MATH 4211/6211 – Optimization

Non-Simplex Methods for LP

Xiaojing Ye
Department of Mathematics & Statistics
Georgia State University

Consider the primal LP given by

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \ge \boldsymbol{b} \\ & \boldsymbol{x} \ge \boldsymbol{0} \end{aligned}$$

and the corresponding dual problem

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{b}^\top \boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{A}^\top \boldsymbol{\lambda} \le \boldsymbol{c} \\ & \boldsymbol{\lambda} \ge \boldsymbol{0} \end{aligned}$$

Due to the duality theory, the primal-dual pair $[\boldsymbol{x}; \boldsymbol{\lambda}] \in \mathbb{R}^{n+m}$ is a solution if and only if

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{x} &= \boldsymbol{b}^\top \boldsymbol{\lambda} \\ \boldsymbol{A}\boldsymbol{x} &\geq \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{\lambda} &\leq \boldsymbol{c} \\ \boldsymbol{x} &\geq \boldsymbol{0} \\ \boldsymbol{\lambda} &\geq \boldsymbol{0} \end{aligned}$$

We can further rewrite $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{b}^\top \boldsymbol{\lambda}$ as $\boldsymbol{c}^\top \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{\lambda} \leq 0$ and $-\boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{\lambda} \leq 0$.

Therefore we obtain a system of inequalities as

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{\lambda} &\le 0 \\ -\boldsymbol{c}^\top \boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{\lambda} &\le 0 \\ \boldsymbol{A}\boldsymbol{x} &\ge \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{\lambda} &\le \boldsymbol{c} \\ \boldsymbol{x} &\ge \boldsymbol{0} \\ \boldsymbol{\lambda} &\ge \boldsymbol{0} \end{aligned}$$

Note that $[\boldsymbol{x}; \boldsymbol{\lambda}]$ solves the system of inequalities above iff $[\boldsymbol{x}; \boldsymbol{\lambda}]$ is a solution to the primal and dual LP.

The system of inequalities can be concisely written as

$$\boldsymbol{P}\boldsymbol{z} \le \boldsymbol{q}$$

where

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{c}^\top & -\boldsymbol{b}^\top \\ -\boldsymbol{c}^\top & \boldsymbol{b}^\top \\ -\boldsymbol{A} & \boldsymbol{0}_{m\times m} \\ -\boldsymbol{I}_n & \boldsymbol{0}_{n\times m} \\ \boldsymbol{0}_{n\times n} & \boldsymbol{A}^\top \\ \boldsymbol{0}_{n\times m} & -\boldsymbol{I}_m \end{bmatrix}, \quad \boldsymbol{z} = \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\lambda} \end{bmatrix}, \quad \boldsymbol{q} = \begin{bmatrix} 0 \\ 0 \\ -\boldsymbol{b} \\ \boldsymbol{0}_n \\ \boldsymbol{c} \\ \boldsymbol{0}_m \end{bmatrix}$$

Now the question becomes solving $\boldsymbol{P}\boldsymbol{z} \le \boldsymbol{q}$.

We introduce the notation of **ellipsoid** in $\mathbb{R}^s$ associated with $\boldsymbol{Q} \in \mathbb{R}^{s\times s}$ centered at $\boldsymbol{z} \in \mathbb{R}^s$ as

$$E_{\boldsymbol{Q}}(\boldsymbol{z}) := \left\{\boldsymbol{z} + \boldsymbol{Q}\boldsymbol{y} : \boldsymbol{y} \in \mathbb{R}^s,\ \|\boldsymbol{y}\| \le 1\right\}$$

If $\boldsymbol{Q}$ is an orthogonal matrix then $E_{\boldsymbol{Q}}(\boldsymbol{z})$ is a unit ball center at $\boldsymbol{z}$.

Khachiyan's method (or **ellipsoid method**) proceeds in the following way: assuming at $\boldsymbol{z}^{(k)}$ we compute $\boldsymbol{Q}_k$ making sure the optimal solution is in the ellipsoid $E_{\boldsymbol{Q}_k}(\boldsymbol{z}^{(k)})$. Then find $\boldsymbol{z}^{(k+1)}$, and so on, until $\boldsymbol{P}\boldsymbol{z}^{(k)} \le \boldsymbol{q}$. However in practice Khachiyan's method is very slow.

Affine Scaling Method

We consider the standard form of LP:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

Suppose we are currently at a feasible point $\boldsymbol{x}^{(0)}$ which is an interior point of $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\}$. That is, $\boldsymbol{A}\boldsymbol{x} = 0$ and $\boldsymbol{x} > \boldsymbol{0}$. Then we seek for a descent direction $\boldsymbol{d}^{(0)}$ and step size α_0, such that

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{d}^{(0)}$$

is still an interior point of Ω but closer to the optimal solution $\boldsymbol{x}^*$.

To make sure $\boldsymbol{x}^{(1)}$ is still in Ω, we need $\boldsymbol{A}\boldsymbol{x}^{(1)} = \boldsymbol{b}$. Hence

$$\boldsymbol{A}(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)}) = \alpha_0 \boldsymbol{A}\boldsymbol{d}^{(0)} = \boldsymbol{0}$$

Hence $\boldsymbol{d}^{(0)}$ is in the null space of $\boldsymbol{A}$.

Note the **orthogonal projector** defined below

$$\boldsymbol{P} = \boldsymbol{I}_n - \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}\boldsymbol{A}$$

has the property that $\boldsymbol{P}\boldsymbol{z}$ is the projection of $\boldsymbol{z}$ onto the null space of $\boldsymbol{A}$. Therefore, we would like to use the projection of the negative gradient $-\nabla f(\boldsymbol{x}) = -\nabla(\boldsymbol{c}^\top\boldsymbol{x}) = -\boldsymbol{c}$ as the descent direction $\boldsymbol{d}^{(0)}$. More specifically,

$$\boldsymbol{d}^{(0)} = \boldsymbol{P}(-\boldsymbol{c}) = -\boldsymbol{P}\boldsymbol{c} = -(\boldsymbol{I}_n - \boldsymbol{A}^\top(\boldsymbol{A}\boldsymbol{A}^\top)^{-1}\boldsymbol{A})\boldsymbol{c}$$

To maximize efficiency, we would like to start from some $\boldsymbol{x}^{(0)}$ near the center of Ω so the step size can be larger.

For simplicity, if $\boldsymbol{A} = [\frac{1}{n}, \dots, \frac{1}{n}] \in \mathbb{R}^{1\times n}$, and $\boldsymbol{b} = 1$, then we choose center $\boldsymbol{x}^{(0)} = \mathbf{1} := [1; \dots; 1] \in \mathbb{R}^n$.

If $\boldsymbol{x}^{(0)} = [x_1^{(0)}; \dots; x_n^{(0)}] \neq \mathbf{1}$, then we apply a diagonal scaling matrix

$$\boldsymbol{D}_0 = \mathrm{diag}(x_1^{(0)}, \dots, x_n^{(0)}) \in \mathbb{R}^{n\times n}$$

to obtain

$$\mathbf{1} = \boldsymbol{D}_0^{-1}\boldsymbol{x}^{(0)}$$

Therefore, as long as $\boldsymbol{x}^{(0)}$ is an interior point ($x_i^{(0)} \neq 0 \;\forall\, i$), we can apply such scaling to get $\mathbf{1}$ close to the center.

Now the LP problem becomes

$$\begin{array}{rl} \text{minimize} & \bar{\boldsymbol{c}}_0^\top \bar{\boldsymbol{x}} \\ \text{subject to} & \bar{\boldsymbol{A}}_0 \bar{\boldsymbol{x}} \geq \boldsymbol{b} \\ & \bar{\boldsymbol{x}} \geq \boldsymbol{0} \end{array}$$

where

$$\bar{\boldsymbol{c}}_0 = \boldsymbol{D}_0 \boldsymbol{c}, \quad \bar{\boldsymbol{A}}_0 = \boldsymbol{A} \boldsymbol{D}_0$$

So the orthogonal projector is

$$\bar{\boldsymbol{P}}_0 = \boldsymbol{I}_n - \bar{\boldsymbol{A}}_0^\top (\bar{\boldsymbol{A}}_0 \bar{\boldsymbol{A}}_0^\top)^{-1} \bar{\boldsymbol{A}}_0$$

and the projection of $-\bar{\boldsymbol{c}}_0$ onto the null space of $\bar{\boldsymbol{A}}_0$ is the descent direction

$$\bar{\boldsymbol{d}}^{(0)} = -\bar{\boldsymbol{P}}_0 \bar{\boldsymbol{c}}_0$$

The scaling idea presented above is applied in every iteration such that

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$$

where

$$\begin{aligned}
\boldsymbol{D}_k &= \mathrm{diag}(x_1^{(k)}, \dots, x_n^{(k)}) \\
\bar{\boldsymbol{A}}_k &= \boldsymbol{A}\boldsymbol{D}_k \\
\bar{\boldsymbol{P}}_k &= \boldsymbol{I}_n - \bar{\boldsymbol{A}}_k^\top (\bar{\boldsymbol{A}}_k \bar{\boldsymbol{A}}_k^\top)^{-1} \bar{\boldsymbol{A}}_k \\
\boldsymbol{d}^{(k)} &= -\boldsymbol{D}_k \bar{\boldsymbol{P}}_k \boldsymbol{D}_k \boldsymbol{c}
\end{aligned}$$

and the step size $\alpha_k = \alpha r_k$ such that $\alpha = 0.9$ or 0.99, and

$$r_k = \min_{i : d_i^{(k)} < 0} -\frac{x_i^{(k)}}{d_i^{(k)}}$$

so that one of the coordinates of $\boldsymbol{x}^{(k+1)}$ is close to 0.

So far we can run the affine scaling method if we were given an interior point as initial.

In practice, we need to solve an artificial problem to find such interior point. This is called the Phase I.

To obtain the artificial problem, we first select an arbitrary positive vector $\boldsymbol{u} > \mathbf{0}$, and check

$$\boldsymbol{v} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}$$

If $\boldsymbol{v} = \mathbf{0}$, then $\boldsymbol{u}$ is an interior point of $\Omega = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \mathbf{0}\}$, and we can just set $\boldsymbol{x}^{(0)} = \boldsymbol{u}$.

If $\boldsymbol{v} \neq \mathbf{0}$, then we introduce the following **artificial problem**:

$$\begin{aligned} \text{minimize} \quad & y \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} + \boldsymbol{v}y = \boldsymbol{b} \\ & \boldsymbol{x} \geq \mathbf{0}, y \geq 0 \end{aligned}$$

Then the artificial problem has a solution $y = 0$ iff Ω in the original problem is nonempty.

Note that $[\boldsymbol{u}; 1] > \mathbf{0}$ is an interior point of the feasible set of the artificial problem, since $\boldsymbol{A}\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{A}\boldsymbol{u} + (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{u}) = \boldsymbol{b}$, so we can use it as the initial and run the affine scaling algorithm. The result we get is $[\boldsymbol{x}; 0^{+}]$ where $\boldsymbol{x} > 0$ is an interior point of the feasible set Ω.

Unlike simplex method, the affine scaling method is an interior-point method and will not stop within finitely many steps. Therefore, we need to impose a stopping criterion, for example,

$$\frac{|\boldsymbol{c}^\top \boldsymbol{x}^{(k+1)} - \boldsymbol{c}^\top \boldsymbol{x}^{(k)}|}{\max\{1, |\boldsymbol{c}^\top \boldsymbol{x}^{(k)}|\}} < \varepsilon$$

for some prescribed $\varepsilon > 0$.

Karmarkar's method

First of all, Karmarkar's method requires to start with the **Karmarkar's canonical form**:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \\ & \boldsymbol{1}^\top \boldsymbol{x} = 1 \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

The last two constraints yields the set called **simplex**

$$\Delta := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1, x_i \geq 0 \right\}$$

We denote $\Omega' = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$. Then the feasible set is $\Omega = \Omega' \cap \Delta$, which is the intersection of the plane Ω containing $\boldsymbol{0}$ and the simplex Δ.

Example. Consider the LP problem

$$\begin{aligned} \text{minimize} \quad & 3x_1 + 3x_2 - x_3 \\ \text{subject to} \quad & 2x_1 - 3x_2 + x_3 = 0 \\ & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

We can see that the problem above is of the Karmarkar's canonical form with

$$\boldsymbol{c} = [3; 3; -1], \quad \boldsymbol{A} = [2, -3, 1]$$

It can be shown that any LP problem can be converted into an equivalent LP problem in Karmarkar's canonical form.

Karmarkar's algorithm also needs the following assumptions:

1. The center of the simplex, $\boldsymbol{a}_0 = \frac{1}{n}\mathbf{1}$, is feasible, i.e., $\boldsymbol{a}_0 \in \Omega'$;

2. The minimum value of the objective function over the feasible set is zero;

3. The matrix $\begin{bmatrix} \boldsymbol{A} \\ \mathbf{1}^\top \end{bmatrix} \in \mathbb{R}^{(m+1)\times n}$ has rank $m+1$;

4. The algorithm terminates when a feasible point $\boldsymbol{x}$ satisfies $\frac{\boldsymbol{c}^\top \boldsymbol{x}}{\boldsymbol{c}^\top \boldsymbol{a}_0} \le 2^{-q}$ for some prescribed $q > 0$.

The last three assumptions are fairly easy to hold.

Now we show how to convert an LP of standard form to the Karmarkar's canonical form (note that any LP can be converted to the standard form).

Recall the standard form of LP is

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$

which is equivalent to

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}'^\top \boldsymbol{z} \\ \text{subject to} \quad & \boldsymbol{A}'\boldsymbol{z} = \boldsymbol{0} \\ & \boldsymbol{z} \geq \boldsymbol{0} \end{aligned}$$

where $\boldsymbol{z} = [\boldsymbol{x}; 1] \in \mathbb{R}^{n+1}$, $\boldsymbol{A}' = [\boldsymbol{A}, -\boldsymbol{b}] \in \mathbb{R}^{m\times(n+1)}$, and $\boldsymbol{c}' = [\boldsymbol{c}; 0] \in \mathbb{R}^{n+1}$.

We need one more step to make the decision variables sum to 1. To this end, let $\boldsymbol{y} = [y_1, \dots, y_n, y_{n+1}]^\top \in \mathbb{R}^{n+1}$, where

$$
\begin{aligned}
y_i &= \frac{x_i}{x_1 + \cdots + x_n + 1}, \quad i = 1, \dots, n \\
y_{n+1} &= \frac{1}{x_1 + \cdots + x_n + 1}
\end{aligned}
$$

i.e., $\boldsymbol{y}$ is the projective transformation of $\boldsymbol{x}$. Note that we can easily get $x_i = \frac{y_i}{y_{n+1}}$ using $y_1, \dots, y_{n+1}$.

Now we have obtained the Karmarkar's canonical form in $\boldsymbol{y} \in \mathbb{R}^{n+1}$:

$$
\begin{aligned}
\text{minimize} \quad & \boldsymbol{c}'^{\top} \boldsymbol{y} \\
\text{subject to} \quad & \boldsymbol{A}' \boldsymbol{y} = \boldsymbol{0} \\
& \boldsymbol{1}^{\top} \boldsymbol{y} = 1 \\
& \boldsymbol{y} \geq \boldsymbol{0}
\end{aligned}
$$

As showed earlier, we can get $\boldsymbol{x}$ from the solution $\boldsymbol{y}$ to the above problem.

Recall that Karmarkar's method requires the simplex's center $\boldsymbol{a}_0$ to be feasible, i.e., $\boldsymbol{A}\boldsymbol{a}_0 = \boldsymbol{0}$. To this end, suppose we have an interior feasible point $\boldsymbol{a}$ (i.e., $\boldsymbol{A}\boldsymbol{a} = \boldsymbol{0}$ and $\boldsymbol{a} > \boldsymbol{0}$), and define the mapping

$$\boldsymbol{T}(\boldsymbol{x}) = [T_1(\boldsymbol{x}); \dots ; T_n(\boldsymbol{x}); T_{n+1}(\boldsymbol{x})] \in \boldsymbol{\Delta}$$

for any $\boldsymbol{x} \geq \boldsymbol{0}$ where

$$\begin{aligned} T_i(\boldsymbol{x}) &= \frac{x_i/a_i}{x_1/a_1 + \cdots + x_n/a_n + 1}, \quad i = 1, \dots, n \\ T_{n+1}(\boldsymbol{x}) &= \frac{1}{x_1/a_1 + \cdots + x_n/a_n + 1} \end{aligned}$$

So we can solve for $\boldsymbol{y} = \boldsymbol{T}(\boldsymbol{x})$ from the Karmarkar's canonical form, where the simplex's center $\boldsymbol{a}_0 = \boldsymbol{T}(\boldsymbol{a})$ is feasible, and then obtain $\boldsymbol{x} = \boldsymbol{T}^{-1}(\boldsymbol{y})$.

Now the question becomes that whether we have an interior point $\boldsymbol{a}$ to start.

Recall the symmetric primal-dual form of LP:

Primal	**Dual**
$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{aligned}$$	$$\begin{aligned} \text{maximize} \quad & \boldsymbol{b}^\top \boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{A}^\top \boldsymbol{\lambda} \leq \boldsymbol{c} \\ & \boldsymbol{\lambda} \geq \boldsymbol{0} \end{aligned}$$

which is equivalent to the primal-dual system

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{x} &= \boldsymbol{b}^\top \boldsymbol{\lambda} \\ \boldsymbol{A}\boldsymbol{x} &\geq \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{\lambda} &\leq \boldsymbol{c} \\ \boldsymbol{x} &\geq \boldsymbol{0} \\ \boldsymbol{\lambda} &\geq \boldsymbol{0} \end{aligned}$$

We introduce slack variable $\boldsymbol{u}$ and surplus variable $\boldsymbol{v}$ to convert the primal-dual system above into

$$\begin{aligned} \boldsymbol{c}^\top \boldsymbol{x} &= \boldsymbol{b}^\top \boldsymbol{\lambda} \\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{v} &= \boldsymbol{b} \\ \boldsymbol{A}^\top \boldsymbol{\lambda} + \boldsymbol{u} &= \boldsymbol{c} \\ \boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{u}, \boldsymbol{v} &\geq \boldsymbol{0} \end{aligned}$$

We know that a solution $[\boldsymbol{x}^*; \boldsymbol{\lambda}^*]$ of the system above gives the optimal primal variable $\boldsymbol{x}^*$ and dual variable $\boldsymbol{\lambda}^*$.

To convert the system above into an LP with easy-to-get initial feasible interior point, we choose arbitrary $\boldsymbol{x}^{(0)}, \boldsymbol{\lambda}_0, \boldsymbol{u}_0, \boldsymbol{v}_0 \geq \mathbf{0}$, e.g., $\boldsymbol{x}^{(0)} = \mathbf{1}_n$ etc.

Then we consider the following **Karmarkar's artificial problem**:

$$
\begin{aligned}
\text{minimize} \quad & z \\
\text{subject to} \quad & \boldsymbol{c}^\top \boldsymbol{x} - \boldsymbol{b}^\top \boldsymbol{\lambda} + (-\boldsymbol{c}^\top \boldsymbol{x}^{(0)} + \boldsymbol{b}^\top \boldsymbol{\lambda}_0) z = 0 \\
& \boldsymbol{A}\boldsymbol{x} - \boldsymbol{v} + (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{v}_0) z = \boldsymbol{b} \\
& \boldsymbol{A}^\top \boldsymbol{\lambda} + \boldsymbol{u} + (c - \boldsymbol{A}^\top \boldsymbol{\lambda}_0) z = \boldsymbol{c} \\
& \boldsymbol{x}, \boldsymbol{\lambda}, \boldsymbol{u}, \boldsymbol{v}, z \geq \mathbf{0}
\end{aligned}
$$

Note that $[\boldsymbol{x}; \boldsymbol{\lambda}; \boldsymbol{u}; \boldsymbol{v}; z] = [\boldsymbol{x}_0; \boldsymbol{\lambda}_0; \boldsymbol{u}_0; \boldsymbol{v}_0; 1] > \mathbf{0}$ is naturally an interior feasible point of the artificial problem above.

In addition, the original LP has a solution iff the artificial problem has a solution with $z^* = 0$.

Therefore, we can solve the Karmarkar's artificial problem (matrix form below) instead of the original LP:

$$\begin{aligned} \text{minimize} \quad & \tilde{\boldsymbol{c}}^\top \tilde{\boldsymbol{x}} \\ \text{subject to} \quad & \tilde{\boldsymbol{A}} \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{b}} \\ & \tilde{\boldsymbol{x}} \geq \boldsymbol{0} \end{aligned}$$

where

$$\begin{aligned} \tilde{\boldsymbol{x}} &= [\boldsymbol{x}; \boldsymbol{\lambda}; \boldsymbol{u}; \boldsymbol{v}; z] \in \mathbb{R}^{2m+2n+1} \\ \tilde{\boldsymbol{c}} &= [\boldsymbol{0}; \boldsymbol{0}; \boldsymbol{0}; \boldsymbol{0}; 1] \\ \tilde{\boldsymbol{A}} &= \begin{bmatrix} \boldsymbol{c}^\top & -\boldsymbol{b}^\top & \boldsymbol{0}_n^\top & \boldsymbol{0}_m^\top & (-\boldsymbol{c}^\top \boldsymbol{x}^{(0)} + \boldsymbol{b}^\top \boldsymbol{\lambda}_0) \\ \boldsymbol{A} & \boldsymbol{0}_{m\times m} & \boldsymbol{0}_{m\times n} & -\boldsymbol{I}_m & (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}_0 + \boldsymbol{v}_0) \\ \boldsymbol{0}_{n\times n} & \boldsymbol{A}^\top & \boldsymbol{I}_n & \boldsymbol{0}_{n\times m} & (\boldsymbol{c} - \boldsymbol{A}^\top \boldsymbol{\lambda}_0) \end{bmatrix} \\ \tilde{\boldsymbol{b}} &= [0; \boldsymbol{b}; \boldsymbol{c}] \in \mathbb{R}^{m+n+1} \end{aligned}$$

Up to this point, we can convert any LP into the Karmarkar's canonical form which also satisfies the four assumptions imposed earlier.

Now we consider how to solve the LP of Karmarkar's canonical form.

Recall the LP problem is given by

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x} \in \Omega' \cap \Delta \end{aligned}$$

where $\Omega' = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$ and $\Delta = \Big\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{1}^\top \boldsymbol{x} = 1, \boldsymbol{x} \geq \boldsymbol{0}\Big\}$.

Karmarkar's algorithm starts from an initial feasible point $\boldsymbol{x}^{(0)}$, and generates a sequence of iterates $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(k)}$ until $\boldsymbol{c}^\top \boldsymbol{x}^{(k)} / \boldsymbol{c}^\top \boldsymbol{x}^{(0)} < 2^{-q}$ for some prescribed $q > 0$.

Karmarkar's algorithm proceeds the following steps iteratively:

1. **Initialize**: Set $k := 0$; $\boldsymbol{x}^{(0)} = \boldsymbol{a}_0 := \frac{1}{n}\mathbf{1}$;

2. **Update**: Set $\boldsymbol{x}^{(k+1)} = \Psi(\boldsymbol{x}^{(k)})$ where Ψ is the update map (see below);

3. **Check stopping criterion**: If the condition $\boldsymbol{c}^\top \boldsymbol{x}^{(k)} / \boldsymbol{c}^\top \boldsymbol{x}^{(0)} < 2^{-q}$ is satisfied then stop; otherwise continue;

4. **Iterate**: Set $k \leftarrow k + 1$, go to 2.

To perform the update map Ψ, we consider the first step with $\boldsymbol{x}^{(0)} = \boldsymbol{a}_0$ and

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha \boldsymbol{d}^{(0)}$$

Recall that the search direction $\boldsymbol{d}^{(0)}$ is better to be the projection of $-\boldsymbol{c}$ onto the null space of $\boldsymbol{B}_0 := [\boldsymbol{A}; \mathbf{1}^\top] \in \mathbb{R}^{(m+1)\times n}$, which is the matrix in defining the equality constraints of the feasible set

$$\Omega = \Omega' \cap \Delta = \{\boldsymbol{x} : \boldsymbol{B}_0 \boldsymbol{x} = [\mathbf{0}; 1], \boldsymbol{x} \geq \mathbf{0}\}$$

Hence, we set $\boldsymbol{d}^{(0)}$ to

$$\boldsymbol{d}^{(0)} = -\frac{1}{\sqrt{n(n-1)}} \frac{\boldsymbol{P}_0 \boldsymbol{c}}{\|\boldsymbol{P}_0 \boldsymbol{c}\|}$$

where $\boldsymbol{P}_0 = \boldsymbol{I}_n - \boldsymbol{B}_0^\top (\boldsymbol{B}_0 \boldsymbol{B}_0^\top)^{-1} \boldsymbol{B}_0$ is the orthogonal projector onto the null space of $\boldsymbol{B}_0$.

A few remarks are in place:

- $\frac{1}{\sqrt{n(n-1)}}$ is the radius of the largest sphere inscribed in the simplex Δ.

- $\alpha \in (0, 1)$, e.g., $\alpha = 1/4$.

- $\boldsymbol{d}^{(0)}$ is in the null space of $\boldsymbol{B}_0$ and hence $\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha \boldsymbol{d}^{(0)}$ is still an interior point of $\Omega = \Omega' \cap \Delta$.

The update map Ψ for general $\boldsymbol{x}^{(k)}$ is very similar, except that $\boldsymbol{x}^{(k)}$ may not be the center of the simplex.

To overcome this issue, we employ the affine scaling idea: given $\boldsymbol{x}^{(k)}$, we define $\boldsymbol{D}_k = \mathrm{diag}(\boldsymbol{x}^{(k)})$ which is diagonal and invertible (since $\boldsymbol{x}^{(k)} > \mathbf{0}$).

Consider the mapping $\boldsymbol{U}_k : \Delta \to \Delta$ given by

$$\boldsymbol{U}_k(\boldsymbol{x}) = \frac{\boldsymbol{D}_k^{-1}\boldsymbol{x}}{\mathbf{1}^\top \boldsymbol{D}_k^{-1}\boldsymbol{x}}$$

Note that $\boldsymbol{U}_k$ is a one-to-one correspondence with inverse $\boldsymbol{U}_k^{-1}(\bar{\boldsymbol{x}}) = \frac{\boldsymbol{D}_k\bar{\boldsymbol{x}}}{\mathbf{1}^\top \boldsymbol{D}_k\bar{\boldsymbol{x}}}$ and $\boldsymbol{U}_k(\boldsymbol{x}^{(k)}) = \boldsymbol{a}_0 = \frac{\mathbf{1}}{n}$. Therefore we can perform the same update Ψ as we did for $\boldsymbol{x}^{(0)}$.

More specifically, we pretend that we are solving the following problem (but just for one iteration):

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^\top \boldsymbol{D}_k \bar{\boldsymbol{x}} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{D}_k \bar{\boldsymbol{x}} = \boldsymbol{0} \\ & \bar{\boldsymbol{x}} \in \boldsymbol{\Delta} \end{aligned}$$

where $\bar{\boldsymbol{x}} = \boldsymbol{U}_k(\boldsymbol{x})$.

We perform the update $\bar{\boldsymbol{x}}^{(k+1)} = \bar{\boldsymbol{x}}^{(k)} + \alpha \boldsymbol{d}^{(k)}$ where $\bar{\boldsymbol{x}}^{(k)} = \boldsymbol{a}_0$ and

$$\boldsymbol{d}^{(k)} = -\frac{1}{\sqrt{n(n-1)}} \frac{\boldsymbol{P}_k(\boldsymbol{D}_k \boldsymbol{c})}{\|\boldsymbol{P}_k(\boldsymbol{D}_k \boldsymbol{c})\|}$$

where $\boldsymbol{P}_k = \boldsymbol{I} - \boldsymbol{B}_k^\top (\boldsymbol{B}_k \boldsymbol{B}_k^\top)^{-1} \boldsymbol{B}_k$ and $\boldsymbol{B}_k = [\boldsymbol{A}\boldsymbol{D}_k; \boldsymbol{1}^\top]$. The step size $\alpha \in (0,1)$.

Then we compute $\boldsymbol{x}^{(k+1)} = \boldsymbol{U}_k^{-1}(\bar{\boldsymbol{x}}^{(k+1)}) = \frac{\boldsymbol{D}_k \bar{\boldsymbol{x}}^{(k+1)}}{\boldsymbol{1}^\top \boldsymbol{D}_k \bar{\boldsymbol{x}}^{(k+1)}} \in \Omega$.

The update map Ψ for general $\boldsymbol{x}^{(k)}$, i.e., $\boldsymbol{x}^{(k+1)} = \Psi(\boldsymbol{x}^{(k)})$, can be summarized below:

1. Compute the matrix $\boldsymbol{D}_k = \text{diag}(\boldsymbol{x}^{(k)})$ and $\boldsymbol{B}_k = [\boldsymbol{A}\boldsymbol{D}_k; \boldsymbol{1}^\top]$;

2. Compute search direction

$$\boldsymbol{d}^{(k)} = -\frac{1}{\sqrt{n(n-1)}} \frac{\boldsymbol{P}_k(\boldsymbol{D}_k\boldsymbol{c})}{\|\boldsymbol{P}_k(\boldsymbol{D}_k\boldsymbol{c})\|}$$

   where $\boldsymbol{P}_k = \boldsymbol{I} - \boldsymbol{B}_k^\top(\boldsymbol{B}_k\boldsymbol{B}_k^\top)^{-1}\boldsymbol{B}_k$.

3. Select $\alpha \in (0, 1)$ and perform update

$$\bar{\boldsymbol{x}}^{(k+1)} = \bar{\boldsymbol{x}}^{(k)} + \alpha\boldsymbol{d}^{(k)}$$

4. Compute $\boldsymbol{x}^{(k+1)} = \boldsymbol{U}_k^{-1}(\bar{\boldsymbol{x}}^{(k+1)}) = \frac{\boldsymbol{D}_k\bar{\boldsymbol{x}}^{(k+1)}}{\boldsymbol{1}^\top\boldsymbol{D}_k\bar{\boldsymbol{x}}^{(k+1)}}$.

Note that we do not need to explicitly write out the projector matrix $\boldsymbol{P}_k$ in Step 2 above.

Instead, we first solve $\boldsymbol{y}$ from the linear system $\boldsymbol{B}_k\boldsymbol{B}_k^\top\boldsymbol{y} = \boldsymbol{B}_k\boldsymbol{D}_k\boldsymbol{c}$ (so that $\boldsymbol{y} = (\boldsymbol{B}_k\boldsymbol{B}_k^\top)^{-1}\boldsymbol{B}_k\boldsymbol{D}_k\boldsymbol{c}$), and set

$$\boldsymbol{P}_k(\boldsymbol{D}_k\boldsymbol{c}) = \boldsymbol{D}_k\boldsymbol{c} - \boldsymbol{B}_k\boldsymbol{y}$$

Note the right-hand side is

$$\boldsymbol{D}_k\boldsymbol{c} - \boldsymbol{B}_k\boldsymbol{y} = \boldsymbol{D}_k\boldsymbol{c} - \boldsymbol{B}_k(\boldsymbol{B}_k\boldsymbol{B}_k^\top)^{-1}\boldsymbol{B}_k\boldsymbol{D}_k\boldsymbol{c} = (\boldsymbol{I} - \boldsymbol{B}_k(\boldsymbol{B}_k\boldsymbol{B}_k^\top)^{-1}\boldsymbol{B}_k)\boldsymbol{D}_k\boldsymbol{c}$$

which is exactly the projection we need.