CS 491I Approximation algorithms

Lecture notes

Jan.25,2001

Gayane Goltukhchyan

Simplex Method

Notation Scheme:

$\vec{x}$ - arrow indicates vector,

X – capital letter indicates matrix,

x – small letter without arrow indicates scalar.

Objective:

$$\begin{aligned} &Max \quad z = \vec{c}\ \vec{x} \\ &s.t. \quad A\ \vec{x} = \vec{b} \\ &\qquad\ \ \vec{x} \geq \vec{0} \end{aligned}$$

where

A - a matrix of rank m x n (n – number of variables, m – number of constraints),

$\vec{c}$ - vector of objective function's coefficients,

$\vec{x}$ - vector of unknown variables,

$\vec{b}$ - vector of constraints' coefficients,

$\vec{0}$ - null vector.

One way to solve this problem is to find all extreme points (basic feasible solutions), put them to objective function and then select the solution, which yields to the maximum value of objective function (z). But it will take in the worst case $\binom{n}{m}$ exponential time.

Simplex

Step1.

Represent matrix A in the following way:

$$\mathbf{A} = (\mathbf{B} \quad \mathbf{N})$$

It is non-trivial problem to get the basis B, so assume we are given the first basis B.

So we have the following:

$\vec{x} = (\vec{x}_B, \vec{x}_N)$ where $\vec{x}_B$ is the vector of basic variables ,

$\vec{x}_N$ is the vector of non basic variables.

$\vec{c} = (\vec{c}_B, \vec{c}_N)$ where $\vec{c}_B, \vec{c}_N$ corresponding vectors of coefficients of basic and non-basic variables.

The expression $A\ \vec{x} = \vec{b}$ can be rewritten in the following way:

$$(B \quad N)\begin{pmatrix} \vec{x}_B \\ \vec{x}_N \end{pmatrix} = \vec{b} \quad \Rightarrow \quad B\,\vec{x}_B + N\ \vec{x}_N = \vec{b}$$

We know that B^{-1} exists since we have basic variables.

$$\vec{x}_B = B^{-1}\,\vec{b} - B^{-1}\ N\ \vec{x}_N$$

$$\vec{x} = \begin{pmatrix} B^{-1}\,\vec{b} \\ \vec{0} \end{pmatrix}$$ here we set $\vec{x}_N$ (all non-basic variables) to zero.

Now objective function can be rewritten in the following way:

$$z = \vec{c}_B\ \vec{x}_B + \vec{c}_N\ \vec{x}_N$$

We wish to check whether we are at optimum or not.

Let's substitute $\vec{x}_B$ into z, we will get:

$$z = \vec{c}_B\,(B^{-1}\,\vec{b} - B^{-1}\ N\ \vec{x}_N) + \vec{c}_N\ \vec{x}_N = \vec{c}_B\ B^{-1}\,\vec{b} - (\vec{c}_B\ B^{-1}\ N - \vec{c}_N)\,\vec{x}_N$$

Let J = {set of non-basic columns}, so we can rewrite z as follows:

$$z = \vec{c}_B\ B^{-1}\,\vec{b} - \sum_{j\in J} (\vec{c}_B\ B^{-1}\ a_j - c_j)\ x_j$$

Observe that the rate of change of objective function with respect to x_j is

$\dfrac{\partial z}{\partial x_j} = -(\vec{c}_B\ B^{-1}\ a_j - c_j)$ which is also (z_j - c_j) and is called reduced cost.

How can we change the basis? We need to pick a variable that will leave the basis and choose the variable that will enter into the basis. If $(\vec{c}_B\ B^{-1}\ a_j - c_j) < 0$, derivative is positive, hence increasing x_j from zero improves objective function. If negative, no need to increase x_j. This leads us to step2.

Step2.

If all non-basic variables cause derivative to be less then or equal to zero, then we are at optimal point (no need to change the basis). If this is not the case then one way to improve objective function is to follow greedy strategy. We can bring in the variables with maximum $\frac{\partial z}{\partial x_j}$, but we have to make sure that we still have basic feasible solution.

<u>Theorem 1.</u> Any vector $\vec{a} \in R^m$ can be expressed as a <u>unique</u> linear combination of the basis vectors $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_m)$.

Proof. Since $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_m)$ is basis, definitely we have

$\vec{a} = \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 \ldots\ldots \lambda_m \vec{b}_m \quad$ where not all $\lambda_i = 0$

Let assume that there is another linear combination, then

$\vec{a} = \mu_1 \vec{b}_1 + \ldots\ldots \mu_m \vec{b}_m \quad \Rightarrow \quad \vec{0} = (\lambda_1 - \mu_1)\vec{b}_1 + \ldots\ldots (\lambda_m - \mu_m)\vec{b}_m$

Since $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_m)$ is a basis, all $(\lambda_i - \mu_i)$ must be equal to $0 \Rightarrow \lambda_i = \mu_i$.

<u>Theorem 2.</u> Let $\vec{a} = \lambda_1 \vec{b}_1 + \lambda_2 \vec{b}_2 \ldots\ldots \lambda_m \vec{b}_m \quad where\ \lambda_m \neq 0,$ then there exists a set of $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_{m-1}, \vec{a})$ that is a basis.

Proof. We need to show that the set $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_{m-1}, \vec{a})$ is linearly independent.

To prove it let's assume that the set $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_{m-1}, \vec{a})$ is not a basis, i.e. it is linearly dependent. To be linearly dependent we should have a constraint like follows:

$\mu_1\vec{b}_1 + \mu_2 \vec{b}_2 \ldots\ldots \mu_{m-1}\vec{b}_{m-1} + \delta \vec{a} = \vec{0}$, where not all $\mu_i, \delta = 0$.

Can $\delta = 0$? No, it cannot, otherwise $\mu_1\vec{b}_1 + \mu_2 \vec{b}_2 \ldots\ldots \mu_{m-1}\vec{b}_{m-1} = \vec{0}$, which is impossible because $(\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_{m-1}) \subseteq (\vec{b}_1, \vec{b}_2 \ldots\ldots \vec{b}_m)$ and a subset of a linearly independent set is always linearly independent.

Substituting $\vec{a}$ by the above expression:

$$\mu_1\vec{b}_1 + \mu_2\,\vec{b}_2 \ldots\ldots \mu_{m-1}\vec{b}_{m-1} + (\lambda_1\,\vec{b}_1 + \lambda_2\,\vec{b}_2 \ldots\ldots \lambda_m\,\vec{b}_m)\delta =$$
$$(\mu_1 + \delta\alpha_1)\vec{b}_1 + (\mu_2 + \delta\alpha_2)\vec{b}_2 + \ldots + \lambda_m\,\delta\,\vec{b}_m = \vec{0}$$

If not linearly independent, then not all $\lambda_m, \delta = 0$. But there is no such combination that makes it 0, because by definition we have $\lambda_m \neq 0$, hence $\lambda_m \delta\, \vec{b}_m \neq 0$.

Pick departing variable:

$$\vec{x}_B = B^{-1}\vec{b} - B^{-1}\,N\,\vec{x}_N \;\Rightarrow$$

$$\vec{x}_B = B^{-1}\vec{b} - \sum_{j\in J}(B^{-1}\,\vec{a}_j)\,x_j$$

Variable that coming in. Let x_k is non basic variable that coming in.

$$B^{-1}\vec{a}_k = \vec{\alpha}_k$$
$$\vec{a}_k = B\,\vec{\alpha}_k$$
$$\vec{\alpha}_k = \begin{pmatrix} \alpha_{1,k} \\ \vdots \\ \\ \vdots \\ \alpha_{m,k} \end{pmatrix}$$
$$\vec{x}_B \geq 0$$

We can only push x_k up until one of the variables becomes negative.

$$\begin{pmatrix} b_1 \\ \vdots \\ \vdots \\ b_k \end{pmatrix} - x_k \begin{pmatrix} \alpha_{1,k} \\ \vdots \\ \vdots \\ \alpha_{m,k} \end{pmatrix} \geq 0$$

$$x_k = \min\left\{ \frac{b_j}{\alpha_{j,k}} \quad , \alpha_{j,k} > 0 \right\}$$

This is called minimum ratio test.

<u>Example.</u> max $z = 2x_1+3x_2$

s.t. $x_1 - 2x_2 \leq 4$

$2x_1 + x_2 \leq 18$

$x_2 \leq 10$

$x_1, x_2 \geq 0$

Rewriting in the standard (canonical) form:

$$\max \ z = 2x_1+3x_2 \tag{1.1}$$

$$\text{s.t.} \quad x_1 - 2x_2 + x_3 = 4 \tag{1.2}$$

$$2x_1 + x_2 + x_4 = 18 \tag{1.3}$$

$$x_2 + x_5 = 10 \tag{1.5}$$

$$x_1, x_2, x_3, x_4, x_5 \geq 0$$

Data for the problem is:

$$A = \begin{pmatrix} 1 & -2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \qquad b = \begin{pmatrix} 4 \\ 18 \\ 10 \end{pmatrix} \qquad c = \begin{pmatrix} 2 & 3 & 0 & 0 & 0 \end{pmatrix}$$

We begin by choosing a starting basis matrix **B**. Since the solution will be determined by $\mathbf{B}^{-1}$, we will choose starting basis matrix **B = I**.

Observe that from matrix **A**,

$$B = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I \qquad x_B = \begin{pmatrix} x_{B,1} \\ x_{B,2} \\ x_{B,3} \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix}$$

Now solving for z and x_B in terms of x_N, we get:

$$z = 2x_1 + 3x_2 \tag{1.6}$$

$$x_3 = 4 - x_1 + 2x_2 \tag{1.7}$$

$$x_4 = 18 - 2x_1 - x_2 \tag{1.8}$$

$$x_5 = 10 - x_2 \tag{1.9}$$

The starting solution, which is obtained by setting the non basic variables equal to zero, can be summarized as follows:

$$z = 0$$

$$x_B = \begin{pmatrix} x_{B,1} \\ x_{B,2} \\ x_{B,3} \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ 18 \\ 10 \end{pmatrix}$$

$$x_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$B = (a_3 \quad a_4 \quad a_5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

From the (1.6) we see that $\frac{\partial z}{\partial x_1} = -(z_1 - c_1) = 2 > 0$ *and* $\frac{\partial z}{\partial x_2} = -(z_2 - c_2) = 3 > 0$ .That is $z_1 - c_1 = -2 < 0$ *and* $z_2 - c_2 = -3 < 0$. So increasing either x_1 or x_2 will increase the value of z, hence the current solution is not optimal. Because $\frac{\partial z}{\partial x_1} < \frac{\partial z}{\partial x_2}$ let us choose variable x_2 as the entering variable. Next we need to find departing variable using the minimum ratio test. As x_2 increased, we must ensure that x_3 , x_4 , x_5 remain nonnegative. From (1.7)-(1.9) the values of the basic variables are given by:

$$x_B = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \beta - x_2\,\alpha_2 = \begin{pmatrix} 4 \\ 18 \\ 10 \end{pmatrix} - x_2 \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \geq 0 \qquad (1.10)$$

From (1.10) $\frac{\partial x_3}{\partial x_2} = 2$, thus x_3 increases 2 units for each unit increase in x_2. $\frac{\partial x_4}{\partial x_2} = -1$ and $\frac{\partial x_5}{\partial x_2} = -1$. From (1.10) x_4 and x_5 will remain nonnegative as long as $x_2 \leq 18/1$ and $x_5 \leq 10/1$ respectively.

Minimal ratio testis min {18, 10} = 10 and x_5 is the departing variable. Equation (1.9) is called *blocking equation* and x_5 is the *blocking variable* or *departing variable*.

New canonical expression now is derived by solving for x_2 in the blocking equation and using this representation of x_2 to eliminate x_2 from the remaining equations.

This process is called *pivot* and results in the following:

$$z = 2x_1 + 2(10 - x_5) = 30 + 2x_1 - 3x_5 \qquad (1.11)$$
$$x_3 = 4 - x_1 + 2(10 - x_5) = 24 - x_1 - 2x_5$$
$$x_4 = 18 - 2x_1 - (10 - x_5) = 8 - 2x_1 + x_5$$
$$x_2 = 10 - x_5$$

The current solution and basis matrix can be summarized as:

$$z = 30$$

$$x_B = \begin{pmatrix} x_{B,1} \\ x_{B,2} \\ x_{B,3} \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \\ x_2 \end{pmatrix} = \begin{pmatrix} 24 \\ 8 \\ 10 \end{pmatrix}$$

$$x_N = \begin{pmatrix} x_1 \\ x_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$B = (a_3 \quad a_4 \quad a_2) = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Note that a_2 has replaced a_5 n the basis matrix **B**.

This solution is not optimal, because from (1.11) we see that $z_1 - c_1 = -2 < 0$, thus x_1 is chosen as the entering variable.

Basis variables can be written as:

$$x_B = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} = \beta - x_1 \alpha_1 = \begin{pmatrix} 24 \\ 8 \\ 10 \end{pmatrix} - x_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \geq 0 \qquad (1.12)$$

From 1.12 x_3 will remain nonnegative s long as $x_2 \leq 24/1 = 24$, similarly, x_4 and x_2 will remain nonnegative as long as $x_4 \leq 8/2 = 4$ and $x_2 \leq \infty$, respectively.

Minimum ratio test yields min {24, 4} = 4 and x_4 is the departing variable.

The pivot operation results in

$$z = 30 + 2\left(4 - \frac{x_4}{2} + \frac{x_5}{2}\right) - 3x_5 = 38 - x_4 - 2x_5$$

$$x_3 = 24 - \left(4 - \frac{x_4}{2} + \frac{x_5}{2}\right) - 2x_5 = 20 + \frac{x_4}{2} - \frac{5x_5}{2}$$

$$x_1 = 4 - \frac{x_4}{2} + \frac{x_5}{2}$$

$$x_2 = 10 - x_5$$

This solution is optimal, because now $z_4 - c_4 = 1 > 0$ and $z_5 - c_5 = 2 > 0$.

The optimal solution can be summarized as follows:

$$z^* = 38$$

$$x_B^* = \begin{pmatrix} x_{B,1}^* \\ x_{B,2}^* \\ x_{B,3}^* \end{pmatrix} = \begin{pmatrix} x_3^* \\ x_1^* \\ x_2^* \end{pmatrix} = \begin{pmatrix} 20 \\ 4 \\ 10 \end{pmatrix}$$

$$x_N^* = \begin{pmatrix} x_4^* \\ x_5^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$B = \left(a_3 \quad a_1 \quad a_2\right) = \begin{pmatrix} 1 & 1 & -2 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$