \notin B$.

If $\mathbf{b} \geqslant \mathbf{0}$ and A already contains I_m as an $m \times m$ submatrix, this is easy!

This *always* happens if $\mathbf{b} \geqslant \mathbf{0}$ and A is created by adding slack variables to make $\leqslant$ inequalities into equalities.

Suppose the constraints are

$$A\mathbf{x} \leqslant \mathbf{b}, \quad \mathbf{x} \geqslant \mathbf{0}$$

where $\mathbf{b} \geqslant \mathbf{0}$. Then an initial BFS is immediate: introducing slack variables $\mathbf{z} = (z_1, \dots, z_m)$,

$$A\mathbf{x} + \mathbf{z} = \mathbf{b}, \quad \mathbf{x}, \mathbf{z} \geqslant \mathbf{0}$$

the initial tableau is

$$\begin{array}{ccc|ccc|c}
x_1 & \cdots & x_n & z_1 & \cdots & z_m & \\
 & A & & & I_m & & \mathbf{b} \\
\hline
 & \mathbf{c}^T & & & \mathbf{0}^T & & 0
\end{array}$$

An initial BFS is $\binom{\mathbf{x}}{\mathbf{z}} = \binom{\mathbf{0}}{\mathbf{b}}$, with the slack variables as basic variables.

Example

$$\begin{array}{llllll}
\max & 6x_1 & + \; x_2 & + \; x_3 & & \\
\text{s.t.} & 9x_1 & + \; x_2 & + \; x_3 & \leqslant 18 \\
 & 24x_1 & + \; x_2 & + \; 4x_3 & \leqslant 42 \\
 & 12x_1 & + \; 3x_2 & + \; 4x_3 & \leqslant 96 \\
 & & x_1, x_2, x_3 \geqslant 0. & & &
\end{array}$$

Add slack variables:

$$\begin{array}{l}
9x_1 + x_2 + x_3 + x_4 = 18 \\
24x_1 + x_2 + 4x_3 + x_5 = 42 \\
12x_1 + 3x_2 + 4x_3 + x_6 = 96 \\
x_1, \dots, x_6 \geqslant 0.
\end{array}$$

Solve by simplex

$$\begin{array}{rrrrrr|r} 9 & 1 & 1 & 1 & 0 & 0 & 18 \\ 24 & 1 & 4 & 0 & 1 & 0 & 42 \\ 12 & 3 & 4 & 0 & 0 & 1 & 96 \\ \hline 6 & 1 & 1 & 0 & 0 & 0 & 0 \end{array}$$

This initial tableau is already in the form we need.

$$\begin{array}{rrrrrr|r} 9 & 1 & 1 & 1 & 0 & 0 & 18 \\ 15 & 0 & 3 & -1 & 1 & 0 & 24 \\ -15 & 0 & 1 & -3 & 0 & 1 & 42 \\ \hline -3 & 0 & 0 & -1 & 0 & 0 & -18 \end{array} \qquad \begin{array}{l} \rho_1' = \rho_1 \\ \rho_2' = \rho_2 - \rho_1 \\ \rho_3' = \rho_3 - 3\rho_1 \\ \rho_4' = \rho_4 - \rho_1 \end{array}$$

This solution is optimal: $x_2 = 18$, $x_1 = 0$, $x_3 = 0$, $f = 18$.

What if A doesn't have this form?

In general we can write the constraints as

$$A\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geqslant \mathbf{0} \tag{4.1}$$

where $\mathbf{b} \geqslant \mathbf{0}$ (if necessary, multiply rows by -1 to get $\mathbf{b} \geqslant \mathbf{0}$).

If there is no obvious initial BFS and we need to find one, we can introduce *artificial variables* $w_1, \dots, w_m$ and solve the LP problem

$$\begin{array}{ll} \min\limits_{\mathbf{x},\mathbf{w}} & \sum_{i=1}^{m} w_i \\ \text{subject to} & A\mathbf{x} + \mathbf{w} = \mathbf{b} \\ & \mathbf{x}, \mathbf{w} \geqslant \mathbf{0}. \end{array} \tag{4.2}$$

(We do not need w_i if the ith unit m-vector is a column of A.)

Two cases arise:

(1) if (4.1) has a feasible solution, then (4.2) has optimal value 0 with $\mathbf{w} = \mathbf{0}$.

(2) if (4.1) has no feasible solution, then the optimal value of (4.2) is > 0.

We can apply the simplex algorithm to determine whether it's case (1) or (2).

- In case (2), stop and check the data!
- In case (1), the optimal BFS for (4.2) with $w_i \equiv 0$ yields a BFS for (4.1).

This leads us to the *two-phase simplex algorithm.*

Two-phase simplex algorithm

Example:

$$\begin{array}{ll} \text{maximize} & 3x_1 \phantom{{}+2x_2} + x_3 \\ \text{subject to} & x_1 + 2x_2 + x_3 = 30 \\ & x_1 - 2x_2 + 2x_3 = 18 \\ & \mathbf{x} \geqslant \mathbf{0} \end{array}$$

With artificial variables:

$$\begin{array}{ll} \text{minimize} & w_1 + w_2 \\ \text{subject to} & x_1 + 2x_2 + x_3 + w_1 = 30 \\ & x_1 - 2x_2 + 2x_3 + w_2 = 18 \\ & \mathbf{x}, \mathbf{w} \geqslant \mathbf{0}. \end{array}$$

Note: To minimise $w_1 + w_2$, we can maximise $-w_1 - w_2$. So start from the 'nearly proper' simplex tableau

$$\begin{array}{rrrrr|r} x_1 & x_2 & x_3 & w_1 & w_2 & \\ 1 & 2 & 1 & 1 & 0 & 30 \\ 1 & -2 & 2 & 0 & 1 & 18 \\ \hline 0 & 0 & 0 & -1 & -1 & 0 \end{array}$$

The objective function row should be expressed in terms of non-basic variables (the entries under the 'identity matrix' should be 0).

So start by adding row 1 + row 2 to objective row:

$$\begin{array}{rrrrr|r} x_1 & x_2 & x_3 & w_1 & w_2 & \\ 1 & 2 & 1 & 1 & 0 & 30 \\ 1 & -2 & 2 & 0 & 1 & 18 \\ \hline 2 & 0 & 3 & 0 & 0 & 48 \end{array}$$

Now start with simplex – pivot on $\overline{a}_{23}$:

$$\begin{array}{rrrrr|r} x_1 & x_2 & x_3 & w_1 & w_2 & \\ \frac{1}{2} & 3 & 0 & 1 & -\frac{1}{2} & 21 \\ \frac{1}{2} & -1 & 1 & 0 & \frac{1}{2} & 9 \\ \hline \frac{1}{2} & 3 & 0 & 0 & -\frac{3}{2} & 21 \end{array}$$

Pivot on $\overline{a}_{12}$:

$$\begin{array}{rrrrr|r} x_1 & x_2 & x_3 & w_1 & w_2 & \\ \frac{1}{6} & 1 & 0 & \frac{1}{3} & -\frac{1}{6} & 7 \\ \frac{2}{3} & 0 & 1 & \frac{1}{3} & \frac{1}{3} & 16 \\ \hline 0 & 0 & 0 & -1 & -1 & 0 \end{array}$$

So we have found a point with $-w_1 - w_2 = 0$, i.e. $\mathbf{w} = \mathbf{0}$.

Phase I is finished. A BFS of the original problem is $\mathbf{x} = (0, 7, 16)^T$.

Deleting the $\mathbf{w}$ columns and replacing the objective row by the original objective function $3x_1 + x_3$:

$$\begin{array}{rrr|r} x_1 & x_2 & x_3 & \\ \frac{1}{6} & 1 & 0 & 7 \\ \frac{2}{3} & 0 & 1 & 16 \\ \hline 3 & 0 & 1 & 0 \end{array}$$

Again we want zeros below the identity matrix – subtract row 2 from row 3:

$$\begin{array}{rrr|r} x_1 & x_2 & x_3 & \\ \frac{1}{6} & 1 & 0 & 7 \\ \frac{2}{3} & 0 & 1 & 16 \\ \hline \frac{7}{3} & 0 & 0 & -16 \end{array}$$

Now do simplex.

Pivot on $\overline{a}_{21}$:

$$
\begin{array}{ccc|c}
x_1 & x_2 & x_3 & \\
0 & 1 & -\frac{1}{4} & 3 \\
1 & 0 & \frac{3}{2} & 24 \\
\hline
0 & 0 & -\frac{7}{2} & -72
\end{array}
$$

Done! Maximum at $x_1 = 24$, $x_2 = 3$, $x_3 = 0$, $f = 72$.

(When (4.1) is feasible as here, $\mathbf{w} = \mathbf{0}$ at the end of phase I. We expect that each artificial variable w_j is non-basic – if not, there is a minor complication, which we shall ignore.)

Shadow prices

Recall the activity analysis $P1$:

$$
\begin{array}{lrcl}
\max & x_1 + x_2 & & \\
\text{s.t.} & x_1 + 3x_2 & \leqslant & 9 \\
 & 2x_1 + x_2 & \leqslant & 8 \\
 & x_1, x_2 & \geqslant & 0
\end{array}
$$

We had initial tableau

$$
\begin{array}{cccc|cl}
1 & 3 & 1 & 0 & 9 & \rho_1 \\
2 & 1 & 0 & 1 & 8 & \rho_2 \\
\hline
1 & 1 & 0 & 0 & 0 & \rho_3
\end{array}
$$

and final tableau

$$
\begin{array}{cccc|cl}
0 & 1 & \frac{2}{5} & -\frac{1}{5} & 2 & \rho_1' \\
1 & 0 & -\frac{1}{5} & \frac{3}{5} & 3 & \rho_2' \\
\hline
0 & 0 & -\frac{1}{5} & -\frac{2}{5} & -5 & \rho_3'
\end{array}
$$

If we increase x_3 from 0 to t (for some small $t > 0$) the profit (objective function) goes down by $\frac{t}{5}$. This suggests that the marginal value of time on machine M_1 is $\frac{1}{5}$.

From the mechanics of the simplex algorithm:

- ρ_1', ρ_2' are created by taking linear combinations of ρ_1, ρ_2
- ρ_3' is $\rho_3 -$ (a linear combination of ρ_1, ρ_2).

Directly from the tableaux (look at columns 3 and 4):

$$
\begin{aligned}
\rho_1' &= \tfrac{2}{5}\rho_1 - \tfrac{1}{5}\rho_2 \\
\rho_2' &= -\tfrac{1}{5}\rho_1 + \tfrac{3}{5}\rho_2 \\
\rho_3' &= \rho_3 - \tfrac{1}{5}\rho_1 - \tfrac{2}{5}\rho_2
\end{aligned}
$$

Suppose we change the constraints to

$$\begin{aligned} x_1 + 3x_2 &\leqslant 9 + \varepsilon_1 \\ 2x_1 + x_2 &\leqslant 8 + \varepsilon_2. \end{aligned}$$

Then the final tableau will change to

$$\begin{array}{cccc|c} 0 & 1 & \frac{2}{5} & -\frac{1}{5} & 2 + \frac{2}{5}\varepsilon_1 - \frac{1}{5}\varepsilon_2 \\ 1 & 0 & -\frac{1}{5} & \frac{3}{5} & 3 - \frac{1}{5}\varepsilon_1 + \frac{3}{5}\varepsilon_2 \\ \hline 0 & 0 & -\frac{1}{5} & -\frac{2}{5} & -5 - \frac{1}{5}\varepsilon_1 - \frac{2}{5}\varepsilon_2 \end{array}$$

This is still a valid tableau as long as

$$\begin{aligned} 2 + \tfrac{2}{5}\varepsilon_1 - \tfrac{1}{5}\varepsilon_2 &\geqslant 0 \\ 3 - \tfrac{1}{5}\varepsilon_1 + \tfrac{3}{5}\varepsilon_2 &\geqslant 0. \end{aligned}$$

In that case we still get an optimal BFS from it, with optimal value

$$5 + \tfrac{1}{5}\varepsilon_1 + \tfrac{2}{5}\varepsilon_2.$$

The profit increases by $\frac{1}{5}$ per extra hour on M_1 and by $\frac{2}{5}$ per extra hour on M_2 (if the changes are 'small enough').

These *shadow prices* can always be read off from the initial tableau.

Termination of simplex algorithm

'Typical situation': each BFS has exactly m non-zero and $n - m$ zero variables.

Then each pivoting operation (moving from one BFS to another) strictly increases the new variable 'entering the basis' and so strictly increases the objective function.

Since there are only finitely many BFSs, we have the following theorem.

Theorem 4.1

If each BFS has exactly m non-zero variables, then the simplex algorithm terminates (i.e. finds an optimal solution or proves that the objective function is unbounded).

What if some BFSs have extra zero variables?

We say that the problem is *degenerate*.

Almost always: this is no problem.

In rare cases, some choices of pivot columns/rows may cause the algorithm to *cycle* (repeat itself). There are various ways to avoid this (e.g. always choosing the leftmost column, and then the highest row, can be proved to work always.)

See e.g. Chvátal's book for a nice discussion.

5 Duality: Introduction

Recall the activity analysis $P1$:

$$\begin{aligned} \text{maximise} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 + 3x_2 \leqslant 9 \qquad (5.1) \\ & 2x_1 + x_2 \leqslant 8 \qquad (5.2) \\ & x_1, x_2 \geqslant 0. \end{aligned}$$

'Obvious' bounds on $f(\mathbf{x}) = x_1 + x_2$:

$$\begin{aligned} x_1 + x_2 \leqslant x_1 + 3x_2 \leqslant 9 \quad & \text{from } (5.1) \\ x_1 + x_2 \leqslant 2x_1 + x_2 \leqslant 8 \quad & \text{from } (5.2). \end{aligned}$$

By combining the constraints we can improve the bound, e.g. $\frac{1}{3}[(5.1) + (5.2)]$:

$$x_1 + x_2 \leqslant x_1 + \tfrac{4}{3}x_2 \leqslant \tfrac{17}{3}.$$

More systematically?

For $y_1, y_2 \geqslant 0$, consider $y_1 \times (5.1) + y_2 \times (5.2)$. We obtain

$$(y_1 + 2y_2)x_1 + (3y_1 + y_2)x_2 \leqslant 9y_1 + 8y_2.$$

Since we want an upper bound for $x_1 + x_2$, we need coefficients $\geqslant 1$:

$$\begin{aligned} y_1 + 2y_2 &\geqslant 1 \\ 3y_1 + y_2 &\geqslant 1. \end{aligned}$$

How to get the best bound by this method?

$$\begin{aligned} D1: \text{minimise} \quad & 9y_1 + 8y_2 \\ \text{subject to} \quad & y_1 + 2y_2 \geqslant 1 \\ & 3y_1 + y_2 \geqslant 1 \\ & y_1, y_2 \geqslant 0. \end{aligned}$$

$P1$ = 'primal problem', $D1$ = 'dual of $P1$'.

Duality: General

In general, given a primal problem

$$P: \quad \text{maximise } \mathbf{c}^T\mathbf{x} \quad \text{subject to } A\mathbf{x} \leqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}$$

the *dual* of P is defined by

$$D: \quad \text{minimise } \mathbf{b}^T\mathbf{y} \quad \text{subject to } A^T\mathbf{y} \geqslant \mathbf{c},\ \mathbf{y} \geqslant \mathbf{0}.$$

Exercise

The dual of the dual is the primal (with suitable recasting).

Weak duality

Theorem 5.1 (Weak duality theorem)

If $\mathbf{x}$ is feasible for P, and $\mathbf{y}$ is feasible for D, then

$$\mathbf{c}^T\mathbf{x} \leqslant \mathbf{b}^T\mathbf{y}.$$

Proof.

Since $\mathbf{x} \geqslant \mathbf{0}$ and $A^T\mathbf{y} \geqslant \mathbf{c}$

$$\mathbf{c}^T\mathbf{x} \leqslant (A^T\mathbf{y})^T\mathbf{x} = \mathbf{y}^T A\mathbf{x}.$$

Since $\mathbf{y} \geqslant \mathbf{0}$ and $A\mathbf{x} \leqslant \mathbf{b}$

$$\mathbf{y}^T A\mathbf{x} \leqslant \mathbf{y}^T\mathbf{b} = \mathbf{b}^T\mathbf{y}.$$

Hence $$\mathbf{c}^T\mathbf{x} \leqslant \mathbf{y}^T A\mathbf{x} \leqslant \mathbf{b}^T\mathbf{y} \qquad \square$$

Comments

Suppose $\mathbf{y}$ is a feasible solution to D. Then any feasible solution $\mathbf{x}$ to P has value bounded above by $\mathbf{b}^T\mathbf{y}$.

So D feasible $\implies$ P has bounded value (or is infeasible).

Similarly P feasible $\implies$ D has bounded value (or is infeasible).

Corollary 5.2 (Optimality Test)

If $\mathbf{x}^*$ is feasible for P, $\mathbf{y}^*$ is feasible for D, and $\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$, then $\mathbf{x}^*$ is optimal for P and $\mathbf{y}^*$ is optimal for D.

Proof.

For all $\mathbf{x}$ feasible for P,

$$\begin{aligned} \mathbf{c}^T\mathbf{x} &\leqslant \mathbf{b}^T\mathbf{y}^* \quad \text{by weak duality, Theorem 5.1} \\ &= \mathbf{c}^T\mathbf{x}^* \end{aligned}$$

and so $\mathbf{x}^*$ is optimal for P.

Similarly, for all $\mathbf{y}$ feasible for D,

$$\mathbf{b}^T\mathbf{y} \geqslant \mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*$$

and so $\mathbf{y}^*$ is optimal for D. □

As an example of applying this result, look at $\mathbf{x}^* = (3, 2)^T$, $\mathbf{y}^* = (\frac{1}{5}, \frac{2}{5})^T$ for $P1$ and $D1$ above.

Both are feasible, both have value 5. So both are optimal.

Does this nice situation always occur?

Strong duality

Theorem 5.3 (Duality Theorem)

Suppose that P and D have feasible solutions. Then both have optimal solutions $\mathbf{x}^*$ and $\mathbf{y}^*$ respectively, and

$$\mathbf{c}^T\mathbf{x}^* = \mathbf{b}^T\mathbf{y}^*.$$

Proof.

Write the constraints of P as $A\mathbf{x} + \mathbf{z} = \mathbf{b}$, $\mathbf{x}, \mathbf{z} \geqslant \mathbf{0}$.

We can start the simplex method on P, and since P has bounded value, the simplex method must terminate with an optimal tableau.

Consider the bottom row in this tableau.

$$
\begin{array}{ccc|ccc|c}
\multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad}^{\mathbf{x}\text{-columns}}} & \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad}^{\mathbf{z}\text{-columns}}} & \\
 & & & & & & \\
 & & & & & & \\
\hline
c_1^* & \cdots & c_n^* & -y_1^* & \cdots & -y_m^* & -f^*
\end{array}
$$

Here f^* is the optimal value of $\mathbf{c}^T\mathbf{x}$,

$$c_j^* \leqslant 0 \quad \text{for all } j \tag{5.3}$$

$$-y_i^* \leqslant 0 \quad \text{for all } i \tag{5.4}$$

By looking at the **z**-columns, we find that to obtain the final bottom row from the initial one we subtract $\sum_i y_i^* \rho_i$, where ρ_i is the ith row of the initial tableau.

(To see this, think first of the case $\mathbf{b} \geqslant \mathbf{0}$, with the slack variables as initial basic variables. The result still holds if we use the two-phase method to get started.)

Thus $-f^* = -\sum_i y_i^* b_i$, so

$$f^* = \mathbf{b}^T\mathbf{y}^*. \tag{5.5}$$

Also $c_j^* = c_j - \mathbf{y}^{*T}\mathbf{a}_j$ (where $\mathbf{a}_j$ is the jth column of A), so

$$\mathbf{c}^* = \mathbf{c} - A^T\mathbf{y}^*.$$

From (5.3) $\mathbf{c}^* \leqslant \mathbf{0}$, hence

$$A^T\mathbf{y}^* \geqslant \mathbf{c}. \tag{5.6}$$

Inequalities (5.4) and (5.6) show that $\mathbf{y}^*$ is feasible for D.

Also, equation (5.5) shows that the objective function of D at $\mathbf{y}^*$ is $\mathbf{b}^T\mathbf{y}^* = f^* =$ optimal value of P.

So by the Optimality Test (Corollary 5.2), $\mathbf{y}^*$ is optimal for D. $\square$

Comments

Note:

- the coefficients $\mathbf{y}^*$ from the bottom row in the columns corresponding to slack variables give us (when negated) the optimal solution to D
- comparing with the shadow prices discussion: these optimal values for the dual variables are the shadow prices!

Example

It is possible that neither P nor D has a feasible solution: consider the problem

$$\begin{array}{ll} \text{maximise} & 2x_1 - x_2 \\ \text{subject to} & x_1 - x_2 \leqslant 1 \\ & -x_1 + x_2 \leqslant -2 \\ & x_1, x_2 \geqslant 0. \end{array}$$

Example

Consider the problem

$$\begin{array}{ll} \text{maximise} & 2x_1 + 4x_2 + x_3 + x_4 \\ \text{subject to} & x_1 + 3x_2 + x_4 \leqslant 4 \\ & 2x_1 + x_2 \leqslant 3 \\ & x_1 + 4x_3 + x_4 \leqslant 3 \\ & x_1, \dots, x_4 \geqslant 0 \end{array}$$

Add slack variables x_5, x_6, x_7 and use the simplex method.

The final tableau is

x_1	x_2	x_3	x_4	x_5	x_6	x_7	
0	1	0	$\frac{2}{5}$	$\frac{2}{5}$	·	·	1
1	0	0	$-\frac{1}{5}$	$-\frac{1}{5}$	·	·	1
0	0	1	$\frac{3}{10}$	$\frac{1}{20}$	·	·	$\frac{1}{2}$
0	0	0	$-\frac{1}{2}$	$-\frac{5}{4}$	$-\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{13}{2}$

(1) By the proof of the Duality Theorem (Theorem 5.3), $\mathbf{y}^* = (\frac{5}{4}, \frac{1}{4}, \frac{1}{4})^T$ is optimal for the dual.

(2) Suppose the RHSs of the original constraints become $4 + \varepsilon_1$, $3 + \varepsilon_2$, $3 + \varepsilon_3$. Then the objective function becomes $\frac{13}{2} + \frac{5}{4}\varepsilon_1 + \frac{1}{4}\varepsilon_2 + \frac{1}{4}\varepsilon_3$.

If the original RHSs of 4, 3, 3 correspond to the amount of raw material i available, then the marginal value of raw material 1, 'the most you'd be prepared to pay per additional unit', is $y_1^* = \frac{5}{4}$ (and similarly for raw material 2 and $y_2^* = \frac{1}{4}$, and so on).

(3) Suppose raw material 1 is available at a price $< \frac{5}{4}$ per unit. How much should you buy? With $\varepsilon_1 > 0$, $\varepsilon_2 = \varepsilon_3 = 0$, the final tableau would be

$$
\begin{array}{c|c}
 & 1 + \frac{2}{5}\varepsilon_1 \\
\cdots & 1 - \frac{1}{5}\varepsilon_1 \\
 & \frac{1}{2} + \frac{1}{20}\varepsilon_1 \\
\hline
 &
\end{array}
$$

For this tableau to represent a BFS, the three entries in the final column must be $\geqslant 0$, giving $\varepsilon_1 \leqslant 5$. So we should buy at least 5 additional units of raw material 1.

(4) The optimal solution $\mathbf{x}^* = (1, 1, \frac{1}{2}, 0)^T$ is *unique* as the entries in the bottom row corresponding to non-basic variables (i.e. the $-\frac{1}{2}$, $-\frac{5}{4}$, $-\frac{1}{4}$, $-\frac{1}{4}$) are < 0.

(5) Suppose now that we can sell the first scarce resource for $\frac{5}{4}$ per unit. Then x_5 has (initial) objective function coefficient $\frac{5}{4}$ not 0.

The reduced profit at the final tableau for x_5 now becomes 0 not $-\frac{5}{4}$.

We could pivot in that column (observe that there would be somewhere to pivot) to get a second optimal BFS $\mathbf{x}^{**}$.

Then $\lambda \mathbf{x}^* + (1 - \lambda)\mathbf{x}^{**}$ is optimal for all $\lambda \in [0, 1]$.

6 Duality: Complementary slackness

Recall

$$\begin{aligned} P: &\quad \text{maximise } \mathbf{c}^T\mathbf{x} \quad \text{subject to } A\mathbf{x} \leqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0} \\ D: &\quad \text{minimise } \mathbf{b}^T\mathbf{y} \quad \text{subject to } A^T\mathbf{y} \geqslant \mathbf{c},\ \mathbf{y} \geqslant \mathbf{0} \end{aligned}$$

The optimal solutions to P and D satisfy 'complementary slackness conditions', that we can use for example to solve one problem when we know a solution of the other.

Theorem 6.1 (Complementary Slackness Theorem)

Suppose $\mathbf{x}$ *is feasible for* P *and* $\mathbf{y}$ *is feasible for* D*. Then* $\mathbf{x}$ *and* $\mathbf{y}$ *are optimal (for* P *and* D *respectively) if and only if*

$$(A^T\mathbf{y} - \mathbf{c})_j x_j = 0 \quad \textit{for all } j \tag{6.1}$$

and

$$(\mathbf{b} - A\mathbf{x})_i y_i = 0 \quad \textit{for all } i. \tag{6.2}$$

Conditions (6.1) and (6.2) are called the *complementary slackness conditions*.

Interpretation

Condition (6.1) says:

if a dual constraint is slack, then the corresponding primal variable is zero

or equivalently

if a primal variable is > 0, then the corresponding dual constraint is tight.

Condition (6.2) says the same except with 'primal' and 'dual' swapped.

Proof.

As in the proof of the weak duality theorem,

$$\mathbf{c}^T\mathbf{x} \leqslant (A^T\mathbf{y})^T\mathbf{x} = \mathbf{y}^T A\mathbf{x} \leqslant \mathbf{y}^T\mathbf{b}. \tag{6.3}$$

From the Duality Theorem,

$$\begin{aligned} & \mathbf{x}, \mathbf{y} \text{ both optimal} \\ \iff & \mathbf{c}^T\mathbf{x} = \mathbf{b}^T\mathbf{y} \\ \iff & \mathbf{c}^T\mathbf{x} = \mathbf{y}^T A\mathbf{x} = \mathbf{b}^T\mathbf{y} \quad \text{from (6.3)} \\ \iff & (\mathbf{y}^T A - \mathbf{c}^T)\mathbf{x} = 0 \quad \text{and} \quad \mathbf{y}^T(\mathbf{b} - A\mathbf{x}) = 0 \\ \iff & \sum_{j=1}^{n} (A^T\mathbf{y} - \mathbf{c})_j x_j = 0 \quad \text{and} \quad \sum_{i=1}^{m} (\mathbf{b} - A\mathbf{x})_i y_i = 0. \end{aligned}$$

But $A^T\mathbf{y} \geqslant \mathbf{c}$ and $\mathbf{x} \geqslant \mathbf{0}$, so $\sum_{j=1}^n (A^T\mathbf{y} - \mathbf{c})_j x_j$ is a sum of non-negative terms. Hence $\sum_{j=1}^n (A^T\mathbf{y} - \mathbf{c})_j x_j = 0$ is equivalent to (6.1).

Similarly, $A\mathbf{x} \leqslant \mathbf{b}$ and $\mathbf{y} \geqslant \mathbf{0}$, so $\sum_{i=1}^m (\mathbf{b} - A\mathbf{x})_i y_i$ is a sum of non-negative terms. Hence $\sum_{i=1}^m (\mathbf{b} - A\mathbf{x})_i y_i = 0$ is equivalent to (6.2).

□

Comments

What's the use of complementary slackness?

Among other things, given an optimal solution of P (or D), it makes finding an optimal solution of D (or P) easy, because we know which the non-zero variables can be and which constraints must be tight.

Sometimes one of P and D is much easier to solve than the other, e.g. with 2 variables, 5 constraints, we can solve graphically, but 5 variables and 2 constraints is not so easy.

Example

Consider P and D with

$$A = \begin{pmatrix} 1 & 4 & 0 \\ 3 & -1 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \qquad \mathbf{c} = \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}.$$

Is $\tilde{\mathbf{x}} = (0, \frac{1}{4}, \frac{13}{4})^T$ optimal? It is feasible. If it is optimal, then

since $\tilde{x}_2 > 0$, for any optimal $\mathbf{y}$ we have $(A^T\mathbf{y})_2 = c_2$, that is $4y_1 - y_2 = 1$; and

since $\tilde{x}_3 > 0$, for any optimal $\mathbf{y}$ we have $(A^T\mathbf{y})_3 = c_3$, that is $0y_1 + y_2 = 3$.

These equations give $\mathbf{y} = (y_1, y_2)^T = (1, 3)^T$.

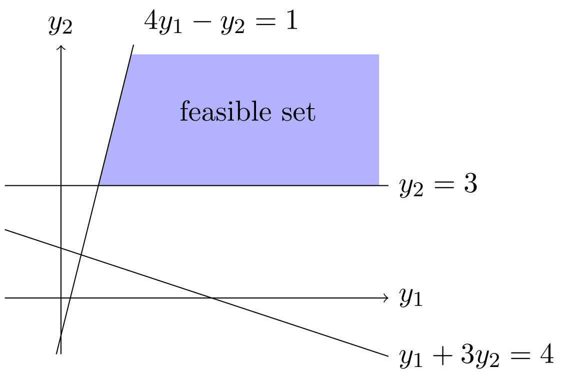
The remaining dual constraint $y_1 + 3y_2 \geqslant 4$ is also satisfied, so $\tilde{\mathbf{y}} = (1, 3)^T$ is feasible for D.

Thus $\tilde{\mathbf{x}} = (0, \frac{1}{4}, \frac{13}{4})^T$ and $\tilde{\mathbf{y}} = (1, 3)^T$ are feasible and satisfy complementary slackness, therefore they are optimal by Theorem 6.1.

Alternatively, at this point we could note that $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are feasible and $\mathbf{c}^T\tilde{\mathbf{x}} = 10 = \mathbf{b}^T\tilde{\mathbf{y}}$, so they are optimal by the Optimality Test (Corollary 5.2).

Example continued

If we don't know the solution to P, we can first solve D graphically.



The optimal solution is at $\tilde{\mathbf{y}} = (1,3)^T$, and we can use this to solve P: for any optimal $\mathbf{x}$

since $\tilde{y}_1 > 0, \quad x_1 + 4x_2 = 1$

since $\tilde{y}_2 > 0, \quad 3x_1 - x_2 + x_3 = 3$

since $\tilde{y}_1 + 3\tilde{y}_2 > 4, \quad x_1 = 0$

and so $\mathbf{x} = (0, \frac{1}{4}, \frac{13}{4})^T$.

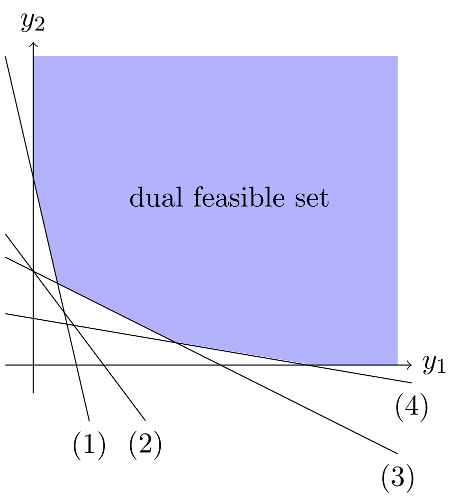
Example

Consider the primal problem

$$\begin{array}{llr} \text{maximise} & 10x_1 + 10x_2 + 20x_3 + 20x_4 & \\ \text{subject to} & 12x_1 + 8x_2 + 6x_3 + 4x_4 \leqslant 210 & \\ & 3x_1 + 6x_2 + 12x_3 + 24x_4 \leqslant 210 & \\ & x_1, \ldots, x_4 \geqslant 0 & \end{array}$$

with dual

$$\begin{array}{llr} \text{minimise} & 210y_1 + 210y_2 & \\ \text{subject to} & 12y_1 + 3y_2 \geqslant 10 & (1) \\ & 8y_1 + 6y_2 \geqslant 10 & (2) \\ & 6y_1 + 12y_2 \geqslant 20 & (3) \\ & 4y_1 + 24y_2 \geqslant 20 & (4) \\ & y_1, y_2 \geqslant 0. & \end{array}$$



The dual optimum is where lines (1) and (3) intersect.

Since the second and fourth dual constraints are slack at the optimum, each optimal $\mathbf{x}$ has $x_2 = x_4 = 0$.

Also, since $y_1, y_2 > 0$ at the optimum,

$$\left.\begin{aligned} 12x_1 + 6x_3 &= 210 \\ 3x_1 + 12x_3 &= 210 \end{aligned}\right\} \text{ and so } \; x_1 = 10, x_3 = 15.$$

Hence the optimal $\mathbf{x}$ is $(10, 0, 15, 0)^T$.

Example continued

Suppose the second 210 is replaced by 421.

The new dual optimum is where (3) and (4) intersect, at which point the first two constraints are slack, so each optimal $\mathbf{x}$ has $x_1 = x_2 = 0$.

Also, since $y_1, y_2 > 0$ at the new optimum,

$$\left.\begin{aligned} 6x_3 + 4x_4 &= 210 \\ 12x_3 + 24x_4 &= 421 \end{aligned}\right\} \text{ and so } \; x_3 = 35 - \tfrac{1}{24}, x_4 = \tfrac{1}{16}.$$

Hence the new optimum is at $\mathbf{x} = (0, 0, 35 - \frac{1}{24}, \frac{1}{16})^T$.

Activity analysis, duality and CS

A firm makes goods G_j for $j = 1, \ldots, n$ using scarce resources R_i for $i = 1, \ldots, m$.

Each unit of G_j requires a_{ij} units of R_i, and yields return of c_j.

Let x_j = number of units of good j made.

Primal P:

$$\max_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{subject to } A\mathbf{x} \leqslant \mathbf{b},\ \mathbf{x} \geqslant \mathbf{0}$$

Dual D:

$$\min_{\mathbf{y}} \mathbf{b}^T\mathbf{y} \quad \text{subject to } A^T\mathbf{y} \geqslant \mathbf{c},\ \mathbf{y} \geqslant \mathbf{0}$$

A competitor wants to buy the firm. She offers $y_i \geqslant 0$ per unit for resource R_i such that, for each good G_j, the return c_j per unit is at most the price of the resources to make it, that is

$$c_j \leqslant a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m$$

(so the firm has no incentive to continue production).
Thus the buyer chooses $\mathbf{y} \geqslant \mathbf{0}$ such that $A^T\mathbf{y} \geqslant \mathbf{c}$.

Subject to this, she aims to minimise the cost $\mathbf{b}^T\mathbf{y}$. So the buyer faces D.

Complementary Slackness. Let $\tilde{\mathbf{x}}$ be optimal in P and let $\tilde{\mathbf{y}}$ be optimal in D. Then:

$(A\tilde{\mathbf{x}})_i < b_i \implies \tilde{y}_i = 0$: 'resources in excess supply are free'

$(A^T\tilde{\mathbf{y}})_j > c_j \implies \tilde{x}_j = 0$: 'unprofitable goods are not made'

7 Two-player zero-sum games (1)

We consider games that are zero-sum in the sense that one player wins what the other loses.

Each player has a list of possible actions.

Players move simultaneously.

Payoff matrix

There is a payoff matrix $A = (a_{ij})$:

Player II plays j

$$\text{Player I plays } i \quad \begin{array}{c} \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{pmatrix} -5 & 3 & 1 & 20 \\ 5 & 5 & 4 & 6 \\ -4 & 6 & 0 & -5 \end{pmatrix} \end{array}$$

If the row player Player I plays i and the column player Player II plays j, then Player I wins a_{ij} from Player II.

The game is defined by the payoff matrix.

Note that our convention is that I wins a_{ij} from II, so $a_{ij} > 0$ is good for the row player, Player I.

Suppose Player I plays conservatively. What's the 'worst that can happen' to him if he chooses row 1? row 2? row 3? (We look at the smallest entry in the appropriate row.)

Similarly, what's the 'worst that can happen' to Player II if Player II chooses a particular column? (We look at the largest entry in that column.)

The matrix above has a special property.

Entry $a_{23} = 4$ is both

- the smallest entry in row 2
- the largest entry in column 3

$(2, 3)$ is a 'saddle point' of A.

We see that:

- Player I can guarantee to win at least 4 by choosing row 2.
- Player II can guarantee to lose at most 4 by choosing column 3.
- Thus both guarantees are best possible.
- The guarantess still hold if either player announces their strategy in advance.

Hence the game is 'solved' and it has 'value' 4.

Mixed strategies

Consider the game of Scissors-Paper-Stone:

- Scissors beats Paper,
- Paper beats Stone,
- Stone beats Scissors.

$$\begin{array}{l} \\ \text{Scissors} \\ \text{Paper} \\ \text{Stone} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Scissors} & \text{Paper} & \text{Stone} \end{array} \\ \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \end{array}$$

No saddle point.

If either player announces a fixed action in advance (e.g. 'play Paper') the other player can take advantage.

So we consider a *mixed strategy* : each action is played with a certain probability. (This is in contrast with a *pure strategy* which is to select a single action with probability 1.)

Suppose Player I plays i with probability p_i, $i = 1, \ldots, m$.

Then Player I's *expected payoff* if Player II plays j is

$$\sum_{i=1}^{m} a_{ij} p_i.$$

Suppose Player I wishes to maximise (over $\mathbf{p}$) his minimal expected payoff

$$\min_j \sum_{i=1}^{m} a_{ij} p_i.$$

Similarly Player II plays j with probability q_j, $j = 1, \ldots, n$, and may look to minimise (over $\mathbf{q}$)

$$\max_i \sum_{j=1}^{n} a_{ij} q_j.$$

This aim for Player II may seem like only one of several sensible aims (and similarly for the earlier aim for Player I).

Soon we will see that they lead to a 'solution' in a very appropriate way, corresponding to the solution for the case of the saddle point.

LP formulation

Consider Player II's problem 'minimise maximal expected payout':

$$\min_{\mathbf{q}} \left\{ \max_i \sum_{j=1}^{n} a_{ij} q_j \right\} \quad \text{subject to} \quad \sum_{j=1}^{n} q_j = 1, \ \mathbf{q} \geqslant \mathbf{0}.$$

This is not yet an LP – look at the objective function.

Equivalent formulation

An equivalent formulation is:

$$\min_{\mathbf{q},v} v \quad \text{subject to } \sum_{j=1}^{n} a_{ij} q_j \leqslant v \quad \text{for } i = 1, \ldots, m$$
$$\sum_{j=1}^{n} q_j = 1$$
$$\mathbf{q} \geqslant \mathbf{0}.$$

since for any given $\mathbf{q}$, when we minimize v it will decrease until it takes the value $\max_i \sum_{j=1}^{n} a_{ij} q_j$.

If v^* is the optimal value then Player II can guarantee expected loss $\leqslant v^*$.

This is an LP but not yet in the most useful form for us.

We could add a constant k to each a_{ij} so that $a_{ij} > 0$ for all i, j. This doesn't change the nature of the game, but guarantees $v > 0$.

So WLOG assume $a_{ij} > 0$ for all i, j.

Now change variables to $x_j = q_j / v$. The problem becomes:

choose $\mathbf{x}$, v to

$$\min v \quad \text{subject to } \sum_{j=1}^{n} a_{ij} x_j \leqslant 1 \quad \text{for } i = 1, \ldots, m$$
$$\sum_{j=1}^{n} x_j = 1/v$$
$$\mathbf{x} \geqslant \mathbf{0}.$$

This transformed problem for Player II is equivalent to

P : choose $\mathbf{x}$ to

$$\max \sum_{j=1}^{n} x_j \quad \text{subject to } A\mathbf{x} \leqslant \mathbf{1}, \ \mathbf{x} \geqslant \mathbf{0}$$

which is now in our 'standard form'. ($\mathbf{1}$ denotes a vector of 1s.)

If $\mathbf{x}^*$ is an optimal solution to P then Player II can guarantee expected loss $\leqslant \frac{1}{\sum_j x_j^*}$.

Doing the same transformations for Player I's problem

$$\max_{\mathbf{p}} \left\{ \min_j \sum_{i=1}^{m} a_{ij} p_i \right\} \quad \text{subject to } \sum_{i=1}^{m} p_i = 1, \ \mathbf{p} \geqslant \mathbf{0}$$

turns it into

D : choose $\mathbf{y}$ to

$$\min \sum_{i=1}^{m} y_i \quad \text{subject to } A^T \mathbf{y} \geqslant \mathbf{1}, \ \mathbf{y} \geqslant \mathbf{0}.$$

(Check: on problem sheet.)

Observe: P and D are *dual* LPs and so have the same optimal value.

Conclusion

Let $\mathbf{x}^*, \mathbf{y}^*$ be optimal for P, D. Then:

- Player I can guarantee an expected gain of at least $v = 1/\sum_{i=1}^{m} y_i^*$, by following strategy $\mathbf{p} = v\mathbf{y}^*$.
- Player II can guarantee an expected loss of at most $v = 1/\sum_{j=1}^{n} x_j^*$, by following strategy $\mathbf{q} = v\mathbf{x}^*$.
- The above is still true if a player announces his strategy in advance.

So the game is 'solved' as in the saddle point case (this was just a special case where the strategies were pure).

v is the *value* of the game (the amount that Player I should 'fairly' pay to Player II for the chance to play the game).

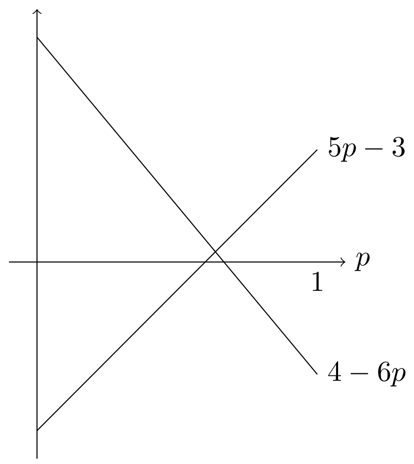
8 Two-player zero-sum games (2)

Some games are easy to solve without the LP formulation, e.g.

$$A = \begin{pmatrix} -2 & 2 \\ 4 & -3 \end{pmatrix}$$

Suppose Player I chooses row 1 with probability p, row 2 with probability $1-p$. Then he should maximise

$$\begin{aligned} &\min(-2p + 4(1-p), 2p - 3(1-p)) \\ =&\min(4-6p, 5p-3) \end{aligned}$$



So the min is maximised when

$$4 - 6p = 5p - 3$$

which occurs when $p = \frac{7}{11}$.

And then $v = \frac{35}{11} - 3 = \frac{2}{11}$.

(We could go on to find Player II's optimal strategy too.)

A useful trick: dominated actions

Consider the game

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 1 & 3 & 4 \\ 3 & 0 & 5 \end{pmatrix}.$$

Player II should never play column 3, since column 2 is always at least as good as column 3 (column 2 dominates column 3.) So we reduce to

$$\begin{pmatrix} 4 & 2 \\ 1 & 3 \\ 3 & 0 \end{pmatrix}$$

(and this makes no difference to player I).

Now Player I will never play row 3 since row 1 is always better, so

$$\begin{pmatrix} 4 & 2 \\ 1 & 3 \end{pmatrix}$$

has the same value (and 'same' optimal strategies) as A.

Final example

Consider the game

$$A = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ -1 & 3 & -1 \end{pmatrix}.$$

Add 1 to each entry

$$\tilde{A} = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 0 & 4 & 0 \end{pmatrix}.$$

The game $\tilde{A}$ has value > 0 (consider e.g. strategy $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$ for Player I).

Solve the LP for Player II's optimal strategy:

$$\max_{x_1,x_2,x_3} x_1 + x_2 + x_3 \quad \text{subject to} \quad \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 0 & 4 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leqslant \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\mathbf{x} \geqslant \mathbf{0}.$$

Initial simplex tableau:

$$\begin{array}{cccccc|c} 0 & 1 & 2 & 1 & 0 & 0 & 1 \\ 2 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 4 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{array}$$

final tableau:

$$\begin{array}{cccccc|c} 0 & 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{8} & \frac{3}{8} \\ 1 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} & \frac{1}{16} & \frac{5}{16} \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} \\ \hline 0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{2} & -\frac{3}{16} & -\frac{15}{16} \end{array}$$

Optimum: $x_1 = \frac{5}{16}$, $x_2 = \frac{1}{4}$, $x_3 = \frac{3}{8}$, and $x_1 + x_2 + x_3 = \frac{15}{16}$.

So value $v = 1/(x_1 + x_2 + x_3) = \frac{16}{15}$.

Player II's optimal strategy: $\mathbf{q} = v\mathbf{x} = \frac{16}{15}(\frac{5}{16}, \frac{1}{4}, \frac{3}{8}) = (\frac{1}{3}, \frac{4}{15}, \frac{2}{5})$.

Dual problem for Player I's strategy has solution $y_1 = \frac{1}{4}$, $y_2 = \frac{1}{2}$, $y_3 = \frac{3}{16}$ (from bottom row of final tableau).

So Player I's optimal strategy:
$\mathbf{p} = v\mathbf{y} = \frac{16}{15}(\frac{1}{4}, \frac{1}{2}, \frac{3}{16}) = (\frac{4}{15}, \frac{8}{15}, \frac{3}{15})$.

The game $\tilde{A}$ has value $\frac{16}{15}$, so the original game A has value $\frac{16}{15} - 1 = \frac{1}{15}$, and the same optimal strategies $\mathbf{p}$ and $\mathbf{q}$.

end