Karmarkar's Algorithm

An Interior Point Method of Linear Programming Problem

AK Dhamija

DIPR, DRDO

November 20, 2009

Overview

Algorithms

Algorithms Complexity

If n is the size of the problem (n may be number of variables in LP Problem)

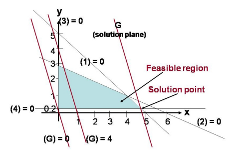
- Time Complexity - Time taken by algorithm
- Space Complexity - Memory needed by algorithm
- Increasing order of complexity : $lg\ n$, $\sqrt{n}$, n, $n\ lg\ n$, n^2, n^3, 2^n, $n!$
- $O(n^2) = c_1 n^2$, $O(2^n) = c_2 2^n$
- Compare $O(n^2) = 100n^2$, $O(2^n) = 2^n$
- 22500 vs 32768 for n = 15
- 1000000 vs 1267650600228229401496703205376 for n = 100
- $O(n^2) = c_1 n^2$ is polynomial time (aka efficient) algorithm and $O(2^n) = c_2 2^n$ is exponential time (aka inefficient) algorithm

LP Problem

Problem Instance

- Set of n real variables e.g. x,y etc
- Set of restrictions in the form of linear inequalities e.g. $x + y \leq 5$, $y \geq 0.2$ etc
- Goal e.g. Maximize $4x + y$
- Solution: $(x, y) = (5, 0)$

Feasible Region



- Feasible region is convex
- Each constraint generates at most one edge in the feasible region
- As you move the goal line, G, to increase it's value, the point that will maximize it is one of the 'corners'.

LP Algorithms

Naive Algorithm

- Find all intersection points, check these points to satisfy all constraints i.e feasibility, and find out the point of optimum goal value
- $O(m^n)$ for n variables and m constraints - exponential in n

Simplex (George B. Dantzig, 1947)

- Start from a feasible corner, and Move to the one that gives the highest value to the goal function
- STOPPING condition : none of the neighboring points gives a higher value than the current one
- Average time is linear; worst case time is exponential (covers all feasible corners)
- Remarkably successful in practice
- Integer Programming is NP-complete

Ellipsoidal Method (Leonid Khachiyan, 1979)

- Runtime: polynomial
- Theoretical value; not useful in practice

Interior point method (Narendra Karmarkar, 1984)

- Runtime: polynomial
- Practical value, Recently these have become competitive in practice with simplex.

An example of Exponential Time

Klee-Minty Example

$$
\begin{array}{llllllllllll}
\max\ 2^{n-1}x_1 & + & 2^{n-2}x_2 & + & \ldots & + & 2x_{n-1} & + & x_n: & & \\
x_1 & & & & & & & & & \leq & 5 \\
4x_1 & + & x_2 & & & & & & & \leq & 25 \\
8x_1 & + & 4x_2 & + & x_3 & & & & & \leq & 125 \\
\vdots & & & & & & & & & & \vdots \\
2^n x_1 & + & 2^{n-1}x_2 & + & \ldots & + & 4x_{n-1} & + & x_n & \leq & 5^n \\
x \geq 0. & & & & & & & & & &
\end{array}
$$

Developments

- The LP has n variables, n constraints and 2^n extreme points
- The elementary simplex method, starting at $x = 0$, goes through each of the extreme points before reaching the optimum solution at $(0, 0, ..., 5^n)$

Klee-Minty Example

Here is the pivot sequence for $n = 3$, which goes through all 8 extreme points, starting at the origin. Let s be the slack variables.

Initial Tableau

Basic	Nonbasic x_1	x_2	x_3	RHS
s_1	1^*			5
s_2	4	1		25
s_3	8	4	1	125
$-z$	4	2	1	0
	↑			

Tableau 1

Basic	Nonbasic s_1	x_2	x_3	RHS
x_1	1			5
s_2	-4	1^*		5
s_3	-8	4	1	85
$-z$	-4	2	1	-20
		↑		

Tableau 2

Basic	Nonbasic s_1	s_2	x_3	RHS
x_1	1^*			5
x_2	-4	1		5
s_3	8	-4	1	65
$-z$	4	-2	1	-30
	↑			

Tableau 3

Basic	Nonbasic x_1	s_2	x_3	RHS
s_1	1			5
x_2	4	1		25
s_3	-8	-4	1^*	25
$-z$	-4	-2	1	-50
			↑	

Tableau 4

Basic	Nonbasic x_1	s_2	s_3	RHS
s_1	1^*			5
x_2	4	1		25
x_3	-8	-4	1	25
$-z$	4	2	-1	-75
	↑			

Tableau 5

Basic	Nonbasic s_1	s_2	s_3	RHS
x_1	1			5
x_2	-4	1^*		5
x_3	8	-4	1	65
$-z$	-4	2	-1	-95
		↑		

Tableau 6

Basic	Nonbasic s_1	x_2	s_3	RHS
x_1	1^*			5
s_2	-4	1		5
x_3	-8	4	1	85
$-z$	4	-2	-1	-105
	↑			

Tableau 7

Basic	Nonbasic x_1	s_2	x_3	RHS
s_1	1^*			5
s_2	4	1		25
x_3	8	4	1	125
$-z$	-4	-2	-1	-125

Exponential Time Algorithm in Worst case

General Introduction

Interior-point methods for linear programming

Components

- A breakthrough development in linear programming during the 1980s
- Started in 1984 by a young mathematician at AT&T Bell Labs, N. Karmarkar
- Polynomial-time algorithm which can solve huge linear programming problems beyond the reach of the simplex method.
- Karmarkar's method stimulated development in both interior-point and simplex methods.

Difference between Interior point methods and the simplex method

The nature of trial solutions and Complexity

Simplex Method

- CPF (Corner Point Feasible) solutions
- Worst Case:No of iterations can increase exponentially in the number of variables $n : O(2^n)$
- Practically remarkably successful except some huge problems (eg Airlines Scheduling etc)

Interior Point Methods

- Interior points (points inside the boundary of the feasible region)
- Polynomial time
- Advantageous in solving huge problems, but cumbersome for not-so-large problems

Karmarkar's Original Algorithm

Karmarkar assumes that the LP is given in Canonical form of the problem

- Min $Z = \mathbf{CX}$
- such that $\mathbf{AX} = \mathbf{0}$, $\mathbf{1X} = 1$, $\mathbf{X} \geq \mathbf{0}$

Assumptions

- $\mathbf{X_0} = (\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n})$ is a feasible solution
- Minimum$(Z) = 0$

To apply the algorithm to LP problem in standard form, a transformation is needed

- Min $Z = \mathbf{CX}$
- such that $\mathbf{AX} \leq \mathbf{b}$, $\mathbf{X} \geq \mathbf{0}$

Examples : Standard Vs Canonical Form

Standard Form

Min $Z = y_1 + y_2$ ($Z = \mathbf{CY}$)
such that

- $y_1 + 2y_2 \leq 2$ ($\mathbf{AY} \leq \mathbf{b}$)
- $y_1, y_2 \geq 0$

Canonical Form

Min $Z = 5x_1 + 5x_2$ ($Z = \mathbf{CX}$)
such that

- $3x_1 + 8x_2 + 3x_3$ - $2x_4 = 0$ ($\mathbf{AX} = \mathbf{0}$)
- $x_1 + x_2 + x_3 + x_4 = 1$ ($\mathbf{1X} = 1$)
- $x_j \geq 0,\ j = 1, 2, 3, 4$

Karmarkar's Original Algorithm

The Principle Idea

Create a sequence of points $x^{(0)}, x^{(1)}, x^{(2)}, ..., x^{(k)}$ having decreasing values of the objective function. In the k^{th} step, the point $x^{(k)}$ is brought into the center of the simplex[a] by projective transformation.

[a] n-dimensional unit simplex S is the set of points$(x_1, x_2, ...x_n)^T$ satisfying $x_1 + x_2 + ... + x_n = 1$ and $x_j \geq 0, j = 1, 2, ...n$

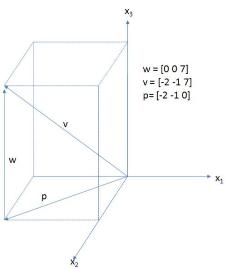
Three key Concepts

- Projection of a vector onto the set of $\mathbf{X}$ satisfying $\mathbf{AX} = \mathbf{0}$
- Karmarkar's Centering Transformation
- Karmarkar's Potential Function

Three Concepts

Projection

- we want to move from a feasible point $\mathbf{X^0}$ to another feasible point $\mathbf{X^1}$, that for some fixed vector $\mathbf{v}$, will have a larger value of $\mathbf{vX}$
- if we choose to move in direction $\mathbf{d} = (d_1, d_2, \ldots d_n)$ that solves the optimization problem
  Max $\mathbf{vd}$ such that
  $\mathbf{Ad} = \mathbf{0}$, $d_1 + d_2 + \ldots d_n = 0$ (so that $\mathbf{Ad}$ remains feasible)
  and $||\mathbf{d}|| = 1$
  then we will be moving in a feasible direction that maximizes the increase in $\mathbf{vX}$ per unit length moved.
- The direction $\mathbf{d}$ that solves this optimization problem is given by the **projection** of $\mathbf{v}$ onto $\mathbf{X}$ satisfying $\mathbf{AX} = \mathbf{0}$ and $x_1 + x_2 + \ldots + x_n = 0$ and is given by $[\mathbf{I} - \mathbf{B}^T(\mathbf{BB}^T)^{-1}\mathbf{B}]\mathbf{v}$, where $\mathbf{B} = \begin{bmatrix} \mathbf{A} \\ \mathbf{1} \end{bmatrix}$



- Geometrically, if we can write $\mathbf{v} = \mathbf{p} + \mathbf{w}$, where $\mathbf{p}$ satisfies $\mathbf{Ap} = \mathbf{0}$ and $\mathbf{w}$ is perpendicular to all vectors $\mathbf{X}$ satisfying $\mathbf{AX} = \mathbf{0}$, then $\mathbf{p}$ is the projection of $\mathbf{v}$ onto the set of $\mathbf{X}$ satisfying $\mathbf{AX} = \mathbf{0}$.
- For Example, $\mathbf{v} = (-2, -1, 7)$ is projected onto a set of 3-d vectors satisfying $x_3 = 0$(ie $x_1 - x_2$ plane), then $\mathbf{v} = (-2, -1, 0) + (0, 0, 7)$. So here, $\mathbf{p} = (-2, -1, 0)$ is the vector in the set of $\mathbf{X}$ satisfying $\mathbf{AX} = 0$ that is **closest** to $\mathbf{v}$.

Three Concepts

Karmarkar's Centering Transformation

- If $\mathbf{x^k}$ is a point in S, then $f([x_1, x_2, \ldots x_n]|\mathbf{x^k})$ transforms a point $[x_1, x_2, \ldots x_n]^T$ in S into a point $[y_1, y_2, \ldots y_n]$ in S, where
$$y_j = \frac{\frac{x_j}{x_j^k}}{\sum_{r=1}^{r=n} \frac{x_r}{x_r^k}}$$
- Consider the LP
  Min $z = x_1 + 3x_2$ - $3x_3$ such that
  x_2 - $x_3 = 0$
  $x_1 + x_2 + x_3 = 1$
  $x_i \geq 0$
  This LP has a feasible solution $[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]^T$ and the optimal value of z is 0.
  The feasible point $[\frac{1}{4}, \frac{3}{8}, \frac{3}{8}]^T$ yields the following transformation
  $$f([x_1, x_2, x_3]|[\frac{1}{4}, \frac{3}{8}, \frac{3}{8}]) =$$
  $$[x_1, x_2, x_3]|[\frac{4x_1}{4x_1 + \frac{8x_2}{3} + \frac{8x_3}{3}}, \frac{\frac{8x_2}{3}}{4x_1 + \frac{8x_2}{3} + \frac{8x_3}{3}}, \frac{\frac{8x_3}{3}}{4x_1 + \frac{8x_2}{3} + \frac{8x_3}{3}}]$$
  For example, $f([\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]|[\frac{1}{4}, \frac{3}{8}, \frac{3}{8}]) = [\frac{12}{28}, \frac{8}{28}, \frac{8}{28}]$
- We now refer to the variables $x_1, x_2, \ldots, x_n$ as being the **original space** and the variables $y_1, y_2, \ldots, y_n$ as being the **transformed space** and the unit simplex involving variables $y_1, y_2, \ldots, y_n$ will be called **transformed unit** simplex.

Three Concepts

Properties of Centering Transformation

- $f(\mathbf{x^k}|\mathbf{x^k}) = [\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}]^T$
  $f(.|\mathbf{x^k})$ maps $\mathbf{x^k}$ into the center of the transformed unit simplex.
- $f(\mathbf{x}|\mathbf{x^k}) \in S$ and For $\mathbf{x} \neq \mathbf{x}'$, $f(\mathbf{x}|\mathbf{x^k}) \neq f(\mathbf{x}'|\mathbf{x^k})$
  Any point in S is transformed into a point in the transformed unit simplex, and no two points in S can yield the same point (i.e. f is one-one mapping)
- For any point $[y_1, y_2, \ldots y_n]^T$ in S, there is a unique point $[x_1, x_2, \ldots x_n]^T$ in S satisfying $f([x_1, x_2, \ldots x_n]^T|\mathbf{x^k}) = [y_1, y_2, \ldots y_n]^T$
  The point $[x_1, x_2, \ldots x_n]^T$ is given by
  $$\frac{x_j^k y_j}{\sum_{r=1}^{r=n} x_r^k y_r}$$
  we can write $f^{-1}([y_1, y_2, \ldots y_n]^T|\mathbf{x^k}) = [x_1, x_2, \ldots x_n]^T$
  The above two equations imply that f is a one-one onto mapping from s to S.
- A point $\mathbf{x}$) in S will satisfy $\mathbf{Ax} = \mathbf{0}$ if $A[diag(\mathbf{x^k})]f(\mathbf{x}|\mathbf{x^k}) = \mathbf{0}$
  Feasible points in the original problem correspond to points $\mathbf{y}$ in the transformed unit simplex that satisfy $A[diag(\mathbf{x^k})]\mathbf{y} = \mathbf{0}$

Karmarkar's Original Algorithm

Step $k = 0$

- Start with the solution point $\mathbf{Y_0} = \mathbf{X_0} = (\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n})^T$
- Compute step length parameters $r = \frac{1}{\sqrt{n(n-1)}}$ and $\alpha = \frac{(n-1)}{3n}$

Step k

Define

- $\mathbf{D_k} = diag\{\mathbf{X_k}\} = diag\{x_{k1}, x_{k2}, ..., x_{kn}\}$
- $\mathbf{P} = \begin{bmatrix} \mathbf{AD_k} \\ \mathbf{1} \end{bmatrix}$

Compute

- $\mathbf{c_p} = [\mathbf{I} - \mathbf{P^T}(\mathbf{PP^T})^{-1}\mathbf{P}](\mathbf{CD_k})^T$
- $\mathbf{Y_{new}} = \mathbf{Y_0}$ - $\alpha r \frac{\mathbf{c_p}}{||\mathbf{c_p}||}$
- $\mathbf{X_{k+1}} = \frac{\mathbf{D_k Y_{new}}}{\mathbf{1 D_k Y_{new}}}$
- $Z = \mathbf{CX_{k+1}}$
- $k = k + 1$
- Repeat iteration k until Z becomes less than prescribed tolerance ϵ

Centering : $\mathbf{X}$ is brought to center by $\mathbf{Y} = \frac{\mathbf{D_k}^{-1}\mathbf{X}}{\mathbf{1D_k^{-1}X}}$

Karmarkar's Original Algorithm

Certain Remarks

- We move from the **center** of the transformed unit simplex in a direction opposite to the projection of $\mathbf{CD_k})^T$ onto the transformation of the feasible region(the set of $\mathbf{y}$ satisfying $\mathbf{A}diag\{\mathbf{X_k}\}\mathbf{y} = \mathbf{0}$). As per our explanation of **projection**, this ensures that we maintain feasibility(in the transformed space) and move in a direction that maximizes the rate of decrease of $\mathbf{CD_k})^T$
- By moving a distance αr from the center of the transformed unit simplex, we ensure that $\mathbf{y^{k+1}}$ will remain in the interior of the transformed simplex
- When we use the inverse of Karmarkar's centering transformation to transform $\mathbf{y^{k+1}}$ back into $\mathbf{x^{k+1}}$, the definition of projection imply that $\mathbf{x^{k+1}}$ will be feasible for the original LP

Karmarkar's Original Algorithm

Potential Function : Why do we project $\mathbf{CD_k})^T$ rather than $\mathbf{C})^T$ onto the transformed space?

- Because we are projecting $\mathbf{CD_k}^T$ rather than $\mathbf{C}^T$, we can not be sure that each iteration will decrease Z. In fact it is possible for $\mathbf{CX^{k+1}} > \mathbf{CX^k}$ to occur.
- **Karmarkar's Potential Function** $f(\mathbf{X})$ is defined as $f(\mathbf{X}) = sum_{j=1}^{j=n} ln(\frac{\mathbf{CX^T}}{x_j}), \mathbf{X} = [x_1, x_2, ..., x_n]^T$
- Karmarkar showed that if we project $\mathbf{CD_k})^T$ (not $\mathbf{C})^T$) onto the feasible region in the transformed space, then for some $\delta > 0$, it will be true that for $k = 0, 1, 2, \ldots$ $f(\mathbf{X^k})$ - $f(\mathbf{X^{k+1}}) \geq \delta$.
- So, each iteration of karmarkar's algorithm decreases the potential function by an amount bounded away from 0.
- Karmarkar showed that if the potential function evaluated at $\mathbf{x^k}$ is small enough, the $Z = \mathbf{CX^k}$ will be near 0. Because $f(\mathbf{x^k})$ is decreased by at least δ per iteration, it follows that by choosing k sufficiently large, we can ensure that the Z-value for $\mathbf{X^k}$ is less than ϵ

Karmarkar's Original Algorithm - Two Iterations

Problem

Min $Z = 2x_2$ - x_3
such that

- x_1 - $2x_2 + x_3 = 0$
- $x_1 + x_2 + x_3 = 1$
- $x_j \geq 0,\ j = 1, 2, 3$

Solution : $z = 0$ at $(0, 0.33334, 0.66667)$

Preliminary Step

- $k = 0$
- $\mathbf{X_0} = (\frac{1}{n}, \frac{1}{n}, \ldots., \frac{1}{n})^T = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$
- $r = \frac{1}{\sqrt{n(n-1)}} = \frac{1}{\sqrt{3(3-1)}} = \frac{1}{\sqrt{6}}$
- $\alpha = \frac{(n-1)}{3n} = \frac{(3-1)}{(3)(3)} = \frac{2}{9}$

Karmarkar's Original Algorithm - Two Iterations

Iteration 0

- $\mathbf{Y_0} = \mathbf{X_0} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$
- $\mathbf{D_0} = diag\{\mathbf{X_0}\} = diag\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}$
- $\mathbf{AD_0} = (1, -2, 1)diag\{\mathbf{X_0}\} = diag\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\} = (\frac{1}{3}, \frac{-2}{3}, \frac{1}{3})$
- $\mathbf{P} = \begin{bmatrix} \mathbf{AD_0} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{-2}{3} & \frac{1}{3} \\ 1 & 1 & 1 \end{bmatrix}$
- $\mathbf{CD_0} = (0, 2, -1)diag\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\} = (0, \frac{2}{3}, \frac{-1}{3})$
- $\mathbf{PP^T} = \begin{bmatrix} 0.667 & 0 \\ 0 & 3 \end{bmatrix}; (\mathbf{PP^T})^{-1} = \begin{bmatrix} 1.5 & 0 \\ 0 & 0.333 \end{bmatrix}$
- $\mathbf{P^T}(\mathbf{PP^T})^{-1}\mathbf{P} = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$
- $\mathbf{c_p} = [\mathbf{I} - \mathbf{P^T}(\mathbf{PP^T})^{-1}\mathbf{P}](\mathbf{CD_0})^T = (0.167, 0, -0.167)^T$
- $||\mathbf{c_p}|| = \sqrt{(0.167)^2 + 0 + (-0.167)^2} = 0.2362$
- $\mathbf{Y_{new}} = \mathbf{Y_0}$ - $\alpha r \frac{\mathbf{c_p}}{||\mathbf{c_p}||} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T$ - $\frac{(\frac{2}{9})(\frac{1}{\sqrt{6}})}{0.2362}(0.167, 0, -0.167)^T =$ $(0.2692, 0.3333, 0.3974)^T$
- $\mathbf{X_1} = Y_{new} = (0.2692, 0.3333, 0.3974)^T$
- $Z = \mathbf{CX_1} = (0, 2, -1)(0.2692, 0.3333, 0.3974)^T = 0.2692$
- $k = 0 + 1 = 1$

Karmarkar's Original Algorithm - Two Iterations

Iteration 1

- $\mathbf{D_1} = diag\{\mathbf{X_1}\} = diag\{0.2692, 0.3333, 0.3974\}$
- $\mathbf{AD_1} = (1, -2, 1)diag\{\mathbf{X_1}\} = diag\{0.2692, 0.3333, 0.3974\} = (0.2692, -0.6666, 0.3974)$
- $\mathbf{P} = \begin{bmatrix} \mathbf{AD_1} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} 0.2692 & -0.6666 & 0.3974 \\ 1 & 1 & 1 \end{bmatrix}$
- $\mathbf{CD_1} = (0, 2, -1)diag\{0.2692, 0.3333, 0.3974\} = (0, 0.6666, -0.3974)$
- $\mathbf{PP^T} = \begin{bmatrix} 0.675 & 0 \\ 0 & 3 \end{bmatrix}; (\mathbf{PP^T})^{-1} = \begin{bmatrix} 1.482 & 0 \\ 0 & 0.333 \end{bmatrix}$
- $\mathbf{P^T}(\mathbf{PP^T})^{-1}\mathbf{P} = \begin{bmatrix} 0.441 & 0.067 & 0.492 \\ 0.067 & 0.992 & -0.059 \\ 0.492 & -0.059 & 0.567 \end{bmatrix}$
- $\mathbf{c_p} = [\mathbf{I} - \mathbf{P^T}(\mathbf{PP^T})^{-1}\mathbf{P}](\mathbf{CD_1})^T = (0.151, -0.018, -0.132)^T$
- $||\mathbf{c_p}|| = \sqrt{(0.151)^2 + (-0.018)^2 + (-0.132)^2} = 0.2014$
- $\mathbf{Y_{new}} = \mathbf{Y_0} - \alpha r \frac{\mathbf{c_p}}{||\mathbf{c_p}||} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})^T - \frac{(\frac{2}{9})(\frac{1}{\sqrt{6}})}{0.2014}(0.151, -0.018, -0.132)^T = (0.2653, 0.3414, 0.3928)^T$
- $\mathbf{D_1 Y_{new}} = diag\{0.2692, 0.3333, 0.3974\}(0.2653, 0.3414, 0.3928)^T = (0.0714, 0.1138, 0.1561)^T; \mathbf{1D_1Y_{new}} = 0.3413$
- $\mathbf{X_2} = \frac{\mathbf{D_1Y_{new}}}{\mathbf{1D_1Y_{new}}} = (0.2092, 0.3334, 0.4574)^T$
- $Z = \mathbf{CX_1} = (0, 2, -1)(0.2092, 0.3334, 0.4574)^T = 0.2094; k = 1 + 1 = 2$

Transforming a LP Problem into Karmarkar's Special Form

Steps for Conversion of the LP problem : Min $Z = \mathbf{CX}$ such that $\mathbf{AX} \geq \mathbf{b}$, $\mathbf{X} \geq \mathbf{0}$

1. Write the dual of given primal problem
   Min $Z = \mathbf{bW}$
   such that $\mathbf{A^TW} \leq \mathbf{C^T}$, $\mathbf{W} \geq \mathbf{0}$

2.
   1. Introduce slack & surplus variables to primal and dual problems
   2. Combine these two problems.

3.
   1. Introduce a bounding constraint $\sum x_i + \sum w_i \leq K)$, where K should be sufficiently large to include all feasible solutions of original problem.
   2. Introduce a slack variable in the bounding constraint and obtain $\sum x_i + \sum w_i + s = K$

4.
   1. Introduce a dummy variable d (subject to condition $d = 1$)to homogenize the constraints.
   2. Replace the equations
      $\sum x_i + \sum w_i + s = K$ and $d = 1$
      with the following equations
      $\sum x_i + \sum w_i + s - Kd = 0$ and $\sum x_i + \sum w_i + s + d = K + 1$

5. Introduce the following transformation so as to obtain one on the RHS of the last equation
   $x_j = (K+1)y_j, j = 1, 2, \ldots m+n$
   $w_j = (K+1)y_{m+n+j}, j = 1, 2, \ldots m+n$
   $s = (K+1)y_{2m+2n+1}$
   $d = (K+1)y_{2m+2n+2}$

6. Introduce an artificial variable $y_{2m+2n+3}$ (to be minimized) in all the equations such that the sum of the coefficients in each homogenous equation is zero and coefficient of the artificial variable in the last equation is one.

Example of transforming a LP Problem into Karmarkar's Special Form

Steps for Conversion of the LP problem : Max $Z = 2x_1 + x_2$ such that $x_1 + x_2 \leq 5$, $x_1 - x_2 \leq 3$, x_1, x_2 are nonnegative

1. Write the dual of given primal problem

   Min $Z = 5w_1 + 3w_2$
   such that
   $w_1 + w_2 \geq 2$
   $w_1 - w_2 \geq 1$
   $w_1, w_2 \geq 0$

2. Introduction of slack & surplus variables and combination of primal and dual problems

   $x_1 + x_2 + x_3 = 5$
   $x_1 - x_2 + x_4 = 3$
   $w_1 + w_2 - w_3 = 2$
   $w_1 - w_2 - w_4 = 1$
   $2x_1 + x_2 = 5w_1 + 3w_2$

3. Addition of boundary constraint with slack variable s

   $\sum_{i=1}^{4} x_i + \sum_{i=1}^{4} w_i + s = K$

4. Homogenized equivalent system with dummy variable d

   $x_1 + x_2 + x_3 - 5d = 0$
   $x_1 - x_2 + x_4 - 3d = 0$
   $w_1 + w_2 - w_3 - 2d = 0$
   $w_1 - w_2 - w_4 - d = 0$
   $2x_1 + x_2 - 5w_1 - 3w_2 = 0$
   $\sum_{i=1}^{4} x_i + \sum_{i=1}^{4} w_i + s - Kd = 0$
   $\sum_{i=1}^{4} x_i + \sum_{i=1}^{4} w_i + s + d = (K+1)$
   with all variables non-negative

Example of transforming a LP Problem into Karmarkar's Special Form

Steps for Conversion of the LP problem : Max $Z = 2x_1 + x_2$ such that $x_1 + x_2 \leq 5$, $x_1 - x_2 \leq 3$, x_1, x_2 are nonnegative

5. Introduction of transformations

$$x_j = (K+1)y_j, j = 1, 2, \ldots 4$$
$$w_j = (K+1)y_{4+j}, j = 1, 2, \ldots 4$$
$$s = (K+1)y_9$$
$$d = (K+1)y_{10}$$

The system of equations thus obtained is as follows

$$y_1 + y_2 + y_3 - 5y_{10} = 0$$
$$y_1 - y_2 + y_4 - 3y_{10} = 0$$
$$y_5 + y_6 - y_7 - 2y_{10} = 0$$
$$y_5 - y_6 - y_8 - y_{10} = 0$$
$$2y_1 + y_2 - 5y_5 - 3y_6 = 0$$
$$\sum_{i=1}^{9} y_i - Ky_{10} = 0$$
$$\sum_{i=1}^{10} y_i = 1$$

with all variables non-negative

6. Introduce an artificial variable y_{11}

Minimize y_{11}

subject to

$$y_1 + y_2 + y_3 - 5y_{10} + 2y_{11} = 0$$
$$y_1 - y_2 + y_4 - 3y_{10} + 2y_{11} = 0$$
$$y_5 + y_6 - y_7 - 2y_{10} + y_{11} = 0$$
$$y_5 - y_6 - y_8 - y_{10} + 2y_{11} = 0$$
$$2y_1 + y_2 - 5y_5 - 3y_6 + 5y_{11} = 0$$
$$\sum_{i=1}^{9} y_i - Ky_{10} + (K-9)y_{11} = 0$$
$$\sum_{i=1}^{11} y_i = 1$$

with all variables non-negative

The Affine Variant of Algorithm

Three Concepts

- **Concept 1**: Shoot through the interior of the feasible region towards an optimal solution
- **Concept 2**: Move in the direction that improves the objective function value at the fastest feasible rate.
- **Concept 3**: Transform the feasible region to place the current trail solution near its center, thereby enabling the fastest feasible rate for Concept 2

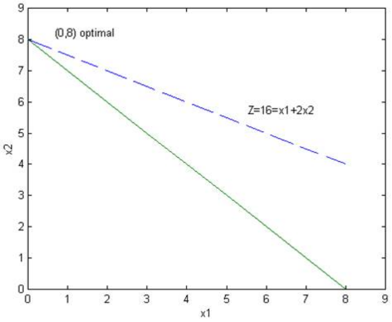
An Example

The Problem

Max $Z = x_1 + 2x_2$
such that

- $x_1 + x_2 \leq 8$
- $x_j \geq 0, \ j = 1, 2$

The optimal Solution is $(x_1, x_2) = (0, 8)$ with $Z = 16$



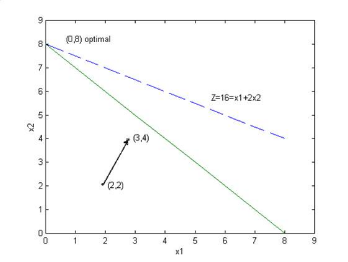
An Example

Concept 1: Shoot through the interior of the feasible region toward an optimal solution

- The algorithm begins with an initial solution that lies in the interior of the feasible region
- Arbitrarily choose $(x_1, x_2) = (2, 2)$ to be the initial solution

Concept 2: Move in a direction that improves the objective function value at the fastest feasible rate

- The direction is perpendiculars to (and toward) the objective function line.
- $(3, 4) = (2, 2) + (1, 2)$, where the vector $(1, 2)$ is the gradient(aka Coefficients) of the Objective Function



An Example

The algorithm actually operates on the augmented form

Max $\mathbf{Z} = \mathbf{CX}$ such that

- $\mathbf{AX} = b$
- $\mathbf{X} \geq 0$

The Transformation

Max $Z = x_1 + 2x_2 \Longrightarrow$ Max $Z = x_1 + 2x_2$ such that

- $x_1 + x_2 \leq 8 \Longrightarrow x_1 + x_2 + x_3 = 8$, where x_3 is the slack
- $x_j \geq 0,\ j = 1, 2 \Longrightarrow x_j \geq 0,\ j = 1, 2, 3$

The optimal Solution is $(x_1, x_2) = (0, 8)$ with $Z = 16$

The Matrices

$$\mathbf{C} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}\ \mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\ \mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}\ b = \begin{bmatrix} 8 \end{bmatrix}\ \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

The Solution

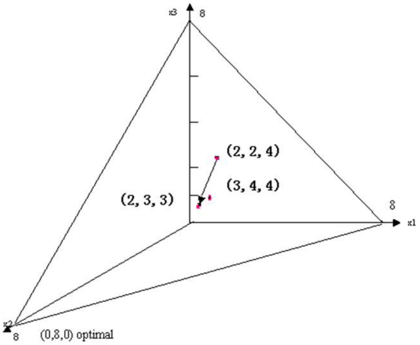
- Initial Solution $(2, 2) \Longrightarrow (2, 2, 4)$, Optimum $(0, 8) \Longrightarrow (0, 8, 0)$
- Gradient of Objective Function $(1, 2) \Longrightarrow (1, 2, 0)$

An Interior-point Algorithm

Gradient of Objective function : $\mathbf{C} = \begin{bmatrix} 1 & 2 & 0 \end{bmatrix}$

Using Projected Gradient to implement Concept 1 & 2

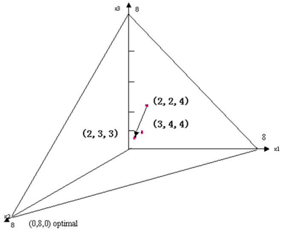
- Adding the gradient to the initial leads to $(3, 4, 4) = (2, 2, 4) + (1, 2, 0) \implies$ Infeasible
- To remain feasible, the algorithm projects the point $(3, 4, 4)$ down onto the feasible tetrahedron
- The next trial solution moves in the direction of projected gradient i.e. the gradient projected onto the feasible region



An Interior-point Algorithm

Using Projected Gradient to implement Concept 1 & 2

- Projection Matrix $\mathbf{P} = \mathbf{I} - \mathbf{A}^{\mathbf{T}}(\mathbf{A}\mathbf{A}^{\mathbf{T}})^{-1}\mathbf{A} = \begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$
- Projected Gradient $\mathbf{c_P} = \mathbf{P}\mathbf{C}^{\mathbf{T}} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$
- Move $(2, 2, 4)$ towards $\mathbf{c_P}$ to $\mathbf{X} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} + 4\alpha\mathbf{c_P} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} + 4\alpha \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$
- α determines how far we move. large $\alpha \Rightarrow$ too close to the boundary and $\alpha \Rightarrow$ more iterations we have chosen $\alpha = 0.5$, so the new trial solution move to: $\mathbf{X} = (2, 3, 3)$



An Interior-point Algorithm

Concept 3: Transform the feasible region to place the current trail solution near its center, thereby enabling a large improvement when concept 2 is implemented

Centering scheme for implementing Concept 3

- Why : The centering scheme keeps turning the direction of the projected gradient to point more nearly toward an optimal solution as the algorithm converges toward this solution
- How : Simply changing the scale for each of the variable so that the trail solution becomes equidistant from the constraint boundaries in the new coordinate system
- x is brought to the center $\tilde{\mathbf{X}} = (1, 1, 1)$ in the new coordinate system by $\tilde{\mathbf{X}} = \mathbf{D}^{-1}\mathbf{X}$, where $\mathbf{D} = diag\{\mathbf{X}\}$

An Interior-point Algorithm

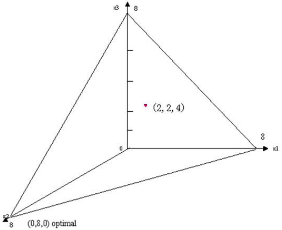
Centering scheme for implementing Concept 3

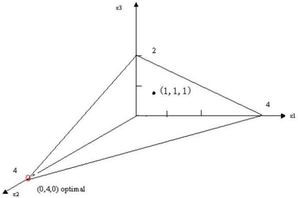
$$\tilde{\mathbf{X}} = \mathbf{D}^{-1}\mathbf{X} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

The Problem in the new coordinate system becomes

Max $Z = 2\tilde{x}_1 + 4\tilde{x}_2$
such that

- $2\tilde{x}_1 + 2\tilde{x}_2 + 4\tilde{x}_3 = 8$
- $\tilde{x}_j \geq 0,\ j = 1, 2, 3$





Complete Illustration of the Algorithm

Iteration 1

- Initial Solution $(x_1, x_2, x_3) = (2, 2, 4)$
- $\mathbf{D} = diag\{x_1, x_2, x_3\} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}$ $\tilde{\mathbf{X}} = \mathbf{D}^{-1}\mathbf{X} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

$$= \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

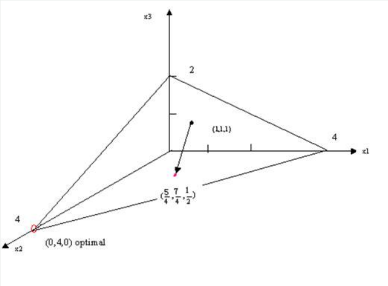
- $\tilde{\mathbf{A}} = \mathbf{AD} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 4 \end{bmatrix}$
- $\tilde{\mathbf{C}^{\mathbf{T}}} = \mathbf{DC^T} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 & 4 & 0 \end{bmatrix}$
- Projection Matrix $\mathbf{P} = \mathbf{I} - \tilde{\mathbf{A}}^T(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T)^{-1}\tilde{\mathbf{A}} = \begin{bmatrix} \frac{5}{6} & -\frac{1}{6} & -\frac{1}{3} \\ -\frac{1}{6} & \frac{5}{6} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$
- Projected Gradient $\mathbf{c_P} = \mathbf{P}\tilde{\mathbf{C}^{\mathbf{T}}} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$

Complete Illustration of the Algorithm

Iteration 1

- Identify v (how far to move) as the absolute value of the negative component of $\mathbf{c_P}$ having the largest value, so that $v = |-2| = 2$
- In this coordinate system the algorithm moves from current solution to $\tilde{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{\alpha}{v}\mathbf{c_P} =$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{0.5}{2}\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} \\ \frac{7}{4} \\ \frac{1}{2} \end{bmatrix}$$

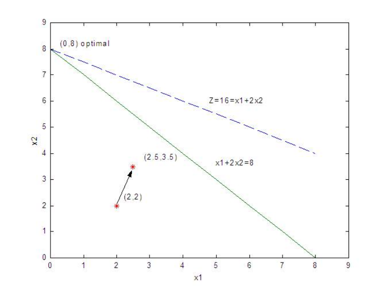


Complete Illustration of the Algorithm

Iteration 1

- In the original coordinate system the solution is $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \mathbf{D}\tilde{\mathbf{X}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

$$\begin{bmatrix} \frac{5}{4} \\ \frac{7}{4} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{2} \\ \frac{7}{2} \\ 2 \end{bmatrix}$$

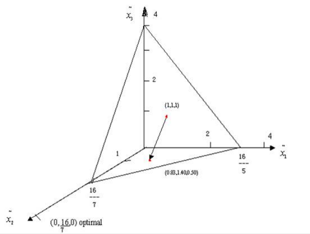


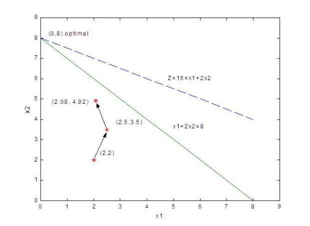
- This completes the iteration 1. The new solution will be used to start the next iteration

Complete Illustration of the Algorithm

Iteration 2

- Current trial Solution $(x_1, x_2, x_3) = (\frac{5}{2}, \frac{7}{2}, 2)$
- $\mathbf{D} = diag\{x_1, x_2, x_3\} = \begin{bmatrix} \frac{5}{2} & 0 & 0 \\ 0 & \frac{7}{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}$ $\tilde{\mathbf{X}} = \mathbf{D}^{-1}\mathbf{X} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- $\mathbf{P} = \begin{bmatrix} \frac{13}{18} & -\frac{7}{18} & -\frac{2}{9} \\ -\frac{7}{18} & \frac{41}{90} & -\frac{14}{45} \\ -\frac{2}{9} & -\frac{14}{45} & \frac{37}{45} \end{bmatrix}$ $\mathbf{c_P} = \begin{bmatrix} -\frac{11}{12} \\ \frac{133}{60} \\ -\frac{41}{15} \end{bmatrix}$
- The current solution becomes $(0.83, 1.4, 0.5)$ which corresponds $(2.08, 4.92, 1.0)$ in the original coordinate system

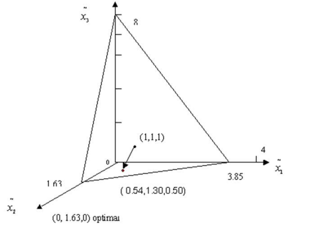


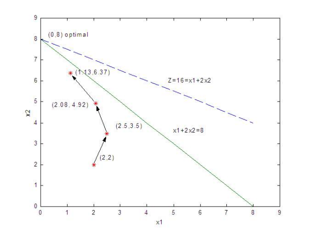


Complete Illustration of the Algorithm

Iteration 3

- Current trial Solution $(x_1, x_2, x_3) = (2.08, 4.92, 1.0)$
- $\mathbf{D} = diag\{x_1, x_2, x_3\} = \begin{bmatrix} 2.08 & 0 & 0 \\ 0 & 4.92 & 0 \\ 0 & 0 & 1.0 \end{bmatrix}$
- $\mathbf{c_P} = \begin{bmatrix} -1.63 \\ 1.05 \\ -1.79 \end{bmatrix}$
- The current solution becomes $(0.54, 1.3, 0.5)$ which corresponds $(1.13, 6.37, 0.5)$ in the original coordinate system

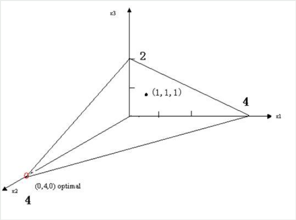




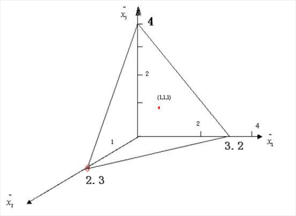
Complete Illustration of the Algorithm

Effect of rescaling of each iteration

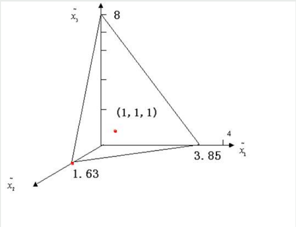
Sliding the optimal solution toward $(1, 1, 1)$ while the other BF solutions tend to slide away



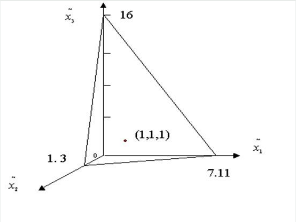
(1)



(2)



(3)

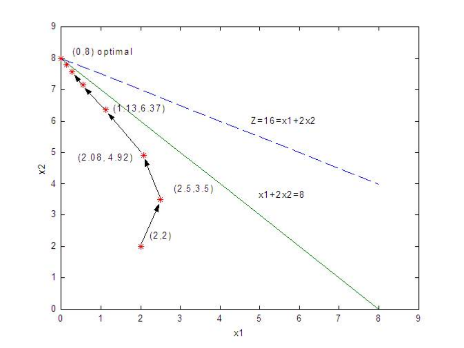


(4)

Complete Illustration of the Algorithm

More iterations

Starting from the current trial solution x, following the steps, x is moving toward the optimum $(0, 8)$. When the trial solution is virtually unchanged from the proceeding one, then the algorithm has virtually converged to an optimal solution. We stop.



Another Example

The Problem

Max $Z = 5x_1 + 4x_2$
such that

- $6x_1 + 4x_2 \leq 24$
- $x_1 + 2x_2 \leq 6$
- $-x_1 + x_2 \leq 1$
- $x_2 \leq 2$
- $x_j \geq 0,\ j = 1, 2$

Augmented Form

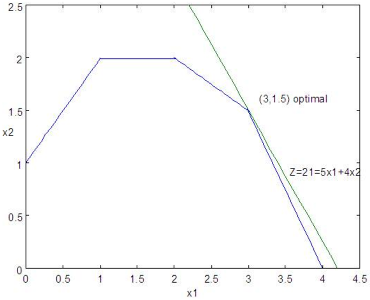
Max $Z = 5x_1 + 4x_2$
such that

- $6x_1 + 4x_2 + x_3 = 24$
- $x_1 + 2x_2 + x_4 = 6$
- $-x_1 + x_2 + x_5 = 1$
- $x_2 + x_6 = 2$
- $x_j \geq 0,\ j = 1, 2, 3, 4, 5, 6$

Another Example

The Matrices

$$\mathbf{A} = \begin{bmatrix} 6 & 4 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 24 \\ 6 \\ 1 \\ 2 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 5 & 4 & 0 & 0 & 0 & 0 \end{bmatrix}$$



Another Example

Starting from an initial solution $\mathbf{X} = (1, 1, 14, 3, 1, 1)$

The Steps

$$\mathbf{D} = diag\{\mathbf{X}\}$$

$$\tilde{\mathbf{A}} = \mathbf{AD}$$

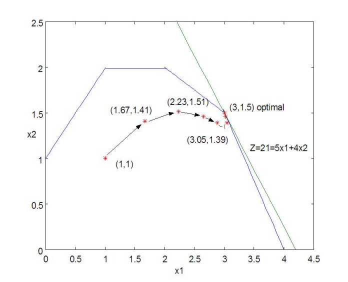
$$\tilde{\mathbf{c}} = \mathbf{DC}$$

$$\mathbf{P} = \mathbf{I} - \tilde{\mathbf{A}}^T(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T)^{-1}\tilde{\mathbf{A}}$$

$$\mathbf{c_P} = \mathbf{P}\tilde{\mathbf{c}}$$

$$\tilde{\mathbf{X}} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \frac{\alpha}{v}\mathbf{c_P}$$

$$\mathbf{X} = \mathbf{D}\tilde{\mathbf{X}}$$



Further Issues

Interior-point methods is designed for dealing with big problems. Although the claim that it's much faster than the simplex method is controversy, many tests on huge LP problems show its outperformance

Future Research

- Infeasible interior points method - remove the assumption that there always exits a nonempty interior
- Methods applying to LP problems in standard form
- Methods dealing with finding initial solution, and estimating the optimal solution
- Methods working with primal-dual problems
- Studies about moving step-long/short steps
- Studies about efficient implementation and complexity of various methods

Karmarkar's paper not only started the development of interior point methods, but also encouraged rapid improvement of simplex methods

References

Karmarkar's Algorithm

AK Dhamija

References

- N. Karmarkar, A New Polynomial - Time Algorithm for Linear Programming, Combinatorica 4 (4), 1984, p. 373-395
- Yanrong Hu & Ce Yu, Paper review of ENGG*6140 Optimization Techniques – Interior-Point Methods for LP, 2003
- I. Lustig, R. Marsten, D. Shanno, Interior-point Methods for Linear Programming: Computational State of the Art, ORSA Journal on Computing, 6 (1), 1994, p. 1-14
- Hillier,Lieberman,Introduction to Operations Research (7th edition) 320-334
- Taha, Operations Research: An Introduction (6th edition) 336-345
- D. Gay, A Variant of Karmarkar's Linear Programming Algorithm for Problems in Standard form, Mathematical Programming 37 (1987) 81-90
- E.R. Barnes, A Variation on Karmarkar's Algorithm for Solving Linear Programming problems, Mathematical Programming 36, 1986, p. 174-182
- V. Klee and G.J. Minty, How Good is the Simplex Algorithm? In O. Shisha, editor, Inequalities, III, pages 159-175. Academic Press, New York, NY, 1972.
- Richard Bronson and Govindasamy Naadimuthu, Schaum's Outlines, Operations Research, Second Edition, Mcgraw Hill.
- Wayne L Winston, Operations Research, Applications and Algorithms, Fourth Edition, Thomson Brooks/Cole, 2007

Presentation available at

`www.geocities.com/a_k_dhamija`.