

15.081J/6.251J Introduction to Mathematical
Programming

Lecture 20: The Affine Scaling Algorithm

1 Outline

SLIDE 1

- History
- Geometric intuition
- Algebraic development
- Affine Scaling
- Convergence
- Initialization
- Practical performance

2 History

SLIDE 2

- In 1984, Karmakar at AT&T “invented” interior point method
- In 1985, Affine scaling “invented” at IBM + AT&T seeking intuitive version of Karmarkar’s algorithm
- In early computational tests, A.S. far outperformed simplex and Karmarkar’s algorithm
- In 1989, it was realised Dikin invented A.S. in 1967

3 Geometric intuition

3.1 Notation

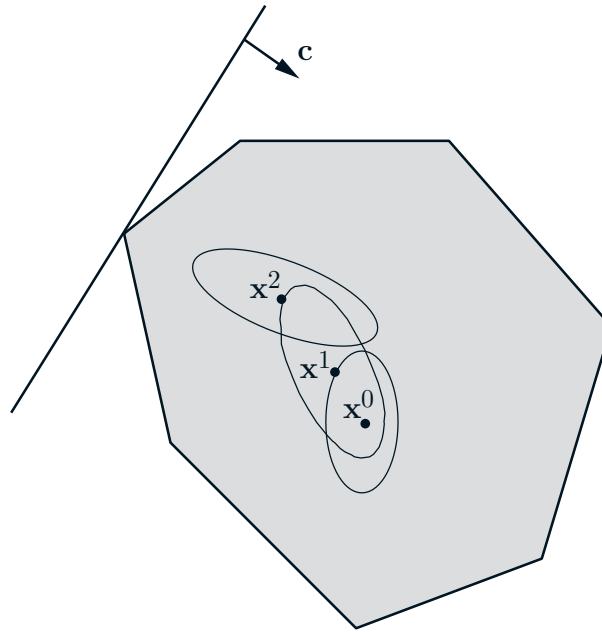
SLIDE 3

$$\begin{array}{ll} \min & \mathbf{c}'\mathbf{x} \\ \text{s.t.} & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{array}$$

and its dual

$$\begin{array}{ll} \max & \mathbf{p}'\mathbf{b} \\ \text{s.t.} & \mathbf{p}'\mathbf{A} \leq \mathbf{c}' \end{array}$$

- $P = \{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$
- $\{\mathbf{x} \in P \mid \mathbf{x} > \mathbf{0}\}$ the *interior* of P and its elements *interior points*



3.2 The idea

SLIDE 4

4 Algebraic development

4.1 Theorem

SLIDE 5

$\beta \in (0, 1)$, $\mathbf{y} \in \mathbb{R}^n$: $\mathbf{y} > \mathbf{0}$, and

$$S = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i=1}^n \frac{(x_i - y_i)^2}{y_i^2} \leq \beta^2 \right\}.$$

Then, $\mathbf{x} > \mathbf{0}$ for every $\mathbf{x} \in S$

Proof

- $\mathbf{x} \in S$
- $(x_i - y_i)^2 \leq \beta^2 y_i^2 < y_i^2$
- $|x_i - y_i| < y_i$; $-x_i + y_i < y_i$, and hence $x_i > 0$

SLIDE 6

$$\mathbf{x} \in S \text{ is equivalent to } \|\mathbf{Y}^{-1}(\mathbf{x} - \mathbf{y})\| \leq \beta$$

Replace original LP:

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \|\mathbf{Y}^{-1}(\mathbf{x} - \mathbf{y})\| \leq \beta. \end{aligned}$$

$$\mathbf{d} = \mathbf{x} - \mathbf{y}$$

$$\begin{aligned} \min \quad & \mathbf{c}'\mathbf{d} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{d} = \mathbf{0} \\ & \|\mathbf{Y}^{-1}\mathbf{d}\| \leq \beta \end{aligned}$$

4.2 Solution

SLIDE 7

If rows of \mathbf{A} are linearly independent and \mathbf{c} is not a linear combination of the rows of \mathbf{A} , then

- optimal solution \mathbf{d}^* :

$$\mathbf{d}^* = -\beta \frac{\mathbf{Y}^2(\mathbf{c} - \mathbf{A}'\mathbf{p})}{\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|}, \quad \mathbf{p} = (\mathbf{A}\mathbf{Y}^2\mathbf{A}')^{-1}\mathbf{A}\mathbf{Y}^2\mathbf{c}.$$

- $\mathbf{x} = \mathbf{y} + \mathbf{d}^* \in P$
- $\mathbf{c}'\mathbf{x} = \mathbf{c}'\mathbf{y} - \beta\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\| < \mathbf{c}'\mathbf{y}$

4.2.1 Proof

SLIDE 8

- $\mathbf{A}\mathbf{Y}^2\mathbf{A}'$ is invertible; if not, there exists some $\mathbf{z} \neq \mathbf{0}$ such that $\mathbf{z}'\mathbf{A}\mathbf{Y}^2\mathbf{A}'\mathbf{z} = 0$
- $\mathbf{w} = \mathbf{Y}\mathbf{A}'\mathbf{z}$; $\mathbf{w}'\mathbf{w} = 0 \Rightarrow \mathbf{w} = \mathbf{0}$
- Hence $\mathbf{A}'\mathbf{z} = \mathbf{0}$ contradiction
- Since \mathbf{c} is not a linear combination of the rows of \mathbf{A} , $\mathbf{c} - \mathbf{A}'\mathbf{p} \neq \mathbf{0}$ and \mathbf{d}^* is well defined
- \mathbf{d}^* feasible

$$\mathbf{Y}^{-1}\mathbf{d}^* = -\beta \frac{\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})}{\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|} \Rightarrow \|\mathbf{Y}^{-1}\mathbf{d}^*\| = \beta$$

$$\mathbf{A}\mathbf{d}^* = \mathbf{0}, \text{ since } \mathbf{A}\mathbf{Y}^2(\mathbf{c} - \mathbf{A}'\mathbf{p}) = \mathbf{0}$$

•

$$\begin{aligned} \mathbf{c}'\mathbf{d} &= (\mathbf{c}' - \mathbf{p}'\mathbf{A})\mathbf{d} \\ &= (\mathbf{c}' - \mathbf{p}'\mathbf{A})\mathbf{Y}\mathbf{Y}^{-1}\mathbf{d} \\ &\geq -\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\| \cdot \|\mathbf{Y}^{-1}\mathbf{d}\| \\ &\geq -\beta\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|. \end{aligned}$$

SLIDE 9

•

$$\begin{aligned} \mathbf{c}'\mathbf{d}^* &= (\mathbf{c}' - \mathbf{p}'\mathbf{A})\mathbf{d}^* \\ &= -(\mathbf{c}' - \mathbf{p}'\mathbf{A})\beta \frac{\mathbf{Y}^2(\mathbf{c} - \mathbf{A}'\mathbf{p})}{\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|} \\ &= -\beta \frac{(\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p}))'(\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p}))}{\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|} \\ &= -\beta\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|. \end{aligned}$$

- $\mathbf{c}'\mathbf{x} = \mathbf{c}'\mathbf{y} + \mathbf{c}'\mathbf{d}^* = \mathbf{c}'\mathbf{y} - \beta\|\mathbf{Y}(\mathbf{c} - \mathbf{A}'\mathbf{p})\|$

4.3 Interpretation

SLIDE 10

- \mathbf{y} be a nondegenerate BFS with basis \mathbf{B}
- $\mathbf{A} = [\mathbf{B} \ \mathbf{N}]$
- $\mathbf{Y} = \text{diag}(y_1, \dots, y_m, 0, \dots, 0)$ and $\mathbf{Y}_0 = \text{diag}(y_1, \dots, y_m)$, then $\mathbf{AY} = [\mathbf{BY}_0 \ \mathbf{0}]$

$$\begin{aligned}\mathbf{p} &= (\mathbf{AY}^2\mathbf{A}')^{-1}\mathbf{AY}^2\mathbf{c} \\ &= (\mathbf{B}')^{-1}\mathbf{Y}_0^{-2}\mathbf{B}^{-1}\mathbf{BY}_0^2\mathbf{c}_B \\ &= (\mathbf{B}')^{-1}\mathbf{c}_B\end{aligned}$$

- Vectors \mathbf{p} dual estimates
- $\mathbf{r} = \mathbf{c} - \mathbf{A}'\mathbf{p}$ becomes reduced costs:

$$\mathbf{r} = \mathbf{c} - \mathbf{A}'(\mathbf{B}')^{-1}\mathbf{c}_B$$

- Under degeneracy?

4.4 Termination

SLIDE 11

\mathbf{y} and \mathbf{p} be primal and dual feasible solutions with

$$\mathbf{c}'\mathbf{y} - \mathbf{b}'\mathbf{p} < \epsilon$$

\mathbf{y}^* and \mathbf{p}^* be optimal primal and dual solutions. Then,

$$\begin{aligned}\mathbf{c}'\mathbf{y}^* &\leq \mathbf{c}'\mathbf{y} < \mathbf{c}'\mathbf{y}^* + \epsilon, \\ \mathbf{b}'\mathbf{p}^* - \epsilon &< \mathbf{b}'\mathbf{p} \leq \mathbf{b}'\mathbf{p}^*\end{aligned}$$

4.4.1 Proof

SLIDE 12

- $\mathbf{c}'\mathbf{y}^* \leq \mathbf{c}'\mathbf{y}$
- By weak duality, $\mathbf{b}'\mathbf{p} \leq \mathbf{c}'\mathbf{y}^*$
- Since $\mathbf{c}'\mathbf{y} - \mathbf{b}'\mathbf{p} < \epsilon$,

$$\begin{aligned}\mathbf{c}'\mathbf{y} &< \mathbf{b}'\mathbf{p} + \epsilon \leq \mathbf{c}'\mathbf{y}^* + \epsilon \\ \mathbf{b}'\mathbf{p}^* &= \mathbf{c}'\mathbf{y}^* \leq \mathbf{c}'\mathbf{y} < \mathbf{b}'\mathbf{p} + \epsilon\end{aligned}$$

5 Affine Scaling

5.1 Inputs

SLIDE 13

- $(\mathbf{A}, \mathbf{b}, \mathbf{c})$;
- an initial primal feasible solution $\mathbf{x}^0 > \mathbf{0}$
- the optimality tolerance $\epsilon > 0$
- the parameter $\beta \in (0, 1)$

5.2 The Algorithm

SLIDE 14

1. (Initialization) Start with some feasible $\mathbf{x}^0 > \mathbf{0}$; let $k = 0$.
2. (Computation of dual estimates and reduced costs) Given some feasible $\mathbf{x}^k > \mathbf{0}$, let

$$\begin{aligned}\mathbf{X}_k &= \text{diag}(x_1^k, \dots, x_n^k), \\ \mathbf{p}^k &= (\mathbf{A}\mathbf{X}_k^2\mathbf{A}')^{-1}\mathbf{A}\mathbf{X}_k^2\mathbf{c}, \\ \mathbf{r}^k &= \mathbf{c} - \mathbf{A}'\mathbf{p}^k.\end{aligned}$$

3. (Optimality check) Let $\mathbf{e} = (1, 1, \dots, 1)$. If $\mathbf{r}^k \geq \mathbf{0}$ and $\mathbf{e}'\mathbf{X}_k\mathbf{r}^k < \epsilon$, then stop; the current solution \mathbf{x}^k is primal ϵ -optimal and \mathbf{p}^k is dual ϵ -optimal.
4. (Unboundedness check) If $-\mathbf{X}_k^2\mathbf{r}^k \geq \mathbf{0}$ then stop; the optimal cost is $-\infty$.
5. (Update of primal solution) Let

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta \frac{\mathbf{X}_k^2\mathbf{r}^k}{\|\mathbf{X}_k\mathbf{r}^k\|}.$$

5.3 Variants

SLIDE 15

- $\|\mathbf{u}\|_\infty = \max_i |u_i|$, $\gamma(\mathbf{u}) = \max\{u_i \mid u_i > 0\}$
- $\gamma(\mathbf{u}) \leq \|\mathbf{u}\|_\infty \leq \|\mathbf{u}\|$
- *Short-step* method.
- *Long-step* variants

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta \frac{\mathbf{X}_k^2\mathbf{r}^k}{\|\mathbf{X}_k\mathbf{r}^k\|_\infty}$$

$$\mathbf{x}^{k+1} = \mathbf{x}^k - \beta \frac{\mathbf{X}_k^2\mathbf{r}^k}{\gamma(\mathbf{X}_k\mathbf{r}^k)}$$

6 Convergence

6.1 Assumptions

SLIDE 16

Assumptions A:

- (a) The rows of the matrix \mathbf{A} are linearly independent.
- (b) The vector \mathbf{c} is not a linear combination of the rows of \mathbf{A} .
- (c) There exists an optimal solution.
- (d) There exists a positive feasible solution.

Assumptions B:

- (a) Every BFS to the primal problem is nondegenerate.
- (b) At every BFS to the primal problem, the reduced cost of every nonbasic variable is nonzero.

6.2 Theorem

SLIDE 17

If we apply the long-step affine scaling algorithm with $\epsilon = 0$, the following hold:

- (a) For the Long-step variant and under Assumptions A and B, and if $0 < \beta < 1$, \mathbf{x}^k and \mathbf{p}^k converge to the optimal primal and dual solutions
- (b) For the second Long-step variant, and under Assumption A and if $0 < \beta < 2/3$, the sequences \mathbf{x}^k and \mathbf{p}^k converge to some primal and dual optimal solutions, respectively

7 Initialization

SLIDE 18

$$\begin{array}{ll} \min & \mathbf{c}'\mathbf{x} + Mx_{n+1} \\ \text{s.t.} & \mathbf{A}\mathbf{x} + (\mathbf{b} - \mathbf{A}\mathbf{e})x_{n+1} = \mathbf{b} \\ & (\mathbf{x}, x_{n+1}) \geq \mathbf{0} \end{array}$$

8 Example

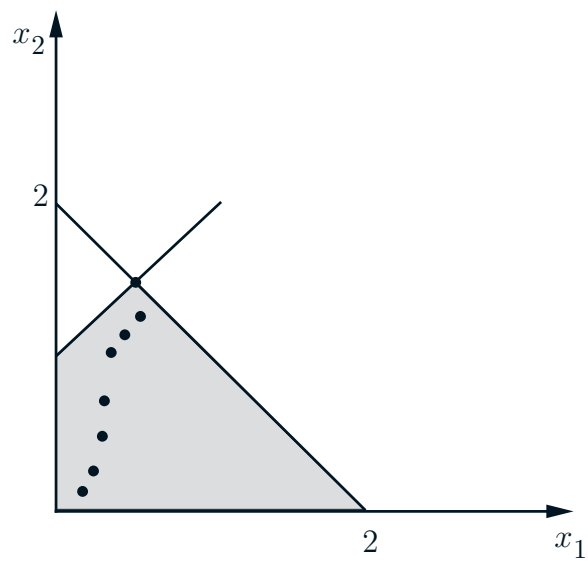
SLIDE 19

$$\begin{array}{ll} \max & x_1 + 2x_2 \\ \text{s.t.} & x_1 + x_2 \leq 2 \\ & -x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0 \end{array}$$

9 Practical Performance

SLIDE 20

- Excellent practical performance, simple
- Major step: invert $\mathbf{A}\mathbf{X}_k^2\mathbf{A}'$
- Imitates the simplex method near the boundary



MIT OpenCourseWare
<http://ocw.mit.edu>

6.251J / 15.081J Introduction to Mathematical Programming
Fall 2009

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.