3 Linear Programming

A special case of convex optimization is **Linear Programming (LP)**. However, this special case is of a great importance for us. It is not a big surprise that LP was one of the first optimization problems, which were carefully studied and exactly solved; LP is associated with linear algebra, one of the richest and most popular tool sets in the entire mathematics. Many later developments in other areas of the optimization theory was inspired by LP. For example, the development of the theory of convex optimization, to great extent owes to the understanding of LPs. Remember that I previously talked about affine functions and half-spaces, being the building blocks of convex functions and convex sets, respectively. In the same manner, linear programming is the building block of all convex optimization problems. Again, I skip more detailed discussion. If you want to know more about the convex optimization theory, you should take a course in convex optimization.

The main reason that we study linear optimization in this course is that it has a profound application in the theory of discrete optimization. From the early days of discrete optimization studies, it became clear that many problems of interest could be formulated as the so-called **Integer Linear Programs (ILP)**. ILPs are similar to LPs, but different in that their variable space is discrete. Later, we will study ILPs and their relation with LPs. We will see how LPs can solve ILPs or lead to approximate algorithms. Now, we focus on LP and its theory.

Let us start by the definition of an LP. An optimization problem in the form of (2) is linear if:

1. The variable domain is $\mathbb{R}^n$ for some n.
2. The cost function $f(\mathbf{x})$ and the constraint functions $g_i(\mathbf{x})$ and $h_j(\mathbf{x})$ are all affine functions.

It is popular to write an LP as

$$\begin{gathered}
\min_{\mathbf{x}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n} c_1x_1 + c_2x_2 + \ldots + c_nx_n \\
\text{subject to} \\
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \le b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \le b_2 \\
\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \le b_m \\
, \\
d_{11}x_1 + d_{12}x_2 + \ldots + d_{1n}x_n = e_1 \\
d_{21}x_1 + d_{22}x_2 + \ldots + d_{2n}x_n = e_2 \\
\vdots \\
d_{p1}x_1 + d_{m2}x_2 + \ldots + d_{mn}x_n = e_p
\end{gathered} \tag{27}$$

where the terms $a_{k,l}, b_k, c_k, d_{k,l}, e_k$ are all arbitrary real coefficients. As seen, it might be difficult to read an optimization, written in the form of (27). Hence, a conciser notation is often used. Let us introduce the following matrix definitions:

$$\begin{gathered}
\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \quad \mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \\
\mathbf{D} = \begin{pmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,n} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{p,1} & d_{p,2} & \cdots & d_{p,n} \end{pmatrix} \quad \mathbf{e} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{pmatrix}
\end{gathered} \tag{28}$$

Now, the optimization in (27) is written as

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \\ &\text{subject to,} \\ &\mathbf{A}\mathbf{x} \leq \mathbf{b} \\ &\mathbf{D}\mathbf{x} = \mathbf{e} \end{aligned} \tag{29}$$

Notice that we used the symbol "$\leq$" to refer to **element-wise inequality**.

The main issue here is that you do not need all the matrices $\mathbf{A}, \mathbf{b}, \mathbf{c}, \mathbf{D}$ and $\mathbf{e}$ to express an LP. In Example 7 we saw that the linear equality constraints can always be eliminated. Then, the resulting LP is in the following form:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \\ &\mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned} \tag{30}$$

On the contrary, one can substantially simplify the inequality constraints by turning them into equality ones. We show this in the following example.

Example 8. Consider the following optimization

$$\begin{aligned} &\min_{(x_1,x_2)\in\mathbb{R}^2} 2x_2 - 3x_1 \\ &\text{subject to} \\ &2x_1 - x_2 \leq 4 \\ &x_1 + 3x_2 \geq 5 \\ &x_2 \geq 0 \end{aligned} \tag{31}$$

First of all, notice that the second and third inequalities can be written in the standard form of (27) by multiplying them by -1. There is also no equality constraint. So, the two first inequality constraints can be written as

$$\begin{aligned} 2x_1 - x_2 &\leq 4 \\ -x_1 - 3x_2 &\leq -5 \end{aligned} \tag{32}$$

Now take the first inequality. The statement that $2x_1 - x_2$ is less than 4 is the same as it is equal to $4 - z_1$ for some non-negative (positive or zero) value of z_1. Also for the second inequality, $-x_1 - 3x_2$ is equal to $-5 - z_2$ for another non-negative value z_2. So, we can write down the two inequalities as

$$\begin{aligned} 2x_1 - x_2 &= 4 - z_1 \\ -x_1 - 3x_2 &= -5 - z_2 \\ z_1 &\geq 0 \\ z_2 &\geq 0 \end{aligned} \tag{33}$$

Hence, our optimization in (31) can be written as

$$\begin{aligned} &\min_{(x_1,x_2,z_1,z_2)\in\mathbb{R}^4} 2x_2 - 3x_1 \\ &\text{subject to} \\ &2x_1 - x_2 + z_1 = 4 \\ &-x_1 - 3x_2 + z_2 = -5 \\ &z_1 \geq 0 \\ &z_2 \geq 0 \\ &x_2 \geq 0 \end{aligned} \tag{34}$$

The variables z_1, z_2 are called **slack variables**. We can go even further by noticing that in (34), the variable x_1 is not included in any inequality constraint. Hence, it can be eliminated by calculating it from one of the equalities and replacing in the others. Take the second equality constrain for example, which gives that $x_1 = 5 + z_2 - 3x_2$. Plugging this expression into other terms yields to

$$\begin{aligned} &\min_{(x_2,z_1,z_2)\in\mathbb{R}^3} 11x_2 - 3z_2 - 15 \\ &\text{subject to} \\ &-7x_2 + 2z_2 + z_1 = -6 \\ &z_1 \geq 0 \\ &z_2 \geq 0 \\ &x_2 \geq 0 \end{aligned} \tag{35}$$

What I showed in the above example is that by introducing slack variables and eliminating some others, you can always bring an LP into the following form:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \\ &\mathbf{D}\mathbf{x} = \mathbf{e} \\ &\mathbf{x} \geq 0 \end{aligned} \tag{36}$$

which means that every variable has to be positive and they should satisfy a number of additional equality constraints. In many books, this is the canonical form of LP and is often referred to as the **augmented form**. It is because some algorithms such as the **simplex method** only admit this form. Moreover, this form is related to a very nice geometrical interpretation of LP, which we should here neglect for simplicity. We also do not worry about the implementation, as we use CVX, which can solve an LP in any arbitrary form. There is another interesting form of LP, which naturally appear in solving many discrete optimization problems:

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \\ &\mathbf{A}\mathbf{x} \leq \mathbf{b} \\ &\mathbf{x} \geq 0 \end{aligned} \tag{37}$$

As seen, this is a special case of (30), where every variable is assumed to be positive.

Now, Let us talk about the solution of a LP. Remember Example 2, where the optimization was not unbounded, but still it could not attain a minimum value. This cannot happen for a LP. So, exactly one the following three cases happen for a LP: It has at least one optimal solution, it is infeasible or it is unbounded.

Example 9. As a simple example, consider the following optimization

$$\min_{\mathbf{x}=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n} c_1x_1 + c_2x_2 + \ldots + c_nx_n$$
$$\text{subject to}$$
$$\mathbf{x} \geq 0 \tag{38}$$

Suppose that one of the elements, say c_{k_0} is negative (not zero!). Then, one can fix all x_ks to zero, except x_{k_0}. The cost value will be $x_{k_0}c_{k_0}$, which is negative. By increasing x_{k_0} to $+\infty$, the cost can get as small as desired. Hence, the optimization is unbounded. On the other hand, if all c_ks are non-negative the cost cannot get negative, while it can be set to zero if we select $\mathbf{x} = \mathbf{0}$. Hence, in this case $\mathbf{x}^* = \mathbf{0}$ is an optimal solution and the optimal cost is zero.

With a similar argument, you should be able to show that a totally unconstrained LP is unbounded, except when $\mathbf{c} = \mathbf{0}$, in which case $\mathbf{x}^* = \mathbf{0}$ is an optimal solution.

LPs can be exactly solved by different algorithms. Developed in 1930s and 1940s, the **simplex** method is one of the first and most popular LP techniques. Although the simplex method has a good performance in practice, the algorithm can have a very high computational complexity in the worst case scenario. The simplex method remained the best alternative for LPs until 1980s, when the ellipsoid method and subsequently the Karmarkar's projective techniques were introduced. The later methods have provable low computational complexity, but more importantly, they led to the development of the interior methods, which are applicable to all convex problems. In this course, I am not going to talk about the details of these techniques. What you need to know is that both the simplex and the interior point techniques are widely used in practice and often have similar performance. In MATLAB, the simplex method is implemented in the `linprog` function. The CVX package uses the interior point method.