CSC 284/484 notes
(Spring 2020)

1 Simplex Method for LP

Let $x = (x_1, \dots, x_n)^\top$ be a vector of variables. Let $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, and $A \in \mathbb{R}^{m\times n}$. We consider the following linear program.

$$\begin{aligned} \max c^\top x& \\ Ax &= b \\ x &\geq 0 \end{aligned} \tag{1}$$

We assume that A has rank m (otherwise the constraints are either non-feasible or redundant). Note that from the assumption we have $n \geq m$. For a set $B \subseteq [n]$ let A_B be the submatrix of A obtained by taking the columns whose indices are in B. Let

$$\mathcal{B} = \{B \subseteq [n] \,|\, |B| = m \text{ and } \mathrm{rank}(A_B) = m\}.$$

Given $B \in \mathcal{B}$ does there exist a feasible solution x that is zero on indices outside B? Such a solution would have to satisfy $A_B x_B = b$ and since $\mathrm{rank}(B) = m$ there is only one candidate: $x_B = A_B^{-1} b$. The value of the solution would be

$$v(B) := c_B^\top A_B^{-1} b.$$

The set of **feasible basis** is

$$\mathcal{B}' = \{B \subseteq [n] \,|\, B \in \mathcal{B} \text{ and } A_B^{-1} b \geq 0\}.$$

For $B \in \mathcal{B}'$ the solution that is 0 on indices outside B and is given by $A_B^{-1} b$ on indices in B is called a **basic feasible solution**. The next lemma shows that it is sufficient to focus our attention to basic feasible solutions.

Lemma 1.1 *Suppose that* (1) *is bounded. Let $\hat{x}$ be an optimal solution of* (1)*. There exists $B \in \mathcal{B}'$ such that $v(B) = c^\top \hat{x}$.*

Proof :
Let $\hat{x}$ be an optimal solution that minimizes the number of non-zero coordinates in $\hat{x}$. Let $P \subseteq [n]$ be the non-zero coordinates of $\hat{x}$, that is, $P = \{i \in [n] \,|\, \hat{x}_i > 0\}$.

Suppose that $\mathrm{rank}(A_P) < |P|$, that is, the columns of A_P are linearly dependent. Then there exist $y_P \neq 0$ such that $A_P y_P = 0$. Let $y \in R^n$ be zero on coordinates outside P and given by y_P on coordinates in P. Note that for all λ sufficiently close to 0 we have $\hat{x} + \lambda y \geq 0$. From this and the optimality of $\hat{x}$ we obtain $c^\top y = 0$. We can now chose λ such that $\hat{x} + \lambda y$ is an optimal solution of (13) with a smaller number of non-zero coordinates than $\hat{x}$, a contradiction. (W.l.o.g., $y_i < 0$ for some $i \in P$; then take $\lambda = \min\{-x_i/y_i \,|\, y_i < 0\}$.)

From the previous paragraph we have $\mathrm{rank}(A_P) = |P|$. Using Steinitz exchange lemma we obtain B such that $P \subseteq B$, $|B| = m$ and $\mathrm{rank}(A_B) = m$. Note that

$$A\hat{x} = A_B \hat{x}_B = A_P \hat{x}_P = b,$$

and hence $\hat{x}_B = A_B^{-1} b$ which implies

$$v(B) = c_B^\top A_B^{-1} b = c_B^\top \hat{x}_B = c^\top \hat{x}.$$

■

Definition 1.2 We say that $B, B' \in \mathcal{B}'$ are **neighbors** if $|B \cap B'| = m - 1$.

1.1 First Attempt at an Algorithm (Degenerate LPs)

Consider the following algorithm.

$B \leftarrow$ an element from $\mathcal{B}'$
while *there exist* $B' \in \mathcal{B}'$ *that is a neighbor of* B *such that* $v(B') > v(B)$ **do**
| $B \leftarrow B'$
end

Algorithm 1: "Simplex" algorithm v1

Remark 1 Algorithm 1, as Example 1.3 shows, is not correct—it is possible that there exist B such that 1) B is not optimum and 2) for all neighboring B' we have $v(B') \leq v(B)$.

Example 1.3

1.2 Second Attempt at an Algorithm

Consider the following algorithm.

$B \leftarrow$ an element from $\mathcal{B}'$
while *there exist* $B' \in \mathcal{B}'$ *that is a neighbor of* B *such that* $v(B') \geq v(B)$ **do**
| $B \leftarrow B'$
end

Algorithm 2: "Simplex" algorithm v2

Remark 2 There are at least two problems with Algorithm 2.

- Termination. Example 1.4
- Cycling. Example 1.5

Example 1.4

Example 1.5

1.3 Simplex Algorithm - a view through equations

We are going to take a slightly different view of the algorithm (we will connect the view with the basic feasible solutions in Section 1.4).

Suppose that we have n variables $x_1, \ldots, x_n$, a list B of m distinct elements of $[n]$, and our linear program has the following form (V, b's, c's, and A's are real numbers and x's are variables).

$$\begin{aligned} \max V + \sum_{j \notin B} c_j x_j \\ x_{B[1]} &= b_1 - \sum_{j \notin B} A_{1,j} x_j \\ &\ldots \\ x_{B[m]} &= b_m - \sum_{j \notin B} A_{m,j} x_j \\ x &\geq 0 \end{aligned} \tag{2}$$

where

$$b_1, \dots, b_m \geq 0. \tag{3}$$

Note that letting

$$x_{B[i]} = b_i \text{ for } i \in [m] \quad \text{and} \quad x_j = 0 \text{ for } j \notin B \tag{4}$$

yields a feasible solution of (2); the value of the solution is V. (Note that the condition (3) ensures $x \geq 0$ in (4).)

Remark 3 Suppose that $c_t \leq 0$ for all $t \notin B$. Then (4) is an optimal solution of (2).

Now suppose that $c_t > 0$ for some $t \notin B$. Consider the following assignment

$$x_{B[i]} = b_i - \lambda A_{i,t} \text{ for } i \in [m], \quad x_t = \lambda \quad \text{and} \quad x_j = 0 \text{ for } j \notin (B \cup \{t\}). \tag{5}$$

Note that (5) is a feasible solution of (2) for any $\lambda \geq 0$ such that $b_i - \lambda A_{i,t} \geq 0$ for all $i \in [m]$; the value of the solution is $V + \lambda c_t$. If $A_{i,t} \leq 0$ for all $i \in [m]$ then any $\lambda \geq 0$ yields a valid solution and (2) is unbounded. If $A_{i,t} > 0$ for some $i \in [m]$ then the largest λ we can take (and still have feasible solution) is

$$\lambda^* := \min\{b_i/A_{i,t} \,|\, i \in [m]; A_{i,t} > 0\}. \tag{6}$$

Let $s \in [m]$ be a minimizer in (6) (that is, $A_{s,t} > 0$ and $b_s/A_{s,t} = \lambda^*$). We will now re-write (2) into a program of the same form with B' that is obtained from B by replacing the s-th element by t. We replace the equation

$$x_{B[s]} = b_s - \sum_{j \notin B} A_{s,j} x_j$$

by an equivalent equation

$$x_t = \frac{b_s}{A_{s,t}} - \frac{1}{A_{s,t}} x_{B[s]} - \sum_{j \notin B \cup \{t\}} \frac{A_{s,j}}{A_{s,t}} x_j. \tag{7}$$

Note that equation (7) has the required form (with B replaced by B'):

$$x_{B'[s]} = b'_s - \sum_{j \notin B'} A'_{s,j} x_j. \tag{8}$$

To bring the remaining equations (including the objective) into the required form we plug in x_t from equation (7); note that this makes the relevant part of the new equations a constant term plus a linear combination of x_j, $j \notin B'$. Note that $V' = V + \lambda^* c_t = V + c_t b_s / A_{s,t}$.

Finally, note that the condition (3) is satisfied. The solution (5) for $\lambda = \lambda^*$ is zero for $j \notin B'$ and non-negative for $j \in B'$ (since it is feasible). Hence $0 \leq x_{B'[i]} = b'_i$.

The simplex algorithm is the process we just described. The termination criterion is given by Remark 3 and a step of the algorithm is the change described in the case $c_t > 0$ for some $t \notin B$.

1.4 Simplex Algorithm - a view through basic feasible solutions

We are going to describe the algorithm from Section 1.3 in the language of basis.

Remark 4 Note that row operations (and equivalently left multiplication by regular matrices) on the system $Ax = b$ do not change the solution space, that is, if $S \in \mathbb{R}^{m \times m}$ is regular then $SAx = Sb$ iff $Ax = b$.

Remark 5 Note that row operations (and equivalently left multiplication by regular matrices) on the linear program $Ax = b, x \geq 0$ do not change the set of basic feasible solutions, that is, 1) if $S \in \mathbb{R}^{m \times m}$ is regular then $\det(SA_B) \neq 0$ iff $\det(A_B) \neq 0$ and 2) $(SA)_B^{-1} Sb = A_B^{-1} b$.

Remark 6 Note that for any $s \in \mathbb{R}^m$ adding $s^\top(Ax-b)$ to the objective $c^\top x$ does not change the objective on the space of feasible solutions.

Remarks 4 and 6 will allow us to bring the linear program into the same form as in Section 1.3 (each variable in the basis is expressed as a constant plus a linear combination of variables outside the basis; the objective is expressed as a constant plus a linear combination of variables outside basis).

A step of the Simplex algorithm starts with a feasible basis and either 1) terminates with an optimal solution or 2) terminates finding the program is unbounded or 3) moves to a new feasible basis. We address the question how to find the initial feasible basis in Section 1.5

Suppose B is a feasible basis. By Remarks 4 and 6 the program is equivalent to

$$\begin{aligned} c_B^\top A_B^{-1} b + \max \hat{c}^\top x \\ A_B^{-1} A x = A_B^{-1} b \\ x \geq 0 \end{aligned} \tag{9}$$

where

$$\hat{c}^\top := \left(c^\top - c_B^\top A_B^{-1} A\right).$$

Recall that basic solution for B is given by $x_B = A_B^{-1} b$ and $x_j = 0$ for $j \notin B$. Note that $\hat{c}_B = 0$. The program (13) is in the same form as (2). The analog of Remark 3 is the following.

Remark 7 If $\hat{c} \leq 0$ then the basic solution for B is an optimal solution.

Suppose that $\hat{c}_t > 0$ for some t (note that we must have $t \notin B$) then we have a candidate to move into the basis. Let j-th element in B be denoted $B[j]$. We have

$$(A_B^{-1} A)\mathrm{e}_{B[j]} = \mathrm{e}_j.$$

Hence we have

$$A_B^{-1} A(\mathrm{e}_t - \sum_{j=1}^{m} (A_B^{-1} A)_{jt} \mathrm{e}_{B[j]}) = 0. \tag{10}$$

To verify (10) note

$$\mathrm{e}_s^\top A_B^{-1} A(\mathrm{e}_t - \sum_{j=1}^{m} (A_B^{-1} A)_{jt} \mathrm{e}_{B[j]}) = \left(A_B^{-1} A\right)_{st} - \sum_{j=1}^{m} (A_B^{-1} A)_{jt} \mathrm{e}_s^\top \left(A_B^{-1} A\right) \mathrm{e}_{B[j]} =$$

$$\left(A_B^{-1} A\right)_{st} - \sum_{j=1}^{m} (A_B^{-1} A)_{jt} \mathrm{e}_s^\top \mathrm{e}_j = \left(A_B^{-1} A\right)_{st} - (A_B^{-1} A)_{st} = 0.$$

Let x be the basic solution for B. For what $\lambda \geq 0$ is

$$y(\lambda) := x + \lambda(\mathrm{e}_t - \sum_{j=1}^{m} (A_B^{-1} A)_{jt} \mathrm{e}_{B[j]})$$

feasible? If $(A_B^{-1} A)_{jt} \leq 0$ for all j then any $\lambda \geq 0$ and the linear program is unbounded. Otherwise let

$$\lambda^* = \min\left\{ \frac{(A_B^{-1} b)_j}{(A_B^{-1} A)_{jt}} \mid (A_B^{-1} A)_{jt} > 0 \right\}. \tag{11}$$

Let s be a minimizer in (11). Then $y(\lambda^*)_s = 0$ and the support of $y(\lambda^*)$ is contained in $B' := B \cup \{t\} \setminus \{s\}$. Since $(A_B^{-1} A)_{st} > 0$ we have that $(A_B^{-1} A)_{B'}$ has full rank (subtracting appropriate multiple of columns in $(A_B^{-1} A)_{B \cap B'}$ from the t-th column of $A_B^{-1} A$ one can obtain a column whose only non-zero entry is in the s-th row; the columns in $(A_B^{-1} A)_{B \cap B'}$ cover the remaining rows).

We have that $y(\lambda^*)$ is the basic solution for B' whose value increased by $\lambda^* c_t$ over the basic solution for B. Again this is what the simplex algorithm does.

1.5 Finding the initial feasible basis

To find the initial feasible basis of program (1) we will construct a new linear program such that 1) the new linear program has an "obvious" feasible basis and 2) solving the new linear program will yield an initial feasible basis of (1).

Multiplying equations in (1) by -1, if necessary, we can ensure $b \geq 0$ in (1). Consider the following linear program.

$$\begin{aligned} \min 1^\top y & \\ Ax + y &= b \\ x &\geq 0 \\ y &\geq 0 \end{aligned} \tag{12}$$

Note that (12) is feasible—one can take $y = b$ and $x = 0$. This is a basic feasible solution (the basis contains variables y) of (12).

Lemma 1.6 *Optimal solution of* (12) *has value* 0 *iff* (1) *is feasible.*

Proof :
If the optimal solution (x, y) of (12) has value 0 then $y = 0$ and hence $Ax = b, x \geq 0$, that is, x is a feasible solution of (1).

If (1) is feasible then taking $y = 0$ yields a feasible solution (x, y) of (12) of value 0. ■

Thus to obtain an initial feasible basis of (1) we solve (12) using the Simplex algorithm (the good news is that for (12) we have an initial feasible solution). If the value of the optimal solution is positive then (1) is not feasible.

Now assume that the value of the optimal solution is 0 and the Simplex algorithm ended up with a basis B for (12). The basis B can still contain variables from y. We are now going to remove the variable in y from B and obtain an initial feasible basis of (1).

Let $A' = (A \,|\, I)$ and $x' = \begin{pmatrix} x \\ y \end{pmatrix}$. Thus the first n coordinates of x' are variables in x (we will call these x-variables) and the last m coordinates are variables in y (we will call these y-variables).

Suppose that there exists a variable y_i that is the j-th variable in the basis B. There are two possibilities:

- There exists $k \in [n]$ such that $(A'^{-1}_B)A'_{jk} \neq 0$, that is, there exists an x-variable that is not in B that can be swapped for y_i. In this case replacing y_i in the basis with x_k yields a basis for the same solution that has one fewer y-variable.
- For all $k \in [n]$ we have $(A'^{-1}_B)A'_{jk} = 0$. This means that the rows of A are linearly dependent and $\text{rank}(A) < m$, a contradiction (we assumed $\text{rank}(A) = m$).

1.6 Cycling

Example 1.7 Consider the following linear program

$$\begin{aligned} \max\ (-3, 80, -2, 24, 0, 0, 0)x & \\ \begin{pmatrix} 1/4 & -8 & -1 & 9 & 1 & 0 & 0 \\ 1/2 & -12 & -1/2 & 3 & 0 & 1 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 1 \end{pmatrix} x &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ x &\geq 0 \end{aligned} \tag{13}$$

- Suppose that the starting feasible basis is $\{2,5,7\}$. The following are the neighboring feasible basis (all have value 0):

$$\{1,2,7\},\{1,5,7\},\{2,3,7\},\{2,4,7\},\{2,6,7\},\{3,5,7\},\{4,5,7\},\{5,6,7\}.$$

- Suppose that the starting feasible basis is $\{2,4,7\}$. The following are the neighboring feasible basis (all have value 0):

$$\{1,2,7\},\{1,4,7\},\{2,3,7\},\{2,5,7\},\{2,6,7\},\{3,4,7\},\{4,5,7\},\{4,6,7\}.$$

- Suppose that the starting feasible basis is $\{4,5,7\}$. The following are the neighboring feasible basis (all have value 0):

$$\{1,4,7\},\{1,5,7\},\{2,4,7\},\{2,5,7\},\{3,4,7\},\{3,5,7\},\{4,6,7\},\{5,6,7\}.$$

1.7 Bland's rule

The notation in this section will differ a little bit from the previous sections (to make the indexing convenient for the proof). We are going to follow the following two rules:

1. Pick the smallest index j such that $c_j > 0$ to enter the base.
2. Pick smallest $k \in B$ tight for j to exit the base.

Lemma 1.8 *If we follow Bland's rule the Simplex algorithm cannot cycle.*

Proof :
Suppose that the algorithm does cycle. Let $B_0, \ldots, B_m = B_0$ be a sequence of feasible basis encountered during a cycle in the algorithm. Let F be the set of indices of variables that are 1) in some of basis and 2) not in some of the basis. (The variables whose indices are in F are called "fickle".) Note that the underlying solution (assignment to x) stays the same throughout the cycle (if a step of the simplex algorithm changes the underlying solution then the value of the solution strictly increases). Since each fickle variable is not in some of the bases, this means that its assigned value must be 0 in the underlying solution.

Let x_t be a fickle variable with the largest index t. Let B be a base where x_t leaves and some x_s enters (since both $t, s \in F$ we have $s < t$ by our choice of t). Following Section 1.3 the state of the algorithm for B is as follows. For $i \in B$ we have

$$x_i = b_i - \sum_{j \notin B} a_{ij} x_j \tag{14}$$

and the objective is

$$z = V + \sum_{j \notin B} c_j x_j, \tag{15}$$

note that $c_j = 0$ for $j \in B$. Note that $x_t = b_t = 0$ (since x_t is fickle).

Consider the parametric solution for x_s entering B and x_t leaving B (see equation (5)), that is,

$$x_i = \begin{cases} \lambda & i = s, \\ 0 & i \notin (B \cup \{s\}), \\ b_i - \lambda a_{is} & i \in B. \end{cases} \tag{16}$$

Let $\hat{B}$ be a base where x_t enters. In the state of the algorithm for $\hat{B}$ we have the objective

$$z = V + \sum_{j \notin \hat{B}} \hat{c}_j x_j, \tag{17}$$

note again $\hat{c}_j = 0$ for $j \in \hat{B}$.

The value of a solution in the linear subspace $Ax = b$ is the same in both views (that is in (15) and in (17)). Evaluating the solution (16) (note that the solution lies in the subspace $Ax = b$ for any λ) in (15) and (17) we obtain

$$V + c_s\lambda = V + \hat{c}_s\lambda + \sum_{j \in B \setminus \hat{B}} \hat{c}_j(b_j - \lambda a_{js}),$$

which simplifies to

$$\lambda(c_s - \hat{c}_s + \sum_{j \in B \setminus \hat{B}} \hat{c}_j a_{js}) = \sum_{j \in B \setminus \hat{B}} \hat{c}_j b_j.$$

Since RHS does not depend on λ we have

$$c_s - \hat{c}_s + \sum_{j \in B \setminus \hat{B}} \hat{c}_j a_{js} = 0. \tag{18}$$

(If equation $\lambda S = T$ is true for every λ then $S = T = 0$.)

Note that $c_s > 0$ since x_s was entering B. Note that $\hat{c}_s \leq 0$ since x_s is not entering $\hat{B}$ (if $\hat{c}_s > 0$ then since $s < t$ we would choose x_s to enter $\hat{B}$ instead of x_t.) Hence, using (18), there exists $r \in B \setminus \hat{B}$ such that

$$\hat{c}_r a_{rs} < 0. \tag{19}$$

Note that $r \in B \setminus \hat{B}$ implies $r \in F$, that is, x_r is fickle. By our choice of t we have $r \leq t$. Note that $a_{ts} > 0$ since x_t is leaving B. Note that $\hat{c}_t > 0$ since x_t is entering $\hat{B}$. Hence $\hat{c}_t a_{ts} > 0$ which (together with (19)) implies $r \neq t$. Thus $r < t$.

Note that $\hat{c}_r \leq 0$ (otherwise x_r would enter $\hat{B}$), which, in turn, implies $a_{rs} > 0$. Since x_r is fickle we have $b_r = 0$ in (14) (for $i = r$). We have $b_r / a_{rs} = 0$ and $a_{rs} > 0$ which means $r \in B$ was a tight for s to exit the base and since $r < t$ by Bland's rule x_r should have left B instead of x_t, a contradiction. ■