

**Instructions:** All your solutions should be prepared in L<sup>A</sup>T<sub>E</sub>X and the PDF and .tex should be submitted to Canvas. Please submit all your files as ONE archive of filetype zip, tgz, or tar.gz. For each question, a well-written and correct answer will be selected a sample solution for the entire class to enjoy. If you prefer that we do not use your solutions, please indicate this clearly on the first page of your assignment.

**1. Simplex example** Consider the following linear program:

$$\begin{aligned} \min \quad & -x_1 + 2x_2 \\ \text{s.t.} \quad & 3x_1 + x_2 + x_3 = 6 \\ & x_1 - 2x_2 + x_4 = 1 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

- Run the simplex algorithm on the linear program to find an optimal point. Start from the vertex  $(x_1, x_2, x_3, x_4) = (0, 0, 6, 1)$ . List the vertex visited at each round.
- What is the optimal value of the primal program?
- Write down the dual of the linear program.
- State an optimal solution to the dual and explain why it proves that the simplex algorithm found an optimal solution.

a. **Iter 1.**

$$\begin{aligned} \bar{\mathbf{c}}_N^T &= \mathbf{c}_N^T - \mathbf{c}_B^T A_B^{-1} A_N \\ &= (-1 \quad 2) - (0 \quad 0) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix} \\ &= (-1 \quad 2) \end{aligned}$$

$$\begin{aligned} \mathbf{x}_B &= \bar{\mathbf{b}} = A_B^{-1} \mathbf{b} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 6 \\ 1 \end{pmatrix} \end{aligned}$$

Select  $1 \in N$ , since  $\bar{c}_1 < 0$ .

$$\begin{aligned}\mathbf{d}_1 &= A_B^{-1} A_1 = \begin{pmatrix} d_{13} \\ d_{14} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 63 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3 \\ 1 \end{pmatrix}\end{aligned}$$

$$\begin{aligned}k &= \operatorname{argmin}\left\{\frac{\bar{b}_3}{d_{13}}, \frac{\bar{b}_4}{d_{14}}\right\} \\ &= \operatorname{argmin}\left\{\frac{6}{3}, \frac{1}{1}\right\} \\ &= 4\end{aligned}$$

So, the next

$$B = \{1, 3\}, N = \{2, 4\}$$

**Iter 2.**

$$\begin{aligned}\bar{\mathbf{c}}_N^\top &= \mathbf{c}_N^\top - \mathbf{c}_B^\top A_B^{-1} A_N \\ &= \begin{pmatrix} 2 & 0 \end{pmatrix} - \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \end{pmatrix}\end{aligned}$$

and hence the optimal solution is,

$$\begin{aligned}\mathbf{x}_B &= \bar{\mathbf{b}} = A_B^{-1} \mathbf{b} \\ &= \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} 6 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ 3 \end{pmatrix}\end{aligned}$$

The vertices visited are  $(0, 0, 6, 1)$  on the first iteration and  $(1, 0, 3, 1)$  on the second iteration.

b.

$$-x_1 + 2x_2 = -1$$

c.

$$\max_{y_1, y_2} 6y_1 + y_2$$

s.t.

$$3y_1 + y_2 \leq -1$$

$$y_1 - 2y_2 \leq 2$$

$$y_1, y_2 \leq 0$$

- d. The optimal solution to the dual is  $(0 \ -1)$ . We know that this is optimal because,  $\mathbf{y}^\top = \mathbf{c}_B^\top A_B^{-1} = (-1 \ 0) \begin{pmatrix} 0 & 1 \\ 1 & -3 \end{pmatrix} = (0 \ -1)$ , with the optimal value being  $-1$ . There are feasible solutions to both the primal and the dual, and hence, we can apply the strong duality theorem. From the theorem, we have that the optimal values of the two LPs are the same. Since  $(0 \ -1)$  gives us the same optimal value as the primal, we know that it is the optimal solution.

**2. Max-Flow Min-Cut.** In this problem we will study *flow networks*. A *flow network* is defined by a finite set of vertices  $V$  with two distinguished vertices, the *source*  $s$  and the *sink*  $t$ . Each edge  $(u, v) \in V \times V$  has a *capacity*  $c_{uv} \in \mathbb{R}^+$ . Finally the source  $s$  has no incoming capacity, i.e.  $c_{us} = 0$  for  $u \in V$  and the sink  $t$  has no outgoing capacity, i.e.  $c_{tu} = 0$  for  $u \in V$ .

A *flow* of the network is a family  $(f_{uv})_{(u,v) \in V \times V}$  of real numbers satisfying the following constraints:

- *Positivity:*  $f_{uv} \geq 0$  for all  $(u, v) \in V \times V$ .
- *Capacity constraint:* for all  $(u, v) \in V \times V$ ,  $f_{uv} \leq c_{uv}$ .
- *Flow conservation:* for all  $u \in V \setminus \{s, t\}$ ,  $\sum_{v \in V} f_{vu} = \sum_{v \in V} f_{uv}$ .

The amount of flow *leaving*  $s$  is  $\sum_{u \in V} f_{su}$  and is called the *value of the flow*  $f$ . The goal of the *maximum flow problem* is to find a flow  $f$  satisfying the three conditions above whose value is maximal.

- Show that the value of the flow  $f$  is equal to  $\sum_{v \in V} f_{vt}$ . Hence, the value of the flow could equivalently be defined as the amount of flow entering  $t$ .
- Write the maximum flow problem as a linear program. Is this LP feasible? Is it bounded?

An  $s - t$  *cut* of  $G$  is a partition of  $V$  into two sets:  $V = S \cup T$  with  $S \cap T = \emptyset$ , such that  $s \in S$  and  $t \in T$ . The *cost*  $c$  of an  $s - t$  cut is defined by  $c(S, T) := \sum_{(u,v) \in S \times T} c_{uv}$ . The goal of the *minimum cut problem* is to find an  $s - t$  cut of minimal cost.

- Formulate a 0/1-integer program for the minimum cut problem (recall that 0/1-integer programs and their relaxations were covered in Section 1). Let us name this problem MINIMUMCUT. Write the LP relaxation of MINIMUMCUT. We will name this relaxation MINIMUMCUT-LP.
- Derive the dual of the maximum flow LP that you wrote in b. and show that it is equivalent to MINIMUMCUT-LP. Show that the optimal value of MINIMUMCUT is an upper bound on the value of the optimal flow.
- Assuming that MINIMUMCUT and MINIMUMCUT-LP have the same optimal value [bonus credits if you prove this fact], prove the famous *max-flow min-cut theorem*:

*The maximum value of a flow is equal to the minimal cost of an  $s - t$  cut.*

a. From flow conservation,

$$\sum_{v \in V \setminus \{s, t\}} \left( \sum_{u \in V} f_{uv} - \sum_{u \in V} f_{vu} \right) = 0$$

Hence,

$$\sum_{u \in V} f_{su} = \sum_{u \in V} f_{su} - \sum_{v \in V \setminus \{s, t\}} \left( \sum_{u \in V} f_{uv} - \sum_{u \in V} f_{vu} \right)$$

On the right hand side, all edges except those entering the sink has both inflow and outflow in the equation. Hence, again from flow conservation,

$$\sum_{u \in V} f_{su} - \sum_{v \in V \setminus \{s, t\}} \left( \sum_{u \in V} f_{uv} - \sum_{u \in V} f_{vu} \right) = \sum_{v \in V} f_{vu}$$

Hence, we proved that

$$\sum_{u \in V} f_{su} = \sum_{v \in V} f_{vu}$$

b.

$$\max f$$

subject to

$$\begin{aligned} f_{uv} &\geq 0, \forall (u, v) \in V \times V \\ f_{uv} &\leq c_{uv}, \forall (u, v) \in V \times V \\ \sum_{v \in V} f_{vu} - \sum_{v \in V} f_{uv} &= 0, \forall u \in V \setminus \{s, t\} \end{aligned}$$

d.

$$\min \sum_e c_e y_e$$

subject to

$$\begin{aligned} \sum_{e \in P} y_e &\geq 1, \forall P \\ y_e &\in \{0, 1\}, \forall e \in V \times V \end{aligned}$$

The linear relaxation of this LP is

$$\min \sum_e c_e y_e$$

subject to

$$\begin{aligned} \sum_{e \in P} y_e &\geq 1, \forall P \\ y_e &\geq 0, \forall e \in V \times V \end{aligned}$$

**Here's why we can make such relaxation:**

The solution of LP only lies in the vertices, so we can make the condition

$$y_e \in \{0, 1\}, \forall e \in V \times V$$

to be

$$1 \geq y_e \geq 0, \forall e \in V \times V$$

Furthermore, the objective

$$\min \sum_e c_e y_e$$

indicates that the smaller the  $y_e$  is, the better. Meaning that if the original integer program has an optimal solution, then there exists a same optimal solution even if the constraint is

$$y_e \geq 0, \forall e \in V \times V$$

.

e. (b) can be written as

$$\max \sum_{p \in P} f_p$$

subject to

$$\begin{aligned} \sum_{P: e \in P} f_p &\leq c_e, \forall e \\ f_p &\geq 0, \forall p \in P \end{aligned}$$

This follows directly from the flow decomposition theorem. The flow decomposition theorem basically says that we can rewrite flow constraints of edges using flow constraints of the paths. I am not going to prove this theorem as this is a known fact in graph theory and providing the proof for this problem is too much.

The dual of this LP is:

$$\min \sum_e c_e y_e$$

subject to

$$\begin{aligned} \sum_{e \in P} y_e &\geq 1, \forall P \\ y_e &\geq 0, \forall e \in V \times V \end{aligned}$$

This is exactly the linear relaxation of the min-cut integer program from (d).

To show that the optimal value of MINIMCUT is an upper bound on the value of the optimal flow, we use the weak duality:  $\sum_{p \in P} f_p \leq \sum_e c_e y_e$  or all feasible  $f_p, y_e$  in (P) and (D). Hence,

$$\sum_{p \in P} f_p^* \leq \sum_e c_e y_e^* \leq \sum_e c_e y_e$$

f. The primal is feasible as above. The dual is also feasible by letting  $y_e = 1$  for all edges. Hence, from strong duality, we have that the objective of the primal is equal to the objective of the dual. Thus, it must be the case that

$$\sum_{p \in P} f_p^* = \sum_e c_e y_e^*$$

**3. Optimality criteria.** Let  $f$  be a differentiable convex function from  $\mathbb{R}^n$  to  $\mathbb{R}$ . We consider the following optimization problem:

$$\min_{\mathbf{x} \in C} f(\mathbf{x})$$

where  $C \subseteq \mathbb{R}^n$  is a closed convex set.

- a. Show that  $\mathbf{x}^* \in C$  is an optimal solution to the above problem if and only if:

$$\forall \mathbf{y} \in C, \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0 \quad (\text{P})$$

- b. Show that when  $\nabla f(\mathbf{x}^*) \neq 0$  this implies that  $\mathbf{x}^*$  lies on the boundary of  $C$  and that  $\nabla f(\mathbf{x}^*)$  defines a supporting hyperplane of  $C$  at  $\mathbf{x}^*$ . Remember that the boundary  $\delta(C)$  of a closed set  $C$  is defined by:

$$\delta(C) \stackrel{\text{def}}{=} \{\mathbf{x} \in C \mid \forall r > 0, B(\mathbf{x}, r) \not\subseteq C\}$$

where  $B(\mathbf{x}, r)$  denotes the  $\ell_2$ -ball of center  $\mathbf{x}$  and radius  $r$ .

- c. Show that when  $C = \mathbb{R}^n$ ,  $\mathbf{x}^* \in \mathbb{R}^n$  satisfies condition (P) of part a. iff:

$$\nabla f(\mathbf{x}^*) = 0$$

Then observe (you don't have to prove it) that by combining part a. and c. we obtain the following theorem.

**Theorem 1.**  $\mathbf{x}^* \in \mathbb{R}^n$  is an optimal solution to the problem:

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

if and only if  $\nabla f(\mathbf{x}^*) = 0$ .

- a. (i)  $\mathbf{x}^* \in C$  is optimal, (ii)  $\forall \mathbf{y} \in C, \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0$ .

**Proof of (i)  $\Rightarrow$  (ii)**

Suppose for contradiction

$$\exists \mathbf{y} \in C, \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) < 0$$

Because  $C$  is convex,

$$\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*) \in C, \forall \lambda \in [0, 1]$$

Let  $g(\lambda) = f(\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*))$ . Then,

$$g'(\lambda) = (\mathbf{y} - \mathbf{x}^*)^\top \nabla f(\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*))$$

So,

$$g'(0) = (\mathbf{y} - \mathbf{x}^*)^\top \nabla f(\mathbf{x}^*)$$

This means  $\exists \epsilon, \epsilon > 0$  small enough such that

$$g(\lambda) < g(0), \forall \lambda \in (0, \epsilon)$$

This is because the gradient is negative around  $\lambda = 0$  and so the function is decreasing within some small interval around  $\lambda = 0$ . Hence, we have

$$f(\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*)) < f(\mathbf{x}^*), \forall \lambda \in (0, \epsilon)$$

This means  $\mathbf{x}^*$  is not as optimal  $\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*)$ , which contradicts with the assumption that  $\mathbf{x}^*$  is optimal. Hence,

$$\forall \mathbf{y} \in C, \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0$$

**Proof of (ii)  $\Rightarrow$  (i)** From the first order characterization of convex functions,

$$\forall \mathbf{y} \in C, f(\mathbf{y}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*)$$

Also, from the assumption, we have

$$\forall \mathbf{y} \in C, \nabla f(\mathbf{x}^*)^\top (\mathbf{y} - \mathbf{x}^*) \geq 0$$

Combining these two, we have

$$\forall \mathbf{y} \in C, f(\mathbf{y}) \geq f(\mathbf{x}^*)$$

Hence,  $\mathbf{x}^*$  is optimal.

b. We prove the contraposition.

We have from (a) that

$$\forall \mathbf{y} \in C, f(\mathbf{y}) \geq f(\mathbf{x}^*)$$

If  $\mathbf{x}^*$  lies on the interior of  $C$ , there exists  $\mathbf{y} = \mathbf{x}^* - \lambda \nabla f(\mathbf{x}^*)$  for a small enough  $\lambda > 0$  such that  $\mathbf{y} \in C$ . Plugging this in to the above inequality, we have

$$-\lambda \nabla f(\mathbf{x}^*)^\top \nabla f(\mathbf{x}^*) \geq 0$$

Since  $-\lambda < 0$  it must be the case that  $\nabla f(\mathbf{x}^*) = 0$ . Hence, we proved the contraposition.

c. Suppose for contradiction that  $\nabla f(\mathbf{x}^*) \neq 0$ . Then, from b., we have that  $\mathbf{x}^*$  lies on the boundary of  $C$ . However, since  $C = \mathbb{R}^n$ , we have  $\delta(C) = \phi$ . Hence, contradiction. Therefore, it must be the case that  $\nabla f(\mathbf{x}^*) = 0$ .