A simple introduction to Karmarkar's Algorithm for Linear Programming

Sanjeev Saxena*
Dept. of Computer Science and Engineering,
Indian Institute of Technology,
Kanpur, INDIA-208 016

December 21, 2017

Abstract

An extremely simple, description of Karmarkar's algorithm with very few technical terms is given.

1 Introduction

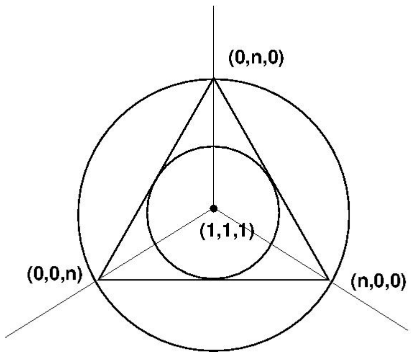
A simple description of Karmarkar's algorithm[5] together with analysis is given in this paper. Only knowledge of simple algebra, vector dot product and matrices is assumed. Even though the method is described in several books [8, 1, 2, 3, 7], analysis is either left out [8] or is fairly complicated. In this paper, it is shown that the description of Roos, Terlaky and Vial[6] can be further simplified. In addition to using only essential notation, a simpler proof of properties of "$\Psi(\ )$" function is given.

Let A be an $m \times n$ matrix of rank m and e is a vector of all ones. Karmarkar's Problem is: $\min c^T x$ subject to constraints $Ax = 0$, $e^T x = 1$ and $x \geq 0$.

Further, it is assumed that optimal value of $c^T x$ is zero and all ones vector is feasible, i.e., $Ae = 0$. We have to either find a point of cost 0 or show that none exist

We first scale the variables: $x' = xn$, then the problem becomes $\min c^T(x'/n)$ or equivalently $\min c^T(x')$ subject to $Ax' = 0$, $e^T x' = n$ and $x' \geq 0$. We will drop the primes, and the problem is[6]:

*E-mail: ssax@iitk.ac.in



$\min c^T x$ subject to constraints $Ax = 0$, $e^T x = n$ and $x \geq 0$.

REMARK Problem is trivial if $c^T e = 0$, hence we assume that $c^T e > 0$.

We will assume that all $m + 1$ equality constraints are linearly independent (else we eliminate redundant rows of A).

We need few definitions. **Standard simplex** consists of points in n dimensions s.t. $e^T x = n$, $x \geq 0$. The centre of the simplex is $e = (1, 1, \ldots, 1)$.

Let R be the radius of **outer sphere**, the smallest sphere containing the standard simplex (circumscribes standard simplex). R is the distance from e to one of the corner point (see Figure for an example in three dimensions) of the standard simplex (say, $(0, 0, \ldots, n)$), thus $R^2 = (0 - 1)^2 + \ldots + (0 - 1)^2 + (n - 1)^2 = (n - 1)^2 + (n - 1) = n(n - 1)$, or $R = \sqrt{n(n-1)}$.

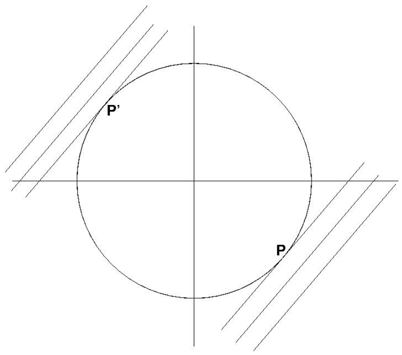
Let r be the radius of another sphere— the largest sphere centred at e and completely inside the standard simplex (inscribed inside the standard simplex). This sphere will be tangent to each face. Each face will have one coordinate as 0. By symmetry, all other coordinates at the point of contact will be same (say w). As the point of contact is on the standard simplex S, $0 + w + w + \ldots + w = n$ or $(n - 1)w = n$ or $w = n/(n - 1)$. Hence,

$$r^2 = (0 - 1)^2 + (w - 1)^2 + \ldots + (w - 1)^2 = 1 + (n - 1)(w - 1)^2 = 1 + \frac{1}{n - 1} = \frac{n}{n - 1}$$

Or

$$r = \sqrt{\frac{n}{n - 1}}$$

We take $e = (1, 1, \ldots, 1)$ as the starting point (which by assumption is feasible). Then, we minimise the objective function over a smaller sphere, which we will call the **inner sphere**, having same centre e, but radius αr, less than r (we will see in Section 5.2 that α can be chosen as $1/(r+1)$). Let us assume that the minimum occurs at point z. Then, in next iteration, we take the starting point as z. The problem of minimisation on sphere is discussed in next section (Section 2). Finally, the point z is mapped to e and the process repeated; the details of mapping are in Section 4. In



Section 3, it is shown that the objective value at next point z is a fraction of that at the initial point e.

2 Mathematical Preliminaries

Let P be a point in (possible hyper) plane and $\hat{n}$ a unit vector normal to it. If P_0 is any point in plane, then the (vector) dot product will be zero, i.e., $PP_0 \bullet \hat{n} = 0$. Writing in full,
$n_1(x_1 - x'_1) + n_2(x_2 - x'_2) + \ldots + n_k(x_k - x'_k) = 0$ or equivalently,
$n_1x_1 + n_2x_2 + \ldots + n_kx_k = n_1x'_1 + n_2x'_2 + \ldots + n_kx'_k = C$ Thus, for $a \bullet x = b$, normal will be in the direction of vector a.

Equation of sphere with centre Q as β is: $(x - \beta) \bullet (x - \beta) = r^2$ Let P be any point in plane $a \bullet x = b$, then we know that $a \bullet x_P = b$. If we want QP to be perpendicular to plane, then $(x_P - \beta) = (\text{constant})a$

Consider the problem [4]: $\min c^T x$ subject to $\sum(x - a)^2 \leq \rho^2$. There are two points on the sphere (see figure for an example in two dimensions), where planes parallel to given plane can be tangent— one corresponding to maxima and other minima. These are $+n$ and $-n$. These points (say "P") are in direction $\pm\hat{n}$ (a unit vector in direction of c) with length $|x_P - \beta| = r$. Thus, to minimise $c^T x$ over our sphere, we start at the centre β and take a step of length r in direction $-c^T$ Less informally, if $c = 0$, all points on the sphere are optimal. If $c \neq 0$, the solution is obtained by taking a step of length ρ (radius of sphere) from the centre a in the direction $-c$; this can be seen by considering "parallel" planes $c^T x$ =constant. The point at which minimum will be attained will be the point of contact on the sphere to a tangent plane.

Next [4] consider the problem: $\min c^T x$ subject to $Ax = b$ and $\sum(x - a)^2 \leq \rho^2$. If $c = 0$, all points common to (i.e., on the intersection of) sphere and the plane are optimal. If $c \neq 0$, let $\bar{c}$ be the orthogonal projection of c onto the plane $Ax = b$. If $\bar{c} = 0$, then c is a linear combination of rows of A and the objective function (on the intersection of sphere and plane) is constant, and all "feasible" points are also optimal. If $\bar{c} \neq 0$, then the solution is obtained by taking a step of length equal to the radius of lower-dimensional sphere (intersection of our sphere and plane $Ax = b$) from the centre a in the direction $-\bar{c}$.

3 Solution on inner Sphere

Let us consider the problem: $\min c^T x$ subject to $Ax = 0$, $\sum_i x_i = n$ and $\sum_i (x_i - 1)^2 < \alpha^2 r^2$, i.e, we are minimising only over points of the inner sphere. As the inner sphere is completely inside the standard simplex, we can drop the constraint $e^T x = n$ (and also $x \geq 0$), and the problem becomes $\min c^T x$ s.t. $Ax = 0$ and $\sum (x_i - 1)^2 < \alpha^2 r^2$.

As the point e is feasible, $Ae = 0$, e lies on the plane $Ax = 0$. Moreover, as e, the centre of inner sphere, lies on $Ax = 0$, the intersection of $Ax = 0$ and the inner sphere will be a "sphere" of the same radius αr but in lower dimension. (e.g., intersection of a sphere and a plane containing the centre is a circle with same radius and centre).

The minimum value of a linear function on sphere will be at a point (lower one) where linear function touches the sphere. The radius vector at that point will be perpendicular to the plane. Let us assume that $\hat{p}$ is a unit vector in that direction. Then the point at which minimum will be attained will be $e - \hat{p}$(radius). Thus, for outer sphere it will be (say) $z_R = e - \hat{p}R$ and for the inner sphere it will be $z_\alpha = e - \hat{p}\alpha r$. We will see how to determine $\hat{p}$ later (see Section 4.2).

As the outer sphere completely contains the solution space, the minimum value (of objective function $c^T x$) will be smaller than (or equal to) the actual optimal value which is zero. Thus, (assuming minimum value is at z_R)

$$0 \geq c^T z_R = c^T (e - \hat{p}R) \text{ or } c^T \hat{p} R \geq c^T e$$

Thus, $c^T \hat{p} \geq \frac{c^T e}{R}$

As the inner sphere is inside the simplex, value of the objective function can not be less than that on the simplex. But, as the optimal value on simplex is zero, optimal value on the inner sphere is non-negative. If the minimum value occurs at z_α on the inner sphere, then the value of objective function is:

$$0 \leq c^T z_\alpha = c^T (e - \hat{p}\alpha r) \text{ or } c^T \hat{p} \alpha r \leq c^T e$$

But as $\frac{c^T e}{R} \leq c^T \hat{p}$, we get

$$c^T z_\alpha = c^T e - c^T \hat{p} \alpha r \leq c^T e - \left(\frac{c^T e}{R} \right) \alpha r = c^T e \left(1 - \frac{\alpha r}{R} \right)$$

Thus, if we start with initial solution e, then z_α (the next solution) is an improvement (in value of objective function) by a factor of $(1 - \frac{\alpha r}{R})$ over the initial solution.

4 Karmarkar Transform and Algorithm

We will like to map the new point z again to e to repeat the process. Thus, for $a = (a_1, a_2, \ldots, a_n)$ and $x = (x_1, x_2, \ldots, x_n)$, we define the transform $y = T_a(x)$ with (for $i = 1, \ldots, n$)

$$y_i = n \frac{a_i x_i}{\sum_j a_j x_j}$$

If all a_is and x_is are non-negative (at least $a_i x_i$ should be non-zero), each component of transform will be less than n and sum of all coordinates will be n, thus the range is again our standard simplex. Moreover, if $b = a^{-1} = \left(\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_n} \right)$, then T_b will be the inverse transformation. Thus this transformation is one-to-one on our standard simplex.

Remark This can also be seen directly. If $T_a(x) = T_a(y)$, then $\frac{x_1}{y_1} = \frac{x_1}{y_2} = \ldots = \frac{x_n}{y_n} = r$ (say). But as $\sum x_i = \sum y_i = 1$, we have $r = 1$.

Moreover, if λ is any number then if $y' = T_a(\lambda x)$ then

$$y'_i = \frac{n \lambda a_i x_i}{\sum \lambda a_j x_j} = \frac{n \cancel{\lambda} a_i x_i}{\cancel{\lambda} \sum a_j x_j} = \frac{n a_i x_i}{\sum a_j x_j} = y_i$$

Thus, $T_a(\lambda x) = T_a(x)$.

4.1 Modified Problem

Assume that we are applying transform T_a (with $a_i = 1/z_i$), then point z will be mapped to e. Let D be a diagonal matrix with diagonal entries: $D = \text{diag}(a_1, \ldots, a_n)$.

If z is any feasible point, then for any positive $x > 0$ satisfying $e^T x = 1$, we saw that there is a unique point ξ (which depends on x) s.t. $x = T_z(\xi)$.

Remark $x = T_z(\xi)$ implies $x_i = z_i \xi_i / (\sum z_i \xi_i)$, thus $\xi_i = \rho \frac{x_i}{z_i}$ where ρ is such that $\sum \xi_i = 1$.

Then, equation $Ax = 0$ is equivalently to $\sum_j A_{ij}x_j = 0$ or $\sum_j A_{ij}\rho x_j = 0$ or equivalently, $\sum A_{ij}\xi_j z_j = 0$ or $A\xi z = 0$.

The objective function $c^T x = \sum c_i x_i = \frac{1}{\rho}\sum c_i\xi_i z_i$. As optimal value of $c^T x$ is zero, it follows that the optimal value of the transformed objective function $\sum c_i\xi_i z_i$ is also 0.

Replacing ξ by x' and using $Z = \text{diag}(z_1, \ldots, z_n)$, the transformed problem is:

$\min(Zc)^T x'$ subject to $AZx' = 0$, $e^T x' = n$ and $x' \geq 0$.

Moreover, as z is a feasible point $Az = 0$ or equivalently $AZe = 0$, thus e is again feasible.

We can thus repeat the previous method with Zc instead of c and AZ instead of A. In other words, we have to minimise the modified objective function over inscribed sphere of radius αr i.e., the "modified" problem is:

$\min(Zc)^T x$ subject to $AZx = 0$, $e^T x = n$ and $||x - e|| \leq \alpha r$.

4.2 Result from Algebra– Projection Matrix

Assume that A is $m \times n$ matrix, then rank of A is said to be m, with $m < n$ iff all m rows of A are linearly independent, i.e., $\beta_1 A_1 + \beta_2 A_2 + \ldots + \beta_m A_n = 0$ (here 0 is a row vector of size n) has only one solution $\beta_i = 0$. Thus, if v is any $1 \times m$ matrix (a column vector of size m), then $vA = 0$ implies $v = 0$. We use the fact that the matrix AA^T has rank m and is invertible[1].

Let A be an $m \times n$ matrix, then $Ax = b$ (hence $Ax = 0$) represents a set of m equations. Each equation will be a hyperplane in n dimensions. Let v be a vector of size n, we will like to "project" v onto the (lower dimension or intersection of) hyperplane $Ax = b$. If $P = pv$ is the projection, then [9], we wish to write $p = Pv$, as best as possible, as $p = Pv = \alpha_1 a_1 + \alpha_2 a_2 + \ldots + \alpha_m a_m$ where $a_1, \ldots, a_m$ are rows of A. Projection p will again be a vector of size n. This can be written as $p = A^T\alpha$.

The error of projection (or rejection) $E = v - p = v - Pv$ will also be a vector of size n. Then as error E can not be in these hyperplane, i.e., it should be "perpendicular" to these hyperplanes, thus we want $AE = 0$. Or $A(v - p) = 0$, or $A(v - A^T\alpha) = 0$, thus $Av = AA^T\alpha$, or $\alpha = (AA^T)^{-1}Av$.

Hence, $p = Pv = A^T\alpha$, or $Pv = A^T(AA^T)^{-1}Av$, or $P = A^T(AA^T)^{-1}A$. Or, $E = v - Pv =$

[1] As A is $m \times n$ matrix, A^T will be $n \times m$ matrix. The product AA^T will be an $m \times m$ square matrix. Let y^T be an $m \times 1$ matrix (or y is a row-vector of size m).

Consider the equation $(AA^T)y^T = 0$. Pre-multiplying by y we get $yAA^Ty^T = 0$ or $(yA)(yA)^T = 0$ or the dot product $< yA, yA >= 0$ which, for real vectors (matrices) means, that each term of yA is (individually) zero, or $yA = 0$, which implies y is identically zero. Thus, the matrix AA^T has rank m and is invertible.

$v - A^T(AA^T)^{-1}Av = \left(I - A^T(AA^T)^{-1}A\right) v$

Thus, to summarise, projection matrix is $P = A^T(AA^T)^{-1}A$ and Rejection matrix (to get the part perpendicular to the hyperplanes) is: $I - A^T(AA^T)^{-1}A$

4.3 Algorithm

Start the algorithm with $x = e$ and find z, the optimal value on the inscribed sphere of radius αr (we will see more details of this step later).

In a typical iteration, we next apply a transformation $T_{z^{-1}}$ to map z to e (we need a transformation T_b such that z gets mapped to e, then $b = 1/z$). Modify A and c and find a new z (say z'), the optimal value on the inscribed sphere of radius αr.

Remark As we know that transformed value of z is e, transform $T_{z^{-1}}$ is not actually applied, only the values of A and c are updated.

Then apply inverse transformation $T_{z^{-1}}^{-1}$ (or T_z) to map z' to original space, to get the value of x' for next iteration.

The formal algorithm is (value of α is fixed to $1/(r+1)$ in Section 5.2):

Initialise: $r = \sqrt{\frac{n}{n-1}}$, $\alpha = \frac{1}{r+1}$, $x = e = (1, \ldots, 1)$

Main step: If $c^T x < \epsilon$, return current x as solution of desired accuracy.

Let $D = \text{diag}(x_1, \ldots, x_n)$

$$P = \begin{pmatrix} AD \\ I \end{pmatrix}$$

Let

$$\begin{aligned}
c_P &= \left(1 - P^T \left(PP^T\right)^{-1} P\right)(cD)^T \\
\hat{p} &= \frac{c_P}{||c_P||} \\
z &= e - \alpha r \hat{p} \\
x &= \frac{nDy}{e^T Dy}
\end{aligned}$$

5 Analysis—Potential Function

For analysis, we define a potential function:

$$\Phi(x) = n \log c^T x - \sum_{i=1}^{n} \log x_i$$

As $e^T x = n$ we have $\frac{1}{n} x_i = 1$. But as Arithmetic mean is greater than or equal to geometric mean[2] $1 = \sum \frac{x_i}{n} \geq \left(\prod x_i\right)^{1/n}$ or taking logs, $\frac{1}{n}\sum \log x_i \leq 0$, or $\sum \log x_i \leq 0$, or $\Phi(x) \leq n \log c^T x$ or

$$\boxed{c^T x \geq \exp\left(\frac{\Phi(x)}{n}\right)}$$

Next observe that

$$\begin{aligned}
\Phi(\lambda x) &= n \log c^T(\lambda x) - \sum_{i=1}^{n} \log(\lambda x_i) \\
&= \left(\log c^T x + \log \lambda\right) - \left(\sum_{i=1}^{n} \log x_i + \sum_{i=1}^{n} \log \lambda\right) \\
&= n \log c^T x - \sum_{i=1}^{n} \log x_i = \Phi(x)
\end{aligned}$$

Let x be a positive vector in our standard simplex and let $y = T_x(z)$, then

$$y_i = \frac{n x_i z_i}{\sum x_i z_i}, \text{ or } \Phi(y) = \Phi(T_z(x)) = \Phi(xz)$$

Moreover, $\Phi(xz) = n \log c^T(xz) - \sum_{i=1}^{n} \log x_i z_i = n \log c^T xz - \sum \log x_i - \sum \log z_i$

Finally,

$$\begin{aligned}
\Delta\Phi &= \Phi(x) - \Phi(y) = \Phi(x) - \Phi(xz) \\
&= n \log c^T x - \cancel{\sum_{i=1}^{n} \log x_i} - \left(n \log c^T xz - \cancel{\sum \log x_i} - \sum \log z_i\right) \\
&= n \log \frac{c^T x}{c^T xz} + \sum \log z_i
\end{aligned}$$

But $\log \frac{c^T x}{c^T xz}$ is the ratio of original and transformed problems, which we saw reduces by at least $(1 - \alpha \frac{r}{R})$. Thus:

[2] This can be very easily seen by induction when n is a power of 2. For basis, as $xy = \frac{1}{2}\left((x+y)^2 - (x-y)^2\right)$, $\sqrt{xy} \leq \frac{x+y}{2}$. Assume that the claim is true till $n/2$. Then, $\left(\prod_{i=1}^{n} x_i\right)^{1/n} = \sqrt{\left(\prod_{i=1}^{n/2} x_i\right)^{2/n} \left(\prod_{i=1+n/2}^{n} x_i\right)^{2/n}} \leq \sqrt{\frac{2}{n}\sum_{i=1}^{n/2} x_i + \frac{2}{n}\sum_{i=1+n/2}^{n} x_i} \leq \frac{1}{2}\left(\frac{2}{n}\sum_{i=1}^{n/2} x_i + \frac{2}{n}\sum_{i=1+n/2}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n/2} x_i$

$$\Delta\Phi \geq -n \log\left(1 - \alpha\frac{r}{R}\right) + \sum \log z_i$$

We will show that Δ is more than a constant. But, before continuing with analysis, we need some more results from algebra.

5.1 Function Ψ

In this section, all logs are to base e. Assume that $|x| < 1$. Recall[3]
$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$ Thus, for $-1 < y < 0$, (let $y = -x$) we have
$\log(1+y) = \log(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \ldots$

Or $y - \log(1+y) = -x - \log(1-x) = \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots > 0$

We define[6] $\Psi(x) = x - \log(1+x)$. Observe that $\Psi(0) = 0$ Hence, $\Psi(x) = \frac{x^2}{2} - \frac{x^3}{3} + \frac{x^4}{4} - \ldots$ And, $\Psi(x) \geq 0$ for $x > -1$

Claim 1 *If $x > 0$ and $a^2 = b^2 + c^2$, then if ax, bx, cx are greater than -1 (for $\log(1+hx)$ to be defined), then*

$$\Psi(-|a|x) > \Psi(bx) + \Psi(cx)$$

Proof: From $a^2 = b^2 + c^2$, we know that $|a| \geq |b|$ and $|a| \geq |c|$. Without loss of generality, assume $a < 0$, then $-|ax| = ax$ and we have

$\Psi(ax) = \frac{a^2x^2}{2} - \frac{a^3x^3}{3} + \frac{a^4x^4}{4} - \ldots$

$\Psi(bx) = \frac{b^2x^2}{2} - \frac{b^3x^3}{3} + \frac{b^4x^4}{4} - \ldots$

$\Psi(cx) = \frac{c^2x^2}{2} - \frac{c^3x^3}{3} + \frac{c^4x^4}{4} - \ldots$

As $a^2 = b^2 + c^2$, coefficients of x^2 on both sides are equal.

We thus need to compare $-\frac{a^3x^3}{3}$ with $-\frac{(b^3+c^3)x^3}{3}$. As $a^2 = b^2 + c^2$, $a^3 = ab^2 + ac^2$. Thus, we are comparing $-x(ab^2 + ac^2)$ and $-x(b^3 + c^3)$ or equivalently 0 and $xb^2(a-b) + xc^2(a-c)$.

As $a < 0$ and $-a > |b|$ and $-a > |c|$, the expressions can be re-written as: 0 and $-xb^2(b-a) - xc^2(c-a)$. Both terms inside brackets are positive, and as $x > 0$, right hand side will be negative. ∎

Corollary 1 *If $a^2 = b^2 + c^2 + \ldots + s^2$, then if $ax, bx, cx, \ldots$ are greater than -1 (for $\log(1+hx)$ to be defined), then if $x > 0$, $\Psi(-|a|x) > \Psi(bx) + \Psi(cx) + \ldots + \Psi(sx)$*

[3] As, $\frac{1}{1+x} = 1 - x + x^2 - x^3 + x^4 - x^5 + \ldots$, integrating both sides (from 0 to x) we get $\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$

Proof: (induction) Let $v^2 = c^2 + \ldots + s^2$. Then $a^2 = b^2 + v^2$. Then from the above claim, $\Psi(-|a|x) > \Psi(bx) + \Psi(-|v|x)$. By induction hypothesis, $\Psi(-|v|x) > \Psi(cx) + \ldots + \Psi(sx)$ Proof follows by adding the two inequalities. ∎

5.2 Analysis Continued

As $-\log(1-x) = x + \frac{x^2}{2} + \frac{x^3}{3} + \ldots > x$, hence,

$$\begin{aligned} \Delta\Phi &\geq -n\log\left(1 - \alpha\frac{r}{R}\right) + \sum \log z_i \\ &= n\alpha\frac{r}{R} + \sum \log z_i \\ &= \alpha r^2 + \sum \log z_i \end{aligned}$$

But $z = e - \alpha r\hat{p}$ or $z_i = 1 - \alpha r p_i$

And, $\log z_i = \log(1 - \alpha r p_i) = -(-\alpha r p_i - \log(1 - \alpha r p_i)) - \alpha r p_i = -\alpha r p_i - \Psi(-\alpha r p_i)$.

But $\hat{p}$ is a unit vector, $\sum p_i^2 = 1$. As z is feasible $\sum z_i = n$, or $\sum p_i = 0$. Hence,

$$-\sum \log z_i = \sum \alpha r p_i + \sum \Psi(-\alpha r p_i) = \sum \Psi(-\alpha r p_i) \leq \Psi(-\alpha r)$$

The last inequality follows from Corollary 1 above.

For $\Psi(-\alpha r)$ to be defined, $\alpha r < 1$, we thus[4] choose $\alpha = \frac{1}{r+1}$, we have

$$\begin{aligned} \Delta\Phi &\geq \alpha r^2 - \Psi(-\alpha r) = \alpha r^2 - ((-\alpha r) - \log(1 - \alpha r)) \\ &= \alpha r^2 + \alpha r + \log(1 - \alpha r) \\ &= \frac{r^2}{1+r} + \frac{r}{1+r} + \log\left(1 - \frac{r}{r+1}\right) \\ &= r + \log\left(1 - \frac{r}{r+1}\right) \\ &= r - \log(1+r) = \Psi(r) \text{ but as } r = \sqrt{n/n-1} > 1 \\ &\geq \Psi(1) = 1 - \ln 2 \approx 0.3 \end{aligned}$$

Hence, potential decreases by a fixed amount after each iteration.

After k, iterations, $\Phi(e) - \Phi(x) > k\Psi(1)$. But as $\Psi(e) = n\log c^T e$, we get $\Phi(x) < n\log c^T e - k\Psi(1)$. As x is inside the standard simplex,

$$c^T x \leq \exp\left(\frac{\Phi(x)}{n}\right) < \exp\left(\frac{n\log c^T e - k\Psi(1)}{n}\right)$$

[4] Differentiating, $f(\alpha) = \alpha r^2 - \Psi(-\alpha r)$ w.r.t. α, $f'(\alpha) = r^2 - \frac{-\alpha r}{1-\alpha r} + n\frac{-r/R}{1-(\alpha r/R)}$ $f'(\alpha) = r^2 + r + (-r)\frac{1}{1-\alpha r}$, equating $f'(\alpha)$ to 0, we get $1 + r = \frac{1}{1-\alpha r}$, or $1 - \alpha r = \frac{1}{1+r}$ or $\alpha r = 1 - \frac{1}{1+r} = \frac{r}{r+1}$ or $\alpha = \frac{1}{1+r}$. Thus, the maximum value is at $\alpha = \frac{1}{1+r}$.

If we stop as soon as error $c^T x \leq \epsilon$, we get

$$\begin{aligned}
\exp\left(\frac{n \log c^T e - k\Psi(1)}{n}\right) &\leq \epsilon \text{ or} \\
\frac{n \log c^T e - k\Psi(1)}{n} &\leq \log \epsilon \text{ or} \\
\log c^T e - k\Psi(1) &\leq n \log \epsilon \text{ or} \\
k &\geq \frac{n}{\Psi(1)} \log \frac{c^T e}{\epsilon}
\end{aligned}$$

Thus, after at most $\frac{n}{\Psi(1)} \log \frac{c^T e}{\epsilon}$ iterations, algorithm finds a feasible point x for which $c^T x \leq \epsilon$

Acknowledgments

Many thanks to students of CS647A (2017-2018 batch) for their valuable feedback, questions and comments on an earlier version.

References

[1] Mokhtar S. Bazaraa, John J. Jarvis, Hanif D. Sherali, Linear Programming and Network Flows, 4th Edition, Wiley 2009

[2] Donald Goldfarb and Michael J.Todd, Linear programming, Chapter II in Optimization Edited by G.L. Nemhauser, A.H.G. Rinooy Kan and M.J. Todd, Handbooks in Operations Research and Management Science Volume 1, 1989, Pages 73-170, Elsevier

[3] H.Karloff, Linear Programming, Birkhauser, 1991.

[4] Katta G. Murty, chapter 11, pp 472-473 of "Linear Complementarity, Linear and Nonlinear Programming", web book `http://www-personal.umich.edu/~murty/books/linear_complementarity_webbook/kat11.pdf`

[5] Narendra Karmarkar, A New Polynomial Time Algorithm for Linear Programming, Combinatorica, Vol 4, nr. 4, pp. 373-395 (1984).

[6] C.Roos,T Terlaky and J-P Vial, Interior Point Methods for Linear Optimization, 2nd Ed, 2006, Springer.

[7] R.Saigal, Linear Programming, A Modern Integrated Analysis, Kluwer, 1995.

[8] Hamdy A. Taha, Operations Research: An Introduction, 7th Edition, Pearson, 2002.

[9] Ling-Hsiao Lyu, Lecture notes on Linear Algebra Lecture 3 (Chap. 4), Projection and Projection Matrix, Institute of Space Science, National Central University, Chung-Li, Taiwan, R. O. C., 2012 Spring `http://www.ss.ncu.edu.tw/~lyu/lecture_files_en/lyu_LA_Notes/Lyu_LA_2012/Lyu_LA_3_2012.pdf`