11

SOME REMARKS ON PROJECTIVE ALGORITHM

IN

LINEAR OPTIMIZATION

GRADUATE SEMINAR REPORT

BY

HALGEYO JILLO

ADVISOR

PROF. DR. RER. NAT. HABIL. R. DEUMLICH

SCHOOL OF GRADUATE STUDIES
ADDIS ABABA UNIVERSITY

June 2003
Addis Ababa

Preface

Although the origin of the linear optimization as a mathematical discipline is quite recent, it is now well established as an important and very active branch of applied mathematics. The wide applicability of linear optimization models and the rich mathematical theory underlying these models and the methods developed to solve them have been the driving forces behind the rapid and continuing evolution of the subject. The recognition of the importance of linear optimization models, especially in the areas of economic analysis and planning, coincide with the development of both an effective method, the 'simplex method' of G. B. Dantzig, for solving linear optimization problems, and a means, the digital computer, for doing so.

All forms of the simplex algorithm reach the optimum by transversing a series of basic solutions. Since each basic solution represents an extreme point of the feasible region, the track followed by the algorithm moves around the boundary of the feasible region. In the worst case, it may be necessary to examine most of, if not all, the extreme points. The worst case here just simply means a worst possible data configuration that would take the simplex algorithm run many steps. This can be cripplingly inefficient given that the number of extreme points grows exponentially with n for fixed m.

The existence of such worst cases suggests that one should look possibly for algorithms other than the simplex algorithms, which are good in such cases.

The first success was attributed to the Russian mathematician, Khachian (1979), who proposed the *ellipsoid algorithm* or sometimes the *Russian Method.* Though theoretically efficient, code developers were never able to realize an implementation that matched the performance of concurrent simplex codes.

Just about the time when the interest in the ellipsoid method was waning, a new technique to solve linear optimization problems was proposed by N. K. Karmarkar (1984) of AT & T Bell Laboratories, namely, the *projective algorithm* or more generally *interior point algorithm.*

Contrary to the simplex algorithm, the projective algorithm moves in the relative interior rather than on the outside of the polyhedron in the n- dimensional space.

The purpose of this paper is to describe the development of this new algorithm and the idea (coming from non-linear optimization and projective geometry) on which it is based. Here we present the original algorithm, which is the approximation to the underlying algorithmic idea because of linearization of what is inherently a non-linear problem due to a projective transformation.

I want to express my deepest gratitude to my adviser Prof. Dr. rer. nat. habil. R. Deumlich who cultivated me highly not only academically but also socially. His constructive comments, advice and material provision are extremely invaluable.

I am especially thankful to my wife, Giloya Lemita, who made my life pleasurable and shouldered all the family responsibilities throughout my study. My heartfelt thanks also goes to my brothers Guyo Jillo and Halege Jillo, and others I did not mention, who, regardless of situations, untiringly have given me a lot of encouragement throughout my study.

Finally, I am grateful to all my friends for their unreserved moral and material support and made my study fruitful. Among the many, Mengistu Goa, Tesfa Yigrem and Workaferahu Seyoum are worth mentioning.

Halgeyo Jillo

Table of Contents

Contents
Pages

1. Mathematical Pre-Requisites

For the proper understanding of the linear optimization problem we require knowledge of some elementary concepts of mathematical analysis as well as linear algebra. In this section we shall summarize the notions and some mathematical tools, which will be needed, for our purpose.

1.1.Euclidean Space

Definition 1.1.1: The n- dimensional (Euclidean) space, symbolized by R^n is defined as the set of all (column) vectors x having n scalar components.

The partial order in the set of all vectors of R^n will be given in the following

Definition 1.1.2: For any two vectors $x \in R^n$ and y $\in$ R^n we define

(i) $x > y$ *if and only if* $x_j > y_j$ *for all* $j \in \{1,\ldots,n\}$

(ii) $x \geq y$ *if and only if* $x_j \geq y_j$ *for all* $j \in \{1,\ldots,n\}$

(iii) $x = y$ *if and only if* $x_j = y_j$ *for all* $j \in \{1,\ldots,n\}$

(iv) $x \geq 0$ *if and only if* $x_j \geq 0$ *for all* $j \in \{1,\ldots,n\}$

Remark:

1. The above definition is also valid if we replace > and $\geq$ by < or $\leq$, respectively.
2. By $x \ngeq y$ we mean there exist an index $j \in \{1,..n\}$ such that $x_j < y_j$.

Definition 1.1.3: For any $x, y \in R^n$, the inner (scalar) product of x and y is given by

$$x^T y = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n} y_i x_i = y^T x$$

We know that we can define different norms on R^n. Among these norms the most used norm is the Euclidean norm, which is defined as follows:

Definition 1.1.4: The real number

$$\|x\| = \sqrt{x^T x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

is called the Euclidean norm of the vector $x \in R^n$.

For our purpose, by the norm of a vector always we mean the Euclidean norm.

Definition 1.1.5:

(i) The rank of a set $S \subseteq R^n$ is the cardinality of the largest subset of the linearly independent vectors in S.

(ii) The rank of a matrix A is the rank of its column vectors (which is equal to the rank of its row vectors).

The rank of a matrix A with size m x n satisfies

$$r(A) \leq \min\{m, n\}$$

where r(A) stands for "rank of A".

We say A is of **full rank** if $r(A) = \min\{m, n\}$.

Definition 1.1.6: A real symmetric matrix A of order n is said to be positive definite if and only if

$x^T Ax > 0$ for all $x \in R^n, x \neq 0$.

Remark:

A positive definite matrix is necessarily non-singular, for if a matrix is singular there is an $x \neq 0$ for which Ax = 0 and hence $x^T Ax = 0$.Since A is positive definite this is a contradiction .

Lemma 1.1: Let $A = (a_1, a_2, ..., a_t)$ be an n×t matrix where the vectors a_i *for* $i = 1,...,t$ are linearly independent. Then AA^T is non singular.

Proof: By the above remark it suffices to show that AA^T is positive definite. To this end let $y \in R^n, y \neq 0$ and $A^T Ay = 0$.

Then

$$y^T (A^T A)y = (Ay)^T Ay = \langle Ay, Ay \rangle = \|Ay\|^2 = 0.$$

This implies $Ay = 0$ *or* $\sum_{i=1}^{t} a_i y^i = 0$.

But this is impossible as the vectors a_i are linearly independent and y is a non zero vector. Thus, we have

$$y^T (A^T A) y > 0 \quad \textit{for all } y \in R^t, y \neq 0.$$

Hence, AA^T is positive definite by definition. //

Remark:

1. The matrix A in Lemma 1.1 is of full row rank. Hence it follows that for any matrix A of full rank the matrix AA^T is positive definite.
2. If A is positive definite, then so is A^{-1}.

1.2. Rank- One Update

Theorem 1.2: Let B be a non-singular n× n matrix and $u, v \in R^n$ be any two vectors such that $v^T B^{-1} u \neq -1$. Then

$$\left(B + uv^T\right)^{-1} = B^{-1} - \frac{1}{1 + v^T B^{-1} u}(B^{-1}u)(v^T B^{-1}).$$

Proof: Since $v^T B^{-1} u \neq -1$, the expression on the right hand side is well defined.

If we set $\omega := \dfrac{1}{1 + v^T B^{-1} u}$, then we get $1 - \omega = \omega(v^T B^{-1} u)$ and since the dyadic product $u v^T$ is an n×n matrix we have that

$$\begin{aligned}
&(B + uv^T)(B^{-1} - \omega(B^{-1}u)(v^T B^{-1})) \\
&= I_n - \omega u(v^T B^{-1}) + u(v^T B^{-1}) - \omega uv^T (B^{-1}u)(v^T B^{-1}) \\
&= I_n - \omega u(v^T B^{-1}) + u(v^T B^{-1}) - \omega (v^T B^{-1} u) u(v^T B^{-1}) \\
&= I_n + [-\omega + 1 - \omega (v^T B^{-1} u)] u(v^T B^{-1}) \\
&= I_n
\end{aligned}$$

Hence the theorem follows. //

1.3. Geometric / Arithmetic Inequality

For any x$\in R^n$, let $m(x)=\left(\frac{1}{n}\right)\sum_{i=1}^{n} x_i$ be the arithmetic mean of x and

for every x$\in R^n$ with $x>0$, let $g(x)=\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$ be the geometric mean

of x. We state without proof the following important inequality, known as Geometric/arithmetic inequality, which we shall use very often.

Theorem 1.3: $\left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}} \le \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)$ for all x$\in R^n$, $x>0$.

1.4. Orthogonal Projections

Definition 1.4.1: Let X be a vector space. An endomorphism $P: X \to X$ is called a projection if and only if

$$P^2 = P$$

We give some important properties of projection.

Property 1: Let P: X→ X be a projection and let Q=I-P. Then

(i) $Q = I - P$ is a projection.

(ii) P.Q =Q.P = 0 (where 0 is the zero map; 0x= 0 for all x)

Proof:

(i) $Q^2 = (I-P)^2 = I^2 - 2I.P + P^2 = I - 2P + P = I - P = Q$

(ii) $P.Q = P.(I-P) = P.I - P.P = P - P^2 = P - P = 0$ and

$Q.P = (I-P).P = I.P - P.P = P - P^2 = P - P = 0$

By means of scalar product in R^n we define the notion of orthogonality as below.

Definition 1.4.2: Let $U \subseteq R^n$. Then

(i) $x, y \in R^n$ are said to be orthogonal if and only if

$$\langle x, y\rangle = y^T x = 0.$$

(ii) $U^{\perp} := \{y \in R^n : \langle x, y\rangle = 0 \text{ for all } x \in U\}$ is called the orthogonal complement of U.

Definition 1.4.3: Let $R^n = U \oplus U^{\perp}$. The projection $P: R^n \to U$ along $U^{\perp}$ is said to be the **orthogonal projection** of R^n on to U.

Since $R^n = U \oplus U^{\perp}$, every $x \in R^n$ can be written uniquely as $x = y + z,\ where\ y \in U\ and\ z \in U^{\perp}$. Accordingly, the above definition says, if $x = y + z,\ where\ y \in U\ and\ z \in U^{\perp}$, then the vector y is the orthogonal projection of the vector x on the subspace U.

Now we give the general approach to find the orthogonal projection of a vector onto a given subspace.

Suppose we are given any subspace U with basis $\{a_1, a_2, ..., a_t\}$ and we want to find an orthogonal projection on to U.
Define the n×t matrix A=$(a_1, a_2, ..., a_t)$. Then a typical element of U is of the form Ax for some $x \in R^t$.
Let $\mathbf{a} \in R^n$. We want to find the vector $Ax \in$U which is the orthogonal projection of **a** onto U.
From $R^n = U \oplus U^{\perp}$we get $\mathbf{a} = Ax + z,\ where\ z \in U^{\perp}$.
Since $a - Ax \in U^{\perp}$, by definition, $a - Ax$ is orthogonal to every member of U. This is the case if and only if $a - Ax$ is orthogonal to each of the basis vector of U.
Thus,

$$\langle a_i, a - Ax \rangle = a_i^T (a - Ax) = 0,\ for\ i = 1, ..., t.$$

which is equivalent to

$$A^T (a - Ax) = 0 \tag{1}$$

Solving for x from (1) and noting that $(AA^T)^{-1}$exists by lemma 1.1 we get

$$x = (A^T A)^{-1} A^T a \tag{2}$$

Let the operator P be defined by

$$\text{P:} = A(A^T A)^{-1} A^T$$

Then we have using (2) the orthogonal projection of a vector **a** on to U to be

$$Pa = Ax = A(A^T A)^{-1} A^T a$$

We call the operator P the orthogonal projection operator onto the subspace U.

The operator P has the following interesting properties.

Property 2:

(i) P is a projection, i.e. $P^2 = P$

(ii) P is a symmetric matrix, i.e. $P^T = P$

(iii) Pa lies in the subspace spanned by the vectors a_i for $i = 1,\ldots,t$,i.e. $Pa \in U$.

Proof:

$$(i)\ \ P^2 = P.P = \left(A(A^T A)^{-1} A^T\right).\left(A(A^T A)^{-1} A^T\right)$$
$$= \left(AA^{-1}(A^T)^{-1} A^T\right).\left(A(A^T A)^{-1} A^T\right)$$
$$= I.\left(A(A^T A)^{-1} A^T\right) = I.P = P$$

$$(ii)\ \ P^T = \left(A(A^T A)^{-1} A^T\right)^T = \left(A^T\right)^T \left((A^T A)^{-1}\right)^T A^T$$
$$= \left(A^T\right)^T \left((A^T A)^T\right)^{-1} A^T = A\left((A^T A)^{-1}\right)A^T = P$$

$$(iii)\ \ Pa = A\left(A^T A\right)^{-1} A^T a$$
$$= Au \quad where \quad u = (A^T A)^{-1} A^T a$$
$$= \sum_{i=1}^{t} a_i u^i$$

This means Pa is a linear combination of the vectors a_i. Hence Pa lies in the subspace spanned by the vectors a_i for $i = 1,\ldots,t$ .//

In the following examples we drive the formulas for the orthogonal projection and their respective orthogonal operators for some special subspaces of R^n which we need very often.

Example 1.4.1: Let $U = \{z \in R^n : Az = 0\}$ be a subspace of R^n where A is an $m \times n$ matrix of rank m. We want to find the orthogonal projection of a vector $c \in R^n$ on to U. Here we follow the above general approach.

Let p be the orthogonal projection of vector c on to U.

Obviously, $Ap = 0$ *since* $p \in U$ by property 2 (iii).

Let $A^T = (a_1, a_2, \ldots, a_n)$ where a_i are the row vectors of A.

Then $c - p \perp U$, i.e. $c - p$ is a linear combination of all normal vectors to the hyperplanes $H_i = \{z \in R^n : \langle a_i, z\rangle = 0\}$, i.e. of a_i's.

Consequently,
$c - p = \lambda_1 a_1 + ... + \lambda_n a_n = A^T \lambda \quad where \ \lambda \in R^n$
or
$$c - p = A^T \lambda \tag{3}$$
Multiplying (3) by A and noting that $Ap = 0$ we get

$$Ac = AA^T \lambda \tag{4}$$
By the remark following Lemma 1.1, AA^T is nonsingular.
Solving for λ we get from (4) that
$\lambda = (AA^T)^{-1} Ac$
Substituting this value of λ we get from (3) that

$p = \lfloor I_n - A^T (AA^T)^{-1} A \rfloor c$
or
$p = Pc \quad where \ P = I_n - A^T (AA^T)^{-1} A$ is the projection operator.

Example 1.4.2: Let $L = \{ z \in R^n : Az = 0, e^T z = 0 \}$ where A is an $m \times n$ matrix of rank m and $e^T = (1,...,1)$ is the n- vector of ones.

We want to find the orthogonal projection of a vector $\mathbf{c} \in R^n$ on to the subspace *L*.

Let q be the orthogonal projection of a vector $\mathbf{c} \in R^n$ on to the subspace *L*.

Obviously, $Aq = 0 \ and \ e^T z = 0$.

Let A^T be as given in example 1.4.1.
Then we have
$$c - q = A^T \lambda + \mu e \quad where \ \lambda, \mu \in R^n . \tag{5}$$
Multiplying (5) by A we get
$Ac = AA^T \lambda$,
where we have used the fact that $Aq = 0 \ and \ \mu Ae = 0$

By the remark following Lemma 1.1 $(AA^T)^{-1}$ exists and so we have
$\lambda = (AA^T)^{-1} Ac$

Again multiplying (5) by e^T we get using $Ae = 0 \ and \ e^T q = 0$ that
$e^T c = \mu e^T e$
or

$$\mu = \frac{1}{n} ee^T .$$

Substituting the values of λ and μ into (5) we get

$$q = \left[I_n - A^T (AA^T)^{-1} A - \frac{1}{n} ee^T \right] c$$

or in short

$q = Qc$ *where* $Q = I_n - A^T (AA^T)^{-1} A - \frac{1}{n} ee^T$ is the projection operator.

2. Linear Optimization Problem: Basic Concepts

Consider the linear optimization problem in standard form

(LP) $\quad \min \ \{ \mathbf{cx} : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \},$

where A is an $m \times n$ matrix
$c \in R^n$ is a row vector and
$b \in R^n$ is a column vector .

We will assume w. l. o. g. that A is of full row rank.

Definition 2.1: The set $\mathcal{X} = \{ \mathbf{x} \in \mathbf{R^n} : \mathbf{Ax} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0} \}$ is called the feasible set of (LP).

Definition 2.2:

(i) Each $x \in \mathcal{X}$ *is* called a feasible solution.

(ii) A feasible solution $\bar{x} \in \mathcal{X}$ is called an optimal solution if $c\bar{x} \leq cx \ for \ all \ \ x \in \mathcal{X}$.

2.1. Properties of the Feasible Set of (LP)

In this section we examine some of the properties of the feasible set of (LP) .To begin with, let us introduce the following definition.

Definition 2.1.1: A set $C \subseteq R^n$ is said to be **convex** if and only if
$\alpha x + (1-\alpha) y \in C$ *for every* $x, y \in C$ *and* $\alpha \in [0,1]$.

It means geometrically, the line segment
$[x,y] = \{ z \in R^n : z = \alpha x + (1-\alpha) y, \ \alpha \in [0,1]$
is entirely contained in C whenever its end points x and y are in C.

Definition 2.1.1: Let $x_1, x_2, \ldots, x_k \in R^n, \lambda_1, \lambda_2, \ldots, \lambda_k \in R$. Let
$z = \sum_{i=1}^{k} \lambda_i x_i$, *where* $\sum_{i=1}^{k} \lambda_i = 1$, $\lambda_i \geq 0$. Then
z is said to be a **convex combination** of the points $x_1, x_2, \ldots, x_k$.

It follows from the above definition that the set of all convex combinations of a finite number of points $x_1, x_2, \ldots, x_n$ is a convex set.

Lemma 2.1: χ is a convex subset of R^n.

Proof: Let $x_1, x_2 \in \chi$. Then by definition

$$x_1 \geq 0,\ x_2 \geq 0; \qquad (i)$$

and $$Ax_1 = b, Ax_2 = b \qquad (ii)$$

Let x be any convex combination of x_1 *and* x_2. Then

$$x = \lambda x_1 + (1-\lambda)x_2\ ,\ \lambda \in [0,1].$$

Obviously, $x \geq 0$ from (i).
Furthermore, using (ii) we get

$$Ax = A[\lambda x_1 + (1-\lambda)x_2] = \lambda A x_1 + (1-\lambda)Ax_2 = \lambda b + (1-\lambda)b = b.$$

This implies $x \in \chi$. Since *the* convex combination of every two feasible solution is in χ, it follows that χ is a convex set. //

Definition 2.1.2: Let $c \subseteq R^n, c \neq 0, \alpha \in R$. The set

$$\{x \in R^n : cx = \alpha\}$$

is called a hyperplane.
The sets $\{x \in R^n : cx < \alpha\}$ and $\{x \in R^n : cx > \alpha\}$ are called open halfspaces.
The sets $\{x \in R^n : cx \leq \alpha\}$ and $\{x \in R^n : cx \geq \alpha\}$ are called closed halfspaces.
Obviously, hyperplanes and halfspaces of R^n are convex sets. An open halfspace is naturally an open set being the interior of the corresponding closed halfspace and a closed halfspace is a closed set being the closure of the corresponding open halfspace. Clearly the boundary of a halfspace is a hyperplane.

Theorem 2.1: χ is a closed convex subset of R^n.

Proof: The convexity of χ follows immediately from Lemma 2.1.
Because χ is the intersection of closed halfspaces it is also closed.

Definition 2.1.3:

A set $C \subseteq R^n$ is said to be <u>compact</u> if and only if it is closed and bounded.

2.2. Measure of the "Size" of a Linear Optimization Problem

In this section we introduce some basic notions about the measure of the size of a linear optimization problem (LP).

I. Descriptive data of (LP)

The descriptive data of a linear optimization problem are:

- n the number of variables
- m the number of linear constraints (other than non-negativity)
- c_j the objective function coefficients
- b_i the right –hand side elements
- a_j^i the technological coefficients(coefficients of linear constraints)

II. Digital Size of (LP)

Let us assume that all data of the linear optimization problem are integers or rational numbers. This is no real restriction since, for instance, $\sqrt{2}$ is represented on a digital computer by its rational approximation to a given number p of "digits " that are prescribed by the "word length" of the "register" of the digital computer that is set aside to store this number.

To store any integer number a on a digital computer

$$\langle a \rangle = 1 + \left\lceil \log_2^{(1+|a|)} \right\rceil$$

bits suffices evidently , where [α] is the smallest integer number greater than or equal to α.
A rational number $\frac{a}{b}$ where GCF (a, b) = 1 can be stored by storing a and b separately.

To store any linear optimization problem with integer data we thus need at least

$$\mathbf{L} \geq \langle n \rangle + \langle m \rangle + \sum_{j=1}^{n} \langle c_j \rangle + \sum_{i=1}^{m} \langle b_i \rangle + \sum_{i=1}^{m} \sum_{j=1}^{n} \langle a_j^i \rangle \tag{2.2.1}$$

bits on a digital computer .

Definition 2.2.1: We call L in (2.2.1) the **digital size** of the linear optimization problem.

Now let us denote

$step\ (m, n)$ the number of steps an iterative procedure (or an algorithm) required to solve a linear optimization problem having n variables and m linear constraints and

$time\ (n, L)$ the number of elementary operations such as $+, -, \times, \div, <, =, < or >$ required by a digital computer to carry out the iterations of an algorithm for a linear optimization problem of size L.

Obviously,

$$time\,(n, L) \le step(m, n)$$

no matter what parameters m, n and L we consider, since $step(m,n)$ simply counts the "step" and not the "work" of the algorithm.

Definition 2.2.1: An algorithm is said to be <u>technically good</u> or <u>polynomially bounded</u> if there exist integers $K > 0, p \ge 0, q \ge 0$ that are independent of the data of the linear optimization problem such that

$$time(n,L) \le Kn^q L^p. \tag{2.2.2}$$

If $time\ (n, L)$ satisfies the relation (2.2.2) we write

$$time\,(n, L) = \mathcal{O}(n^q L^p)$$

for short which is read as

the time complexity of $time(n, L)$ is of order of magnitude of $n^q L^p$. So, $time\ (n, L) = \mathcal{O}(n^q L^p)$ means, in particular, that $time\ (n, L)$ is bounded by a polynomial function of its parameters.

Likewise, we can write for the step complexity

$$step(m,n) = \mathcal{O}(m^p n^q)$$

to express the existence of data -independent integers $K > 0, p \ge 0$ and $q \ge 0$ such that

$$step\ (m, n) \le Km^p n^q,$$

i.e. the polynomial boundedness of the number of steps of the algorithm.

We will show in section 3.2.3 that in the worst case, unlike the simplex algorithm, the projective algorithm is polynomially bounded.

2.3. Lagrangean Multiplier Technique

The Lagrangean multiplier technique (or simply the Lagrange method), so named after Joseph Louis de Lagrange (1736-1813), is a well-known procedure for determining the minimum or maximum of a function subject to constraints. The basic idea of the method is to replace a given optimization problem with constraints by an optimization problem without constraints.

A given optimization problem can be either with equality constraints or with inequality constraints or both. In the following section we consider how to apply the Lagrange method in each case.

2.3.1. Lagrange method for optimization problem with equality constraints

Let $U \subseteq R^n, S \subseteq U$ *and* $f : U \to R$. Let $h : U \to R^m$
Consider the optimization problem

(P) $\quad \min\{f(x) : x \in S\}$,
where $S = \{x \in U : h(x) = 0\} = \{x \in U : h_j(x) = 0 \text{ for } j = 1,\ldots,m\}$

The function

$$L(x,\lambda) = f(x) + \sum_{j=1}^{m} \lambda_j h_j, \quad \text{where } (x,\lambda) \in R^n \times R^m$$

is said to be Lagrangean or Lagrange function with regard to (P).
The vector $\lambda \in R^n$ is said to be the vector of Lagrange multiplier.

Now we consider a new optimization problem without constraints:

$(P_\lambda) \quad \min\{L(x,\lambda) : x \in U\}$

The connection between (P) and (P_λ) is described in the following

Theorem 2.3.1: (Lagrange-Lemma for equality constraints)
Let (P) be given and let $S = \{x \in U : h_j(x) = 0 \text{ for } j = 1,\ldots,m\}$.
If there exist $\lambda^* \in R^n$ *such that* $x_0 \in S$ is a solution of (P_{λ^*}), then x_0 is a solution of (P).

Necessary conditions for $x_0 \in S$ to be a solution of (P_{λ^*}) are given in the following

Corollary 1: Let (P) be given and let $S = \{x \in U : h_j(x) = 0 \text{ for } j = 1,...,m\}$ and $U \subseteq R^n$ be an open set. Furthermore, let f and h_j be partially differentiable on U. If $x_0 \in S$ is a solution of (P_{λ^*}), then

1. $h_j(x_0) = 0 \text{ for all } j \in \{1,...,m\}$,
2. $\frac{\partial}{\partial x_k} f(x_0) + \sum \lambda_j^* \frac{\partial}{\partial x_k} h_j(x_0) = 0 \text{ for all } k \in \{1,...,n\}$

Obviously the conditions 1 and 2 of the above corollary form the system of n + m equations with m + n variables.

2.3.2. Lagrange method for optimization problem with inequality constraints

We consider the optimization problem

$$(P) \qquad \min\{f(x) : x \in S\},$$

$$\begin{aligned} &\text{where } S = \{x \in U : g(x) \le 0, h(x) = 0\}, \\ &g : U \to R^m, h : U \to R^p, \\ &g(x) = (g_1(x),..., g_m(x))^T \in R^m \text{ and} \\ &h(x) = (h_1(x),..., h_p(x))^T \in R^p \end{aligned}$$

We define the Lagrange function for (P)

$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x), \quad \text{where } (x, \lambda, \mu) \in U \times R_+{}^m \times R^p.$

The numbers $\lambda_i, i \in \{1,...,m\} \text{ and } \mu_j, j \in \{1,....,p\}$ are called Lagrange - multipliers.

For fixed $(\lambda, \mu) \in R_+^m \times R^p$ we consider the optimization problem,

$$(P_\lambda, \mu) \qquad \min\{L(x, \lambda, \mu) : x \in U\}$$

As a connection between (P) and $(P_{\lambda,\mu})$ we give the following

Theorem 2.3.2:(Lagrange –Lemma for inequality constraints)

Let (P) be given, and $\lambda \in R^m, \mu \in R^p, x_0 \in S$, *with* $\langle \lambda, g(x_0) \rangle = 0$.

If x_0 is a solution of ($P_{\lambda,\mu}$), then x_0 is a solution of (P).

Now let f, g, and h be differentiable on the open set $U \subseteq R^n$. Let x^* be a minimum point of $L(x, \lambda^*, \mu^*)$ on U and $x^* \in \{x \in S : \langle \lambda^*, g(x) \rangle = 0\}$. Then the following necessary conditions are satisfied.

Kuhn-Tucker Conditions:

(1) $h_j(x^*) = 0, j \in \{1,..., p\}$

(2) $\lambda_i^* g_i(x^*) = 0, \lambda_i^* \geq 0, g_i(x^*) \leq 0, i \in \{1,..., m\}$,

(3) $\frac{\partial}{\partial x_k} f(x^*) + \sum_{i=1}^{m} \lambda_i^* \frac{\partial}{\partial x_k} g_i(x^*) + \sum_{j=1}^{p} \mu_j^* \frac{\partial}{\partial x_k} h_j(x^*) = 0, k \in \{1,..., n\}$

(4) $\lambda_i \geq 0$

2.3.3. Reduction of inequality constraints to equality constraints

Let the optimization problem be given by

(P) $\min\{f(x) : x \in S\}$

where $S = \{x \in U : g(x) \leq 0, h(x) = 0\}$

Then we introduce the so-called Slack- variables

$z = \{z_1,..., z_m\} \in R^m$ *such that* $g_1(x) + z_1^2 = 0,..., g_m(x) + z_m^2 = 0, x \in S$

Then we define

$\bar{f}(x,z) := f(x), \bar{h}(x,z) := h(x), \ \bar{g}(x,z) := (g_1(x) + z_1^2,..., g_m(x) + z_m^2)^T$

and consider the following optimization problem :

$(\bar{P})$ $\bar{f}(x,z) \to \min, (x,z) \in \bar{S}\}$

$\bar{S} = \{(x,z) \in U \times R^m / \bar{g}(x,z) = 0, \bar{h}(x,z) = 0\}$

Obviously this problem is equivalent to (P).

To solve this problem we can use the Lagrange–lemma for optimization problem with equality constraints.

3. Projective Algorithm

3.1. Introduction

Consider the linear optimization problem in standard form

(LP) $$\min \{\mathbf{cx} : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\},$$

where A is an $m \times n$ matrix of rank m,
$c \in R^n$ is a row vector, and
$b \in R^m$ is a column vector .

Let us denote as usual the set of all feasible solutions of (LP) by

$$\chi = \{\mathbf{x} \in \mathbf{R}^n : \mathbf{Ax} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$$

Furthermore, we assume that

(i) there exists $x^0 \in \chi$ such that $x^0 > 0$, i.e. the relative interior of χ is nonempty.

(ii) the objective function $\mathbf{cx}$ over χ is not constant, i.e., in particular, $\mathbf{c} \neq 0$.

Definition 3.1.1: We call a feasible solution x^k to (LP) a relative interior feasible point if each components of x^k is positive, i.e. $x^k > 0$.

The basic idea of projective algorithms for (LP) is *to generate a sequence of relative interior feasible points that converges towards an optimal solution of* (LP). This is done -in essence-by solving a sequence of restricted problems that involve no inequalities and which are amenable to solution by classical methods of calculus. The projective transformations that are employed correspond to an iterated change of variables that permit to implement this idea.

Let us rewrite the problem (LP) as follows

(LP*) $$\min \{\mathbf{cx} : \mathbf{Ax} = \mathbf{bx}_{n+1}, \mathbf{x}_{n+1} = 1, \mathbf{x} \geq \mathbf{0}\},$$

where x_{n+1} is a new variable.

Geometrically we are embedding R^n into R^{n+1} by identifying a point $x \in R^n$ with the point $(x,1) \in R^{n+1}$.

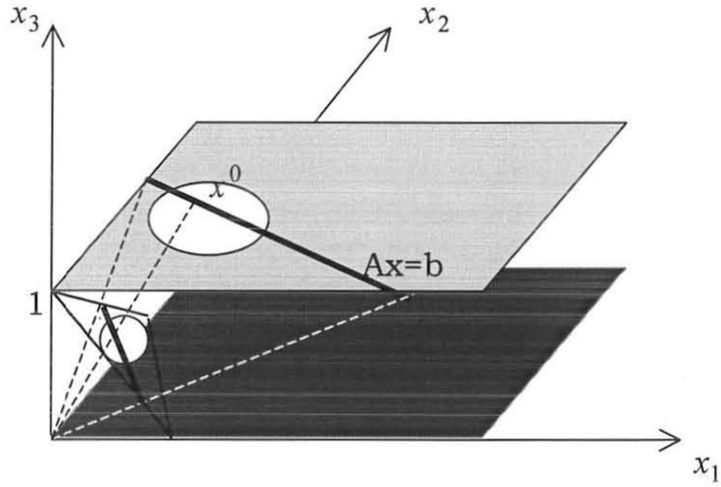


Fig .3.1. Embedding of R^n into R^{n+1} for n = 2.

By the assumption that we have made above, there is a feasible point $(x^0,1) \in R^{n+1}$ to (LP*) such that $x^0 > 0$.

Now we consider the projective transformation T_0 given by

$$(T_0) \quad \begin{aligned} y_j &= \frac{\dfrac{x_j}{x_j^0}}{1 + \sum_{j=1}^{n} \dfrac{x_j}{x_j^0}} \quad \textit{for} \quad j \in \{1,\dots,n\} \\ y_{n+1} &= \frac{1}{1 + \sum_{j=1}^{n} \dfrac{x_j}{x_j^0}} \end{aligned}$$

Since $(x^0,1) > 0$, T_0 is well defined.

Obviously, T_0 maps the nonnegative orthant $\{(x,1) \in R^{n+1} : x \geq 0\}$ in to the n - dimensional simplex

$$S^{n+1} = \left\{ y \in R^{n+1} : \sum_{j=1}^{n+1} y_j = 1, y \geq 0 \right\}.$$

Furthermore, the point $(x^0,1) \in R^{n+1}$ is mapped into the center

$$y^0 = \left(\frac{1}{n}\right)\mathbf{f}$$

of the simplex S^{n+1} where $\mathbf{f}^T = (1,\ldots,1)$ is the n+1- vector of ones.

Let us rewrite $y = T_0(x)$ to denote the image of $y \in R^{n+1}$ of the point $(x^0,1) \in R^{n+1}$, where $x \in R^n$.

Thus T_0 can be regarded as a mapping from R^n into R^{n+1} in the natural way.

If we consider T_0 as the embedding of R^n into R^{n+1}, then it possesses an inverse that is given by

$$(T_0^{-1}) \qquad \begin{aligned} x_j &= \frac{x_j^0 y_j}{y_{n+1}} \qquad \textit{for } j = 1,\ldots,n \\ x_{n+1} &= 1 \end{aligned}$$

Let D = $\{x_1^0,\ldots,x_n^0\}$ be the $n \times n$ matrix with diagonal elements x_i^0 for $i = 1,\ldots,n$ and zero elsewhere. The projective transformation T_0 in matrix form is then written as

$$(T_0) \qquad \begin{pmatrix} y_N \\ y_{n+1} \end{pmatrix} = \frac{1}{1+e^T D^{-1} x} \begin{pmatrix} D^{-1}x \\ 1 \end{pmatrix}, \qquad (3.1.1)$$

where $\mathbf{e}^T = (1,\ldots,1)$ is the n-vector of ones and

$y_N = (y_1,\ldots,y_n)$ is the n-vector whose components are the first n components of the vector $y \in R^{n+1}$.

Using the equality of matrices we get from (3.1.1) that

$$y_N = \frac{1}{1+e^T D^{-1} x}(D^{-1}x) \quad \text{and} \quad y_{N+1} = \frac{1}{1+e^T D^{-1} x}.$$

Obviously, we have $y_N = y_{n+1} D^{-1} x$ from which we get $x = \left(\frac{1}{y_{n+1}} \right) D y_N$

Hence the inverse of T_0 in matrix form becomes

$$(T_0^{-1}) \qquad \begin{aligned} x &= \left(\frac{1}{y_{n+1}} \right) D\, y_N \\ x_{n+1} &= 1 \end{aligned} \qquad (3.1.2)$$

For (LP*) we have the corresponding set of all feasible solutions given by

$$\chi^* = \left\{ \mathbf{x} \in \mathbf{R}^{\mathbf{n+1}} : \mathbf{Ax} = \mathbf{bx}_{\mathbf{n+1}}, \mathbf{x}_{\mathbf{n+1}} = \mathbf{1}, \mathbf{x} \geq \mathbf{0} \right\},$$

which is the intersection of the nonnegative orthant

$$O = 0\left\{ (x,1) \in R^{n+1}: x \geq \right\}$$

and the subspace

$$L = \left\{ (x,1) \in R^{n+1} : Ax - bx_{n+1} = 0 \right\}.$$

Obviously, T_0 (O)= S^{n+1}.

From (3.1.2) and the equation $Ax - bx_{n+1} = 0$ we have that

$$ADy_N - by_{n+1} = 0$$

or

$$(AD, -b) \begin{pmatrix} y_N \\ y_{n+1} \end{pmatrix} = 0$$

which implies

$$(AD, -b) y = 0 .$$

Thus we have

$$T_0 \text{ (L)} = \left\{ y \in R^{n+1} : (AD, -b) y = 0 \right\}.$$

Hence it follows that the image $T_0(\chi)$ of a feasible set χ is given by

$$T_0 (\chi) = \left\{ \mathbf{y} \in \mathbf{R}^{\mathbf{n+1}} : (\mathbf{AD}, -\mathbf{b})\mathbf{y} = \mathbf{0}, \mathbf{f}^{\mathbf{T}} \mathbf{y} = \mathbf{1}, \mathbf{y} \geq \mathbf{0} \right\}$$

or

$$T_0 (\chi) = \left\{ \mathbf{y} \in \mathbf{R}^{\mathbf{n+1}} : (\mathbf{AD}, -\mathbf{b})\mathbf{y} = \mathbf{0}, \mathbf{y} \in \mathbf{S}^{\mathbf{n+1}} \right\} \qquad (3.1.3)$$

Trivially, we have $T_0(\chi) \subseteq S^{n+1}$ and $y^0 \in T_0(\chi)$.

Hence the linear optimization problem becomes the non-linear optimization problem

(FLP) $$\min\left\{\frac{\mathbf{cDy}_{\mathbf{N}}}{\mathbf{y}_{\mathbf{n+1}}} : \mathbf{y} \in \mathbf{T_0}(\chi)\right\}$$

Note that the non- linearity of the problem (FLP) is the result of the variable y_{n+1} in the denominator of the objective function of (FLP).

Let us denote by

$$B_\rho^{n+1} = \left\{ y \in R^{n+1} : \sum_{j=1}^{n+1} y_j = 1 \, , \, \sum_{j=1}^{n+1}\left(y_j - \frac{1}{n+1}\right)^2 \le \rho^2 \right\},$$

the *intersection* of the n+1 - dimensional ball with radius ρ and center y^0 with the hyper plane $f^T y = 1$.

Remark:

From the definition of B_ρ^{n+1} it follows that if the radius ρ is small enough, then B_ρ^{n+1} becomes an n- dimensional ball in the simplex S^{n+1}.

Now we prove a very important relation between B_ρ^{n+1} and S^{n+1} (in the sense of inclusion) in the following

Lemma 3.1: Let $r^2 = \frac{1}{n(n+1)}$. Then

(i) $B_\rho^{n+1} \subseteq S^{n+1}$ if and only if $0 \le \rho \le r$.

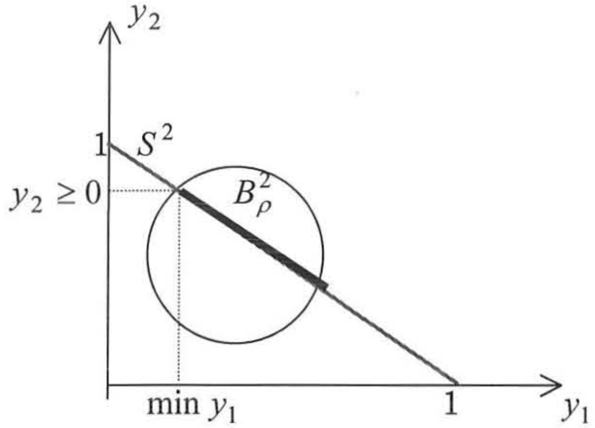
(ii) $S^{n+1} \subseteq B_\rho^{n+1}$ if and only if $\rho \ge \sqrt{\frac{n}{n+1}}$ = nr

Proof:

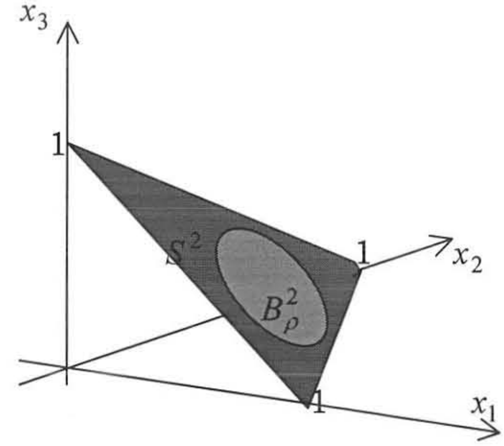
Let $y \in B_\rho^{n+1}$. Then $y \in S^{n+1}$ if and only if $y \ge 0$. By definition 1.1.2, $y \ge 0$ if and only if $y_j \ge 0$ *for all* $j \in \{1,\ldots,n+1\}$. This will be the case if and only if

$\min\left\{y_i : y \in B_\rho^{n+1}\right\} \ge 0$ for an arbitrary but fixed $i \in \{1,\ldots, n+1\}$.

We illustrate this by the following diagrams for the cases n= 1and n= 2.



For n = 1



For n = 2

For an arbitrary but fixed $i \in \{1,\ldots,n+1\}$ now we consider the optimization problem

(P*) $$\min\{y_i : y \in B_\rho^{n+1}\}$$

or equivalently

(P*) $$\min \left\{ y_i : \sum_{j=1}^{n+1} y_j = 1 \, , \sum_{j=1}^{n+1} \left(y_j - \frac{1}{n+1} \right)^2 \leq \rho^2 \right\}$$

The Lagrange function for (P*) is

$$L(y,\lambda,\mu) = y_i + \lambda \left[\sum_{i=1}^{n+1} \left(y_j - \frac{1}{n+1} \right)^2 - \rho^2 \right] + \mu \left[\sum_{i=1}^{n+1} y_j - 1 \right],$$

where $(y,\lambda,\mu) \in R^{n+1} \times R_+ \times R.$

Then we have the following optimization problem

($P^*_{\lambda,\mu}$) $$\min\{ L(y,\lambda,\mu) : (y,\lambda,\mu) \in R^{n+1} \times R_+ \times R\}$$

As necessary condition (Kuhn-Tucker conditions) we get

(1) $$\frac{\partial L}{\partial y_i} = 1 + 2\lambda \left(y_i - \frac{1}{n+1} \right) + \mu = 0$$

(2) $$\frac{\partial L}{\partial y_j} = 2\lambda\left(y_j - \frac{1}{n+1}\right) + \mu = 0 \quad for\ all\ i \neq j$$

(3) $$\lambda\left[\sum_{i=1}^{n+1}\left(y_j - \frac{1}{n+1}\right)^2 - \rho^2\right] = 0$$ (4) $$\sum_{i=1}^{n+1} y_j = 1$$

(5) $$\sum_{i=1}^{n+1}\left(y_j - \frac{1}{n+1}\right)^2 \leq \rho^2$$ (6) $$\lambda \geq 0$$

In order to solve this system of equations and inequalities we distinguish some cases:

Case1: $\lambda = 0$. Then from (1) and (2) we get $1 + \mu = 0\ and\ \mu = 0$. But this is obviously a contradiction

Case2: $\lambda > 0$. Then from (1) and (2) we get

(7) $$y_i = \frac{1}{n+1} - \frac{\mu+1}{2\lambda}$$ and

(8) $$y_j = \frac{1}{n+1} - \frac{\mu}{2\lambda} \quad for\ all\ \ j \neq i\ .$$

From (4) we have that

$$1 = \sum_{i=1}^{n+1} y_j = \sum_{i \neq j} y_j + y_i = n\left(\frac{1}{n+1} - \frac{\mu}{2\lambda}\right) + \left(\frac{1}{n+1} - \frac{\mu+1}{2\lambda}\right) = 1 - \frac{1+(n+1)\mu}{2\lambda}$$

which is equivalent to

$$(n+1)\mu+1=0$$

or $$\mu = \frac{-1}{n+1}$$

From (3) we get by $\lambda > 0$

$$\rho^2 = \sum_{i=1}^{n+1}\left(y_j - \frac{1}{n+1}\right)^2 = \sum_{j \neq i}\left(y_j - \frac{1}{n+1}\right)^2 + y_i = \left(-\frac{n}{2\lambda(n+1)}\right)^2 + \left(\frac{1}{2\lambda(n+1)}\right)^2$$
$$= \frac{n}{4\lambda^2(n+1)}$$

or $$\rho^2 = \frac{n}{4\lambda^2(n+1)}$$

from which it follows that

$$\lambda = \frac{1}{2\rho}\sqrt{\frac{n}{n+1}}$$

Using the values of μ and λ we obtain from (7) and (8) that

$$y_i = \frac{1}{n+1} - \rho\sqrt{\frac{n}{n+1}} \qquad \text{and} \qquad y_j = \frac{1}{n+1} + \rho\sqrt{\frac{1}{n(n+1)}}$$

Obviously, we have $y_j \geq 0$ for all $j \neq i$.

Thus, from $y_i \geq 0$ we get

$$y_i = \frac{1}{n+1} - \rho\sqrt{\frac{n}{n+1}} \geq 0$$

i.e. $\frac{1}{n+1} \geq \frac{n\rho}{\sqrt{n(n+1)}}$

or $\rho \leq \frac{1}{\sqrt{n(n+1)}} = r$

Hence $B_\rho^{n+1} \subseteq S^{n+1}$ if and only if $0 \leq \rho \leq r$ .//

(ii) The proof is similar to that of part (i).

But for this case we have to consider for arbitrary but fixed $i \in \{1,\ldots,n+1\}$ the optimization problem

$$(\widetilde{P}) \qquad max\ \{y_i : y \in B_\rho^{n+1}\}$$

or equivalently

$$(\widetilde{P}) \qquad \min\{-y_i : y \in B_\rho^{n+1}\}$$

If we follow the same procedure like (i) we get, as a solution to $(\widetilde{P})$

$$y_i = \frac{1}{n+1} + \frac{\rho n}{\sqrt{n(n+1)}} \quad \text{and} \quad y_j = \frac{1}{n+1} - \frac{\rho}{\sqrt{n(n+1)}}$$

Obviously $y_i \geq 0$. From $y_j \leq 0$ we get

$$y_j = \frac{1}{n+1} - \frac{\rho}{\sqrt{n(n+1)}} \leq 0 \quad \text{or} \quad \frac{\rho}{\sqrt{n(n+1)}} \geq \frac{1}{n+1}, \qquad \text{i.e.} \quad \rho \geq \sqrt{\frac{n}{n+1}}.$$

Hence we have $S^{n+1} \subseteq B_\rho^{n+1}$ if and only if $\rho \geq \sqrt{\frac{n}{n+1}}$ = nr. //

From Lemma 3.1 (i) it follows that if $0 \leq \rho \leq r$, then $B_\rho^{n+1} \subseteq S^{n+1}$.

Thus, we can replace (FLP) by

$$(\mathbf{FLP}_\rho) \qquad \min\left\{\frac{\mathbf{(cD,0)y}}{\mathbf{y_{n+1}}} : \mathbf{(AD,-b)y} = \mathbf{0}, \mathbf{y} \in \mathbf{B}_\rho^{n+1}\right\},$$

where $0 \leq \rho < r$ in order to ensure that $y > 0$.

So, we have for all $0 \le \rho < r$,

$$\min \ \{ \mathbf{cx} : \mathbf{x} \in \chi \} \le \min \left\{ \frac{(\mathbf{cD},0)\mathbf{y}}{\mathbf{y}_{n+1}} : \mathbf{y} \in \mathbf{T}_0(\chi) \cap \mathbf{B}_{\rho}^{n+1} \right\}.$$

From the above discussions one can easily observe that the problem (FLP_ρ) is the restriction of the problem (FLP) and is a classical non linear optimization problem .We have different approaches to the solution of this non linear optimization problem .Two of them are:

1." Linearizing " the objective function of (FLP_ρ) which leads to a particularly simple solutions.
2. Solving (FLP_ρ) exactly.

However, in either case, once a solution to (FLP_ρ) or an approximation to it has been obtained, one can use the inverse transformation T_0^{-1} of the projective transformation T_0 to obtain a new interior feasible point $x^1 \in \chi$ which gives rise to a new projective transformation T_1, etc. (See fig .3.2 below.)

In this paper we restrict ourselves to discuss only the first approach to the solution of (FLP_ρ).

3.2. A Basic Algorithm

In the first approach to the nonlinear optimization problem (FLP$_\rho$) we approximate (FLP) by the auxiliary linear optimization problem

$$\text{(ALP)} \qquad \min \left\{ (\mathbf{cD},\mathbf{0})\mathbf{y} : \mathbf{y} \in \mathbf{T_0}(\chi) \right\}$$

We make the following additional assumptions:

(iii) χ is bounded and

(iv) the optimal objective function value of (LP) equals zero.

Remark:

1. From the additional assumption (iv), it follows that the optimal objective function value of (ALP) equals zero as well no matter what interior point $x^0 \in \chi$ is used in the projective transformation T_0.
2. From the additional assumption (iii) and Theorem 2.1, it follows that χ is compact.

Like we did above, we replace the problem (FLP$_\rho$) by the auxiliary optimization problem

$$(\mathbf{ALP_\rho}) \qquad \min \left\{ (\mathbf{cD},\mathbf{o})\mathbf{y} : (\mathbf{AD},-\mathbf{b})\mathbf{y} = \mathbf{0}, \mathbf{y} \in \mathbf{B}_\rho^{n+1} \right\}$$

By the above remark since χ is compact an optimal solution to (ALP$_\rho$) exist and its optimal objective function value is nonnegative for all $0 \le \rho \le r$.

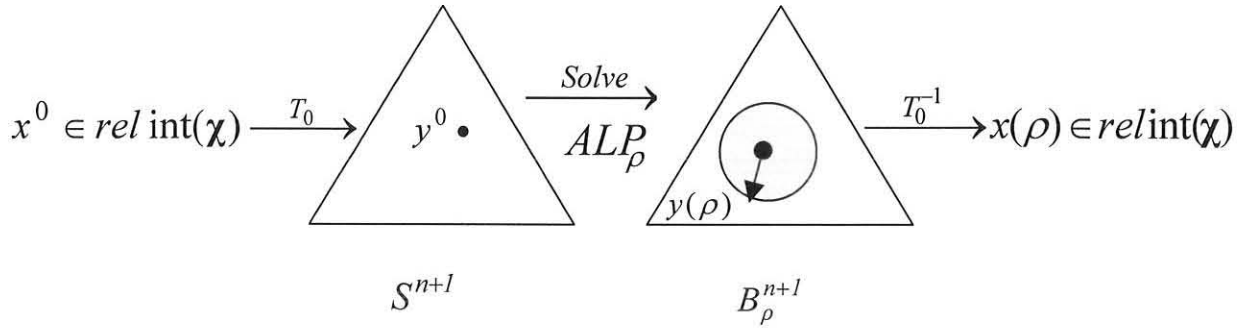


Fig. 3.2. The iterative step of projective algorithms.

We summarize- where we change momentarily our notation- the key facts about the solution of (ALP_ρ) in the following

Theorem 3.1.1: Consider the linear optimization problem

$$\min \left\{ cz{:}Az = 0, e^T z = 1, z \geq 0 \right\},$$

where A is an $m \times n$ matrix of rank m, $z^0 = \left(\frac{1}{n}\right)\mathbf{e}$ is a non optimal feasible solution and the optimal function value equals to zero, i.e. $\mathbf{c}z^0 > 0$.

Then for all $\rho \geq 0$ an optimal solution for the problem

$$(P_\rho) \qquad \min \left\{ cz{:}Az = 0, z \in B^n_\rho \right\}$$

is given by

$$z(\rho) = \left(\frac{1}{n}\right)\mathbf{e} - \frac{\rho\, p}{\|p\|} ,$$

where $p = \left[I_n - A^T (AA^T)^{-1} A - \left(\frac{1}{n}\right) ee^T \right] c^T$ is the orthogonal projection of $\mathbf{c}$ on the subspace

$$\left\{ z \in R^n{:}Az = 0,\, e^T z = 0 \right\}$$

Moreover, for all $\rho \geq 0$ the optimal solution $z(\rho)$ satisfies

$$\frac{cz(\rho z}{cz^0} \leq 1 - \rho\sqrt{\frac{n}{n-1}} .$$

Proof: Since the feasible set is compact (by the above remark) an optimal solution to (P_ρ) exists. So, we can apply the Lagrange multiplier technique to calculate it.

The given optimization problem (P_ρ) is equivalent to

$$(P_\rho) \qquad \min\left\{cz: Az=0, \sum_{i=1}^{n+1} z_j = 1, \sum_{i=1}^{n+1}\left(z_i - \frac{1}{n+1}\right)^2 \le \rho^2\right\}$$

which is the optimization problem with inequality constraints .Now we introduce slack variable z_{n+1}^2 and consider the optimization problem with equality constraints

$$(P^*_\rho) \qquad \min\left\{cz: Az=0, \sum_{i=1}^{n+1} z_j = 1, \sum_{i=1}^{n+1}\left(z_i - \frac{1}{n+1}\right)^2 + z_{n+1}^2 = \rho^2\right\}.$$

Obviously this is equivalent to (P_ρ).

The Lagrange function for (P^*_ρ) is given by

$$L(z,\mu,\lambda) = cz + \mu Az + \lambda_1\left(\sum_{j=1}^{n} z_j - 1\right) + \lambda_2\left[\sum_{j=1}^{n}\left(z_j - \frac{1}{n}\right)^2 + z_{n+1}^2 - \rho^2\right]$$

where $(z,\mu,\lambda) \in R^n \times R^m \times R^2$.

As a necessary condition we have that any optimal solution to (P_ρ) satisfies

$$c + \mu A + \lambda_1 e^T + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T = 0 \tag{1}$$

First we show that $\lambda_2 \neq 0$. Suppose to the contrary $\lambda_2 = 0$.

Then (1) becomes $c + \mu A + \lambda_1 e^T = 0$.

Multiplying by z and since by assumption $Az = 0$ *and* $e^T z = 1$ from (1) we get

$$cz + \lambda_1 = 0$$

So, we have

$cz = -\lambda_1$ for all feasible solution z.

This is impossible since by assumption we have a non-optimal solution.

Hence we have $\lambda_2 \neq 0$.

Now multiplying (1) by e we get

$$ce + \mu Ae + \lambda_1 e^T e + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T e = 0$$

Since by assumption $Ae = 0$ and $e^T z = 1$ we have that

$$ce + \lambda_1 n = 0$$

which gives $\lambda_1 = -cz^0$.

By the remark following Lemma 1.1 AA^T is positive definite and hence nonsingular. That means $(AA^T)^{-1}$ exists.

Again multiplying (1) by A^T we obtain

$$cA^T + \mu AA^T + \lambda_1 e^T A^T + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T A^T = 0$$

But by assumption we have A e = 0 and A z = 0.
So we get

$$cA^T + \mu AA^T = 0$$

from which it follows that

$$\mu = -cA^T\left(AA^T\right)^{-1}$$

Substituting λ_1 and μ we get from (1) that

$$c + cA^T\left(AA^T\right)^{-1} - cz^0 e^T + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T = 0$$

or

$$c\left[I_n - A^T\left(AA^T\right)^{-1}A - \left(\frac{1}{n}\right)ee^T\right] + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T = 0$$

which can be written as

$$cP^T + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T = 0 \quad \text{where } P = I_n - A^T(AA^T)^{-1}A - cz^0 e^T$$

Equivalently we have

$$(Pc^T)^T + \lambda_2\left(z - \left(\frac{1}{n}\right)e\right)^T = 0 \tag{2}$$

Solving for z from (2) we get

$$z = \left(\frac{1}{n}\right)e - \left(\frac{1}{\lambda_2}\right)p \quad \text{where } p = Pc^T$$

Consequently,

$$\frac{1}{\lambda_2} = \frac{\rho}{\|p\|} \quad \text{or}$$

$$\lambda_2 = \pm\frac{\|p\|}{\rho}$$

Since $\lambda_2 \neq 0$ we have that $\|p\| \neq 0$ and thus

$$z(\rho) = \left(\frac{1}{n}\right)e - \frac{\rho\, p}{\|p\|} \tag{3}$$

is the minimizer for (P_ρ) .

Multiplying (3) by vector c we get

$$cz(\rho) = cz^0 - \frac{\rho}{\|p\|} cp$$

But using $P^2 = P$ *and* $P^T = P$ we have

$$cp = c(Pc^T) = c(P^2 c^T) = cPPc^T = (Pc^T)^T Pc^T = p^T p = \|p\|^2$$

Hence

$$cz(\rho) = cz^0 - \rho\|p\|$$

This proves the first part of the theorem.

Next we prove the second part of the theorem.

Since $S^n \subseteq B^n_{\rho^*}$ for $\rho^* = \sqrt{\frac{n-1}{n}}$ by part (ii) of Lemma 3.1 it follows from the assumption that the optimal objective function value of the linear optimization equals zero that $cz(\rho^*) \le 0$ .For, if not then $cz(\rho^*) > 0$ and since $S^n \subseteq B^n_{\rho^*}$ the optimal objective function value of the linear optimization problem is positive which is a contradiction.

Hence

$$cz(\rho^*) \le 0$$

This implies

$$cz^0 - \|p\|\sqrt{\frac{n-1}{n}} \le 0 \quad or \quad \frac{\|p\|}{cz^0} \ge \sqrt{\frac{n}{n-1}}$$

Now

$$\frac{cz(\rho)}{cz^0} = 1 - \frac{\rho\|p\|}{cz^0} \le 1 - \rho\sqrt{\frac{n}{n-1}}$$

That means

$$\frac{cz(\rho)}{cz^0} \le 1 - \rho\sqrt{\frac{n}{n-1}}$$

which completes the proof. //

3.2.1. The Solution to the Approximate Problem (ALP_ρ)

It follows from Theorem 3.1 that in order to solve the problem (ALP_ρ) we have to calculate the orthogonal projection of the vector (c D, 0) on the subspace

$$\{y \in R^{n+1} : (AD,-b)y = 0, f^T y = 0\}. \tag{3.2.1}$$

Let U be the subspace (3.2.1) and q be the orthogonal projection of a vector $(cD,0)$ on the subspace U.

Obviously, since $q \in U$, $(AD,-b)q = 0$ and $f^T q = 0$

Then we have

$$(cD,0)^T - q = (AD,-b)^T \lambda + f\nu \quad \text{, where } \lambda, \nu \in R^{n+1}. \tag{1}$$

Multiplying (1) by $(AD,-b)$ we get

$$(AD,-b)\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} - (AD,-b)q = (AD,-b)(Ad,-b)^T \lambda + (AD,-b) f\nu$$

or

$$(AD,-b)\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} = (AD,-b)(Ad,-b)^T \lambda + (AD,-b)\begin{pmatrix} e \\ 1 \end{pmatrix}\nu \quad \text{since } (AD,-b)q = 0 .$$

Since $ADe = Ax^0 = b$ by assumption we have that

$$(AD,-b)\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} = (AD^2 A^T + bb^T)\lambda .$$

Let us denote $G = AD^2 A^T$.

Since A is of full row rank and D is non-singular, G is positives definite and hence G^{-1} exists. Applying the Rank-One Update formula (Theorem 1.2) we obtain

$$(AD^2 A^T + bb^T)^{-1} = (G + bb^T)^{-1} = G^{-1} - (1+\beta)^{-1}(G^{-1}b)(b^T G^{-1}),$$

where $\beta = b^T G^{-1} b \geq 0$, since G^{-1} is positive definite as well.

So, we have

$$\lambda = \left(G^{-1} - (1+\beta)^{-1}(G^{-1}b)(b^T G^{-1})\right)(AD,-b)\begin{pmatrix} Dc^T \\ 0 \end{pmatrix}$$

Again multiplying (1) by f^T we have

$$f^T \begin{pmatrix} Dc^T \\ 0 \end{pmatrix} - f^T q = f^T (Ad,-b)^T \lambda + f^T f\nu$$

Since $ADe = Ax^0 = b$ by assumption using $f^T q = 0$ we get

$$f^T\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} = f^T f \nu \quad ,$$

or

$$\nu = \frac{1}{n+1} f^T \begin{pmatrix} Dc^T \\ 0 \end{pmatrix} \text{ since } f^T f = n+1 \quad .$$

Substituting the values of λ and ν from (1) we get

$$q = \left[I_{n+1} - \begin{pmatrix} DA^T \\ -b^T \end{pmatrix} \left(G^{-1} - (1+\beta)^{-1}(G^{-1}b)(b^T G^{-1})\right)(AD, -b) - \frac{1}{n+1} ff^T \right] \begin{pmatrix} Dc^T \\ 0 \end{pmatrix}$$

If we denote the projection operator on the subspace (3.2.1) by Q then we have

$$q = Q\begin{pmatrix} Dc^T \\ -b^T \end{pmatrix}$$

where $Q = I_{n+1} - \begin{pmatrix} DA^T \\ -b^T \end{pmatrix} \left(G^{-1} - (1+\beta)^{-1}(G^{-1}b)(b^T G^{-1})\right)(AD, -b) - \frac{1}{n+1} ff^T$

Let us further simplify the expression for Q.
Now

$$Q = I_{n+1} - \begin{pmatrix} DA^T \\ -b^T \end{pmatrix} \left(G^{-1} - (1+\beta)^{-1}(G^{-1}b)(b^T G^{-1})\right)(AD, -b) - \frac{1}{n+1} ff^T$$

$$= \begin{pmatrix} I_n & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} DA^T G^{-1} AD - (1+\beta)^{-1} DA^T (G^{-1}b)(b^T G^{-1}) AD & -DA^T G^{-1} b + (1+\beta)^{-1} DA^T (G^{-1}b)(b^T G^{-1}) b \\ -b^T G^{-1} b + (1+\beta)^{-1} b^T (G^{-1}b)(b^T G^{-1}) AD & b^T G^{-1} b - (1+\beta)^{-1} b^T (G^{-1}b)(b^T G^{-1}) b \end{pmatrix}$$

$$- \frac{1}{n+1} f f^T$$

$$= \begin{pmatrix} I_n - DA^T G^{-1} AD & DA^T G^{-1} b \\ -b^T G^{-1} b & -b^T G^{-1} b \end{pmatrix} + (1+\beta)^{-1} \begin{pmatrix} DA^T (G^{-1}b)(b^T G^{-1}) AD & -DA^T (G^{-1}b)(b^T G^{-1}) b \\ -b^T (G^{-1}b)(b^T G^{-1}) AD & b^T (G^{-1}b)(b^T G^{-1}) b \end{pmatrix} - \frac{1}{n+1} ff^T$$

Let us denote

$$P = I_n - DA^T G^{-1} AD \text{ , } p = PDc^T \textit{ and } d = Pe \qquad (3.2.2)$$

Obviously, **p** is the orthogonal projection of the vector $\mathbf{Dc^T}$, and d the orthogonal projection of the vector **e** on the subspace

$$\{x \in R^n : ADx = 0\} \qquad (3.2.3)$$

Then we have

$e-d=e-Pe=(I_n-P)e=DA^TG^{-1}ADe=DA^TG^{-1}b$ and $e^T-d^T=b^TG^{-1}AD$

Using these values together with $\beta=b^TG^{-1}b$ we get

$$\begin{aligned}
Q&=\begin{pmatrix} P & e-d \\ e^T-d^T & 1-\beta \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} (e-d)(e^T-d^T) & -\beta(e-d) \\ -\beta(e^T-d^T) & \beta^2 \end{pmatrix}-\frac{1}{n+1}ff^T \\
&=\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}+\begin{pmatrix} 0 & e-d \\ e^T-d^T & 1-\beta \end{pmatrix}+\quad(1+\beta)^{-1}\begin{pmatrix} (e-d)(e^T-d^T) & -\beta(e-d) \\ -\beta(e^T-d^T) & \beta^2 \end{pmatrix}-\frac{1}{n+1}ff^T \\
&=\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} (e-d)(e^T-d^T) & e-d \\ e^T-d^T & 1 \end{pmatrix}-\frac{1}{n+1}ff^T \\
&=\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} e-d \\ 1 \end{pmatrix}\left(e^T-d^T,\ 1\right)-\frac{1}{n+1}ff
\end{aligned}$$

Thus we have

$$\begin{aligned}
q&=\left[\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} e-d \\ 1 \end{pmatrix}\left(e^T-d^T,\ 1\right)-\frac{1}{n+1}ff^T\right]\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} \\
&=\begin{pmatrix} P & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} Dc^T \\ 0 \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} e-d \\ 1 \end{pmatrix}\left(e^T-d^T,\ 1\right)\begin{pmatrix} Dc^T \\ 0 \end{pmatrix}-\frac{1}{n+1}ff^T\begin{pmatrix} Dc^T \\ 0 \end{pmatrix} \\
&=\begin{pmatrix} PDc^T \\ 0 \end{pmatrix}+(1+\beta)^{-1}\begin{pmatrix} e-d \\ 1 \end{pmatrix}(e^TDc^T-d^TDc^T)-\frac{1}{n+1}\begin{pmatrix} e \\ 1 \end{pmatrix}(e^TDc^T)
\end{aligned}$$

Setting $z_0=cx^0$ *and* $\gamma=p^Td$ we get

$$\begin{aligned}
e^TDc^T-d^TDc^T&=(cDe)^T-(cDPe)^T=(cDe)^T-(cDP^2e)^T \\
&=(cDe)^T-(cDP^T.Pe)^T=(cx^0)^T-(p^Td)^T \\
&=(z_0)^T-(\gamma)^T=z_0-\gamma
\end{aligned}$$

and $e^TDc^T=(cDe)^T=(cx^0)^T=(z_0)^T=z_0$.

Thus, we have

$$q=\begin{pmatrix} p \\ 0 \end{pmatrix}+\frac{z_0-\gamma}{1+\beta}\begin{pmatrix} e-d \\ 1 \end{pmatrix}-\frac{z_0}{n+1}\begin{pmatrix} e \\ 1 \end{pmatrix} \qquad (3.2.4)$$

Note that necessarily $\|p\|\neq 0$ if x^0 is a non-optimal solution to (LP).

By the definitions of β and **d** and using the properties of orthogonal projection we have the following:

$$\gamma = p^T d = (PDc^T)^T (Pe) = cDP^T Pe = cDP^2 e = cDPe$$
$$= cDP^T e = (PDc^T)^T e = p^T e$$

Also

$$\gamma = p^T d = (PDc^T)^T (Pe) = cDP^T Pe = cDP^2 e = cDPe = cDd$$

Furthermore, by definition of the norm of the vector

$$\| d\|^2 = d^T d = (Pe)^T (Pe) = e^T P^T Pe = e^T P^2 e = e^T Pe = e^T d$$

$$\| p\|^2 = p^T p = \left(PDc^T\right)^T \left(PDc^T\right) = \left(cDP^T\right)\left(PDc^T\right) = cDP^2 Dc^T$$
$$= cD(PDc^T) = cDp$$

For β we also have

$$n - \| d\|^2 = n - e^T d = e^T e - e^T d = e^T (e-d) = e^T (e - Pe)$$
$$= e^T \left[e - \left(I_n - DA^T G^{-1} AD\right)e\right] = e^T DA^T G^{-1} ADe$$
$$= e^T DA^T G^{-1} ADe = (ADe)^T G^{-1} ADe = b^T G^{-1} b = \beta$$

and

$$\| e-d\|^2 = (e-d)^T (e-d) = (e^T - d^T)(e-d) = e^T e - e^T d - d^T e + d^T d$$
$$= n - 2\| d\|^2 + \| d\|^2 = n - \| d\|^2 = \beta$$

Since $\beta \ge 0$ we have $n - \| d\|^2 \ge 0$ or $\| d\|^2 \le n$.

Hence we have

$$\begin{aligned} \gamma &= p^T d = P^T e = cDd \\ \beta &= n - \|d\|^2 = \|e - d\|^2 \\ \| d\|^2 &= e^T d \\ \| p\|^2 &= cDp \end{aligned} \qquad (3.2.5)$$

Now using (3.2.5) we get from (3.2.4) that

$$\|q\|^2 = qq^T = \left[(p^T,0)+\frac{z_0-\gamma}{1+\beta}(e^T-d^T,1)-\frac{z_0}{n+1}(e,1)\right]\left[\begin{pmatrix}p\\0\end{pmatrix}+\frac{z_0-\gamma}{1+\beta}\begin{pmatrix}e-d\\1\end{pmatrix}-\frac{z_0}{n+1}\begin{pmatrix}e\\1\end{pmatrix}\right]$$

$$= p^Tp+\frac{z_0-\gamma}{1+\beta}\left(p^T(e-d)\right)-\frac{z_0}{n+1}(p^Te)+\frac{z_0-\gamma}{1+\beta}\left[\left(e^T-d^T\right)p\right]+\left(\frac{z_0-\gamma}{1+\beta}\right)^2\left[\left(e^T-d^T\right)(e-d)+1\right]$$

$$-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left[(e^T-d^T)e+1\right]-\frac{z_0}{n+1}(e^Tp)-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left[e^T(e-d)+1\right]-+\left(\frac{z_0}{n+1}\right)^2(e^Te+1)$$

$$= p^Tp+\frac{z_0-\gamma}{1+\beta}(p^Te-p^Td)-\frac{z_0}{n+1}(p^Te)+\frac{z_0-\gamma}{1+\beta}\left[\left(p^Te\right)^T-\left(p^Td\right)^T\right]+\left(\frac{z_0-\gamma}{1+\beta}\right)^2\left(\|e-d\|^2+1\right)$$

$$-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left[e^Te-\left(e^Td\right)^T+1\right]-\frac{z_0}{n+1}\left(p^Te\right)^T-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left[e^Te-e^Td+1\right]+\left(\frac{z_0}{n+1}\right)^2(e^Te+1)$$

$$=\|p\|^2-\frac{z_0\gamma}{n+1}+\left(\frac{z_0-\gamma}{1+\beta}\right)^2\left(n-\|d\|^2+1\right)-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left(n-\|d\|^2+1\right)-\frac{z_0\gamma}{n+1}-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}\left(n-\|d\|^2+1\right)$$

$$+\left(\frac{z_0}{n+1}\right)^2(n+1)$$

$$=\|p\|^2-\frac{z_0\gamma}{n+1}+\left(\frac{z_0-\gamma}{1+\beta}\right)^2(1+\beta)-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}(1+\beta)-\frac{z_0\gamma}{n+1}-\frac{z_0(z_0-\gamma)}{(n+1)(1+\beta)}(1+\beta)+\frac{z_0^2}{n+1}$$

$$=\|p\|^2-\frac{z_0\gamma}{n+1}+\frac{(z_0-\gamma)^2}{1+\beta}-\frac{z_0(z_0-\gamma)}{n+1}-\frac{z_0\gamma}{n+1}-\frac{z_0(z_0-\gamma)}{n+1}+\frac{z_0^2}{n+1}$$

$$=\|p\|^2-\frac{2z_0\gamma}{n+1}+\frac{(z_0-\gamma)^2}{1+\beta}-\frac{2z_0(z_0-\gamma)}{n+1}+\frac{z_0^2}{n+1}$$

$$=\|p\|^2+\frac{(z_0-\gamma)^2}{1+\beta}-\frac{z_0^2}{n+1}$$

and using $cDe = cx_0 = z_0$ we also have

$$(cD,0)q = (cD,0)\begin{pmatrix}p\\0\end{pmatrix}+\frac{z_0-\gamma}{1+\beta}\begin{pmatrix}e-d\\1\end{pmatrix}-\frac{z_0}{n+1}\begin{pmatrix}e\\1\end{pmatrix}$$

$$= cDp+\frac{z_0-\gamma}{1+\beta}\left[cD(e-d)\right]-\frac{z_0}{n+1}(cDe)$$

$$= cDp+\frac{z_0-\gamma}{1+\beta}\left[cDe-cDd)\right]-\frac{z_0}{n+1}(cDe)$$

$$= \|p\|^2+\frac{(z_0-\gamma)^2}{1+\beta}-\frac{z_0^2}{n+1}$$

Thus we have

$$\|q\|^2 = \|p\|^2 + \frac{(z_0-\gamma)^2}{1+\beta} - \frac{z_0^2}{n+1} = (cD,0)q$$

Hence it follows from Theorem 3.1 that the solution to (ALP $_\rho$) is given by

$$y^K(\rho) = y^0 - \frac{\rho q}{\|q\|}. \tag{3.2.6}$$

Multiplying (3.2.4) by vector (c D, 0) we get

$$\begin{aligned}
(cD,0)y^K(\rho) &= (cD,0)\left[y^0 - \frac{\rho q}{\|q\|}\right] \\
&= (cD,0)y^0 - \frac{\rho}{\|q\|}(cD,0)q \\
&= \frac{(cD,0)}{n+1}f - \rho\|q\| \quad \text{using } y^0 = \frac{1}{n+1}\mathbf{f} \text{ and } (cD,0)q = \|q\|^2 \\
&= \left(\frac{1}{n+1}\right)(cD,0)\binom{e}{1} - \rho\|q\| = \left(\frac{1}{n+1}\right)(cDe) - \rho\|q\| \\
&= \left(\frac{1}{n+1}\right)(cx^0) - \rho\|q\| = \frac{z_0}{n+1} - \rho\|q\|
\end{aligned}$$

That is

$$(cD,0)y^K(\rho) = \frac{z_0}{n+1} - \rho\|q\| \tag{3.2.7}$$

From the second part of theorem 3.1 we have that

$$\frac{(cD,0)y^K(\rho)}{(cD,0)y^0} \le 1 - \rho\sqrt{\frac{n+1}{n}} \tag{3.2.8}$$

Since we can write

$$y^K(\rho) = \begin{pmatrix} y_N^K \\ y_{n+1}^K \end{pmatrix}$$

using (3.2.4) we have that

$$\begin{pmatrix} y_N^K \\ y_{n+1}^K \end{pmatrix} = \begin{pmatrix} \frac{1}{n}e - \frac{\rho}{\|q\|}\left(p + \frac{z_{0-\gamma}}{1+\beta}(e-d) - \frac{z_0}{n+1}e \right) \\ \frac{1}{n} - \frac{\rho}{\|q\|}\left(\frac{z_{0-\gamma}}{1+\beta}) - \frac{z_0}{n+1} \right) \end{pmatrix}$$

from which it follows that

$$y_N^K(\rho) = \frac{1}{n}e - \frac{\rho}{\|q\|}\left(p + \frac{z_{0-\gamma}}{1+\beta}(e-d) - \frac{z_0}{n+1}e \right) \quad \text{and}$$

$$y_{n+1}^K(\rho) = \frac{1}{n} - \frac{\rho}{\|q\|}\left(\frac{z_{0-\gamma}}{1+\beta} - \frac{z_0}{n+1} \right).$$

Reversing the projective transformation (3.1.2) we have

$$\begin{aligned} x^K(\rho) &= \frac{1}{y_{n+1}^K(\rho)} D\, y_{n+1}^K(\rho) \\ &= \frac{\frac{1}{n}e - \frac{\rho}{\|q\|}\left(p + \frac{z_{0-\gamma}}{1+\beta}(e-d) - \frac{z_0}{n+1}e \right)}{\frac{1}{n} - \frac{\rho}{\|q\|}\left(\frac{z_{0-\gamma}}{1+\beta} - \frac{z_0}{n+1} \right)} \\ &= \frac{(1+\beta)\|q\|De - (1+\beta)(n+1)D(p - \frac{z_0-\gamma}{1+\beta}d) + \rho(\gamma(n+1) - (n-\beta)z_0)De}{(1+\beta)\|q\| + \rho(\gamma(n+1) - (n-\beta)z_0)} \\ &= De - \frac{(1+\beta)(n+1)\rho D(p - \frac{z_0-\gamma}{1+\beta}d)}{(1+\beta)\|q\| + \rho(\gamma(n+1) - (n-\beta)z_0)} \\ &= x^0 - \frac{(1+\beta)(n+1)\rho D(p - \frac{z_0-\gamma}{1+\beta}d)}{(1+\beta)\|q\| + \rho(\gamma(n+1) - (n-\beta)z_0)} \end{aligned}$$

Let us denote

$$t(\rho) = \frac{(1+\beta)(n+1)\rho}{(1+\beta)\|q\| + \rho(\gamma(n+1) - (n-\beta)z_0)} \tag{3.2.9}$$

Then we have

$$x^K(\rho) = x^0 - t(\rho)D\left(p - \frac{z_0 - \gamma}{1+\beta}d\right) \tag{3.2.10}$$

and that the objective function value of $x^K(\rho)$ becomes

$$\begin{aligned} cx^K(\rho) &= cx^0 - t(\rho)cD\left(p - \frac{z_0 - \gamma}{1+\beta}d\right) \\ &= cx^0 - t(\rho)\left[cDp - \frac{z_0 - \gamma}{1+\beta}(cDd)\right] \\ &= z_0 - t(\rho)\left[\|p\|^2 - \left(\frac{\gamma(z_0 - \gamma)}{1+\beta}\right)\right] \end{aligned}$$

That is

$$cx^K(\rho) = z_0 - t(\rho)\left[\|p\|^2 - \left(\frac{\gamma(z_0 - \gamma)}{1+\beta}\right)\right] \tag{3.2.11}$$

It follows from its derivation that

$$t(\rho) = \rho\left(\|q\|y_{n+1}^K(\rho)\right)^{-1}$$

Consequently, from $y_{n+1}^K > 0$ we get that

$$t(\rho) = \rho\left(\|q\|y_{n+1}^K(\rho)\right)^{-1} \geq 0 \ \ for \ \ all \ \ 0 \leq \rho < r = \frac{1}{\sqrt{n(n+1)}}.$$

From (3.2.9) we have since $\|q\| \neq 0$

$$\frac{d}{d\rho}t(\rho) = \frac{(1+\beta)^2(n+1)\|q\|}{\left[(1+\beta)\|q\| + \rho(\gamma(n+1) - (n-\beta)z_0)\right]^2} > 0 \ \ for \ \ all \ \ 0 \leq \rho < r.$$

Thus, the change of parameter from ρ to $t(\rho)$ is a change that preserves strict monotonicity.

3.2.2. Convergence of the approximate iterates

Like the solution (3.2.6) to (ALP_ρ), the loci of $x^K(\rho)$ given by (3.2.10) forms a line in R^n (which is the consequences of the fact that projective transformations always maps lines into lines). From (3.2.11) we infer that depending on the sign of the term in the square bracket the objective function value $cx^K(\rho)$ may increase or decrease along this line contrary to the monotonic behavior of $(cD,0)y^K(\rho)$ which decreases linearly in ρ.

Using (3.2.4) and (3.2.5) we have

$$(cD,-z_0)q = (cD,-z_0)\left[\begin{pmatrix} p \\ 0 \end{pmatrix} + \frac{z_0-\gamma}{1+\beta}\begin{pmatrix} e-d \\ 1 \end{pmatrix} - \frac{z_0}{n+1}\begin{pmatrix} e \\ 1 \end{pmatrix}\right]$$

$$= cDp + \frac{z_0-\gamma}{1+\beta}(cDe - cDd - z_0) - \frac{z_0}{n+1}(cDe - z_0)$$

$$= cDp - \frac{\gamma(z_0-\gamma)}{1+\beta} = \|p\|^2 - \frac{\gamma(z_0-\gamma)}{1+\beta}$$

Thus (3.2.11) becomes

$$cx^K(\rho) = z_0 - t(\rho)(cD,-z_0)q \tag{3.2.12}$$

Taking the derivative of (3.2.12) gives

$$\frac{d}{d\rho}(cx^K(\rho) = -(cD,-z_0)q\frac{d}{d\rho}t(\rho)$$

Since $\frac{d}{d\rho}t(\rho) > 0$ *for all* $0 \le \rho < r$ we have

$$\frac{d}{d\rho}(cx^K(\rho) \ge 0 \text{ if and only if } (cD,-z_0)q \le 0.$$

This implies that the objective function value $cx^K(\rho)$ increases if the vector $(cD,-z_0)^T$ and the orthogonal projection q of $(cD,0)^T$ on to the subspace (3.2.1) form an obtuse angle.

Since **p** and **d** are the orthogonal projections of the vectors **c** and **e** onto the subspace (3.2.3) we have

$$ADp = ADd = 0$$

Also since $x^0 \in \chi$, $Ax^0 = b$
Consequently, from (3.2.10) we get

$$Ax^K(\rho) = A\left[x^0 - t(\rho)D\left(p - \frac{z_0 - \gamma}{1+\beta}d\right)\right] = Ax^0 - t(\rho)\left[ADp - \frac{z_0 - \gamma}{1+\beta}(ADd)\right] = b.$$

Hence by construction $x^K(\rho) \in \chi$ and $x^K(\rho) > 0$ *for all* $0 \le \rho < r$, and thus $x^K(\rho)$ can serve as a **new "iterate"** in an algorithmic application of the basic idea.

In the following section we prove the convergence of the sequence of points generated by an iterative application of the basic idea.
But first let us consider the following direct estimation.
We have using (3.1.2) and from (3.2.8)

$$\frac{cx^K(\rho)}{cx^0} = \frac{(cD,0)y^K(\rho)}{(cD,0)y^0}\frac{y_{n+1}^0}{y_{n+1}^K(\rho)} \le \left(1 - \rho\sqrt{\frac{n+1}{n}}\right)\frac{1}{(n+1)y_{n+1}^K(\rho)}.$$

To estimate the last term we need to know the minimum value of y_{n+1} for all $y \in B_\rho^{n+1}$.
Using the Lagrange Multiplier technique we calculate that

$$(n+1)y_{n+1} \ge 1 - \rho\sqrt{n(n+1)} \ \ \textit{for all} \ \ y \in B_\rho^{n+1}$$

and from the above estimation

$$\frac{cx^K(\rho)}{cx^0} \le \left(1 - \rho\sqrt{\frac{n+1}{n}}\right)\left(1 - \rho\sqrt{n(n+1)}\right)^{-1}.$$

This estimation is of no use to us since the term on the right side is greater than or equal to one for all $0 \le \rho < \frac{1}{\sqrt{n(n+1)}}$, but this shows that $cx^K(\rho)$ cannot increase by "too much" if ρ is small. It also shows us that a "direct' estimation of the relative change $\frac{cx^K(\rho)}{cx^0}$ does not work.

To prove the convergence of the sequence of points generated by the above method, it suffices, of course, to show that some measure other than the objective function gives a sufficiently large decrease.

To this end, we consider the function

$$h(x) = \mathbf{c}x\left(\prod_{j=1}^{n} x_j\right)^{\frac{-1}{n+1}},$$

which is the objective function divided by the geometric means of the point $(x,1) \in R^{n+1}$. The function h(x) is well defined for all $x > 0$.

Using the geometric /arithmetic mean inequality we have that

$$h(x) \geq (n+1)cx\left(1+\sum_{j=1}^{n} x_j\right)^{-1} \geq 0 \tag{3.2.13}$$

By the additional assumption (iii) χ is bounded.
So,

$$\sum_{j=1}^{n} x_j \leq K \quad \text{for all } x \in \chi \text{ where } K > 0 \text{ is some constant.}$$

From (3.2.13) we obtain

$$h(x) \geq (n+1)cx\left(1+\sum_{j=1}^{n} x_j\right)^{-1} \geq (n+1)cx(1+K)^{-1} = \left(\frac{n+1}{1+K}\right)cx$$

Hence minimizing $h(x)$ achieves the goal of minimizing the objective function cx.

Now we calculate using (3.2.6)

$$\frac{h(x^K(\rho))}{h(x^0)} = \frac{cx^K(\rho)\left(\prod_{j=1}^{n} x_j^K(\rho)\right)^{\frac{-1}{n+1}}}{cx^0\left(\prod_{j=1}^{n} x_j^0(\rho)\right)^{\frac{-1}{n+1}}} = \frac{cx^K(\rho)}{cx^0}\left(\left(\prod_{j=1}^{n}\frac{x_j^0}{x_j^K(\rho)}\right)^{\frac{1}{n+1}}\right)$$

$$= \frac{(cD,0)y^K(\rho)}{(cD,0)y^0} \frac{y_{n+1}^0}{y_{n+1}^K(\rho)} \left(\prod_{j=1}^{n} \frac{y_{n+1}^K(\rho)}{y_j^K(\rho)} \right)^{\frac{1}{n+1}} = \frac{(cD,0)y^K(\rho)}{(cD,0)y^0} \frac{1}{(n+1)y_{n+1}^K(\rho)} \left(\prod_{j=1}^{n} \frac{y_{n+1}^K(\rho)}{y_j^K(\rho)} \right)^{\frac{1}{n+1}}$$

$$= \frac{(cD,0)y^K(\rho)}{(cD,0)y^0} \left(\prod_{j=1}^{n} \frac{1}{(n+1)y_j^K(\rho)} \right)^{\frac{1}{n+1}} = \frac{(cD,0)y^K(\rho)}{(cD,0)y^0} \left(\prod_{j=1}^{n} (n+1)y_j^K(\rho) \right)^{\frac{-1}{n+1}}$$

$$\leq \left(1 - \rho\sqrt{\frac{n+1}{n}} \right) \left(\prod_{j=1}^{n} (n+1)y_j^K(\rho) \right)^{\frac{-1}{n+1}}$$

That is

$$\frac{h(x^K(\rho))}{h(x^0)} \leq \left(1 - \rho\sqrt{\frac{n+1}{n}} \right) \left(\prod_{j=1}^{n} (n+1)y_j^K(\rho) \right)^{\frac{-1}{n+1}} \tag{3.2.14}$$

Now we are left with estimating the last term for $y \in B_\rho^{n+1}$.

In order to do so, we set $\rho = \alpha r$ *where* $0 \leq \alpha \leq 1$ and prove the following theorem where we change our notation to simplify the exposition:

Theorem 3.2.1: Let $\rho = \alpha r$ where $r^2 = \frac{1}{n(n-1)}$. Then for all $0 < \alpha < 1$

$$\max \left\{ \left(\prod_{j=1}^{n} n z_j \right)^{\frac{-1}{n}} : z \in B_\rho^n \right\} = \left(\frac{1 + \frac{\alpha}{n-1}}{1-\alpha} \right)^{\frac{1}{n}} \left(1 + \frac{\alpha}{n-1} \right)^{-1}$$

Proof:

Since B_ρ^n is a compact subset of R^n and z > 0 for all $z \in B_\rho^n$ and for all $0 < \rho < r$ the maximum exists. The assertion is equivalent to proving

$$\min \left\{ \prod_{j=1}^{n} z_j : \sum_{j=1}^{n} z_j^2 = \alpha^2 r^2 + \frac{1}{n}, \sum_{j=1}^{n} z_j = 1 \right\} = \left(\frac{1-\alpha}{n^n} \right) \left(1 + \frac{\alpha}{n-1} \right)^{n-1}$$

Using the Lagrange multiplier technique we get the equation

$\mu z_j^2 + \nu z_j + 1 = 0 \; for \;\; j = 1,\ldots,n$

after absorbing the product term into the multiplier s μ and ν .

It follows that $\mu \neq 0$ since otherwise $z_j = \frac{1}{n}$ for all j which contradicts $\alpha > 0$.

Consequently, the components of every solution z to the Lagrangean equation are of the form $z_j = a \pm b$ where a and b are scalars satisfying

$0 < b < a$ since

$z > 0$. Hence because of the symmetries in both in the objective function and the constraints we can assume w. l. o. g. that in the natural indexing there exists $\ell \in \{1,\ldots,n-1\}$ such that

$$z_1 = \ldots = z_\ell < z_{\ell+1} = \ldots = z_n.$$

From the constraint $\sum_{j=1}^{n} z_j = 1$ we get

$$z_1 + \ldots + z_\ell + z_{\ell+1} + \ldots + z_n = 1$$

By taking $z_j = a - b \; for \; 1 \le j \le \ell \;\; and \; z_j = a + b \; for \; j > \ell$ we have

$$\ell(a-b) + (n-\ell)(a+b) = 1$$

or

$$na + (n-2\ell)b) = 1 \tag{i}$$

And from the other constraint $\sum_{j=1}^{n} z_j^2 \;\; = \alpha^2 r^2 + \frac{1}{n}$ we get

$$\ell(a-b)^2 + (n-\ell)(a+b)^2 = \alpha^2 r^2 + \frac{1}{n}$$

which is equivalent to

$$na^2 + nb^2 + 2nab - 4\ell ab = \alpha^2 r^2 + \frac{1}{n} \tag{ii}$$

Squaring (i) we get

$$n^2a^2 + n^2b^2 + 2nab - 4n\ell ab - 4n\ell ab^2 + 4\ell^2 b^2 = 1 \tag{iii}$$

Multiplying (ii) by n and comparing the result with (iii) we obtain

$$\alpha^2 r^2 n - 4n\ell b^2 + 4\ell^2 b^2 = 0$$

from which it follows that

$$b = \frac{\alpha}{2}\sqrt{\frac{1}{\ell(n-1)(n-\ell)}} \quad \text{since b > 0.}$$

Substituting the value of b and solving for a from (i) we get

$$a = \frac{1}{n} - \frac{\alpha}{2n}(n-2\ell)\sqrt{\frac{1}{\ell(n-1)(n-\ell)}} \quad \text{or} \quad a = \frac{1}{n}\left(1 - (n - 2\ell)b\right)$$

Substituting the values of a and b obtained above and using $z_j = a-b \; for \; 1 \le j \le \ell \; and \; z_j = a+b \; for \; j > \ell$ we get after some simplifications that

$$z_j = \frac{1}{n}\left(1 - \alpha\sqrt{\frac{n-\ell}{\ell(n-1)}}\right) \; for \; 1 \le j \le \ell \qquad \text{and}$$

$$z_j = \frac{1}{n}\left(1 + \alpha\sqrt{\frac{\ell}{(n-1)(n-\ell)}}\right) \; for \; j > \ell$$

Now let us denote

$$\xi_\ell(\alpha) = \left(1 - \alpha\sqrt{\frac{n-\ell}{\ell(n-1)}}\right)^{\ell}\left(1 + \alpha\sqrt{\frac{\ell}{(n-1)(n-\ell)}}\right)^{n-\ell}.$$

Then we have

$$\xi_1(\alpha) = (1-\alpha)\left(1 + \frac{\alpha}{n-1}\right)^{n-1}$$

We need to show that $\xi_\ell(\alpha) \le \xi_1(\alpha) \; for \; all \;\; 0 \le \alpha \le 1 \; and \; \ell \ge 2.$

Here we use indirect proof.

Suppose to the contrary that $\xi_\ell(\alpha) < \xi_1(\alpha) \;\; for \;\; some \;\; 0 < \alpha < 1 \;\; and \;\; \ell \ge 2.$

Since $\xi_1(0) = \xi_\ell(0) = 1, \xi_1(1) = 0 \; and \; \xi_\ell(1) > 0$ the function $\xi_\ell(\alpha) - \xi_1(\alpha)$ has a minimum in the interval [0,1]. For k∈{1,...,n-1}we have

$$\frac{d}{d\alpha}(\xi_k(\alpha)) = -k\sqrt{\frac{n-k}{k(n-1)}}\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)^{k-1}\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)^{n-k}$$

$$+(n-k)\sqrt{\frac{k}{(n-k)(n-1)}}\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)^{n-k-1}\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)$$

$$= \frac{-k\sqrt{\frac{n-k}{k(n-1)}}\xi_k(\alpha)}{1-\alpha\sqrt{\frac{n-k}{k(n-1)}}} + \frac{(n-k)\sqrt{\frac{k}{(n-k)(n-1)}}\xi_k(\alpha}{1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}}$$

$$= \frac{\left[\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)\left(-\sqrt{\frac{k(n-k)}{n-1}}\right)+\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)\sqrt{\frac{k(n-k)}{n-1}}\right]\xi_k(\alpha)}{\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)}$$

$$= \frac{\alpha n\xi_k(\alpha)}{(n-1)\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)}$$

after some simplifications.

Similarly, we have

$$\frac{d}{d\alpha}(\xi_1(\alpha)) = \frac{-\alpha n\xi_1(\alpha)}{(n-1)(1-\alpha)\left(1+\frac{\alpha}{n-1}\right)}$$

At a minimum of $\xi_\ell(\alpha)-\xi_1(\alpha)$ we have

$$\frac{d}{d\alpha}(\xi_\ell(\alpha)-\xi_1(\alpha))=0$$

That means

$$\frac{d}{d\alpha}(\xi_\ell(\alpha)=\frac{d}{d\alpha}(\xi_1(\alpha))$$

which is equivalent to

$$\frac{\alpha n\xi_k(\alpha)}{(n-1)\left(1-\alpha\sqrt{\frac{n-k}{k(n-1)}}\right)\left(1+\alpha\sqrt{\frac{k}{(n-k)(n-1)}}\right)} = \frac{-\alpha n\xi_1(\alpha)}{(n-1)(1-\alpha)\left(1+\frac{\alpha}{n-1}\right)}$$

Thus we have

$$\frac{\xi_1(\alpha)}{\xi_\ell(\alpha)} = \frac{(1-\alpha)\left(1+\dfrac{\alpha}{n-1}\right)}{\left(1-\alpha\sqrt{\dfrac{n-\ell}{\ell(n-1)}}\right)\left(1+\alpha\sqrt{\dfrac{\ell}{(n-\ell)(n-1)}}\right)}$$

So, for $\ell \geq 2$ *and* $0 < \alpha < 1$ we have $\dfrac{\xi_1(\alpha)}{\xi_\ell(\alpha)} < 1$ and thus $\xi_1(\alpha) < \xi_\ell(\alpha)$ since $\xi_\ell(\alpha) > 0$, which is the contradiction. //

Using Theorem 3.2.1, (3.1.14) becomes

$$\frac{h(x^K(\rho))}{h(x^0)} \leq \left(1-\rho\sqrt{\frac{n+1}{n}}\right)\left(\prod_{j=1}^{n}(n+1)y_j^K(\rho)\right)^{\frac{-1}{n+1}}$$

For $\rho = \alpha r$ we get

$$\begin{aligned}
\frac{h(x^K(\alpha r))}{h(x^0)} &\leq \left(1-\alpha r\sqrt{\frac{n+1}{n}}\right)\left(\prod_{j=1}^{n}(n+1)y_j^K(\alpha r)\right)^{\frac{-1}{n+1}} \\
&\leq \left(1-\alpha r\sqrt{\frac{n+1}{n}}\right)\left(1+\frac{\alpha}{n}\right)^{-1}\left(\frac{1+\dfrac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} \\
&= \left(1-\alpha\sqrt{\frac{1}{n(n+1)}}\sqrt{\frac{n+1}{n}}\right)\left(1+\frac{\alpha}{n}\right)^{-1}\left(\frac{1+\dfrac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} \\
&= \left(1-\frac{\alpha}{n}\right)\left(1+\frac{\alpha}{n}\right)^{-1}\left(\frac{1+\dfrac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} = \left(\frac{1-\dfrac{\alpha}{n}}{1+\dfrac{\alpha}{n}}\right)\left(\frac{1+\dfrac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}}
\end{aligned}$$

Therefore, we have

$$\frac{h(x^K(\alpha r))}{h(x^0)} \leq \left(\frac{1-\dfrac{\alpha}{n}}{1+\dfrac{\alpha}{n}}\right)\left(\frac{1+\dfrac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} \qquad (3.2.15)$$

Let us denote the right –hand side of (3.2.15) by

$$\tilde{g}(\alpha, n) = \left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} \tag{3.2.16}$$

Now we estimate $\tilde{g}(\alpha,n)$ conveniently as follows.

We recall that $1-x \le e^{-x}$ for all x.

Define the function $f(x) = e^{-2x} - \dfrac{1-x}{1+x}$.

Since $e^{x} \ge 1+x$ we have for all $x \ge 0$

$$f'(x) = -2\left[\frac{1}{(x+1)^2} - \frac{1}{(e^x)^2}\right] = \left(\frac{1}{x+1} - \frac{1}{e^x}\right)\left(\frac{1}{x+1} + \frac{1}{e^x}\right) \ge 0.$$

This implies f is an increasing function and from $f(0) = 0$ we have

$$\frac{1-x}{1+x} \le e^{-2x} \; \textit{for all } x \ge 0. \tag{3.2.17}$$

Consequently, we have

$$\tilde{g}(\alpha,n) = \left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}} = \left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(1+\frac{\alpha}{n}\right)^{\frac{1}{n+1}}(1-\alpha)^{\frac{-1}{n+1}}$$

Since we can write $(1-\alpha)^{\frac{-1}{n+1}} = (1-\alpha)^{\frac{1}{n(n+1)} - \frac{1}{n}} = \dfrac{(1-\alpha)^{\frac{1}{n(n+1)}}}{(1-\alpha)^{\frac{1}{n}}}$ using (3.2.17)

$$\tilde{g}(\alpha,n) = \frac{\left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(1+\frac{\alpha}{n}\right)^{\frac{1}{n+1}}(1-\alpha)^{\frac{1}{n(n+1)}}}{(1-\alpha)^{\frac{1}{n}}} \le \frac{\left(e^{\frac{-2\alpha}{n}}\right)\left(e^{\frac{\alpha}{n}}\right)^{\frac{1}{n+1}}\left(e^{\frac{-\alpha}{n}}\right)^{\frac{1}{n+1}}}{(1-\alpha)^{\frac{1}{n}}}$$

$$= \frac{e^{\frac{-2\alpha}{n} + \frac{\alpha}{n(n+1)} - \frac{\alpha}{n(n+1)} -}}{(1-\alpha)^{\frac{1}{n}}} = \frac{e^{\frac{-2\alpha}{n}}}{(1-\alpha)^{\frac{1}{n}}} = \left(\frac{e^{-2\alpha}}{1-\alpha} \right)^{\frac{1}{n}}$$

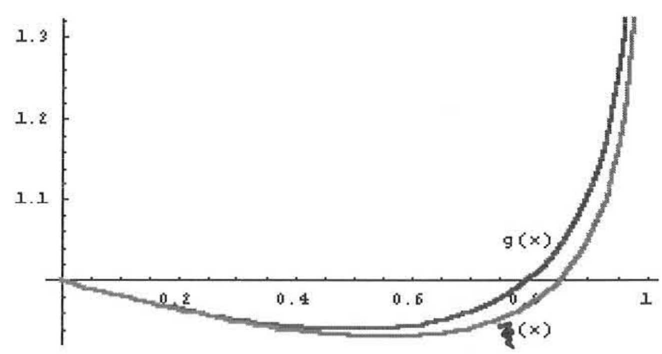
That is

$$\tilde{g}(\alpha, n) \le \left(\frac{e^{-2\alpha}}{1-\alpha} \right)^{\frac{1}{n}} \tag{3.2.18}$$

Let us denote the right side of (3.2.18) by

$$g(\alpha, n) = \left(\frac{e^{-2\alpha}}{1-\alpha} \right)^{\frac{1}{n}} \tag{3.2.19}$$

Now we can illustrate the behavior of the functions $\tilde{g}$ *and* g graphically (for n=5) as shown below.



As we can see from the graph g crosses the line y =1 for $\alpha = 0.796812...$ (using Mathematica).

Therefore, we have

$$g(\alpha, n) < 1 \quad \textit{for all} \quad 0 < \alpha < \alpha_0 = 0.7968 \ldots .$$

Thus - despite the nonmonotonicity of $cx^K(\rho)$ - we get a decrease in the auxiliary test function h(x) that does not depend up on the initial interior point x^0.

Consequently, the iterative application of the algorithmic idea produces a geometric convergence rate in terms of the test function h(x) and for any fixed "step- size" α satisfying

$$0 < \alpha < \alpha_0 = 0.7968 \ldots .$$

Now we are ready to summarize the basic steps of a projective algorithm. The following inputs are assumed:

(i) the descriptive data ,(m ,n, A, c),of the problem (LP);

(ii) a feasible interior starting point $x^0 > 0$ such that $cx^0 > 0$;

(iii) input parameters, α for "step size", p for the desired "precision" in terms of the relative error.

(Note that the optimum of (LP) is assumed to be zero.)

Basic Algorithm (α, p, m, n, **A, c, x⁰)**

Step 0: Set $\mathbf{D_0 = (x_1^0, \ldots, x_n^0)}$, $\mathbf{z = cx^0}$, and $\mathbf{k = 0}$

Step 1: Compute

$$\mathbf{G := AD_k^2A^T \;, G^{-1} \;\; \text{and} \;\; P := I_n - D_kA^TG^{-1}AD_k}$$

Step 2: Compute

$$\mathbf{p := PD_kc^T \;, d = Pe \;, \gamma = p^Td \;, \beta = n - \|d\|^2} \;,$$

$$\mathbf{\|q\| := \sqrt{\|p\|^2 + \frac{(z-\gamma)^2}{1+\beta} - \frac{z^2}{n+1}}} \qquad \textbf{and}$$

$$\mathbf{t := \frac{\alpha(1+\beta)(1+n)}{(1+\beta)\sqrt{n(n+1)}\,\|q\| + \alpha(\gamma(n+1) - (n-\beta)z)}}$$

Step 3: Set

$$\mathbf{x^{k+1} := x^k - tD_k\left(p - \frac{z-\gamma}{1+\beta}d\right)},$$

$$\mathbf{D_{k+1} := diag(x_1^{k+1}, \ldots, x_n^{k+1})}$$

Step 4: if $\mathbf{\frac{cx^{k+1}}{cx^0} < 2^{-p}}$, stop "$\mathbf{x^{k+1}}$ is a p–optimal solution to (LP)".

Set $\mathbf{Z := cx^{k+1}}$; replace **k+1** by **k**; go to step 1.

It is noteworthy that the right hand side vector b of (LP) has " disappeared " from the input of the basic algorithm: given the matrix A its " informational content " has been absorbed into the starting point x⁰ of the algorithm.

3.2.3. Correctness and finiteness of Basic Algorithm

In this section we shall show that the projective algorithm is polynomially bounded.

Theorem 3.3: Let $K > 0$ be such that

$$\chi \subseteq \left\{ x \in R^{n} : 0 \leq x_{j} \leq K \;\; for \;\; j = 1,\ldots, n \right\}.$$

Then for every $0 < \alpha < 0.7968...$ and $p \geq \log_2^K$ the basic algorithm iterates at most $\mathcal{O}(n\,p)$ times.

Proof: By construction the sequence of points generated by the algorithm satisfies $x^k > 0$ since $0<\alpha<1$ and $Ax^k = b$.
Suppose that the algorithm executes $k \geq 1$ iterations.

Now we estimate $\dfrac{cx^{k}}{cx^{0}}$.

$$\frac{cx^{k}}{cx^{0}} = \frac{h(x^{k})}{h(x^{k-1})} \cdots\cdots \frac{h(x^{1})}{h(x^{01})} \left(\prod_{j=1}^{n} \frac{x_j^k}{x_j^0} \right)^{\frac{1}{n+1}} \leq g(\alpha, n) \ldots\ldots g(\alpha, n) \left(\prod_{j=1}^{n} \frac{x_j^k}{x_j^0} \right)^{\frac{1}{n+1}} \qquad by \;\; (3.2.9)$$

$$= \left(g(\alpha,n)\right)^{k} \left(\prod_{j=1}^{n} \frac{x_j^k}{x_j^0} \right)^{\frac{1}{n+1}} = \left(g(\alpha,n)\right)^{k} \frac{\left(\prod_{j=1}^{n} x_j^k \right)^{\frac{1}{n+1}}}{\left(\prod_{j=1}^{n} x_j^0 \right)^{\frac{1}{n+1}}}$$

$$\leq \left(g(\alpha,n)\right)^{k} \frac{\left(1 + \sum_{j=1}^{n} x_j^k \right)}{\left(\prod_{j=1}^{n} x_j^0 \right)^{\frac{1}{n+1}}} \qquad (by \;\; geometric \;\; / \; arthmetic \;\; inequality \;)$$

$$= \frac{\left(\frac{e^{-2\alpha}}{1-\alpha}\right)^{\frac{k}{n}}\left(1+\sum_{j=1}^{n} x_j^k\right)}{\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}}$$

$$\le K\theta^{\frac{k}{n}}\left(\prod_{j=1}^{n} x_j^0\right)^{-\frac{1}{n+1}}, \qquad \textit{where} \quad 0<\theta=\frac{e^{-2\alpha}}{1-\alpha}<1.$$

$$=\mu$$

If $k > \dfrac{n\left(-2p+\left(\frac{1}{n+1}\right)\sum_{j=1}^{n}\log_2^{x_j^0}\right)}{\log_2^{\theta}}$ then we have

$$\mu = K\theta^{\left(\frac{-2p+\left(\frac{1}{n+1}\right)\sum_{j=1}^{n}\log_2^{x_j^0}}{\log_2^{\theta}}\right)}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$=K\theta^{\left(\frac{-2p}{\log_2^{\theta}}\right)}\theta^{\left(\frac{\left(\frac{1}{n+1}\right)\sum_{j=1}^{n}\log_2^{x_j^0}}{\log_2^{\theta}}\right)}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$= K\theta^{\left(-2p\log_\theta^2\right)}\theta^{\left(\log_\theta^{\left(\sum_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}}\right)}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$= K\left(2^{-2p}\right)\left(\sum_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

By assumption, $p \ge \log_2^K$ which is equivalent to

$$\log_2^{2^p} \geq \log_2^K \quad or \quad 2^p \geq K .$$

Thus we have

$$\mu \leq K\left(2^{-2p}\right)\left(\sum_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$\leq \left(2^p\right)\left(2^{-2p}\right)\left(\sum_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$= 2^{-p}\left(\sum_{j=1}^{n} x_j^0\right)^{\frac{1}{n+1}}\left(\prod_{j=1}^{n} x_j^0\right)^{\frac{-1}{n+1}}$$

$$< 2^{-p}$$

by geometric/arithmetic inequality.
Consequently, the basic algorithm stops after at most $\mathcal{O}(n\,p)$ iterations. //

The best "step" size α that the analysis suggests is evidently the value of α that minimizes the function $\tilde{g}(\alpha,n)$ in the interval [0,1].

From (3.2.16) we get

$$\frac{d}{d\alpha}(\tilde{g}(\alpha,n)) = \left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}}\left[\frac{-\frac{1}{n}\left((1+\frac{\alpha}{n})+(1-\frac{\alpha}{n})\right)}{\left(1+\frac{\alpha}{n}\right)^2}\right] + \frac{1}{n+1}\left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{-n}{n+1}}\left[\frac{\frac{1}{n}(1-\alpha)+\left(1+\frac{\alpha}{n}\right)}{\left(1-\frac{\alpha}{n}\right)^2}\right]$$

$$= \left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{1}{n+1}}\left(\frac{-2}{n\left(1+\frac{\alpha}{n}\right)^2}\right) + \frac{1}{n\left(1-\frac{\alpha}{n}\right)^2}\left(\frac{1-\frac{\alpha}{n}}{1+\frac{\alpha}{n}}\right)\left(\frac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{-n}{n+1}}$$

$$= \frac{\dfrac{-2}{n(1-\alpha)\left(1+\frac{\alpha}{n}\right)} + \dfrac{1-\frac{\alpha}{n}}{n\left(1+\frac{\alpha}{n}\right)(1-\alpha)^2}}{\left(\dfrac{1+\frac{\alpha}{n}}{1-\alpha}\right)^{\frac{n}{n+1}}} = \frac{-1+\alpha(2-\frac{1}{n})}{n\left(1+\frac{\alpha}{n}\right)^{2n+1}(1-\alpha)^{\frac{n+2}{n+1}}}$$

As a necessary condition for a minimum point of $\tilde{g}(\alpha,n)$ we have

$$\frac{d}{d\alpha}\left(\tilde{g}(\alpha,n)\right)=0 \quad \text{or} \quad \frac{-1+\alpha\left(2-\frac{1}{n}\right)}{n\left(1+\frac{\alpha}{n}\right)^{2n+1}\left(1-\alpha\right)^{\frac{n+2}{n+1}}}=0$$

which gives $-1+\alpha(2-\frac{1}{n})=0$

,i.e. $\alpha = \dfrac{n}{2n-1}$.

Obviously, $\tilde{g}(0,n) = 1 \ \ for \ \ all \ \ n \in N$.

Also

$$\lim_{\alpha \to 1^-} \tilde{g}(\alpha,n) = +\infty \quad for \ \ all \ \ n \geq 2.$$
$$\lim_{n \to \infty} \tilde{g}(\alpha,n) = 1 \ \ for \ \ all \ \ 0 \leq \alpha < 1 \qquad and$$
$$\tilde{g}(\alpha,n) < 1 \ \ for \ \ all \ \ 0 < \alpha < \alpha_0 = 0.7968 \ ....$$

For $\alpha = \dfrac{n}{2n-1}$ we get using the fact $1-x \leq e^{-x}$ for all x that

$$\tilde{g}\left(\frac{n}{2n-1},n\right)=\left(\frac{1-\frac{\frac{n}{2n-1}}{n}}{1+\frac{\frac{n}{2n-1}}{n}}\right)\left(\frac{1+\frac{n}{2n-1}}{1-\frac{n}{2n-1}}\right)^{\frac{1}{n+1}}=\left(\frac{1-\frac{1}{2n-1}}{1+\frac{1}{2n-1}}\right)\left(\frac{1+\frac{1}{2n-1}}{1-\frac{n}{2n-1}}\right)^{\frac{1}{n+1}}$$
$$=\left(\frac{n-1}{n}\right)\left(\frac{2n}{n-1}\right)^{\frac{1}{n+1}}=\left(\frac{n-1}{n}\right)^{\frac{n}{n+1}}2^{\frac{1}{n+1}}=\left(1-\frac{1}{n}\right)^{\frac{n}{n+1}}2^{\frac{1}{n+1}}$$
$$\leq 2^{\frac{1}{n+1}}\left(e^{\frac{-1}{n}}\right)^{\frac{n}{n+1}}=\left(\frac{2}{e}\right)^{\frac{1}{n+1}}$$

Thus- ignoring the dependency upon the constant K and starting point x^0 and estimating the (remaining) relative error directly from (3.2.14) – we get that about $3p(n+1)$ iterations suffice for the basic algorithm to come to a halt, using that $\ln 2 \leq \frac{3}{4}$. This means that the constant in the $\mathcal{O}(n\,p)$ estimation is reasonable for choice of α around the value $\frac{1}{2}$.

When the basic algorithm comes to a halt, we thus have a feasible solution $x \in \chi$ such that $0 \leq cx \leq z_0 2^{-p}$.

Let L be as defined in (2.2.1). Choosing a precision $p > \log_2^{z_0} + L$ for the relative error we get $cx < 2^{-L}$. Hence it follows that $\mathcal{O}(n\,L)$ steps are required by the basic algorithm to come to a halt, i.e. the step complexity of the basic algorithm is polynomially bounded and its optimal objective function equals zero. The time complexity of the basic algorithm is dominated by Step1 and Step 2, where we have to find the projected vectors **p** and **d**.

References

[1] Deumlich, R.: Optimization and Theory of Approximation, Text Book for the Lecture, Addis Ababa, 1997

[2] Deumlich, R.: Functional Analysis I, Text Book for the Lecture, Addis Ababa, 1997

[3] Padberg, M.: Linear Optimization and Extensions, Springer-Verlag Berlin Heidelberg, New York, 1995

[4] Nemhauser, G.L.: Optimization: Handbooks in Operational Research and Management Science, ELSEVIER SCIENCE Publishers B.V., The Netherlands, 1989

[5] Walsh, G.R.: Methods of Optimization, London, New York, Sydney, Toronto, 1975