LECTURE 6: INTERIOR POINT METHOD

1. Motivation
2. Basic concepts
3. Primal affine scaling algorithm
4. Dual affine scaling algorithm
Motivation

- **Simplex method** works well in general, but suffers from exponential-time computational complexity.
- Klee-Minty example shows simplex method may have to visit every vertex to reach the optimal one.
- **Total complexity** of an iterative algorithm
  \[\text{Total complexity} = \# \text{ of iterations} \times \# \text{ of operations in each iteration}\]
- Simplex method
  - Simple operations: Only check adjacent extreme points
  - May take many iterations: Klee-Minty example

Question: any fix?
Complexity of the simplex method

- Total # of elementary operations
  \[ = (\text{# of elementary operations at each iteration}) \times (\text{# of iterations}).\]

- # of elementary operations at each iteration of the revised simplex method \( O(mn) \).

- From practical experience, the simplex method takes about \((\alpha m)\) iterations where
  \[ e^\alpha < \log_2(2 + n/m). \] Hence it is of \( O(m^2n) \).

- From the worst-case analysis, Klee and Minty [1972] showed a class of examples (in the
dimensional space) which \( 2^d - 1 \) iterations for the simplex method.
Worst case performance of the simplex method

Klee-Minty Example:


\[
\begin{align*}
(2 \text{ dim}) & \quad \text{min} \quad -x_2 \\
\text{s. t.} & \quad x_1 \geq 0 \\
& \quad x_1 \leq 1 \\
& \quad x_2 \geq \varepsilon x_1 \quad (0 < \varepsilon < \frac{1}{2}) \\
& \quad x_2 \leq 1 - \varepsilon x_1 \\
& \quad x_1, x_2 \geq 0
\end{align*}
\]

\[x^0 \rightarrow x^1 \rightarrow x^1 \rightarrow x^3 \text{ (optimal)}\]
\[2^2 - 1 = 3 \text{ iterations}\]
Klee-Minty Example

\begin{align*}
(3 \text{ dim}) \quad & \min \quad -x_3 \\
\text{s. t.} \quad & x_1 \geq 0 \\
\quad & x_1 \leq 1 \\
\quad & x_2 \geq \epsilon x_1 \\
\quad & x_2 \leq 1 - \epsilon x_1 \\
\quad & x_3 \geq \epsilon x_2 \\
\quad & x_3 \leq 1 - \epsilon x_2 \\
\quad & x_1, x_2, x_3 \geq 0
\end{align*}

\[ 2^3 - 1 = 7 \text{ iterations} \]
Klee-Minty Example

\[(d \dim) \min -x_d\]
\[\text{s. t. } x_1 \geq 0\]
\[x_1 \leq 1\]
\[x_2 \geq \epsilon x_1\]
\[x_2 \leq 1 - \epsilon x_1\]
\[\vdots\]
\[x_d \geq \epsilon x_{d-1}\]
\[x_d \leq 1 - \epsilon x_{d-1}\]
\[x_i \geq 0\]

Hence, in theory, the simplex method is not a polynomial-time algorithm. It is an exponential time algorithm!

\[2^d - 1 \text{ iterations}\]
Karmarkar’s (interior point) approach

• Basic idea: approach optimal solutions from the interior of the feasible domain

• Take more complicated operations in each iteration to find a better moving direction
• Require much fewer iterations
General scheme of an interior point method

- An iterative method that moves in the interior of the feasible domain

Step 1: Start with an interior solution.

Step 2: If current solution is good enough, STOP. Otherwise,

Step 3: Check all directions for improvement and move to a better interior solution. Go to Step 2.
Interior movement (iteration)

• Given a current interior feasible solution $x^k$, we have

\[
A x^k = b \\
\quad x^k > 0
\]

An interior movement has a general format

\[
x^{k+1} = x^k + \alpha d_x^k
\]

\[
\begin{align*}
\alpha & \geq 0 : \text{Step - length} \\
\quad d_x^k & \in \mathbb{R}^n : \text{moving direction}
\end{align*}
\]
Key knowledge

1. Who is in the interior?
   - Initial solution

2. How do we know a current solution is optimal?
   - Optimality condition

3. How to move to a new solution?
   - Which direction to move? (good feasible direction)
   - How far to go? (step-length)
Q1 - Who is in the interior?

• Standard for LP

\[
\begin{align*}
\text{Min } c^T x \\
(\text{LP}) \quad \text{s. t. } Ax &= b \\
\quad x &\geq 0
\end{align*}
\]

• Who is at the vertex?
• Who is on the edge?
• Who is on the boundary?
• Who is in the interior?
What have learned before

Learning from example

Minimize $x_1 - 2x_2$

subject to

$x_1 + x_2 + x_3 = 40$

$2x_1 + x_2 + x_4 = 60$

$x_1, x_2, x_3, x_4 \geq 0$

What’s special?

- Vertices

$v^1 = \begin{pmatrix} 0 \\ 0 \\ 40 \\ 60 \end{pmatrix}, v^2 = \begin{pmatrix} 30 \\ 10 \\ 0 \\ 0 \end{pmatrix}, v^3 = \begin{pmatrix} 20 \\ 0 \\ 0 \\ 0 \end{pmatrix}, v^4 = \begin{pmatrix} 0 \\ 40 \\ 0 \\ 20 \end{pmatrix}$

- Edge

$v^5 = \begin{pmatrix} 20 \\ 0 \\ 20 \\ 20 \end{pmatrix}$ - one zero $x_i$

$v^6 = \begin{pmatrix} 15 \\ 15 \\ 10 \\ 15 \end{pmatrix}$ - no zero $x_i$

$n - 4, m - 2, n - m - 2$
Who is in the interior?

- **Two criteria** for a point $\mathbf{x}$ to be an interior feasible solution:
  
  1. $\mathbf{A}\mathbf{x} = \mathbf{b}$ (every linear constraint is satisfied)
  2. $\mathbf{x} > \mathbf{0}$ (every component is positive)

- **Comments:**
  
  1. On a hyperplane $H = \{ \mathbf{x} \in \mathbb{R}^n | \mathbf{a}^T \mathbf{x} = \beta \}$, every point is interior relative to $H$.
  2. For the first orthant $K = \{ \mathbf{x} \in \mathbb{R}^n | \mathbf{x} \geq \mathbf{0} \}$, only those $\mathbf{x} > \mathbf{0}$ are interior relative to $K$. 
Example

\[
\begin{align*}
\text{min} & \quad -2x_1 + x_2 \\
\text{s.t.} & \quad x_1 - x_2 \leq 15 \\
& \quad x_2 \leq 15 \\
& \quad x_1, x_2 \geq 0
\end{align*}
\]
How to find an initial interior solution?

- Like the simplex method, we have
  - Big M method
  - Two-phase method

(to be discussed later!)
Key knowledge

1. Who is in the interior?
   - Initial solution

2. How do we know a current solution is optimal?
   - Optimality condition

3. How to move to a new solution?
   - Which direction to move? (good feasible direction)
   - How far to go? (step-length)
Q2 - How do we know a current solution is optimal?

- Basic concept of optimality:
  A current feasible solution is optimal if and only if “no feasible direction at this point is a good direction.”

- In other words, “every feasible direction is not a good direction to move!”
Feasible direction

• In an interior-point method, a feasible direction at a current solution is a direction that allows it to take a small movement while staying to be interior feasible.

• Observations:

\[ x^{k+1} = x^k + \alpha d_x^k \]

\[ A x^k = b \]

\[ x^k > 0 \]

- There is no problem to stay interior if the step-length is small enough.
- To maintain feasibility, we need

\[ \frac{A x^{k+1}}{A x^k + \alpha A d_x^k} = \frac{b}{b} \implies A d_x^k = 0 \]

\[ i.e. \ d_x^k \in \mathcal{N}(A) \ null \ space \ of \ A. \]
Good direction

• In an interior-point method, a **good direction** at a current solution is a direction that leads it to a new solution with a **lower objective value**.

• Observations:

\[
\frac{c^T x^{k+1}}{c^T x^k + \alpha c^T d_x^k} \leq c^T x^k \leq c^T x^k \implies c^T d_x^k \leq 0
\]
Optimality check

• Principle:
  “no feasible direction at this point is a good direction.”

• At a current solution, we check that
  \[ \text{No } d_x^k \in \mathbb{R}^n \text{ with } A d_x^k = 0 \]
  can make
  \[ c^T d_x^k < 0 \]
Key knowledge

• 1. Who is in the interior?
  - Initial solution

• 2. How do we know a current solution is optimal?
  - Optimality condition

• 3. How to move to a new solution?
  - Which direction to move? (good feasible direction)
  - How far to go? (step-length)
Q3 – How to move to a new solution?

1. Which direction to move?
   - a good, feasible direction
     “Good” requires
     \[ c^T d_x^k \leq 0 \]
     “Feasible” requires
     \[ A d_x^k = 0 \]
     \[ d_x^k \in \mathcal{N}(A) : \text{null space of } A \]

Question: any suggestion?
A good feasible direction

• Reduce the objective value

\[ c^T d^k_x \leq 0 \]

**Candidate:** \[ d^k_x = -c \]

(negative gradient)
(Steepest descent)

• Maintain feasibility

\[ A d^k_x = 0 \]

**Candidate:** projected negative gradient

\[ d^k_x = (I - A^T (AA^T)^{-1} A)(-c) \]
Projection mapping

• A projection mapping projects the negative gradient vector $-c$ into the null space of matrix $A$

Formula for projection: $v = v_p + v_q$

$\mathcal{N}(M) = \{ x \mid Mx = 0 \}$

$v_p = [I - M^T(MMT)^{-1}M]v$

$v_q = MT(MMT)^{-1}Mv$

$d_k^x = (I - A^T(AA^T)^{-1}A)(-c)$
Q3 – How to move to a new solution?

2. How far to go?
   - To **satisfy every linear constraint**
     
     Since
     
     \[
     \mathbf{A}\mathbf{d}^k_x = 0
     \]
     \[
     \mathbf{d}^k_x \in \mathcal{N}(\mathbf{A}) : \text{null space of } \mathbf{A}
     \]
     
     \[
     \mathbf{A}\mathbf{x}^{k+1} = \mathbf{A}\mathbf{x}^k + \alpha \mathbf{A}\mathbf{d}^k_x = \mathbf{b}
     \]
     
     the step-length can be real number.

   - To **stay to be an interior solution**, we need
     
     \[
     \mathbf{x}^{k+1} > 0.
     \]
How to choose step-length?

- One easy approach
  - in order to keep
    \[ x^{k+1} = x^k + \alpha d^k_x \] > 0
  
  we may use the "minimum ratio test" to determine the step-length.

Observation:
- when \( x^k \) is close to the boundary, the step-length may be very small.

Question: then what?
Observations

• If a current solution is near the center of the feasible domain (polyhedral set), in average we can make a decently long move.

• If a current solution is not near the center, we need to re-scale its coordinates to transform it to become “near the center”.

Question: but how?
Where is the center?

- We need to know where is the “center” of the non-negative/first orthant \( \{ x \in \mathbb{R}^n \mid x \geq 0 \} \).

- Concept of equal distance to the boundary

If \( x^k = e \), then

1. \( x^k \) is one-unit away from the boundary
2. As long as \( \alpha < 1 \), \( x^{k+1} > 0 \)

Question: If not, what to do?
Concept of scaling

- **Scale** $x^k$ to be $e$
- Define a **diagonal scaling matrix**

\[ X_k = \text{diag}(x^k) = \begin{pmatrix} x_1^k & 0 & \cdots & 0 \\ 0 & x_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^k \end{pmatrix} \]

then \[ X_k^{-1}x^k = e \]
Transformation – affine scaling

• Affine scaling transformation

• The transformation is
  1. one-to-one
  2. onto
  3. invertible
  4. boundary to boundary
  5. interior to interior
Transformed LP

\[ x = X_k y \]

\[
\begin{align*}
\text{Min} & \quad c^T x \\
\text{s.t.} & \quad A x = b \\
& \quad x \geq 0
\end{align*}
\]

\[ x^k > 0 \]

\[
\begin{align*}
\text{Min} & \quad c^T X_k y \\
\text{s.t.} & \quad A X_k y = b \\
& \quad y \geq 0
\end{align*}
\]

\[ y^k = e \]

\[
d_y^k = [I - X_k A^T (A X_k^2 A^T)^{-1} A X_k] (-X_k c)
\]

\[
x^{k+1} = X_k y^{k+1}
\]

\[
x^{k+1} = X_k y^k + \alpha_k X_k \frac{d_y^k}{\|d_y^k\|}
\]

\[
x^{k+1} = x^k + \frac{\alpha_k}{\|d_y^k\|} d_x^k
\]

\[ \alpha_k = 0.99 \ (\text{say}) \quad 0 < \alpha_k < 1 \]

\[
\therefore \quad d_x^k = -X_k [I - X_k A^T (A X_k^2 A^T)^{-1} A X_k] X_k c
\]
Step-length in the transformed space

- Minimum ratio test in the y-space

In order to make sure that $y^{k+1} > 0$ we need

$$y^k + \alpha_k d_y^k > 0$$

\[ e \]

**Case 1:** $d_y^k \geq 0$ then $\alpha_k \in (0, \infty)$

**Case 2:** $(d_y^k)_i < 0$ for some $i$

$\alpha_k = \min_i \left\{ \frac{1}{-(d_y^k)_i} \mid (d_y^k)_i < 0 \right\}$

or

$\alpha_k = \min_i \left\{ \frac{\alpha}{-(d_y^k)_i} \mid (d_y^k)_i < 0 \right\}$ for some $\alpha \in (0, 1)$
Property 1

• Iteration in the $x$-space

\[ x^{k+1} = X_k y^{k+1} \]
\[ = X_k (e + \alpha_k d_y) \]
\[ = x^k + \alpha_k X_k d_y \]
\[ = x^k + \alpha_k X_k (-P_k X_k c) \]
\[ = x^k + \alpha_k [-X_k [I - X_k A^T (A X_k^2 A^T)^{-1} A X_k] X_k c] \]
\[ = x^k + \alpha_k [-X_k^2 [c - A^T (A X_k^2 A^T)^{-1} A X_k^2 c]] \]
\[ = x^k + \alpha_k [-X_k^2 [c - A^T w^k]] \]
\[ = x^k + \alpha_k d_x^k \]
Property 2

- Feasible direction in x-space

\[
x^{k+1} = X_k y^{k+1}
= X_k y^k + \alpha_k X_k \frac{d^k_y}{\|d^k_y\|}
= x^k + \frac{\alpha_k}{\|d^k_y\|} d^k_x
\]

Since \( d^k_y = P_k(-X_k c) \)

\[\therefore A X_k d^k_y = 0 \text{ and } A d^k_x = 0\]

i.e. \( d^k_x \in \mathcal{N}(A) \) null space of \( A \).
Property 3

- Good direction in x-space

\[
\begin{align*}
    c^T x^{k+1} &= c^T (x^k + \alpha_k X_k d_y^k) \\
    &= c^T x^k + \alpha_k c^T X_k (-P_k X_k c) \\
    &= c^T x^k - \alpha_k \| -P_k X_k c \|_2^2 \\
    &= c^T x^k - \alpha_k \| d_y^k \|_2^2 \\
\end{align*}
\]

Hence, \( c^T x^{k+1} \leq c^T x^k \)

and \( c^T x^{k+1} < c^T x^k \) if \( d_y^k \neq 0 \)

**Lemma 7.1** If \( \exists x^k \in P, \ x^k > 0 \) with \( d_y^k > 0 \), then the standard LP is unbounded below.
Property 4

- Optimality check (Lemma 7.2)

For \( x^k \in P^0 = \{ x \in \mathbb{R}^n \mid Ax = b, \ x > 0 \} \) if \( d^k_y = -P_k X_k c = 0 \) then \( X_k c \) falls in the orthogonal space of \( N(A X_k) \), i.e.

\[ X_k c \in \text{row space of } (A X_k) \]

\[ \Rightarrow \exists u^k \text{ s.t. } (A X_k)^T u^k = X_k c \]

or \( (u^k)^T A X_k = c^T X_k \)

\[ \Rightarrow (u^k)^T A = c^T \]

For any \( x \in P \),

\[ c^T x = (u^k)^T A x = (u^k)^T b \text{ (constant)} \]

\[ \therefore \text{ Any feasible solution is optimal !!} \]

In particular, \( x^k \) is optimal!
Property 5

- Well-defined iteration sequence (Lemma 7.3)

From properties 3 and 4, if the standard form LP is bounded below and $c^T x$ is not a constant, then the sequence $\{c^T x^k \mid k = 1, 2, \ldots\}$ is well-defined and strictly decreasing.
Property 6

- Dual estimate, reduced cost and stopping rule

We may define

$$w^k \equiv (AX_k^2 A^T)^{-1}AX_k^2 c$$ dual estimate

$$r^k \equiv c - A^T w^k$$ reduced cost

If $r^k \geq 0$, then $w^k$ is dual feasible

and $(x^k)^T r^k = e^T X_k r^k$ becomes the duality gap, i.e.,

Therefore, if $r^k \geq 0$ and $e^T X_k r^k = 0$

(Stopping rule)

then $x^k \leftarrow x^*$, $w^k \leftarrow w^*$
Property 7

- Moving direction and reduced cost

\[
d^k_y = \left[ I - X_k A^T (A X_k^2 A^T)^{-1} A X_k \right] ( - X_k c )
= - X_k ( c - A^T (A X_k^2 A^T)^{-1} A X_k^2 c )
= - X_k ( c - A^T w^k )
= - X_k r^k
\]
Primal affine scaling algorithm

**Step 1** Set $k \leftarrow 0$, $\varepsilon > 0$, $0 < \alpha < 1$
find $x^0 > 0$ and $Ax^0 = b$

**Step 2** Compute
$w^k = (AX_k^2A^T)^{-1}AX_k^2c$
$r^k = c - A^Tw^k$
If $r^k \geq 0$, and $e^TX_kr^k \leq \varepsilon$
then STOP! $x^* \leftarrow x^k$, $w^* \leftarrow w^k$
Otherwise,

**Step 3** Compute $d_y^k = -X_kr^k$
If $d_y^k \neq 0$, then STOP! Unbounded.
If $d_y^k = 0$, then STOP! $x^* \leftarrow x^k$
 Otherwise,

**Step 4** Find
$\alpha_k = \min_i \left\{ \frac{\alpha}{-(d_y^k)_i} \mid (d_y^k)_i < 0 \right\}$

$x^{k+1} = x^k + \alpha_kX_kd_y^k$
$k \leftarrow k + 1$
Go to Step 2.
Example

min \ -2x_1 + x_2
s.t. \ x_1 - x_2 \leq 15
      \ x_2 \leq 15
      \ x_1, x_2 \geq 0

Reformulate to standard form

min \ -2x_1 + x_2
s.t. \ x_1 - x_2 + x_3 = 15
      \ x_2 + x_4 = 15
      \ x_1, x_2, x_3, x_4 \geq 0

x^0 = \begin{pmatrix} 10 \\ 2 \\ 7 \\ 13 \end{pmatrix} \text{ is feasible} \quad X_0 = \begin{bmatrix} 10 & 0 \\ 2 & 0 \\ 0 & 7 \\ 0 & 13 \end{bmatrix}
Example

\[
X_0 = \begin{bmatrix}
10 & 0 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 7 & 0 \\
0 & 0 & 0 & 13
\end{bmatrix}
\quad \text{and} \quad \mathbf{w}^0 = (AX_0^2A^T)^{-1}AX_0^2\mathbf{c} = [-1.33353, -0.00771]^T
\]

Moreover,

\[
\mathbf{r}^0 = \mathbf{c} - A^T\mathbf{w}^0 = [-0.66647, -0.32582, 1.33535, -0.00771]^T
\]

Since some components of \( \mathbf{r}^0 \) are negative and \( \mathbf{c}^T X_0 \mathbf{r}^0 = 2.1187 \), we know that the current solution is nonoptimal. Therefore we proceed to synthesize the direction of translation with

\[
d_y^0 = -X_0 \mathbf{r}^0 = [6.6647, 0.6516, -9.3475, 0.1002]^T
\]

Suppose that \( \alpha = 0.99 \) is chosen, then the step-length

\[
\alpha_0 = \frac{0.99}{9.3475} = 0.1059
\]

Therefore, the new solution is

\[
x^1 = x^0 + \alpha_0 X_0 d_y^0 = [17.06822, 2.13822, 0.07000, 12.86178]^T
\]

Notice that the objective function value has been improved from \(-18\) to \(-31.99822\).

The reader may continue the iterations further and verify that the iterative process converges to the optimal solution \( x^* = [30, 15, 0, 0]^T \) with optimal value \(-45\).
How to find an initial interior feasible solution?

- **Big-M method**
  
  Idea: add an artificial variable with a big penalty

  \[
  \begin{align*}
  (LP) & \quad \min \ c^T x \\
  & \text{s.t.} \quad Ax = b \\
  & \quad x \geq 0
  \end{align*}
  \]

  \[
  \begin{align*}
  (\text{big-M}) & \quad \min \ c^T x + Mx^a \\
  & \text{s.t.} \quad Ax + (b - Ae)x^a = b \\
  & \quad x \geq 0, \ x^a \geq 0
  \end{align*}
  \]

- **Objective**

  
  To make \( e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \) be feasible, i.e., \( Ae = b \)?
Properties of (big-M) problem

(1) It is a standard form LP with $n+1$ variables and $m$ constraints.

(2) $e$ is an interior feasible solution of (big-M).

(3) If $x^a > 0$ in $(x^*, x^a)$ then (LP) is infeasible. Otherwise, either (LP) is unbounded or $x^*$ is optimal to (LP).
Two-phase method

\[(LP) \begin{align*}
\text{min} & \quad c^T x \\
\text{s.t.} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}\]

Choose any \( x^0 > 0 \), calculate \( v = b - Ax^0 \).

If \( v = 0 \), then \( x^0 \) is interior feasible.

Otherwise, consider

\[\text{(Phase - I) } \begin{align*}
\text{min} & \quad u \\
\text{s.t.} & \quad Ax + vu = b \\
& \quad x \geq 0, \; u \geq 0
\end{align*}\]
Properties of (Phase-I) problem

1. (Phase-I) is a standard form LP with $n + 1$ variables and $m$ constraints.

2. $\bar{x}^0 = \begin{pmatrix} x^0 \\ u^0 \end{pmatrix} = \begin{pmatrix} x^0 \\ 1 \end{pmatrix}$ is interior feasible for (Phase-I).

3. (Phase-I) is bounded below by 0.

4. Apply primal-affine scaling to (Phase-I) will generate $\begin{pmatrix} x^* \\ u^* \end{pmatrix}$ for (Phase-I).
   
   If $u^* > 0$, (LP) is infeasible.
   Otherwise, $x^* > 0$ for (Phase-II) as an initial feasible solution.
Facts of the primal affine scaling algorithm

(1) The convergence proof, i.e.,
\[ \{x^k\} \rightarrow x^* \]
under Non-degeneracy assumption (Theorem 7.2) is given by Vanderbei/Meketon/Freedman in (1985).

(2) Convergence proof without Non-degeneracy assumption,
T. Tsuchiya (1991)
P. Tseng/ Z. Luo (1992)

(3) The computational bottleneck is to find
\[ (AX_k^2A^T)^{-1} \]

(4) No polynomial-time proof
- J. Lagarias showed primal affine scaling is only of super-linear rate.
- N. Megiddo/M. Shub showed that primal affine scaling might visit all vertices if it moves too close to the boundary.
More facts

(5) In practice, VMF reported

<table>
<thead>
<tr>
<th></th>
<th># iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simplex</td>
<td>$0.7159 , m^{0.9522} , n^{0.3109}$</td>
</tr>
<tr>
<td>Affline Scaling</td>
<td>$7.3385 , m^{-0.0187} , n^{0.1694}$</td>
</tr>
</tbody>
</table>

(6) It may lose primal feasibility due to machine accuracy (Phase-I again).

(7) May be sensitive to primal degeneracy.
Improving performance – potential push

- Idea: (Potential push method)
  - Stay away from the boundary by adding a potential push.

\[
\begin{align*}
\min & \quad -\sum_{j=1}^{n} \log_e x_j \\
\text{s.t.} & \quad Ax = b, \ x > 0 \\
& \quad c^T x = c^T x^k
\end{align*}
\]

Use \((x^k)'\) to replace \(x^k\)
Improving performance – logarithmic barrier

- Idea: (Logarithmic barrier function method)
  Consider

\[
\begin{align*}
\min & \quad c^T x - \mu \sum_{j=1}^{n} \log x_j \\
\text{s.t.} & \quad Ax = b \\
& \quad x > 0
\end{align*}
\]

Properties:

(1) \( \{ x^*(\mu) \mid \mu > 0 \} \rightarrow x^* \)

(2) \[
\begin{align*}
d_k^\mu &= X_k[I - X_k A^T (AX_k^2 A^T)^{-1} A X_k](-X_k c + \mu e) \\
&= X_k P_k (-X_k c) + \mu X_k P_k e \\
&= d_k^x + \underbrace{\mu X_k P_k e}_{\text{centering force}}
\end{align*}
\]

(3) Polynomial-time proof, i.e., terminates in \( O(\sqrt{nL}) \) iterations.

C. Gonzaga (1989) (Problems in Proof !!)
C. Roos/ J. Vial (1990)
- Total complexity \( O(n^3L) \)!
Dual affine scaling algorithm

• Affine scaling method applied to the dual LP

\[
\begin{align*}
\max \quad & b^T w \\
(D) \quad \text{s.t.} \quad & A^T w + s = c \\
\quad & s \geq 0
\end{align*}
\]

Given \((w^k, s^k)\) dual interior feasible, i.e.,

\[
A^T w^k + s^k = c \\
\quad s^k > 0
\]

Objective find \((d^k_w, d^k_s)\) and \(\beta_k > 0\) such that

\[
\begin{align*}
w^{k+1} &= w^k + \beta_k d^k_w \\
s^{k+1} &= s^k + \beta_k d^k_s
\end{align*}
\]

is still dual interior feasible, and

\[
b^T w^{k+1} \geq b^T w^k
\]
Key knowledge

- Dual scaling (centering)
- Dual feasible direction
- Dual good direction – increase the dual objective value
- Dual step-length
- Primal estimate for stopping rule
Observation 1

- Dual scaling (centering)
Observation 2

- Dual feasibility (feasible direction)

\[
\frac{A^T w^{k+1} + s^{k+1}}{c} = A^T (w^k + \beta_k d_w^k) + (s^k + \beta_k d_s^k)
\]

\[
= (A^T w^k + s^k)
\]

\[
+ \beta_k (A^T d_w^k + d_s^k)
\]

\[
> 0
\]

\[
\Rightarrow A^T d_w^k + d_s^k = 0 \text{ is required!}
\]

\[
\Leftrightarrow \quad S_k^{-1} A^T d_w^k + S_k^{-1} d_s^k = 0
\]

\[
\Leftrightarrow \quad A S_k^{-1} (S_k^{-1} A^T d_w^k + d_s^k) = 0
\]

\[
\Leftrightarrow \quad (A S_k^{-2} A^T) d_w^k + A S_k^{-1} d_u^k = 0
\]

\[
\Leftrightarrow \quad d_w^k = - (A S_k^{-2} A^T)^{-1} A S_k^{-1} d_u^k
\]
Observation 3

- Increase dual objective function (good direction)

\[ b^T w^{k+1} = b^T w^k + \beta_k b^T d_w^k \geq b^T w^k \]

\[ b^T d_w^k = -b^T Q d_u^k \geq 0 \]

We can choose

\[ d_u^k = -Q^T b \]

then

\[ b^T d_w^k = b^T Q Q^T b = \|Q^T b\|^2 \geq 0 !! \]

Thus

\[ d_w^k = -Q d_u^k \]

\[ = QQ^T b \]

\[ = (A S_k^{-2} A^T)^{-1} A S_k^{-1} S_k^{-1} A^T (A S_k^{-2} A^T)^{-1} b \]

\[ = (A S_k^{-2} A^T)^{-1} b \]

and

\[ d_s^k = -A^T d_w^k = -A^T (A S_k^{-2} A^T)^{-1} b \]
Observation 4

- Dual step-length

\[ s^{k+1} = s^k + \beta_k d_g^k > 0 \]

(i) \( d_g^k = 0 \), problem (D) has a constant objective value and \((w^k, s^k)\) optimal.

(ii) \( d_g^k \neq 0 \), \( \beta_k \in (0, \infty) \)

problem (D) is unbounded

(iii) some \((d_g^k)_i < 0\)

\[ \beta_k = \min_i \left\{ \frac{\alpha s_i^k}{-(d_g^k)_i} | (d_g^k)_i < 0 \right\} \]

for \( \alpha \in (0, 1) \)
Observation 5

- **Primal estimate**

We define

\[ x^k \triangleq -S_k^{-2}d_s \]

then

\[
Ax^k = -AS_k^{-2}(-A^Td_w) \\
= AS_k^{-2}A^Td_w \\
= (AS_k^{-2}A^T)(AS_k^{-2}A^T)^{-1}b \\
= b
\]

If \( c^T x^k - b^T w^k = 0 \), then

\[
x^k \leftarrow x^* \\
w^k \leftarrow w^* \\
s^k \leftarrow s^*
\]

Hence \( x^k \) is a primal estimate, once \( x^k \geq 0 \), then \( x^k \) is primal feasible.
**Dual affine scaling algorithm**

**Step 1:** Set $k = 0$ and find $(w^0, s^0)$ s.t.

$$A^T w^0 + s^0 = c, \quad s^0 > 0$$

**Step 2:** Set $S_k = \text{diag} (s^k)$

Compute

$$d_w^k = (AS_k^{-2} A^T)^{-1} b$$

$$d_s^k = -A^T d_w^k$$

**Step 3:** If $d_s^k = 0$, STOP! $w^k \leftarrow w^*$, $s^k \leftarrow s^*$

If $d_s^k \neq 0$, STOP! (D) is unbounded

**Step 4:** Compute

$$x^k = -S_k^{-2} d_s^k$$

If $x^k \geq 0$ and $c^T x^k - b^T w^k \leq \varepsilon$

STOP!

$w^k \leftarrow w^*$, $s^k \leftarrow s^*$, $x^k \leftarrow x^*$

**Step 5:** Compute

$$\beta_k = \min_i \left\{ \frac{\alpha s^k_i}{-(d_s^k)_i} \mid (d_s^k)_i < 0 \right\}$$

**Step 6:**

$$w^{k+1} = w^k + \beta_k d_w^k$$

$$s^{k+1} = s^k + \beta_k d_s^k$$

Set $k \leftarrow k + 1$ Go to Step 2.
Find an initial interior feasible solution

Find \((w^0, s^0)\) s.t.

\[
A^T w^0 + s^0 = c
\]

\[
s^0 > 0
\]

If \(c > 0\), then \(w^0 = 0, s^0 = c\) will do.

(Big-M Method)

Define \(p \in R^n, \quad p_i = \begin{cases} 1 & \text{if } c_i \leq 0 \\ 0 & \text{if } c_i > 0 \end{cases} \)

Consider, for a large \(M > 0\),

\[
\begin{align*}
\max & \quad b^T w + M w^a \\
\text{s.t.} & \quad A^T w + p w^a + s = c \\
& \quad w, w^a \text{ unrestricted} \\
& \quad s \geq 0
\end{align*}
\]

(Big-M Problem)
Properties of (big-M) problem

(a) (Big-M) is a standard LP with \( n \) constraints and \( m + 1 + n \) variables.

(b) Define \( \bar{c} = \max_i |c_i| \) and \( \theta > 1 \) then

\[
\begin{align*}
    w &= 0 \\
    w^a &= -\theta \bar{c} \\
    s &= c + \theta \bar{c} p > 0
\end{align*}
\]

is an initial interior feasible solution for problem (D).

(c) \( (w^a)^0 = -\theta \bar{c} < 0 \)

Since \( M > 0 \) is large

\( (w^a)^k \to 0 \) as \( k \to +\infty \)

if \( (w^a)^k \) does not approach or cross zero, then problem (D) is infeasible.
Performance of dual affine scaling

- No polynomial-time proof!
- Computational bottleneck

\[(AS_k^{-2}A^T)^{-1}\]

- Less sensitive to primal degeneracy and numerical errors, but sensitive to dual degeneracy.
- Improves dual objective value very fast, but attains primal feasibility slowly.
Improving performance

1. Logarithmic barrier function method

\[
\begin{align*}
\{ & \max \ b^T w + \mu \sum_{j=1}^{n} \ln[e_j - A_j^T w] \\
& \text{s.t.} \quad A^T w < c \\
\} \\
\end{align*}
\]

\[
\Delta w = \frac{1}{\mu} (A S_K^{-2} A^T)^{-1} b - (A S_K^{-2} A^T) A S_K^{-1} e
\]

\[
d_w^k \quad \text{centering force}
\]

as \( \mu \to 0, \quad w^k(\mu) \to w^* \)

Polynomial-time proof

J. Renegar \( O(n^{3.5} L) \)
P. Vaidya \( O(n^3 L) \)
C. Roos/ J. Vial \( O(n^3 L) \)
Improving performance

- Power series method
  - Basic idea: following a higher order trajectory

$$\begin{align*}
\frac{dw(\beta)}{d\beta} &= \lim_{\beta_k \to 0} \frac{w^{k+1} - w^k}{\beta_k} \\
\frac{ds(\beta)}{d\beta} &= -A^T \frac{dw(\beta)}{d\beta}
\end{align*}$$

Initial condition

$$w(0) = w^0, \quad s(0) = s^0$$

where

$$S(\beta) = \text{diag}(s^0 + \beta d_s)$$
Power series expansion

\[ w(\beta) = w^0 + \sum_{i=1}^{\infty} \beta^j \left[ \frac{1}{j!} \right] \left[ \frac{d^j w(\beta)}{d \beta^j} \right]_{\beta=0} \]

\[ s(\beta) = s^0 + \sum_{i=1}^{\infty} \beta^j \left[ \frac{1}{j!} \right] \left[ \frac{d^j s(\beta)}{d \beta^j} \right]_{\beta=0} \]

(a) As long as 
\[ \left[ \frac{d^j w(\beta)}{d \beta^j} \right]_{\beta=0} \] and \[ \left[ \frac{d^j s(\beta)}{d \beta^j} \right]_{\beta=0}, \ j = 1, 2, \ldots, n \]
are known, \[ w(\beta), \ s(\beta) \] are known.

(b) Dual Affine Scaling is the case of first-order approximation

\[ w(\beta) = w^0 + \beta \left[ \frac{d w(\beta)}{d \beta} \right]_{\beta=0} \]

\[ s(\beta) = s^0 + \beta \left[ \frac{d s(\beta)}{d \beta} \right]_{\beta=0} \]

(c) A power-series approximation of order 4 or 5 cuts total \# of iterations by \( 1/2 \).