LARGE-SCALE OPTIMIZATION

Interior-point methods

Types of interior-point methods
The primal affine-scaling method

Jordi Castro

CONTENTS

PROBLEM FORMULATION

$$\begin{array}{ll} \min\limits_{x} & c^\top x \\ \text{s. to} & Ax = b \\ & x \geq 0 \end{array} \qquad (P)$$

- $x \in \mathbb{R}^n$ is vector of variables.
- $A \in \mathbb{R}^{m\times n}$ $(n \geq m)$ is constraints matrix.
- $b \in \mathbb{R}^m$ is right-hand-side vector.
- $c \in \mathbb{R}^n$ is cost vector.
- $\Omega = \{x \mid Ax = b, x \geq 0\}$ is the *feasible set*.
- if $x^0 \in \Omega$ then is a *feasible point*.

HYPOTHESIS ABOUT (P)

- Matrix A is full row rank (r(A)= m).
- (P) is nondegenerate (any feasible point has at least m nonzero components).

BACKGROUND

- **Dual problem**

$$\begin{array}{ll} \max\limits_{y\in\mathbb{R}^m, z\in\mathbb{R}^n} & b^\top y \\ \text{s. to} & A^\top y + z = c \\ & z \geq 0 \end{array} \qquad (D)$$

- **Strong Duality theorem**

  If (P) has an optimal solution x^*, then (D) also has an optimal solution y^* such that

$$c^\top x^* = b^\top y^*$$

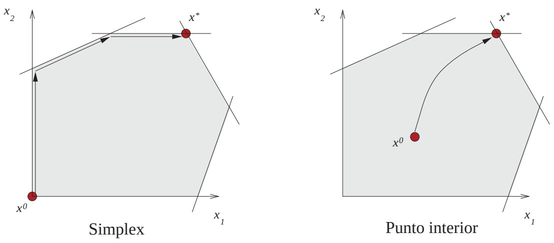
- **Complementary Slackness theorem**

  If x is feasible for (P), and (y, z) is feasible for (D), then they are optimal solutions if and only if

$$x_i z_i = 0 \quad \text{for all } i = 1, \ldots, n$$

INTERIOR-POINT METHODS

- Family of non-simplex methods for Linear Programming, which appeared in 1984 with Karmarkar algorithm.
- Unlike simplex, they iterate through the interior of the feasible region.



- Some interior-point methods have polynomial time complexity. The most efficient have complexity $O(\sqrt{n}L)$.

CLASSIFICATION OF INTERIOR-POINT METHODS

Broadly, they can be classified in 4 categories:

- Affine-scaling methods.
$$\tilde{x}_i = \beta_i x_i$$
- Methods based on projective transformations (e.g., Karmarkar algorithm).
$$\tilde{x}_i = \frac{\beta_i x_i}{\sum_{j=1}^{n} \beta_j x_j}$$
- *Path-following* methods (e.g., primal-dual path-following method). Currently, the most efficient.
- Potential-reduction methods. The computed directions do not follow the central path, instead they measure the quality of the directions through the reduction of a potential function.

FEATURES OF AFFINE-SCALING

- Introduced by Russian mathematician I.I. Dikin in 1967. Rediscovered in the West by 1986.
- It is the *simplest* interior-point method.
- It can be easily implemented.
- In general, it provides good computational results.

OUTLINE OF PRESENTATION

- Basic algorithm.
- Computing the direction of movement.
- Scaling.
- Computing the step length.
- A digression: short-step affine-scaling.
- Stopping criterion.
- Computing initial feasible solution.
- The affine-scaling algorithm.
- Convergence results.
- Computational performance.

STARTING POINT

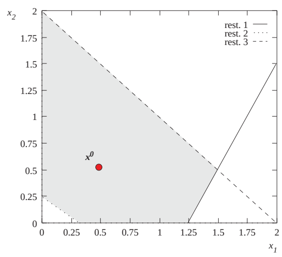
- Starting point x^0 must be feasible and interior:
$$Ax^0 = b \qquad \text{and} \qquad x^0 > 0.$$

EXAMPLE 1

Given problem

$$\begin{array}{ll} \min & -3x_1 - 2x_2 \\ \text{s. to} & \\ & \begin{array}{rrcl} 4x_1 & -2x_2 & \le & 5 \\ 3x_1 & +4x_2 & \ge & 1 \\ x_1 & +\ x_2 & \le & 2 \end{array} \\ & x_1 \ge 0 \quad x_2 \ge 0 \end{array} \Longrightarrow \begin{array}{ll} \min & -3x_1 - 2x_2 \\ \text{s. to} & \\ & \begin{array}{rrrcl} 4x_1 & -2x_2 & +x_3 & = & 5 \\ 3x_1 & +4x_2 & -x_4 & = & 1 \\ x_1 & +\ x_2 & +x_5 & = & 2 \end{array} \\ & x_i \ge 0, \quad i = 1, \dots, 5 \end{array}$$

point $x^0 = (1/2 \;\; 1/2 \;\; 4 \;\; 5/2 \;\; 1)^\top$ is interior and feasible.



ITERATIVE PROCEDURE

- Compute sequence of points $\{x^k\}$ through the iterative procedure
$$x^{k+1} = x^k + \alpha \Delta x, \quad \alpha \geq 0$$
where

  Δx is the *direction of movement*

  α is the *step length*.
- Two questions to be answered:

  i) How to compute Δx?

  ii) How to compute α?

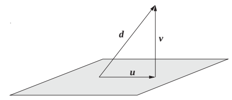
CONDITIONS FOR Δx

1. Preserving feasibility

$$\left.\begin{aligned} b &= Ax^{k+1} \\ &= A(x^k + \alpha\Delta x) \\ &= Ax^k + \alpha A \Delta x \\ &= b + \alpha A \Delta x \end{aligned}\right\} \Rightarrow \boxed{A\Delta x = 0}$$

- Therefore $\Delta x \in N(A) = \{v \mid Av = 0, v \in \mathbb{R}^n\}$ (null space of A).
- For all $d \in \mathbb{R}^n$ we have that $Pd \in N(A)$, where P is the orthogonal projection matrix (if A is full row rank)
$$P = I_n - A^\top (AA^\top)^{-1} A$$

Exercise 1. Derive the expression of P using the orthogonal decomposition of any vector d: $d = u + v$, $u \in N(A)$.



Exercise 2. Check that P satisfies:

i) $AP = \mathbf{0}$ ii) $P = P^\top$ (symmetric) iii) $P^2 = P$ (idempotent).

EXAMPLE 2

Consider the problem

$$\begin{aligned} \min \quad & -1/3\, x_1 - 2/3\, x_2 \\ \text{s. to} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0 \end{aligned}$$

with current point $x^k = (1/3 \;\; 1/3 \;\; 1/3)^\top$. In this problem $A = (1 \;\; 1 \;\; 1)$.

The projection matrix P is:

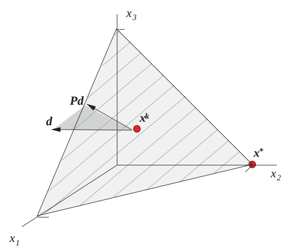
$$P = I_3 - A^\top (AA^\top)^{-1} A = 1/3 \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$$

Given some particular vector $d = (0 \quad -1/2 \;\; 0)^\top$, its projection is:

$$Pd = (\,1/6 \quad -1/3 \quad 1/6\,)^\top .$$

Check that $x^{k+1} = x^k + Pd$ is feasible:

$$x^k + Pd = (\,1/2 \quad 0 \quad 1/2\,)^\top \qquad 1/2 + 0 + 1/2 = 1.$$



2. Decreasing the objective function

In x^{k+1} we must guarantee:

$$c^\top x^{k+1} \leq c^\top x^k$$

Therefore Δx must satisfy the *descent condition*

$$c^\top x^{k+1} = c^\top (x^k + \alpha \Delta x) = c^\top x^k + \alpha c^\top \Delta x \leq c^\top x^k$$
$$\Longrightarrow \quad \boxed{c^\top \Delta x \leq 0}$$

We can use the negative *projected gradient* direction

$$\boxed{\Delta x = -Pc}$$

since

$$c^\top \Delta x = -c^\top P c = -c^\top P^2 c = -c^\top P^\top P c = -||Pc||^2 \leq 0$$

Exercise 3. Show that if $||Pc||^2 = 0$ then any feasible point is optimal.

EXAMPLE 3

Consider again the problem

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

with current point $x^k = (1/3 \quad 1/3 \quad 1/3)^\top$. The projection of $-c = (1/3 \quad 2/3 \quad 0)^\top$ is

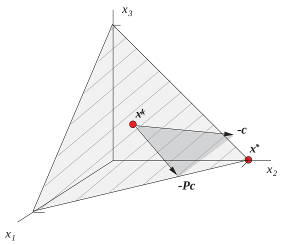
$$-Pc = (\,0 \quad 1/3 \quad -1/3\,)^\top$$

And the new point

$$x^k - Pc = (\,1/3 \quad 2/3 \quad 0\,)^\top, \quad 1/3 + 2/3 + 0 = 1$$

We decreased the objective function:

$$c^\top(x^k - Pc) = -5/9 \leq c^\top x^k = -1/3$$



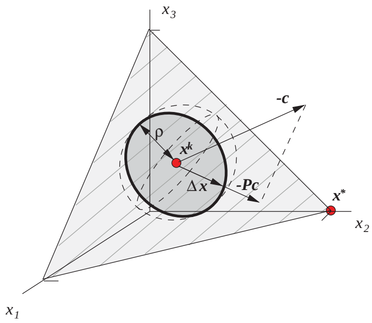
JUSTIFICATION OF $\Delta x = -Pc$

Computing the feasible direction of maximum descent:

$$\begin{array}{ll} \min\limits_{\Delta x} & c^\top(x^k + \Delta x) \\ \text{s. to} & A(x^k + \Delta x) = b \\ & ||\Delta x||^2 = \rho^2 \end{array}$$

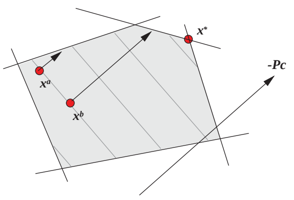
Solution:

$$\Delta x = -\frac{\rho}{||Pc||} Pc$$



MOTIVATION FOR SCALING

- From $x^{k+1} = x^k + \alpha\Delta x$, *well centered* points allow for better α: $\alpha \gg 0$.



A THREE-STAGES PROCEDURE

1. Scale the problem such that current point is far from bound constraints $x \geq 0$. Current point x^k is mapped to $\tilde{x}^k$ in the new space of variables $\tilde{x}$.
2. Compute the direction $\Delta\tilde{x}$ of minus the projected gradient in the new scaled problem.
3. Map the direction $\Delta\tilde{x}$ in the original problem, reverting the scaling, and computing Δx.

THE AFFINE-SCALING CONSIDERED

- Map each component of x^k to 1:

$$\tilde{x}_i = \frac{x_i}{x_i^k} \quad i = 1, \ldots, n$$

- The inverse mapping is easily computed:

$$x_i = \tilde{x}_i x_i^k \quad i = 1, \ldots, n.$$

- In matrix form:

$$X^k = \begin{pmatrix} x_1^k & & & \\ & x_2^k & & \\ & & \ddots & \\ & & & x_n^k \end{pmatrix}$$

$$\tilde{x} = (X^k)^{-1} x$$
$$x = X^k \tilde{x}$$

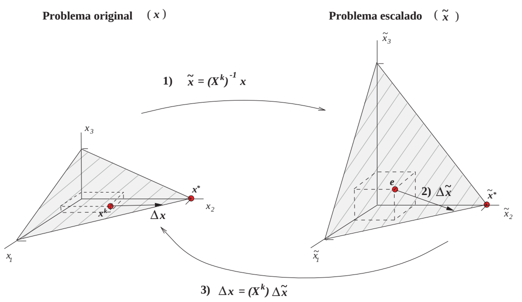
STAGE 1. Scaling: $x \to \tilde{x}$

$$
\begin{array}{ll}
\min\limits_{x} & c^\top x \\
\text{s. to} & Ax = b \\
 & x \geq 0
\end{array}
\longrightarrow
\begin{array}{ll}
\min\limits_{\tilde{x}} & c^\top X^k \tilde{x} \\
\text{s. to} & AX^k\tilde{x} = b \\
 & X^k\tilde{x} \geq 0
\end{array}
\longrightarrow
\begin{array}{ll}
\min\limits_{\tilde{x}} & \tilde{c}^\top \tilde{x} \\
\text{s. to} & \widetilde{A}\tilde{x} = b \\
 & \tilde{x} \geq 0
\end{array}
$$

$$x = X^k\tilde{x} \qquad \tilde{c} = X^k c \quad \widetilde{A} = AX^k$$

STAGE 2. Compute $\Delta\tilde{x}$

$$\Delta\tilde{x} = -\widetilde{P}\tilde{c} = -(I_n - \widetilde{A}^\top(\widetilde{A}\widetilde{A}^\top)^{-1}\widetilde{A})\tilde{c}$$

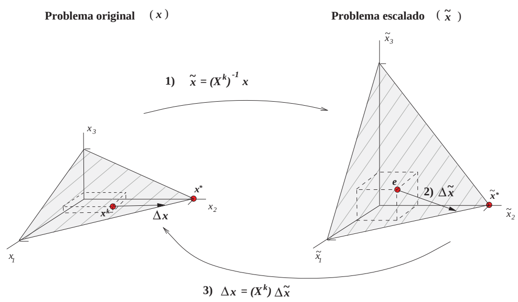


STAGE 3. Back to original problem: $\tilde{x} \to x$

Direction Δx computed from $\Delta\tilde{x}$ with the inverse mapping:

$$\Delta x = X^k\Delta\tilde{x} = -Dz \quad \text{where} \quad \begin{array}{l} D = (X^k)^2 \\ z = c - A^\top y \\ y = (ADA^\top)^{-1}ADc \end{array}$$

Matrix $ADA^\top = A(X^k)^2A^\top$ is symmetric and positive semidefinite ($x^k \geq 0$). It is also nonsingular since A is full rank and problem is nondegenerate.



Exercise 4. Show that $\Delta x = -Dz$ is a feasible ($A\Delta x = 0$) and descent ($c^\top\Delta x \leq 0$) direction.

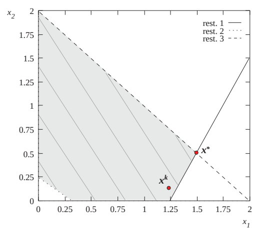
EXAMPLE 4

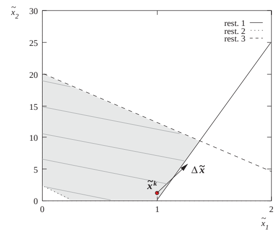
Given problem

$$
\begin{array}{lrcl}
\min & -3x_1 - 2x_2 & & \\
\text{s. to} & & & \\
 & 4x_1 - 2x_2 & \leq & 5 \\
 & 3x_1 + 4x_2 & \geq & 1 \\
 & x_1 + x_2 & \leq & 2 \\
 & x_1 \geq 0 \quad x_2 \geq 0 & &
\end{array}
\Longrightarrow
\begin{array}{lrcl}
\min & -3x_1 - 2x_2 & & \\
\text{s. to} & & & \\
 & 4x_1 - 2x_2 + x_3 & = & 5 \\
 & 3x_1 + 4x_2 - x_4 & = & 1 \\
 & x_1 + x_2 + x_5 & = & 2 \\
 & x_i \geq 0, \quad i = 1, \ldots, 5 & &
\end{array}
$$

and current point $x^k = (\,1.25 \quad 0.1 \quad 0.2 \quad 3.15 \quad 0.65\,)^\top$, the scaled problem is

$$
\begin{array}{lrcl}
\min & -3.75\,\tilde{x}_1 - 0.2\,\tilde{x}_2 & & \\
\text{s. to} & & & \\
 & 5\,\tilde{x}_1 - 0.2\,\tilde{x}_2 + 0.2\,\tilde{x}_3 & = & 5 \\
 & 3.75\,\tilde{x}_1 + 0.4\,\tilde{x}_2 - 3.15\,\tilde{x}_4 & = & 1 \\
 & 1.25\,\tilde{x}_1 + 0.1\,\tilde{x}_2 + 0.65\,\tilde{x}_5 & = & 2 \\
 & \tilde{x}_i \geq 0 \text{ for all } i = 1, \ldots, 5 & &
\end{array}
$$





(EXAMPLE (cont.) 4)

- Direction $\Delta\tilde{x}$

$$\Delta\tilde{x} = -\widetilde{P}\tilde{c} = (\,0.0183 \quad 0.319 \quad -0.140 \quad 0.0623 \quad -0.0844\,)^\top$$

- Direction Δx

$$\Delta x = X^k\Delta\tilde{x} = -Dz = (\,0.0229 \quad 0.0318 \quad -0.0280 \quad 0.196 \quad -0.0548\,)^\top$$

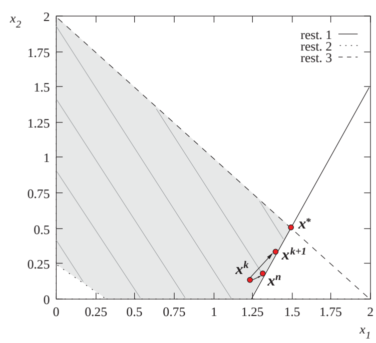
- New point x^{k+1}

$$x^{k+1} = (\,1.414 \quad 0.328 \quad 5\cdot 10^{-7} \quad 4.55 \quad 0.258\,)^\top$$

- Point x^n without scaling, and comparison of objective functions:

$$x^n = (\,1.32 \quad 0.153 \quad 5\cdot 10^{-7} \quad 3.59 \quad 0.520\,)^\top$$

$$c^\top x^n = -4.2870 > c^\top x^{k+1} = -4.8984$$



STEP LENGTH α

- Goal: move from x^k through Δx as much as possible

$$x^{k+1} \geq 0 \quad \Longrightarrow \quad x^k + \alpha \Delta x \geq 0$$

- Since $x^k > 0$ and $\alpha \geq 0$, we just focus on components $\Delta x_i < 0$.
- Maximum step length $\overline{\alpha}$:

$$\overline{\alpha} = \min \left\{ -\frac{x_i^k}{\Delta x_i} \text{ for all } i \text{ such that } \Delta x_i \leq 0 \right\}$$

- Definition α: move away 0 the null components by reducing $\overline{\alpha}$

$$\alpha = \rho \cdot \overline{\alpha} \qquad \rho \in [0.95, 0.9995]$$

Exercise 5. Show that if $\Delta x \geq 0$ (and $\Delta x \neq 0$) then the problem is unbounded.

EXAMPLE 5

In previous example we iterated from point

$$x^k = (\, 1.25 \quad 0.1 \quad 0.2 \quad 3.15 \quad 0.65 \,)^\top$$

obtaining the direction

$$\Delta x = (\, 0.0229 \quad 0.0318 \quad -0.0280 \quad 0.196 \quad -0.0548 \,)^\top .$$

Using a value $\rho = 0.95$, we have

$$\begin{aligned} \alpha &= \rho \cdot \overline{\alpha} \\ &= \rho \cdot \min \left\{ -\frac{x_3^k}{\Delta x_3}, -\frac{x_5^k}{\Delta x_5} \right\} \\ &= 0.95 \cdot \min \left\{ -\frac{0.2}{-0.0280}, -\frac{0.65}{-0.0548} \right\} \\ &= 0.95 \cdot \min \{7.143, 11.861\} \\ &= 0.95 \cdot 7.143 = 6.7885. \end{aligned}$$

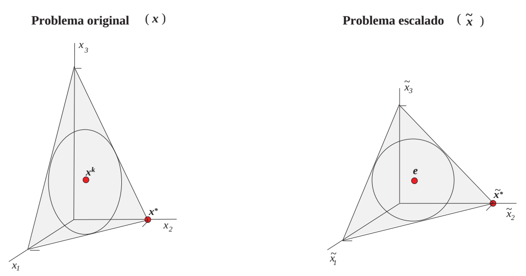
The new point is

$$x^{k+1} = x^k + \alpha \Delta x = \begin{pmatrix} 1.25 \\ 0.1 \\ 0.2 \\ 3.15 \\ 0.65 \end{pmatrix} + 6.7885 \begin{pmatrix} 0.0229 \\ 0.0318 \\ -0.0280 \\ 0.196 \\ -0.0548 \end{pmatrix} = \begin{pmatrix} 1.406 \\ 0.316 \\ 0.0101 \\ 4.483 \\ 0.278 \end{pmatrix}.$$

SHORT-STEP AFFINE-SCALING

- Consider the same affine scaling $\tilde{x} = (X^k)^{-1} x$ $(x^k \to e)$.
- Consider feasible sphere and ellipsoid:

$$\begin{array}{ll} (\tilde{x}) & \widetilde{B}(\tilde{x}, \rho \leq 1) = \{\tilde{x} \mid ||\tilde{x} - e|| \leq \rho, \widetilde{A}\tilde{x} = b\} \\ (x) & B(x, \rho \leq 1) = \{x \mid ||(X^k)^{-1}x - e|| \leq \rho, Ax = b\} \end{array}$$



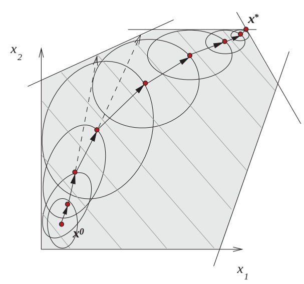
- Obtain the new point (on the ellipsoid) by solving the problem:

$$\begin{array}{llcll} & (x) & & & (\tilde{x}) \\ \min\limits_{\Delta x} & c^\top (x^k + \Delta x) & & \min\limits_{\Delta \tilde{x}} & \tilde{c}^\top (\tilde{x}^k + \Delta \tilde{x}) \\ \text{s. to} & A(x^k + \Delta x) = b & \Longleftrightarrow & \text{s. to} & \widetilde{A}(\tilde{x}^k + \Delta \tilde{x}) = b \\ & ||(X^k)^{-1} \Delta x||^2 = \rho^2 & & & ||\Delta \tilde{x}||^2 = \rho^2 \end{array}$$

- The computed direction is

$$\Delta x = -\frac{\rho}{||X^k z||} D z.$$

- No need to compute α (it is directly $\alpha = -\rho / ||X^k z||$).
- Sequence of points:



STOPPING CRITERIA

- When x^k is close enough to x^*?
- Reminder: dual problem

$$\begin{aligned} \max_{y,z} \quad & b^\top y \\ \text{s. to} \quad & A^\top y + z = c \\ & z \geq 0 \end{aligned} \qquad (D)$$

- Stopping criteria: small *duality gap*.

$$\frac{|gap\ dual^k|}{1+|c^\top x^k|} = \frac{|c^\top x^k - b^\top y|}{1+|c^\top x^k|} \leq \varepsilon \qquad \varepsilon \in [10^{-6}, 10^{-8}]$$

- Requirement: we need an estimation of y.

ESTIMATION OF y

- By Complementary Slackness theorem, in the optimum of (P) and (D), $x_i z_i = 0,\ i = 1, \ldots, n$.
- Estimation of y and z by solving

$$\begin{aligned} \min_{z,y} \quad & \frac{1}{2}||X^k z||^2 \\ \text{s. to} \quad & z = c - A^\top y \end{aligned}$$

- Solution:

$$\begin{aligned} y &= (A(X^k)^2 A^\top)^{-1} A(X^k)^2 c \\ z &= c - A^\top y \end{aligned}$$

- Reminder: direction of movement Δx

$$\Delta x = X^k \Delta \tilde{x} = -Dz \quad \text{where} \quad \begin{aligned} D &= (X^k)^2 \\ z &= c - A^\top y \\ y &= (ADA^\top)^{-1} ADc \end{aligned}$$

- Conclusion: estimations of y and z already computed for Δx. There is no extra computational effort.

Exercise 6. Show that the estimation of the dual variables coincide with those of the simplex algorithm as we approach an extreme point of the feasible set.

THE BIG-M METHOD

- Consider any point $x^0 > 0$ (e.g., $x^0 = (1, \ldots, 1)^\top$). In general, it will not be feasible (e.g., $Ax^0 \neq b$).
- Compute the vector of infeasibilities:

$$r = b - Ax^0$$

- Using some $M \gg 0$, $M \in \mathbb{R}$, we define

$$\begin{aligned} \min_{x} \quad & [c^\top\ M] \begin{bmatrix} x \\ x_{n+1} \end{bmatrix} \\ \text{s. to} \quad & [A\ r] \begin{bmatrix} x \\ x_{n+1} \end{bmatrix} = b \\ & \begin{bmatrix} x \\ x_{n+1} \end{bmatrix} \geq 0 \end{aligned} \qquad (P_M)$$

- An immediate feasible point for (P_M) is

$$(\, x_1^0 \quad x_2^0 \quad \ldots \quad x_n^0 \quad 1\,)^\top$$

$$[A\ r] \begin{bmatrix} x^0 \\ 1 \end{bmatrix} = Ax^0 + r = Ax^0 + (b - Ax^0) = b$$

- Since $M \gg 0$ it is expected that in the optimal solution of (P_M) $x_{n+1}^* = 0$, obtaining the solution of (P). If $x_{n+1}^* \neq 0$, (P) is infeasible.

INCONVENIENCES OF BIG-M METHOD

- Choosing the proper M: neither too big (otherwise, numerical stability problems) nor too small (otherwise we may wrongly conclude (P) is infeasible).

- The infeasibilities vector $r = b - Ax^0$ is usually dense. Matrix of (P_M) is $A_M = [A\ r]$, and then $A_M A_M^\top = [A\ r] \begin{bmatrix} A^\top \\ r^\top \end{bmatrix}$ is dense $\Longrightarrow$ the factorization of $A_M A_M^\top$ at each iteration is computationally expensive.

PRIMAL AFFINE-SCALING ALGORITHM

INITIALIZATIONS

Compute infeasibilities: $r = b - Ae_n$, where $e_l = (1_1, \ldots,\ 1_l)^\top$

Extend A with additional column r: $A \leftarrow [A\ r]$, $n \leftarrow n+1$

Extend cost vector: $c \leftarrow [c^\top\ M]^\top$, $M \in \mathbb{R}$ big

$x^0 = e_{n+1}$; x^0 is interior ($x^0 > 0$) and feasible ($Ax^0 = b$)

$D = I$, $y = (ADA^\top)^{-1}ADc$

$k = 0$, $\varepsilon = 10^{-6}$, $\rho \in [0.95, 0.9995]$

ITERATIVE PROCEDURE

while $\dfrac{|c^\top x^k - b^\top y|}{1 + |c^\top x^k|}\varepsilon$ **do**

  Compute z: $z = c - A^\top y$

  Compute Δx: $\Delta x = -Dz$

  if ($\Delta x \geq 0$) **then** STOP: Unbounded problem

  Compute α: $\alpha = \rho \cdot \min\left\{-\dfrac{x_i^k}{\Delta x_i}\ \forall i\ \Delta x_i \leq 0\right\}$

  Compute x^{k+1}: $x^{k+1} = x^k + \alpha\Delta x$

  $k \leftarrow k+1$

  Compute X^k: $X^k = \mathrm{diag}(x_1^k, \ldots, x_n^k)$

  Compute D: $D = (X^k)^2$

  Compute y: $(ADA^\top)y = ADc$

end_while

END

if $x_{n+1}^k \neq 0$ **then**

  STOP: Infeasible problem

else

  STOP: Optimal solution found: $x^* \leftarrow (x_1^k, \ldots, x_n^k)^\top$

end_if

EXAMPLE 6

Given the problem

$$\begin{array}{ll} \min & -3x_1 - 2x_2 \\ \text{s. to} & \\ & 4x_1 - 2x_2 + x_3 = 5 \\ & 3x_1 + 4x_2 - x_4 = 1 \\ & x_1 + x_2 + x_5 = 2 \\ & x_i \geq 0 \text{ for all } i = 1, \ldots, 5 \end{array}$$

and $x^0 = (1/2\ \ 1/2\ \ 4\ \ 5/2\ \ 1)^\top$.

- Objective function: $c^\top x^0 = -2.5$
- Compute y: $(ADA^\top)y = ADc$, where $D = (X^k)^2$, $X^k = \mathrm{diag}(0.5, 0.5, 4, 2.5, 1)$

$$\begin{pmatrix} 21 & 1 & 0.5 \\ 1 & 12.5 & 1.75 \\ 0.5 & 1.75 & 1.5 \end{pmatrix} y = \begin{pmatrix} -2 \\ -4.25 \\ -1.25 \end{pmatrix} \implies y = \begin{pmatrix} -0.070718 \\ -0.264115 \\ -0.501627 \end{pmatrix}$$

- Check optimality condition:

$$\text{gap dual} = \frac{|c^\top x^0 - b^\top y|}{1 + |c^\top x^0|} = \frac{|-2.5 - (-1.620)|}{3.5} = 0.25114$$

- Compute $z = c - A^\top y$ y $\Delta x = -Dz$

$$z = c - A^\top y = \begin{pmatrix} -1.423158 \\ -0.583349 \\ 0.070718 \\ -0.264115 \\ 0.501627 \end{pmatrix} \qquad \Delta x = -Dz = \begin{pmatrix} 0.35579 \\ 0.14584 \\ -1.13148 \\ 1.65072 \\ -0.50163 \end{pmatrix}$$

- Compute α

$$\begin{aligned} \alpha &= \rho \cdot \min\{-x_i^0/\Delta x_i\ \forall i \text{ tal que } \Delta x_i < 0\} \\ &= 0.95 \cdot \min\{4/1.13148, 1.0/0.50163\} = 0.95 \cdot 1.9935 \end{aligned}$$

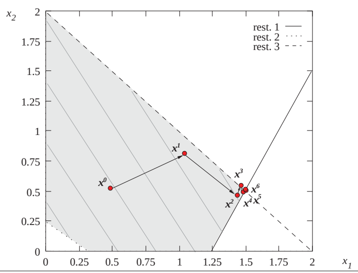
(EXAMPLE (cont.) 6)

- New point $x^1 = x^0 + \alpha\Delta x$

$$x^1 = \begin{pmatrix} 0.5 \\ 0.5 \\ 4 \\ 2.5 \\ 1 \end{pmatrix} + 0.95 \cdot 1.9935 \begin{pmatrix} 0.35579 \\ 0.14584 \\ -1.13148 \\ 1.65072 \\ -0.50163 \end{pmatrix} = \begin{pmatrix} 1.173803 \\ 0.776190 \\ 1.857169 \\ 5.626170 \\ 0.050007 \end{pmatrix}$$

- Objective function $c^\top x_1 = -5.0738 < c^\top x^0 = -2.5$
- Iterations:

k	x_1^k	x_2^k	x_3^k	x_4^k	x_5^k	gap dual
0	0.5	0.5	4	2.5	1	0.25114
1	1.173808	0.776192	1.857154	5.626192	0.050000	0.06753
2	1.475381	0.497191	0.092858	5.414905	0.027429	0.01181
3	1.487639	0.510988	0.071418	5.506874	0.001371	0.00235
4	1.499064	0.499914	0.003570	5.496851	0.001020	0.00045
5	1.4995	0.50042	0.002723	5.500300	0.000051	$8.82 \cdot 10^{-5}$
6	1.5000	0.50000	0.000103	5.500000	0.000038	$1.72 \cdot 10^{-5}$



Exercise 7. Implement the primal affine-scaling algorithm using some high-level language for matrix computations, such as Matlab or Octave.

Exercise 8. Solve the linear problem

$$\begin{array}{ll} \min & 3x_1 + x_2 \\ \text{s. to} & \\ & 2x_1 + x_2 \geq 2 \\ & 3x_1 + 4x_2 \leq 12 \\ & x_1 \geq 0 \quad x_2 \geq 0. \end{array}$$

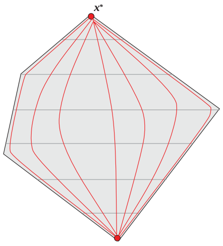
using the primal affine-scaling algorithm.

CONVERGENCE OF ALGORITHM

- If $(x^k, y^k, z^k)_{k\to\infty} \longrightarrow (x^*, y^*, z^*)$ then x^* is an optimizer of (P) and (y^*, z^*) is an optimizer of (D).
- If (P) and (D) are nondegenerate and no-unbounded, then for all $\rho < 1$ the sequences (x^k, y^k) are convergent.
- If $\rho < 2/3$ then the sequences (x^k, y^k) are convergent, independently of the nondegeneracy of (P) and (D).

COMPLEXITY OF THE ALGORITHM

- It is believed that it is not a polynomial-time algorithm

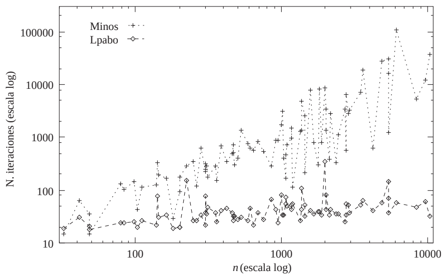


COMPUTATIONAL PERFORMANCE

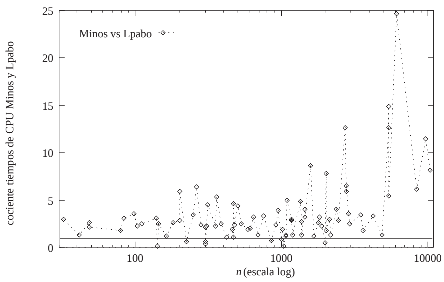
- We compared Lpabo and MINOS
- We used 90 problems of Netlib collection (from 25 to 2400 constraints, and from 32 to 10500 variables).
- Summary of results with Lpabo y MINOS:

	Lpabo		MINOS	
solved problems	86		89	
	N. Iter	CPU	N. Iter	CPU
mean	46.3	32.8	4126.3	86.3
st. dev.	41.1	161.9	13009.0	279.5
max.	350	1422.7	106649	1876.3
min.	18	0.04	15	0.12

- Number of iterations vs number of variables.



- (CPU MINOS)/(CPU Lpabo) vs number of variables.



REFERENCES

Basic

- A. Arbel, *Exploring Interior-Point Linear Programming: Algorithms and Software*, The MIT Press, Cambridge MA, 1993.
- D. Bertsimas y J.N. Tsitsiklis, *Introduction to Linear Optimization*, Athena Scientific, Belmont MA, 1997.
- R.J. Vanderbei, *Linear Programming: Foundations and Extensions*, Kluwer Academic Publishers, Boston MA, 1996.

Complementary

- E.D. Andersen, J. Gondzio, C. Mészáros y X. Xu, "Implementation of interior point methods for large scale linear programming", *Interior Point Methods of Mathematical Programming*, ed. T. Terlaky, Kluwer Academic Publishers, The Netherlands, 1996.
- T. Tsuchiya, "Affine Scaling Algorithm" en *Interior Point Methods of Mathematical Programming*, ed. T. Terlaky, Kluwer Academic Publishers, The Netherlands, 1996.

CONCLUSIONS

- Affine-scaling algorithm is an alternative to simplex method.
- It belongs to the family of interior-point methods.
- It considers the negative projected gradient direction using a scaling of the variables.
- It can be easily implemented, and in practice it provides decent results.
- It is believed not to be a polynomial-time algorithm.
- There are other more complex, yet more efficient, interior-point methods.