Note: Your TA probably will not cover all the problems. This is totally fine, the discussion worksheets are not designed to be finished in an hour. They are deliberately made long so they can serve as a resource you can use to practice, reinforce, and build upon concepts discussed in lecture, readings, and the homework.

1 LP Basics

Linear Program. A *linear program* is an optimization problem that seeks the optimal assignment for a linear objective over linear constraints. Let $x \in \mathbb{R}^n$ be the set of variables and $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$. The canonical form of a linear program is

$$\begin{aligned} &\text{minimize } c^\top x \\ &\text{subject to } Ax \geq b \\ &x \geq 0 \end{aligned}$$

Any linear program can be written in canonical form.

Let's check this is the case:

(i) What if the objective is maximization?

(ii) What if you have a constraint $Ax \leq b$?

(iii) What about $Ax = b$?

(iv) What if the constraint is $x \leq 0$?

(v) What about unconstrained variables $x \in \mathbb{R}$?

Solution:

(i) Take the negative of the objective.

(ii) Negate both sides of the inequality.

(iii) Write both $Ax \leq$ and $Ax \geq b$ into the constraint set.

(iv) Change of variable: replace every x by $-z$, and add constraint $z \geq 0$.

(v) Replace every x by $x^+ - x^-$, add constraints $x^+, x^- \geq 0$. Note that for every solution to the original LP, there is a solution to the transformed LP (with the same objective value). Similarly, if there is a feasible solution for the transformed problem, then there is a feasible solution for the original problem with the same objective value.

Dual. The dual of the canonical LP is

$$\begin{aligned} &\text{maximize } b^\top y \\ &\text{subject to } A^T y \leq c \\ &y \geq 0 \end{aligned}$$

Weak duality: The objective value of any feasible dual $\leq$ objective value of any feasible primal

Strong duality: The *optimal* objective values of these two are equal.

Both are solvable in polynomial time by the Ellipsoid or Interior Point Method.

2 Huffman and LP

Consider the following Huffman code for characters a, b, c, d: $a = 0, b = 10, c = 110, d = 111$.

Let f_a, f_b, f_c, f_d denote the fraction of characters in a file (only containing these characters) that are a, b, c, d respectively. Write a linear program with variables f_a, f_b, f_c, f_d to solve the following problem: What values of f_a, f_b, f_c, f_d that can generate this Huffman code result in the Huffman code using the most bits per character?

Solution:

Our objective is to maximize the bits per character used:

$$\max f_a + 2f_b + 3f_c + 3f_d$$

We know the fractions must add to 1 and be non-negative:

$$f_a + f_b + f_c + f_d = 1, f_a, f_b, f_c, f_d \geq 0$$

We know the frequencies of the characters must satisfy $f_a \geq f_b \geq f_c, f_d$. We also know that $f_c + f_d \leq f_a$, since we chose to merge (c, d) with b instead of merging a. So we get the following constraints:

$$f_c \leq f_b, f_d \leq f_b, f_b \leq f_a, f_c + f_d \leq f_a$$

3 Job Assignment

There are I people available to work J jobs. The value of person i working 1 day at job j is a_{ij} for $i = 1, \dots, I$ and $j = 1, \dots, J$. Each job is completed after the sum of the time of all workers spend on it add up to be 1 day, though partial completion still has value (i.e. person i working c portion of a day on job j is worth $a_{ij}c$). The problem is to find an optimal assignment of jobs for each person for one day such that the total value created by everyone working is optimized. No additional value comes from working on a job after it has been completed.

(a) What variables should we optimize over? I.e. in the canonical linear programming definition, what is x?

Solution: An assignment x is a choice of numbers x_{ij} where x_{ij} is the portion of person i's time spent on job j.

(b) What are the constraints we need to consider? Hint: there are three major types.

Solution: First, no person i can work more than 1 day's worth of time.

$$\sum_{j=1}^{J} x_{ij} \leq 1 \qquad \text{for } i = 1, \dots, I.$$

Second, no job j can be worked past completion:

$$\sum_{i=1}^{I} x_{ij} \leq 1 \qquad \text{for } j = 1, \dots, J.$$

Third, we require positivity.

$$x_{ij} \geq 0 \qquad \text{for } i = 1, \dots, I, j = 1, \dots, J.$$

(c) What is the maximization function we are seeking?

Solution: By person i working job j for x_{ij}, they contribute value $a_{ij}x_{ij}$. Therefore, the net value is

$$\sum_{i=1,j=1}^{I,J} a_{ij}x_{ij} = A \bullet x.$$

4 (Simplex) Understanding convex polytopes

So far in this class we have seen linear programming defined as

$$(\mathcal{P}) = \begin{cases} \max & c^T x \\ \text{s.t.} & Ax \leq b. \end{cases}$$

Today, we explore the different properties of the region $\Omega = \{x : Ax \leq b\}$ – i.e. the region that our linear program maximizes over.

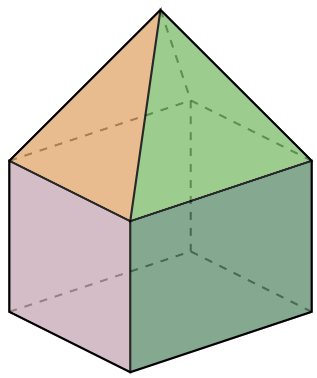

Figure 1: An example of a convex polytope. We can consider each face of the polytope as an affine inequality and then the polytope is all the points that satisfy each inequality. Notice that an affine inequality defines a half-plane and therefore is also the intersection of the half-planes.

(a) The first property that we will be interested in is *convexity*. We say that a space X is convex if for any $x, y \in X$ and $\lambda \in [0, 1]$,

$$\lambda x + (1 - \lambda)y \in X.$$

That is, the entire line segment $\overline{xy}$ is contained in X. Prove that Ω is indeed convex.

Solution: Let $x, y \in \Omega$. We need to show that

$$A(\lambda x + (1 - \lambda)y)) \leq b.$$

We apply the only facts we know, namely $Ax \leq b$ and $Ay \leq b$.

$$\begin{aligned} A(\lambda x + (1 - \lambda)y) &= \lambda Ax + (1 - \lambda)Ay \\ &\leq \lambda b + (1 - \lambda)b \\ &= b. \end{aligned}$$

(b) The second property that we will be interested in is showing that linear objective functions over convex polytopes achieve their maxima at the vertices. A vertex is any point $v \in \Omega$ such that v **cannot** be expressed as a point on the line $\overline{yz}$ for $v \neq y, v \neq z$, and $y, z \in \Omega$.

Prove the following statement: Let Ω be a convex space and f a linear function $f(x) = c^T x$. Show that the for a line $\overline{yz}$ for $y, z \in \Omega$ that $f(x)$ is maximized on the line at either y or z. I.e. show that

$$\max_{\lambda \in [0,1]} f(\lambda y + (1 - \lambda)z)$$

achieves the maximum at either $\lambda = 0$ or $\lambda = 1$. *Hint: Assume without loss of generality that* $f(y) \geq f(z)$.

Solution: Assume without loss of generality that $f(y) \geq f(z)$. (Otherwise, swap their names). Then $c^T y \geq c^T z$. We now aim to show the maximum is achieved at $\lambda = 1$. Then,

$$\begin{aligned} f(\lambda y + (1 - \lambda)z) &= c^T(\lambda y + (1 - \lambda)z) \\ &= \lambda c^T y + (1 - \lambda)c^T z \\ &\leq \lambda c^T y + (1 - \lambda)c^T y \\ &= f(y). \end{aligned}$$

(c) Now, prove that global maxima will be achieved at vertices. For simplicity, you can assume there is a unique global maximum. Hint: Use the definition of a vertex presented above. (*Side note:* This argument is the basis of the Simplex algorithm by Dantzig to solve linear programs.)

Solution: Assume, for contradiction, that the maximum was not achieved at a vertex and was instead achieved at a point x that was *not* a vertex. Then, there exists a line y, z containing x such that $x \neq y$ and $x \neq z$. But by the previous argument, the function achieves a maximum at either y or z. then, the maximum isn't unique. A contradiction.