Part II
Linear Constrained Optimization
Chapter 18
NONSIMPLEX METHODS

Chia-Hui Chang (張嘉惠)

National Central University

2008.12.08

Introduction

- The simplex algorithm for solving LP problem has complexity $O(2^n-1)$, where *n* is the number of variables
- Khachiyan (also translated as Hacijan) proposed an algorithm (called the ellipsoid algorithm) with complexity $O(n^4L)$, where L represents the number of bits used in the computations.
- Another nonsimplex algorithm for solving LP was proposed in 1984 by Karmarkar which has complexity of $O(n^{3.5}L)$.

Khachiyan's Method

- Primal LP + Dual LP

$$\begin{aligned} &\text{minimize } c^T x \\ &\text{subject to } Ax \geq b \\ &\qquad\qquad x \geq 0. \end{aligned} \qquad\qquad \begin{aligned} &\text{maximize } b^T \lambda \\ &\text{subject to } A^T \lambda \leq c \\ &\qquad\qquad \lambda \geq 0. \end{aligned}$$

- Using Theorem 17.1

$$\mathrm{c}^{\mathrm{T}}\mathrm{x} = b^T \lambda \Leftrightarrow \begin{cases} \mathrm{c}^{\mathrm{T}}\mathrm{x} - b^T \lambda \leq 0, \\ -\mathrm{c}^{\mathrm{T}}\mathrm{x} + b^T \lambda \leq 0. \end{cases}$$

$$\begin{aligned} &c^T x = b^T \lambda, \\ &Ax \geq b, \\ &A^T \lambda \leq c, \\ &x \geq 0, \\ &\lambda \geq 0. \end{aligned}$$

$$\underbrace{\begin{bmatrix} c^T & -b^T \\ -c^T & b^T \\ -A & 0 \\ -I_n & 0 \\ 0 & A^T \\ 0 & -I_m \end{bmatrix}}_{P} \begin{bmatrix} x \\ \lambda \end{bmatrix} \leq \underbrace{\begin{bmatrix} 0 \\ 0 \\ -b \\ 0 \\ c \\ 0 \end{bmatrix}}_{q}.$$

Ellipsoid Method

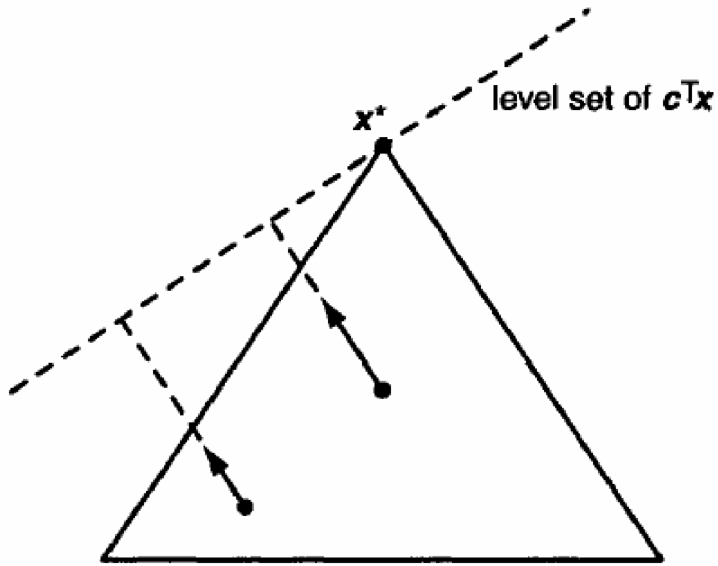
- Let $z \in R^{m+n}$ be a given vector and let Q be an $(m+n) \times (m+n)$ nonsingular matrix. The ellipsoid associated with Q centered at z is defined as the set

$$E_Q(z) = \{z + Qy : y \in R^{m+n}, \|y\| \leq 1\}.$$

- Assume the entries in P and q are all integers.
- At each iteration, the associated ellipsoid contains a solution to the given system of $Pz \leq q$.
- The algorithm updates z and Q in such a way that the ellipsoid at the next step is smaller than the current step.
- The number of iterations N is computed based on L and $m+n$.
- The algorithm inspired other researches.

Interior Point Method

- Recall: Simplex method
  - Jumps from vertex to vertex of the feasible set seeking an optimal vertex
- Interior-point method
  - Starts inside the feasible set and moves within it toward an optimal vertex



Affine Scaling Method

- Basic Algorithm
  - Suppose we have a feasible point $x^{(0)}$ that is strictly interior.
  - Search in a direction $d^{(0)}$ to decrease the objective value while remains feasible. $x^{(1)} = x^{(0)} + \alpha_0 d^{(0)}$

    $\Rightarrow d^{(0)}$ must be a vector in the nullspace of A.
  - Choose $d^{(0)}$ to be the orthogonal projection of the negative gradient –c.

    $\Rightarrow P(v)=v-A^T(AA^T)^{-1}Av=[I_n-A^T(AA^T)^{-1}A]v.$

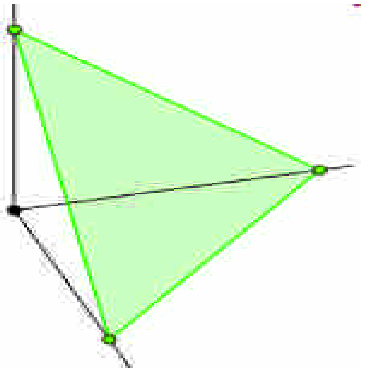
    $\ker(A) \perp \text{im}(A^T)$

Affine Scaling

- Observation
  - The initial point should be chosen close to the center of the feasible set such that we can take a larger step in the search direction.
- Solution
  - Transform a feasible interior point to the center by applying affine scaling:
  - Ex: the center for $\underbrace{\frac{1}{n}\begin{bmatrix}1 & \cdots & 1\end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix}1\end{bmatrix}}_{b}$ is e=[1,…,1]

    To transform $x^{(0)}$ to *e*, we use the scaling transformation D_0^{-1}

$$e = \underbrace{\begin{bmatrix} x_1^{(0)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^{(0)} \end{bmatrix}}_{D_0}^{-1} x^{(0)}$$



New Formulation

- New coordinate system: $\bar{x} = D_0^{-1} x_0$

$$\begin{aligned} \text{minimize} \quad & \bar{c}_0^T \bar{x} \\ \text{subject to} \quad & \bar{A}_0 \bar{x} = b \\ & \bar{x} \geq 0, \end{aligned}$$

where
$$\begin{aligned} \bar{c}_0 &= D_0 c \\ \bar{A}_0 &= A D_0. \end{aligned}$$

set direction to be $\bar{d}^{(0)} = -\bar{P}_0 \bar{c}_0.$

where $\bar{P}_0 = I_n - \bar{A}_0^T (\bar{A}_0 \bar{A}_0^T)^{-1} \bar{A}_0$

compute $\bar{x}^{(1)}$ using

$$\bar{x}^{(1)} = \bar{x}^{(0)} - \alpha_0 \bar{P}_0 \bar{c}_0,$$

obtain the point in the original coordinates:

$$x^{(1)} = D_0 \bar{x}^{(1)}.$$

Final Format

- Iteration step:

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

$$\begin{aligned} D_k &= \text{diag}[x_1^{(k)}, \ldots, x_n^{(k)}] \\ \bar{A}_k &= A D_k \\ \bar{P}_k &= I_n - \bar{A}_k^T (\bar{A}_k \bar{A}_k^T)^{-1} \bar{A}_k \\ d^{(k)} &= -D_k \bar{P}_k D_k c. \end{aligned}$$

  - $Ax^{(k)}=b$
  - $x^{(k)}>0$
  - choosing α_k such that $x_i^{(k+1)} = x_i^{(k)} + \alpha_k d_i^{(k)} > 0$ for $i = 1, \ldots, n$.

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

$\alpha_k = \alpha r_k$, where $\alpha \in (0, 1)$.

$\alpha = 0.9$ or 0.99

- Stopping criteria: $\dfrac{|cx^{(k+1)} - cx^{(k)}|}{\max(1, |cx^{(k)}|)} < \varepsilon$

Two Phase Method

- Phase I
  - Let ***u*** be an arbitrary vector with positive components
  - Let ***v***=***b***-A***u***.
  - If ***v*** = 0, let $\boldsymbol{x}^{(0)}=\boldsymbol{u}$.
  - Else solve the following LP

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

  - The objective function is bounded below by 0, thus the affine scaling method will terminate with some optimal solution.

Karmarkar's Canonical Form

- (all entries in A and c are integers)

$$\begin{aligned} &\text{minimize} && \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to} && \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0} \\ & && \sum_{i=1}^{n} x_i = 1 \\ & && \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned}$$

- Nullspace of A: $\Omega = \{x \in R^n : Ax = 0\}$
- Simplex Δ in R^n: $\Delta = \{x \in R^n : e^T x = 1, x \geq 0\}$.
- Center of the simplex Δ: $a_0 = \frac{e}{n} = \left[\frac{1}{n}, \cdots, \frac{1}{n}\right]$

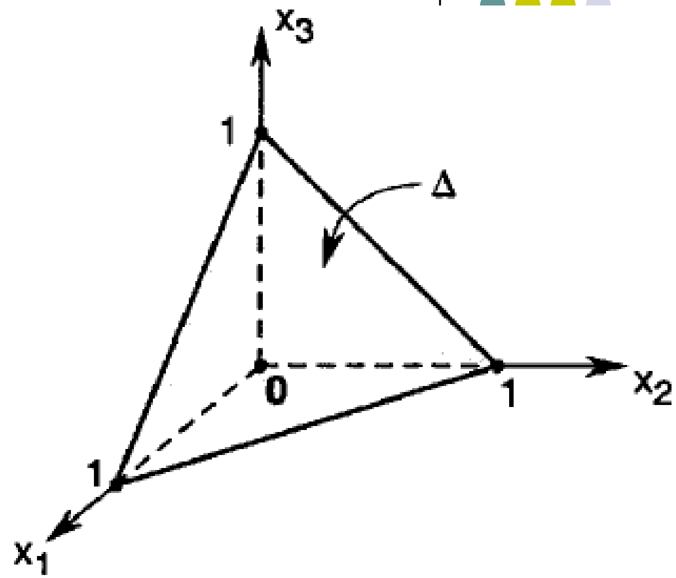
n-simplex:

$$\det\begin{bmatrix} p_0 & p_1 & \cdots & p_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \neq 0.$$

$$\begin{aligned} \Omega \cap \Delta &= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},\ \boldsymbol{e}^T\boldsymbol{x} = 1,\ \boldsymbol{x} \geq \boldsymbol{0}\} \\ &= \left\{\boldsymbol{x} \in \mathbb{R}^n : \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{e}^T \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix},\ \boldsymbol{x} \geq \boldsymbol{0}\right\}. \end{aligned}$$

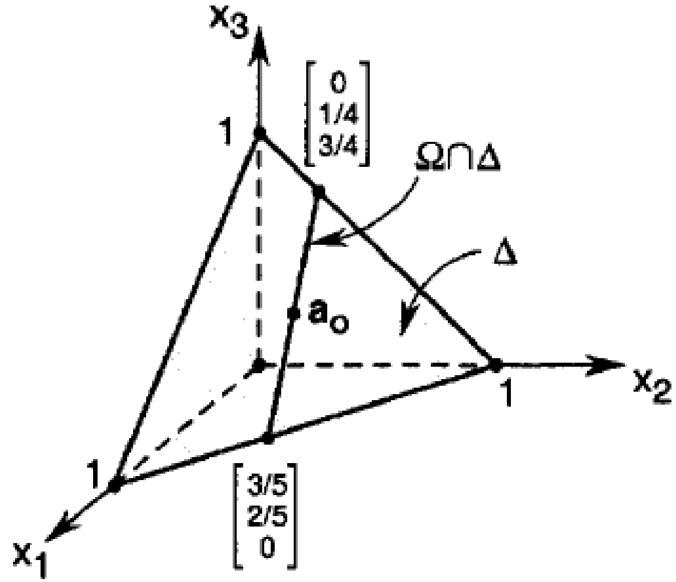
Example 18.1 Consider the following LP problem, taken from [90]:

$$\begin{array}{ll} \text{minimize} & 5x_1 + 4x_2 + 8x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0. \end{array}$$



Example 18.2 Consider the following LP problem, taken from [80]:

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



Karmarkar's Restricted Problem

Karmarkar's algorithm solves LP problems in Karmarkar's canonical form, with the following assumptions:

A. The center $\boldsymbol{a}_0$ of the simplex Δ is a feasible point, that is, $\boldsymbol{a}_0 \in \Omega$;

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

C. The $(m+1) \times n$ matrix

$$\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{e}^T \end{bmatrix}$$

has rank $m+1$;

D. We are given a termination parameter $q > 0$, such that if we obtain a feasible point $\boldsymbol{x}$ satisfying

$$\frac{\boldsymbol{c}^T \boldsymbol{x}}{\boldsymbol{c}^T \boldsymbol{a}_0} \leq 2^{-q},$$

then we consider the problem solved.

How to satisfy the assumptions?

- Assumption A can be achieved when we convert an LP into Karmarkar's canonical form
- Assumption B can be achieved if we know beforehand the minimum value of its objective function value.

$$f(\boldsymbol{x}) = \boldsymbol{c}^T\boldsymbol{x} - M = \boldsymbol{c}^T\boldsymbol{x} - M\boldsymbol{e}^T\boldsymbol{x} = (\boldsymbol{c}^T - M\boldsymbol{e}^T)\boldsymbol{x} = \tilde{\boldsymbol{c}}^T\boldsymbol{x},$$

Example 18.3 Recall the LP problem in Example 18.1:

$$\begin{array}{ll} \text{minimize} & 5x_1 + 4x_2 + 8x_3 \\ \text{subject to} & x_1 + x_2 + x_3 = 1 \\ & x_1, x_2, x_3 \geq 0. \end{array}$$

From Standard Form to Karmarkar's Canonical Form

$$\begin{array}{lll} \text{minimize} & c^T x, & x \in \mathbb{R}^n \\ \text{subject to} & Ax = b & \\ & x \geq 0. & \end{array}$$

$$\begin{array}{lll} \text{minimize} & c'^T z, & z \in \mathbb{R}^{n+1} \\ \text{subject to} & A'z = 0 & \\ & z \geq 0. & \end{array}$$

where $c' = [c^T, 0]^T$ and $A' = [A, -b]$. $z = \begin{bmatrix} x \\ 1 \end{bmatrix}$.

let $y = [y_1, \ldots, y_n, y_{n+1}]^T \in \mathbb{R}^{n+1}$,

projective transformation

$$y_i = \frac{x_i}{x_1 + \cdots + x_n + 1}, \qquad i = 1, \ldots, n$$

$$y_{n+1} = \frac{1}{x_1 + \cdots + x_n + 1}.$$

$$\begin{array}{lll} \text{minimize} & c'^T y, & y \in \mathbb{R}^{n+1} \\ \text{subject to} & A'y = 0 & \\ & e^T y = 1 & \\ & y \geq 0. & \end{array}$$

Ensuring Assumption A ($a_0 \in \Omega$)

Center of simplex Δ

- Suppose we are given a point **a**=[a_1, a_2, …, a_n] that is a strictly interior feasible point: A***a***=***b*** and ***a***>0.
- P_+: positive orthant of R^n: $P_+ = \{\mathbf{x} \in R^n : \mathbf{x} \geq \mathbf{0}\}$.
- Δ: the simplex in R^{n+1}: $\Delta = \{\mathbf{z} \in R^{n+1} : \mathbf{e}^T \mathbf{z} = 1, \mathbf{z} \geq 0\}$.
- Define T: $P_+ \rightarrow \Delta$ by

$$\boldsymbol{T}(\boldsymbol{x}) = [T_1(\boldsymbol{x}), \ldots, T_{n+1}(\boldsymbol{x})]^T$$

with

$$T_i(\boldsymbol{x}) = \frac{x_i/a_i}{x_1/a_1 + \cdots + x_n/a_n + 1}, \qquad i = 1, \ldots, n$$

$$T_{n+1}(\boldsymbol{x}) = \frac{1}{x_1/a_1 + \cdots + x_n/a_n + 1}.$$

minimize	$\boldsymbol{c}'^T \boldsymbol{y}$,	$\boldsymbol{y} \in \mathbb{R}^{n+1}$
subject to	$\boldsymbol{A}'\boldsymbol{y} = \boldsymbol{0}$	
	$\boldsymbol{e}^T \boldsymbol{y} = 1$	
	$\boldsymbol{y} \geq \boldsymbol{0}$.	

- T(***a***) is the center of the simplex and is feasible.

Karmarkar's Algorithm

- Restricted Karmarkar problem

$$\begin{array}{lll} \text{minimize} & \boldsymbol{c}^T\boldsymbol{x}, & \boldsymbol{x} \in \mathbb{R}^n \\ \text{subject to} & \boldsymbol{x} \in \Omega \cap \Delta, & \end{array}$$

- Steps:
  1. Initialize: Set k:=0, $\boldsymbol{x}^{(0)}=\boldsymbol{a}_0=\boldsymbol{e}/n$.
  2. Update: Set $\boldsymbol{x}^{(k+1)}=\Psi(\boldsymbol{x}^{(k)})$
  3. Check the stopping criterion: $\boldsymbol{c}^T\boldsymbol{x}^{(k)}/\boldsymbol{c}^T\boldsymbol{x}^{(0)}\leq 2^{-q}$
  4. Iterate: Set k:=k+1; go to step 2.

Update for $x^{(1)}$: $\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha \boldsymbol{d}^{(0)}$,

- Constraints: $\Omega \cap \Delta$

$$\boldsymbol{B}_0 = \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{e}^T \end{bmatrix}.$$

$$\begin{aligned} \Omega \cap \Delta &= \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}, \boldsymbol{e}^T\boldsymbol{x} = 1, \boldsymbol{x} \geq \boldsymbol{0}\} \\ &= \left\{\boldsymbol{x} \in \mathbb{R}^n : \begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{e}^T \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}, \boldsymbol{x} \geq \boldsymbol{0}\right\} \\ &= \left\{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{B}_0\boldsymbol{x} = \begin{bmatrix} \boldsymbol{0} \\ 1 \end{bmatrix}, \boldsymbol{x} \geq \boldsymbol{0}\right\}, \end{aligned}$$

- Choose $d^{(0)}$ to be the orthogonal projection of –c onto the nullspace of B_0. $\boldsymbol{P}_0 = \boldsymbol{I}_n - \boldsymbol{B}_0^T(\boldsymbol{B}_0\boldsymbol{B}_0^T)^{-1}\boldsymbol{B}_0.$

- Let $d^{(0)} = -r\hat{c}^{(0)}$.

  where $\hat{\boldsymbol{c}}^{(0)} = \dfrac{\boldsymbol{P}_0\boldsymbol{c}}{\|\boldsymbol{P}_0\boldsymbol{c}\|}$,

  and $r = 1/\sqrt{n(n-1)}$

Update for $x^{(k)}$, k>1

- Since $\boldsymbol{x}^{(k)}$ is not in the center of the simplex, we need to transform this point to the center.

$$D_k^{-1} = \begin{bmatrix} 1/x_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1/x_n^{(k)} \end{bmatrix}.$$

- Let U_k: $\Delta \rightarrow \Delta$ be defined by $U_k(\mathbf{x}) = D_k^{-1}\mathbf{x} / e^T D_k^{-1}\mathbf{x}$
- Note that $U_k(\boldsymbol{x}^{(k)}) = \boldsymbol{e}/n = \boldsymbol{a}_0$.
- We need to state the original LP in the new coordinates:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}^T \boldsymbol{D}_k \bar{\boldsymbol{x}} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{D}_k \bar{\boldsymbol{x}} = \boldsymbol{0} \\ & \bar{\boldsymbol{x}} \in \Delta. \end{array}$$

$$\boldsymbol{B}_k = \begin{bmatrix} \boldsymbol{A}\boldsymbol{D}_k \\ \boldsymbol{e}^T \end{bmatrix}.$$

$$\hat{\boldsymbol{c}}^{(k)} = \frac{\boldsymbol{P}_k \boldsymbol{D}_k \boldsymbol{c}}{\|\boldsymbol{P}_k \boldsymbol{D}_k \boldsymbol{c}\|}.$$

- Apply the update step as for $\boldsymbol{x}^{(1)}$

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

- Finally apply the inverse transformation U_k^{-1} to obtain $\boldsymbol{x}^{(k+1)}$

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

The update of $x^{(k+1)}$

1. Compute the matrices:

$$\boldsymbol{D}_k = \begin{bmatrix} x_1^{(k)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_n^{(k)} \end{bmatrix}$$

$$\boldsymbol{B}_k = \begin{bmatrix} \boldsymbol{A}\boldsymbol{D}_k \\ \boldsymbol{e}^T \end{bmatrix}.$$

2. Compute the orthogonal projector onto the nullspace of $\boldsymbol{B}_k$:

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

3. Compute the normalized orthogonal projection of $\boldsymbol{c}$ onto the nullspace of $\boldsymbol{B}_k$:

$$\hat{\boldsymbol{c}}^{(k)} = \frac{\boldsymbol{P}_k\boldsymbol{D}_k\boldsymbol{c}}{||\boldsymbol{P}_k\boldsymbol{D}_k\boldsymbol{c}||}.$$

4. Compute the direction vector:

$$\boldsymbol{d}^{(k)} = -r\hat{\boldsymbol{c}}^{(k)},$$

   where $r = 1/\sqrt{n(n-1)}$.

5. Compute $\bar{\boldsymbol{x}}^{(k+1)}$ using:

$$\bar{\boldsymbol{x}}^{(k+1)} = \boldsymbol{a}_0 + \alpha\boldsymbol{d}^{(k)},$$

   where α is the prespecified step size, $\alpha \in (0, 1)$.

6. Compute $\boldsymbol{x}^{(k+1)}$ by applying the inverse transformation $\boldsymbol{U}_k^{-1}$:

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