

# Chapter 3: Interior point methods

## 3.1. The affine scaling algorithm

Consider a pair of LPs :  $A \in \mathbb{R}^{m \times n}$ ,  $c \in \mathbb{R}^n$ ,  $b \in \mathbb{R}^m$

$$\min c^T x$$

$$\text{s.t. } \boxed{\begin{matrix} Ax = b \\ x \geq 0 \end{matrix}}$$

$$\max p^T b$$

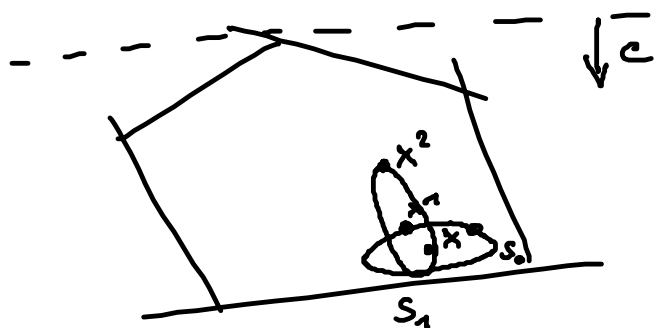
$$\text{s.t. } p^T A \leq c^T$$

Let  $P = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}$  and

$\{x \in P \mid x > 0\}$  the interior of P. A point  $x \in P$  with  $x > 0$  is called interior point of P.

Main idea:

- start with interior point  $x^0$
- Form ellipsoid  $S_0$  centered at  $x^0$  and contained in the interior of  $P$ .
- Optimize  $c^T x$  over all  $x \in S_0$  and find opt. sol.  $x^1$
- Form ellipsoid  $S_1$  centered at  $x^1$ .....



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

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

Then  $x > 0$  for all  $x \in S$ .

Proof:  $x \in S$ ,  $i \in \{1, \dots, n\}$ :

$$(x_i - y_i)^2 \leq \beta^2 \cdot y_i^2 \Rightarrow |x_i - y_i| < y_i$$

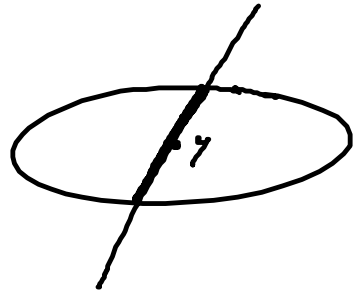
$$\Rightarrow -x_i + y_i < y_i \Rightarrow x_i > 0. \quad \square$$

Fix some  $y \in \mathbb{R}^n$  with  $y > 0$  and  $Ay = b$ .

$$Y := \text{diag}(y_1, \dots, y_n)$$

Then:

$$x \in S \iff \|\underbrace{Y^{-1} \cdot (x-y)}_{\text{Euclidean norm}}\| \leq \beta$$



$$\iff (x-y)^T \cdot (Y^{-1} \cdot Y^{-1}) \cdot (x-y) \leq \beta^2$$

In particular,  $S$  is an ellipsoid centered at  $y$ .

Let  $S_0 := \{x \in S \mid Ax = b\}$  "section" of ellipsoid  $S$ .

$S_0$  is itself an ellipsoid contained in the interior of  $P$ .

Optimize  $c^T x$  over all  $x \in S_0$

$$\begin{aligned} \min c^T x \\ \text{s.t. } Ax = b \end{aligned}$$

$$\|Y^{-1} \cdot (x-y)\| \leq \beta$$

Reformulate this by setting  $d := x - y$ :

$$\begin{aligned} \min c^T d \\ \text{s.t. } Ad = 0 \end{aligned}$$

$$\|Y^{-1} \cdot d\| \leq \beta$$

Lemma: Assume that rows of  $A$  are linearly independent and  $c$  is not a linear combination of rows of  $A$ . An optimal solution  $d^*$  is given by

$$d^* := -\beta \cdot \frac{Y^2 \cdot (c - A^T \cdot p)}{\|Y \cdot (c - A^T \cdot p)\|}$$

$$\text{where } p := (A \cdot Y^2 \cdot A^T)^{-1} \cdot A \cdot Y^2 \cdot c.$$

Moreover,  $x := y + d^* \in P$  and

$$c^T x = c^T y - \beta \cdot \|Y \cdot (c - A^T p)\| < c^T y.$$

Proof: see book.  $\square$

### Remarks:

(i) If  $d^* \geq 0$ , the LP is unbounded since  $A \cdot d^* = 0$  and  $y + \alpha \cdot d^* > 0$  for  $\alpha > 0$  and  $c^T \cdot d^* < 0$ .

(ii) Assume that  $y$  is a nondegenerate basic feasible sol. (contradiction to  $y > 0$ ). Let  $B$  be the corresponding basic matrix and  $A = (B, N)$  (w.l.o.g.)

Let  $Y = \text{diag}(y_1, \dots, y_m, 0, \dots, 0)$  and

$$Y_0 = \text{diag}(y_1, \dots, y_m)$$

Then  $A \cdot Y = (B, N) \cdot \begin{pmatrix} Y_0 & 0 \\ 0 & 0 \end{pmatrix} = (B \cdot Y_0, 0)$  and

$$p = (A Y^2 \cdot A^T)^{-1} \cdot A \cdot Y^2 \cdot c$$

$$= (B^T)^{-1} \cdot Y_0^{-2} \cdot B^{-1} \cdot B \cdot Y_0^2 \cdot c_B = (B^T)^{-1} \cdot c_B$$

is the corresponding dual basic solution.

The vector  $p$  corresponding to an arbitrary primal solution  $y$  is called dual estimate, even if  $y$  is not basic.

The vector  $r := c - A^T p$  becomes

$$r = c - A^T \cdot (B^T)^{-1} \cdot c_B$$

the reduced cost vector.

If  $y$  is degenerate, the matrix  $A \cdot Y^2 \cdot A^T$  is singular and this interpretation breaks down.

(iii) If  $r = c - A^T \cdot p \geq 0$ , then  $p$  is a dual feasible solution and  $r^T \cdot y = (c - A^T \cdot p)^T \cdot y = c^T y - p^T \cdot A \cdot y = c^T y - p^T \cdot b$  is the difference between the primal and dual objective function value (called the duality gap).

If  $r^k = 0$  then complementary slackness cond. hold and  $y$  and  $p$  are both optimal.

### The affine scaling algorithm

1) Start with a feasible  $x^0 > 0$ ; set  $k := 0$

2) Let  $X_k = \text{diag}(x_1^k, \dots, x_n^k)$

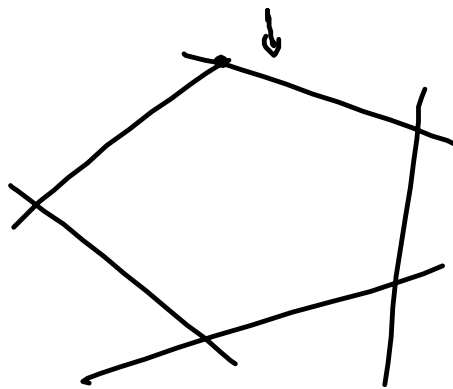
$$p^k = (A \cdot X_k^2 \cdot A^T)^{-1} \cdot A \cdot X_k^2 \cdot c$$

$$r^k = c - A^T \cdot p$$

3) If  $r^k \geq 0$  and  $(r^k)^T \cdot x^k < \epsilon$ , then stop;  $x^k$  is primal  $\epsilon$ -opt. and  $p^k$  is dual  $\epsilon$ -opt.

4) If  $-X_k^2 \cdot r^k \geq 0$ , then stop; the optimal cost is  $-\infty$ .

5) Let  $x^{k+1} = x^k - \rho \cdot \frac{X_k^2 \cdot r^k}{\|X_k \cdot r^k\|}$ ;  $k := k+1$ ; goto 2)



### Convergence

#### Assumptions:

(i) rows of  $A$  are linearly indep.

(ii)  $c$  is not a linear combin. of rows of  $A$ .

(iii) There exists an opt. sol.

(iv) There exists a positive feasible sol.

(v) every basic feasible solution to the primal LP is non-deg.

(vi) At every basic feasible sol. to the primal LP, the reduced costs of non-basic variables are non-zero.

Theorem: Under these assumptions, for  $\epsilon = 0$ , the algorithm converges to a pair of primal and dual opt. solutions.

Initialization: Consider auxiliary problem:

$$\min c^T x + M \cdot x_{n+1}$$

$$\text{s.t. } A \cdot x + (b - A \cdot \mathbb{1}) \cdot x_{n+1} = b$$

$$(x, x_{n+1}) \geq 0$$

and notice that  $(\mathbb{1}, 1)$  is a positive feasible solution.

Computational performance:

The running time of one iteration of the algorithm is dominated by the computation of  $p$ . This takes  $O(m^2 \cdot n)$  steps.