7. Constrained Optimization

7.1 Linear Programming

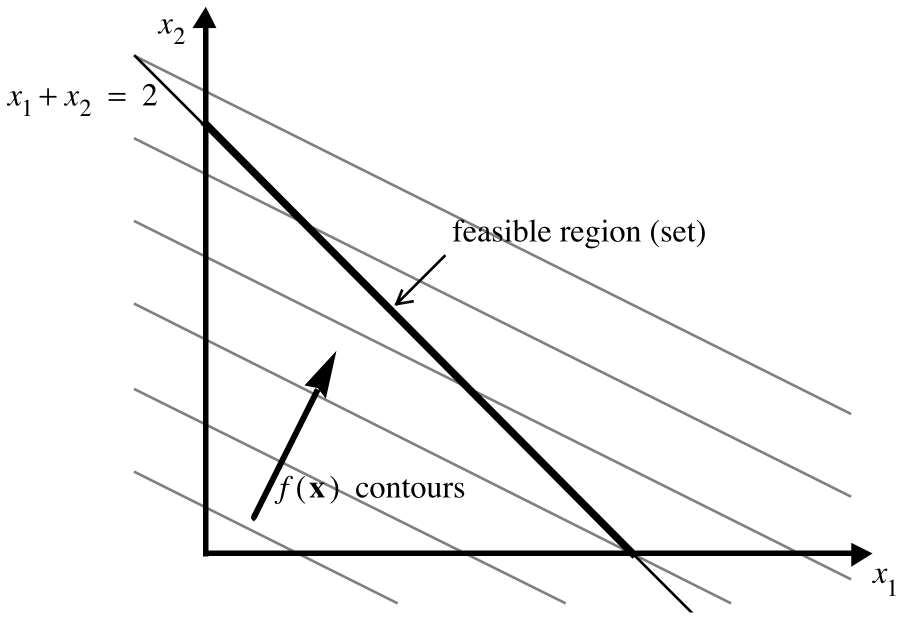
An important class of constrained optimization problems are those for which the *problem functions* (objective and constraints) are *linear* functions of the control variables:

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}\ , & \mathbf{x} = (x_1, x_2, \ldots, x_n)^T \\ \text{subject to} & \mathbf{a}_i^T\mathbf{x} - b_i = 0\ , & i = 1, 2, \ldots, m \\ \text{and} & x_j \geq 0\ , & j = 1, 2, \ldots, n\ . \end{array} \tag{7.1}$$

A number of methods have been developed for tackling such problems, of which one of the best known is the *Simplex Method*.

7.1.1 Example: A Simple Linear Program

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) = x_1 + 2x_2\ , \\ \text{subject to} & x_1 + x_2 = 2\ , \\ \text{and} & x_j \geq 0\ . \end{array}$$



- The equality constraint $x_1 + x_2 = 2$ and the bounds $x_j \geq 0$ define the feasible region (set).
- The objective function $f(\mathbf{x}) = x_1 + 2x_2$ defines a series of parallel lines (contours of f). The minimum value of f is the first of these lines to intersect the feasible set, as f increases.
- In this case, the minimum of f is $x_1 + 2x_2 = 2$, and the optimal solution given by $(x_1, x_2) = (2, 0)$.

This example illustrates several important points about linear programming:

- The constraint equations $\mathbf{a}_i^T\mathbf{x} - b_i = 0$ define a plane in $\mathbb{R}^n$. In this example, this was a line.
- The bounds $x_j \geq 0$ cut out a portion of that plane, called the *feasible region* (or feasible set).
- The feasible region will have a number of corners (or *extreme points*). At these corners a number of control variables will be zero. In this example, the extreme points were $(2, 0)$ and $(0, 2)$.
- The objective function contours defines a series of planes (lines in this example). As f increases, the first plane to touch the feasible set will inevitably do so at one of the extreme points of the feasible set.

7.1.2 Example: Setting up a Linear Program

The following is an example of a linear problem that we want to reduce to the form described above:

An oil refinery has three sources of crude oil:

- a heavy crude that costs $10/barrel;
- a medium crude that costs $17/barrel;
- a light crude that costs $20/barrel.

The refinery produces petrol and heating oil from crude in the following amounts per barrel:

	Petrol	Heating Oil
Heavy Crude	0.3	0.7
Medium Crude	0.5	0.5
Light Crude	0.7	0.3

The refinery is contracted to supply 200000 barrels of petrol and 100000 barrels of heating oil. How much of each type of crude should the refinery buy to minimize its costs?

Give the following variable names to the amount of each crude bought:

Heavy crude: x_1; medium crude: x_2; light crude: x_3.

The contract to supply then becomes two constraints:

$$0.3x_1 + 0.5x_2 + 0.7x_3 = 2 \times 10^5$$
$$0.7x_1 + 0.5x_2 + 0.3x_3 = 1 \times 10^5$$

While the objective function (cost) to be minimized is:

$$f(\mathbf{x}) = 10x_1 + 17x_2 + 20x_3$$

Thus, the problem can be written as:

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} \,, & \\ \text{subject to} & \boldsymbol{A}\mathbf{x} = \mathbf{b} \,, & \qquad (7.2) \\ \text{and} & x_j \geq 0 \,, & \end{array}$$

where

$$\boldsymbol{A} = \begin{bmatrix} 0.3 & 0.5 & 0.7 \\ 0.7 & 0.5 & 0.3 \end{bmatrix}; \ \mathbf{b} = \begin{bmatrix} 2 \times 10^5 \\ 1 \times 10^5 \end{bmatrix}; \ \mathbf{c}^T = \begin{bmatrix} 10 & 17 & 20 \end{bmatrix}.$$

We know that the optimum must lie at one of the extreme points. For this simple problem we can find these points by hand:

$$x_1 = 0 \Rightarrow \quad 0.5x_2 + 0.7x_3 = 2\times10^5$$
$$0.5x_2 + 0.3x_3 = 1\times10^5$$
$$\Rightarrow \ 0.4x_3 = 1\times10^5 \Rightarrow x_3 = 2.5\times10^5$$
$$x_2 = 0.5\times10^5$$
$$f(x) = 58.5\times10^5$$

$$x_2 = 0 \Rightarrow \quad 0.3x_1 + 0.7x_3 = 2\times10^5$$
$$0.7x_1 + 0.3x_3 = 1\times10^5$$
$$\Rightarrow \ 2.1x_1 + 4.9x_3 = 14\times10^5$$
$$2.1x_1 + 0.9x_3 = 3\times10^5$$
$$x_3 = 2.75\times10^5$$
$$x_1 = 0.25\times10^5$$
$$f(x) = 57.5\times10^5$$

$$x_3 = 0 \Rightarrow \quad 0.3x_1 + 0.5x_2 = 2\times10^5$$
$$0.7x_1 + 0.5x_2 = 1\times10^5$$
$$\Rightarrow \ -0.4x_1 = 1\times10^5 \Rightarrow x_1 = -2.5\times10^5$$

infeasible

$\therefore$ no solution

$\therefore$ optimum is $(0.25, 0, 2.75)\times10^5$

7.2 The Simplex Method

The feasible set in the example in section 7.2.1 was simply a line, but in higher dimension problems the feasible set is a complicated polyhedron with edges, faces and vertices. The *Simplex Method* exploits the fact that the optimum of a linear program is at one of the extreme points of the feasible set, i.e. one of its vertices:

- The Simplex Method first finds one of the vertices, and then moves from vertex to vertex along the edges of the feasible set.

- For each move the Simplex Method moves along an edge that reduces the value of the objective function.
- Eventually it finds one of the extreme points where the objective function increases along every edge leading away from it — this is the optimum.

Finding the first vertex is called *Phase 1* of the Simplex Method, which will be explained later. First, we will look at the remainder of the algorithm, known as *Phase 2*.

7.2.1 Example: Simplex Method Phase 2

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T\mathbf{x} = 9x_1 + 3x_2 + x_3 + x_4 \,, \\ \text{subject to} & \begin{cases} 2x_1 + x_2 + x_4 = 4 \\ x_1 + x_3 - x_4 = 2 \end{cases}, \text{ i.e. } \boldsymbol{A}\mathbf{x} = \mathbf{b} \,, \\ \text{and} & x_j \ge 0 \,. \end{array}$$

The starting vertex is the point $\mathbf{x}^T = (0, 4, 2, 0)$ which has $m = 2$ positive entries, and $n - m = 2$ zeros.

- The starting point is a *particular solution* of the equation $\boldsymbol{A}\mathbf{x} = \mathbf{b}$.
- x_2 and x_3 are the *basic variables* (they are the columns with pivots).
- The other variables are the *free variables.*
- The objective function at this point $\mathbf{c}^T\mathbf{x} = 14$.

To move away from this particular vertex either x_1 or x_4 will have to increase. There are two edges leading away from this vertex and we have to decide which to follow.

Consider first increasing x_1 while leaving $x_4 = 0$, then the constraint equations $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ can be rearranged to give:

$$x_2 = 4 - 2x_1$$
$$x_3 = 2 - x_1$$

Substituting these into the objective function gives:

$$9x_1 + 3(4 - 2x_1) + (2 - x_1) + 0 = 14 + 2x_1$$

Thus, the objective increases with x_1 because $+2x_1 > 0$. The coefficient $r_1 = +2$ is called the *reduced cost*, and its sign is important. In this case, because the reduced cost is positive, following this particular edge makes things worse — the objective increases.

Alternatively, we could try increasing x_4 while leaving $x_1 = 0$. The constraint equations now become:

$$x_2 = 4 - x_4$$
$$x_3 = 2 + x_4$$

Substituting these into the objective function gives:

$$9 \times 0 + 3(4 - x_4) + (2 + x_4) + x_4 = 14 - x_4$$

Now the reduced cost $r_4 = -1$, and the objective function decreases as x_4 increases. Because the reduced cost is negative, following this edge decreases the objective.

So, the next question is how far to follow this edge. Obviously, we have to follow it until we reach another vertex, when either x_2 or x_3 become zero. Looking again at the equations above:

$$x_2 = 4 - x_4$$
$$x_3 = 2 + x_4$$

So, as x_4 increases, only x_2 decreases (moves towards zero). x_2 reaches 0 when $x_4 = 4$, at which point $x_3 = 6$.

Thus, the new vertex $\mathbf{x}^T = (0, 0, 6, 4)$ and the objective function $\mathbf{c}^T\mathbf{x} = 10$, reduced from the original 14.

There is now a new set of basic variables, x_3 and x_4. As x_4 became a basic variable it is called the *entering variable*, and as x_2 is no longer a basic variable it is called the *leaving variable*.

To check whether this new vertex is the optimum, we now repeat the whole process. But there is a problem. The original form of the constraint equations were given by:

$$\begin{aligned} 2x_1 + x_2 \phantom{{}+x_3} + x_4 &= 4 \\ x_1 \phantom{{}+x_2} + x_3 - x_4 &= 2 \end{aligned}$$

Note that, for the first step, the coefficients of the initial basic variables formed an identity matrix:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

This was crucially important to being able to calculate the reduced costs, and is called the *canonical form* of the problem.

Now the basic variables are x_3 and x_4, we need to convert the problem to canonical form. We can do this by making the coefficients of these variables the identity matrix by rearranging the initial constraint equations.

If the second equation is replaced by the first plus the second (i.e. we perform *Gaussian elimination*), then the constraint equations become:

$$\begin{aligned} 2x_1 + x_2 \quad + x_4 &= 4 \\ 3x_1 + x_2 + x_3 \quad &= 6 \end{aligned}$$

which is in the form required.

We now want to be able to calculate the reduced costs associated with x_1 or x_2 increasing from 0. If x_2 was increased, with $x_1 = 0$, then this would force:

$$\begin{aligned} x_4 &= 4 - x_2 \\ x_3 &= 6 - x_2 \end{aligned}$$

The objective function then becomes:

$$9 \times 0 + 3x_2 + (6 - x_2) + (4 - x_2) = 10 + x_2$$

Thus, the coefficient of x_2, the reduced cost $r_2 = +1$. As this is positive, we do not follow this edge.

If x_1 was increased, with $x_2 = 0$, then this would force:

$$\begin{aligned} x_4 &= 4 - 2x_1 \\ x_3 &= 6 - 3x_1 \end{aligned}$$

The objective function then becomes:

$$9x_1 + 3 \times 0 + (6 - 3x_1) + (4 - 2x_1) = 10 + 4x_1$$

So, the coefficient of x_1, the reduced cost $r_1 = +4$.

Thus, both reduced costs are positive, so moving away from this vertex will increase the objective. Therefore this vertex is the optimum.

7.2.2 The Tableau

The Simplex Method can be written down in a systematic way by generating a matrix called the *tableau*, which contains all the information about the linear program.

At the start, when x_2 and x_3 were the basic variables, the tableau formed from $\boldsymbol{A}$, $\mathbf{b}$ and $\mathbf{c}$ is as follows:

$$\boldsymbol{T}_0 = \begin{bmatrix} \boldsymbol{A} & \mathbf{b} \\ \mathbf{c}^T & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 & 1 & 4 \\ 1 & 0 & 1 & -1 & 2 \\ 9 & 3 & 1 & 1 & 0 \end{bmatrix}$$

The current objective function value can be found by eliminating the components of the basic variables, x_2 and x_3, from the final row.

Subtracting $3\times$ the first equation and $1\times$ the second equation from the third:

$$\boldsymbol{T}_1 = \begin{bmatrix} 2 & 1 & 0 & 1 & 4 \\ 1 & 0 & 1 & -1 & 2 \\ 2 & 0 & 0 & -1 & -14 \end{bmatrix}$$

You should recognise all the entries in the final row:

- The values 2 and −1 are the reduced costs for the corresponding variables (as found earlier), $r_1 = +2$ and $r_4 = -1$.
- The current objective function value appears in the bottom right-hand corner (multiplied by −1!).

It is easily verified that the elimination steps performed replicate the calculations we did earlier.

To generate the next tableau we need to decide which is the *entering variable* and which the *leaving variable*.

The *entering variable* is decided by examining the reduced costs. The most negative reduced cost tells us which path will most quickly reduce the objective function. In this case the entering variable is x_4.

The *leaving variable* is decided by identifying which of the current basic variables will reach zero most quickly. This can be found by comparing the current values of the basic variables, which are shown in the final column (4 and 2 in this case), with how quickly these will change as the entering variable, x_4, increases, given by the coefficients corresponding to entering variable (here the coefficients for x_4 are 1 and −1). Thus, these ratios are:

For x_2 : $\frac{4}{1} = 4$ i.e. $x_2 = 0$ when $x_4 = 4$

For x_3 : $\frac{2}{-1} = -2$ i.e. $x_3 = 0$ when $x_4 = -2$ (impossible)

These ratios give the value of x_4 when each basic variable reaches zero. The negative ratio can be ignored because x_4 is not allowed (by the bounds) to become negative. The leaving variable will thus be the variable that has the lowest non-negative ratio; in this case x_2.

We have now decided that our basic variables for the next tableau will be x_3 and x_4. The tableau is then generated by making sure, using elimination, that the coefficients of these variables are the identity matrix.

In this case, we add the first row of $\boldsymbol{T}_1$ to the second, and the first row to the third, giving:

$$\boldsymbol{T}_2 = \begin{bmatrix} 2 & 1 & 0 & 1 & 4 \\ 3 & 1 & 1 & 0 & 6 \\ 4 & 1 & 0 & 0 & -10 \end{bmatrix}$$

Again we can recognise:

- The current values of the basic variables ($x_3 = 6$, $x_4 = 4$) in the final colum.
- The reduced costs $r_1 = +4$ and $r_2 = +1$ in the final row.
- The current cost 10 (multiplied by -1) in the bottom right corner.

Because all the reduced costs are positive, this is the optimum.

7.2.3 Organisation of a Simplex Step

1. Convert the problem to canonical form. In the tableau this is done by elimination from the matrix $\boldsymbol{A}$ so that the coefficients of the basic variables form an identity matrix. It is important that this step is possible, i.e. that the constraint equations are independent.
2. Calculate the reduced costs r_i associated with each of the free variables. In the tableau this is done by eliminating from the objective function any component that corresponds to a basic variable. Steps 1 and 2 are usually done at the same time.
3. If $r_i \geq 0 \ \forall\ i$, stop. The current solution is the optimum. Otherwise find the most negative component r_i and let the corresponding x_i increase from 0. This is the *entering variable*.
4. Calculate the ratio of the final column of the tableau (the current values of the basic variables) to the ith column of the tableau — this shows how quickly each variable will decrease as x_i increases. The smallest of these ratios (ignoring the negative ones) will identify the first basic variable to reach 0 as x_i increases. This is the *leaving variable*. (If all the ratios are negative, the problem is *unbounded*, and the minimum objective function value will be $-\infty$.)

This completes a step of the Simplex Method

7.2.4 Slack Variables

The standard form for linear programming (equation (7.2)) assumes that all the constraints are *equality constraints*. This is, of course, not necessarily the case. However, it is possible to convert linear *inequality constraints* of the form

$$\mathbf{a}_i^T\mathbf{x} \geq b_i \tag{7.3}$$

into a linear *equality constraint* by introducing a slack variable s_i defined by

$$s_i = \mathbf{a}_i^T\mathbf{x} - b_i\ , \tag{7.4}$$

and as long as $s_i \geq 0$ the original constraint is satisfied. Thus, by the introduction of another control variable, the inequality constraint can be converted into the standard form:

$$\mathbf{a}_i^T\mathbf{x} - s_i = b_i\ ;\quad s_i \geq 0\ . \tag{7.5}$$

Similarly, an inequality constraint of the form

$$\mathbf{a}_i^T\mathbf{x} \le b_i \tag{7.6}$$

can be converted into standard form as:

$$\mathbf{a}_i^T\mathbf{x} + s_i = b_i \ ; \quad s_i \ge 0 \ . \tag{7.7}$$

7.2.5 Phase 1 of the Simplex Method

Let us now return to the question of finding an initial feasible solution — Phase 2 of the Simplex Method assumes one has been found. Phase 1 finds an initial feasible solution by creating and solving a new problem.

Consider the earlier example. Phase 1 of the Simplex Method creates $m = 2$ new variables, x_5 and x_6, and inserts them into $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ so that they are the *initial basic variables* of the new problem (after ensuring the signs are reversed in any equation that has a negative right hand side):

$$\begin{aligned} 2x_1 + x_2 \quad + x_4 + x_5 \quad &= 4 \\ x_1 + \quad x_3 - x_4 \quad + x_6 &= 2 \end{aligned}$$

Phase 1 also creates a new objective function:

$$f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} = x_5 + x_6 \text{ ; i.e. } \mathbf{c}^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

This problem is then solved using Simplex Method Phase 2.

The optimum for this new problem will be $f(\mathbf{x}) = 0$ at $x_5 = x_6 = 0$, and the remaining variables will now be a feasible solution to the original problem. If the optimum is reached when $f(\mathbf{x}) > 0$, then there is no feasible solution to the original problem.

The initial tableau for this Phase 1 problem (which is already in canonical form) is:

$$\boldsymbol{T}_0 = \begin{bmatrix} 2 & 1 & 0 & 1 & 1 & 0 & 4 \\ 1 & 0 & 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

Note that x_5 and x_6 are the basic variables. The reduced costs can be found by elimination:

$$\boldsymbol{T}_1 = \begin{bmatrix} 2 & 1 & 0 & 1 & 1 & 0 & 4 \\ 1 & 0 & 1 & -1 & 0 & 1 & 2 \\ -3 & -1 & -1 & 0 & 0 & 0 & -6 \end{bmatrix}$$

Obviously x_1 must be the entering variable, as its reduced cost is the most negative. The ratios of the current values of x_5 and x_6 to column 1 are then:

$$\begin{bmatrix} 4/2 \\ 2/1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

This shows that both x_5 and x_6 will reach 0 at the same time, when $x_1 = 2$, and so either can be the leaving variable. If we choose x_6, then the new basic variables are x_1 and x_5, and we are ready to start a new Simplex step.

The new canonical form of the tableau, with the reduced costs already calculated, is:

$$\boldsymbol{T}_2 = \begin{bmatrix} 0 & 1 & -2 & 3 & 1 & -2 & 0 \\ 1 & 0 & 1 & -1 & 0 & 1 & 2 \\ 0 & -1 & 2 & -3 & 0 & 3 & 0 \end{bmatrix}$$

Now x_4 will be the entering variable, and the ratio of the columns gives:

$$\begin{bmatrix} 0/3 \\ 2/-1 \end{bmatrix} = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$$

The 0 in the first row is interesting. Should this variable (x_5) leave the tableau? The answer comes from looking at the meaning of this number. It comes from the equation:

$$3x_4 + x_5 = X = 0$$

Thus, x_5 reaches 0 (from its current value X which just happens to be 0) when $x_4 = X/3$. This is a perfectly legitimate solution, so x_5 is the leaving variable.

Note that if the coefficient of x_5 was negative, i.e.

$$3x_4 - x_5 = X ,$$

this implies that x_5 increases as x_4 increases from 0, and in this case x_5 cannot be the leaving variable.

With x_5 as the leaving variable and x_1 and x_4 the new basic variables, we are ready to start a new Simplex step:

$$\boldsymbol{T}_3 = \begin{bmatrix} 0 & 1/3 & -2/3 & 1 & 1/3 & -2/3 & 0 \\ 1 & 1/3 & 1/3 & 0 & 1/3 & 1/3 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

However, none of the reduced costs are now negative, so this is the optimum, with $f(\mathbf{x}) = 0$ as expected.

This basic solution can now be used in Phase 2 of the Simplex Method for the original problem, and can be inserted directly into the starting tableau (obviously without the added variables x_5 and x_6):

$$
\boldsymbol{T}_0 = \begin{bmatrix} 0 & 1/3 & -2/3 & 1 & 0 \\ 1 & 1/3 & 1/3 & 0 & 2 \\ 9 & 3 & 1 & 1 & 0 \end{bmatrix}
$$

Note that the original vector $\mathbf{c}^T$ has been restored in the tableau.