Linear Programming

Aaron Mills

February 5, 2023

This document is a summary of core concepts related to linear programming. I am not taking credit for inventing any of the concepts conveyed. I am simply collecting information from other sources and presenting it here. For more information on the sources used, please see the references section of this document.

If you find any mistakes, please let me know so that I can amend them.

1 Introduction

Linear programs are an important class of optimization problems in which a linear function is optimized (maximized or minimized) subject to a system of linear constraints (linear inequalities). Linear programs are important because they can be solved efficiently, meaning their running time is bounded from above by some polynomial [1]. There are many problems in computer science (and other fields) that can be modeled as linear programs and can thus be solved efficiently, for example, the maximum-flow problem from flow networks [2]. Before defining a linear program formally, we will first discuss some important background in order to better understand the linear constraints and what the set of feasible solutions to our system looks like.

2 Definitions

Hyperplanes and Half-spaces

A *hyperplane* is defined as the solution set of a linear equation. A hyperplane defined in $\mathbb{R}^n$ has dimension $n-1$. Let H denote some hyperplane, then:

$$H = \{x \mid c^T x = b\} \tag{1}$$

for some nonzero $c \in \mathbb{R}^n$, and $b \in \mathbb{R}$. A hyperplane splits the space, $\mathbb{R}^n$, defining two *half-spaces*. Let $H_\leq$ and $H_\geq$ denote the half-spaces formed by the hyperplane H defined above.

$$H_\leq = \{x \mid c^T x \leq b\} \tag{2}$$

$$H_{\geq} = \{x \mid c^T x \geq b\} \tag{3}$$

These definitions are important in the context of linear programs because each of the constraints in a linear program is itself a half-space, and the intersection of these half-spaces defines the feasible set of solutions [3].

Polyhedra

A *polyhedron* is defined as the solution set of a finite number of linear inequalities in $\mathbb{R}^n$ (i.e. the intersection of half-spaces). Let P denote some polyhedron, then:

$$P = \{x \mid Ax \leq b\} \tag{4}$$

for $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. Note that both above and in the remainder of the document, an inequality between vectors denotes an element-wise inequality. Each row in the matrix A, along with its corresponding element in vector b, define a linear inequality: $a_i^T x \leq b_i$. The constraints of our linear program define a polyhedron [1, 3].

3 Linear Programs

Now that we have defined hyperplanes, half-spaces, and polyhedra, we will write a formal definition for each type of linear program (one for maximization and one for minimization). The form in which the linear programs are written below is called *standard form*.

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \\ & x \geq 0. \end{aligned} \tag{5}$$

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \geq b \\ & x \geq 0. \end{aligned} \tag{6}$$

For the above we have $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$. The linear function $c^T x$ is the *objective function* we wish to optimize. The matrix A and vector b formulate our constraints which define the region in $\mathbb{R}^n$ containing the points that satisfy these constraints.

3.1 Converting From General to Standard Form

The linear programs above have been shown have been in standard form. Linear programs do not have to be in this form. Consider the following general forms for linear programs:

$$
\begin{aligned}
\text{maximize} \quad & c^T x \\
\text{subject to} \quad & A_1 x \leq b_1 \\
& A_2 x = b_2 \\
& A_3 x \geq b_3
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & A_1 x \leq b_1 \\
& A_2 x = b_2 \\
& A_3 x \geq b_3
\end{aligned} \tag{8}
$$

Conversion to standard form is simple. Suppose WLOG we would like to convert a maximization problem from general form to standard form. For each inequality constraint with a $\geq$ rather than a $\leq$, $a^T x \geq b$, we simply multiply the expression by -1 to obtain $-a^T x \leq -b$. For each equality constraint, $a^T x = b$, we replace it with two inequality constraints, $a^T x \leq b$ and $-a^T x \leq -b$. The standard form also has a nonnegativity requirement on the variables $x \in \mathbb{R}^n$. However, our general form has no such requirement. Suppose x_0 is a free variable in our linear program (i.e. $x_0 \in \mathbb{R}$). To enforce nonnegativity we introduce two new variables $x_1, x_2 \in \mathbb{R}$ such that $x_1 \geq 0$ and $x_2 \geq 0$. We then perform the substitution: $x_0 = x_1 - x_2$. This allows us to enforce nonnegativity on our variables, but the expression $x_1 - x_2$ can evaluate to any value in $\mathbb{R}$.

3.2 Minimax and Maximin Problems

Consider the problem of minimizing the maximal value or maximizing the minimal value of a set of m affine functions: $a_i^T x + b_i$, $\forall i \in \{1, 2, ..., m\}$. Formally, the problems can be written as follows.

$$
\text{minimize} \max_{i=1,2,\ldots,m} a_i^T x + b_i \tag{9}
$$

$$
\text{maximize} \min_{i=1,2,\ldots,m} a_i^T x + b_i \tag{10}
$$

Clearly, neither of these problems is a linear program, as the max and min functions are nonlinear. However, if we can convert such a problem into an equivalent linear program, then we have a guarantee that we can solve it efficiently. Consider the following two linear programs.

$$
\begin{aligned}
\text{minimize} \quad & z \\
\text{subject to} \quad & a_i^T x + b_i \leq z \quad \forall i \in \{1, 2, ..., m\}
\end{aligned} \tag{11}
$$

$$
\begin{aligned}
\text{maximize} \quad & z \\
\text{subject to} \quad & a_i^T x + b_i \geq z \quad \forall i \in \{1, 2, ..., m\}
\end{aligned} \tag{12}
$$

Realize that the problems in equations 11 and 12 are linear programs that are equivalent to the problems in equations 9 and 10, respectively. These problems can then be solved using one of the various known techniques to solve linear programs [4].

3.3 Feasibility and Optimal Solutions

A *feasible solution* of a linear program is a point $x \in \mathbb{R}^n$ that satisfies all the constraints of the linear program (i.e. lies in the polyhedron defined by the linear constraints). Any of these points in the set of feasible solutions that yield the maximum value of the linear objective function $c^T x$ over all the feasible solutions is referred to as an *optimal solution*. A linear program that has no feasible solution (this occurs when there is no point that can satisfy all the linear constraints) is called *infeasible*. In this case, there is clearly no optimal solution. Even with a nonempty feasible set, a linear program can have no optimal solution when it is *unbounded*. This means the objective function can reach arbitrarily large negative values in the case of a minimization problem and arbitrarily large positive values in the case of a maximization problem.

4 Duality

WLOG, suppose we have a maximization problem in the form of a linear program:

$$\begin{aligned} \text{maximize} \quad & c^T x \\ \text{subject to} \quad & Ax \le b \\ & x \ge 0. \end{aligned} \tag{13}$$

We want to obtain an upper bound on our objective function $c^T x$. Let $y \in \mathbb{R}^m$ with $y \ge 0$, then:

$$\begin{aligned} & Ax \le b \\ \implies \quad & y^T Ax \le y^T b \end{aligned} \tag{14}$$

To ensure an upper bound on our objective function, we require:

$$\begin{aligned} & c \le A^T y = (y^T A)^T \\ \implies \quad & c^T x \le y^T Ax \le y^T b = b^T y \end{aligned} \tag{15}$$

Now consider the following linear program:

$$\begin{aligned} \text{minimize} \quad & b^T y \\ \text{subject to} \quad & A^T y \ge c \\ & y \ge 0. \end{aligned} \tag{16}$$

This linear program, called the *dual*, clearly describes an upper bound on the original linear program called the *primal*. If the primal program was a minimization problem, it's dual would be a maximization problem that defines a lower bound on the primal. Observe that the number of variables that we have in the dual is equal to the number of constraints in the primal, and the number of constraints in the dual is equal to the number of variables in the primal.

4.1 Weak Duality

The derivation of the dual linear program above leads us to a theorem regarding the relationship between the objective function of the primal and the objective function of the dual.

Theorem (Weak Duality) For any $x \in \mathbb{R}^n$ that is a feasible solution for the primal and $y \in \mathbb{R}^m$ that is a feasible solution for the dual we have:

$$c^T x \leq b^T y \tag{17}$$

From this theorem, we can conclude that if there exists an $x \in \mathbb{R}^n$ that is a feasible solution for the primal and a $y \in \mathbb{R}^m$ that is a feasible solution for the dual such that $c^T x = b^T y$ then x is an optimal solution for the primal and y an optimal solution for the dual. Furthermore, we can also conclude that the dual is infeasible if the primal is unbounded from above (the optimal value of the primal is ∞), and the primal is infeasible if the dual is unbounded from below (the optimal value of the dual is $-\infty$) [5].

4.2 Strong Duality

The weak duality theorem only ensures that the primal objective function is lesser than or equal to the dual objective function but does not guarantee that there exists a feasible solution for the primal and a feasible solution for the dual such that we have equality between the values of the objective functions. The strong duality theorem gives this guarantee.

Theorem (Strong Duality) Assume the primal and dual are bounded. The primal has an optimal solution if and only if the dual has an optimal solution. Suppose x^* is an optimal solution for the primal and y^* is an optimal solution for the dual, then:

$$c^T x^* = b^T y^* \tag{18}$$

This statement is very significant. It means that if we have a bounded linear program we wish to solve, we can solve its dual to obtain the optimal value [5].

4.3 Converting Between the Primal and the Dual

Given a primal, we want to be able to derive the dual and vice-versa. There are a simple set of rules that can be followed, and they are summarized as follows. Once again, WLOG, we will be taking the primal to be a maximization problem and the dual a minimization problem. The key aspect to keep in mind is that characteristics of the constraints in the primal translate to characteristics of the variables in the dual, and the characteristics of the variables in the primal translate to characteristics of the constraints in the dual. Consider the following table:

Primal				Dual
Constraints	$\leq b$	$\longleftrightarrow$	≥ 0	Variables
	$= b$	$\longleftrightarrow$	$\mathbb{R}$	
	$\geq b$	$\longleftrightarrow$	≤ 0	
Variables	≥ 0	$\longleftrightarrow$	$\geq c$	Constraints
	$\mathbb{R}$	$\longleftrightarrow$	$= c$	
	≤ 0	$\longleftrightarrow$	$\leq c$	

Recall that each constraint in the primal has a corresponding variable in the dual, and each variable in the primal has a corresponding constraint in the dual. The table displays translations between these associated constraints and variables. For example, if we have an equality constraint in the primal $a_i^T x = b_i$, its corresponding variable in the dual $y_i \in \mathbb{R}$ is free from any constraint. If we have a nonpositive variable in the primal $x_j \in \mathbb{R}$ such that $x_j \leq 0$, that translates to a $\leq$ constraint for the corresponding constraint in the dual, $a_j^T y \leq c_j$ [5].

4.4 Complementary Slackness

Before stating the complementary slackness theorem, we must first understand the concept of a *slack variable*. Given an arbitrary inequality, $a^T x \leq b$, we can convert it to equality by introducing a new variable, $s \geq 0$, such that $a^T x + s = b$. The relationship between the primal and the dual discussed above in section 4.3 is known as *complementary slackness*. The term *complementary* refers to the association of the constraints in a linear program to the variables in its dual. In general, a constraint is said to have slack if its slack variable is positive. The term *slackness* further refers to how the slackness in the constraints of a linear program translates to the slackness of the variables in its dual.

Theorem (Complementary Slackness) Let $x^* \in \mathbb{R}^n$ be a feasible solution for the primal, and $y^* \in \mathbb{R}^m$ a feasible solution for the dual. x^* is an optimal solution for the primal and y^* and optimal solution for the dual if and only if:

1. If $x_i^* > 0$ then $a_i^T y^* = c_i$
2. If $a_i^T y^* > c_i$ then $x_i^* = 0$
3. If $y_j^* > 0$ then $a_j^T x^* = b_j$
4. If $a_j^T x^* < b_j$ then $y_j^* = 0$

The complementary slackness theorem is stating a necessary relationship between the values of optimal solutions and the slackness of their corresponding constraints. To demonstrate how these facts arise mathematically consider equation 15 from section 4:

$$c^T x \leq y^T A x \leq y^T b = b^T y \tag{19}$$

Now suppose we have $x^* \in \mathbb{R}^n$ as the optimal value for the primal and $y^* \in \mathbb{R}^m$ as the optimal value for the dual. By the strong duality theorem, this implies equality in the above expression:

$$c^T x^* = y^{*T} A x^* = y^{*T} b = b^T y^* \tag{20}$$

In particular, consider the equality:

$$\begin{aligned} & c^T x^* = y^{*T} A x^* \\ \implies \quad & c^T x^* - y^{*T} A x^* = 0 \\ \implies \quad & (c^T - y^{*T} A) x^* = 0 \end{aligned} \tag{21}$$

This implies that at least one of two things must be true. Either $x_i^* = 0$ or $a_i^T y^* = c_i$, $\forall i \in \{1, 2, ..., m\}$. This shows statements 1 and 2 of the complementary slackness theorem. A similar derivation can be done from the inequality: $b^T y \geq x^T A^T y \geq x^T c = c^T x$ to show statements 3 and 4 [6, 7].

References

[1] "Jiˇrí Matoušek - EPFL." [Online]. Available: https://blogs.epfl.ch/extrema/documents/Maison%2020.05.10.pdf.

[2] "Maximum flow problem," Wikipedia, 29-Jan-2023. [Online]. Available: https://en.wikipedia.org/wiki/Maximum_flow_problem

[3] S. Boyd and L. Vandenberghe, "Convex Optimization," Convex optimization – Boyd and Vandenberghe. [Online]. Available: https://web.stanford.edu/ boyd/cvxbook/.

[4] "CS/ECE/Isye 524: Introduction to optimization - Laurent Lessard," Laurent Lessard -, 17-Dec-2022. [Online]. Available: https://laurentlessard.com/teaching/524-intro-to-optimization/.

[5] "Chapter 4 Duality - Stanford University." [Online]. Available: https://www.web.stanford.edu/ ashishg/msande111/notes/chapter4.pdf.

[6] "Linear programming notes VI Duality and complementary slackness." [Online]. Available: https://econweb.ucsd.edu/ jsobel/172aw02/notes6.pdf.

[7] O. A. Camarena, "Complementary Slackness," Complementary slackness. [Online]. Available: https://www.matem.unam.mx/ omar/math340/comp-slack.html.