ORF307 – Optimization

19. Linear optimization review

Bartolomeo Stellato — Spring 2022

Ed Forum

- How do mathematicians know when to use "heuristics," or arbitrary measures to make the algorithm work better?
- Why is not being able to warm start a problem if we can start interior point methods with an infeasible solution?

Today’s lecture

Linear optimization review

- Formulations
- Piecewise linear optimization
- Duality
- Sensitivity analysis
- Simplex method
- Interior point methods

Formulations

Linear optimization

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & Dx = f \end{array}$$

- Minimization
- Less-than ineq. constraints
- Equality constraints

x is **feasible** if it satisfies the constraints $Ax \leq b$ and $Dx = f$

The **feasible set** is the set of all feasible points

$x^\star$ is **optimal** if it is feasible and $c^T x^\star \leq c^T x$ for all feasible x

The **optimal value** is $p^\star = c^T x^\star$

Unbounded problem: $c^T x$ is unbounded below on the feasible set ($p^\star = -\infty$)

Infeasible problem: feasible set is empty ($p^\star = +\infty$)

Feasibility problems

$$
\begin{array}{ll}
\text{find} & x \\
\text{subject to} & Ax \leq b \\
& Dx = f
\end{array}
\quad \longrightarrow \quad
\begin{array}{ll}
\text{minimize} & 0 \\
\text{subject to} & Ax \leq b \\
& Dx = f
\end{array}
$$

Possible results

- $p^\star = 0$ if constraints are feasible (consistent).
  (Every feasible x is optimal)
- $p^\star = \infty$ otherwise

Standard form

Definition

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

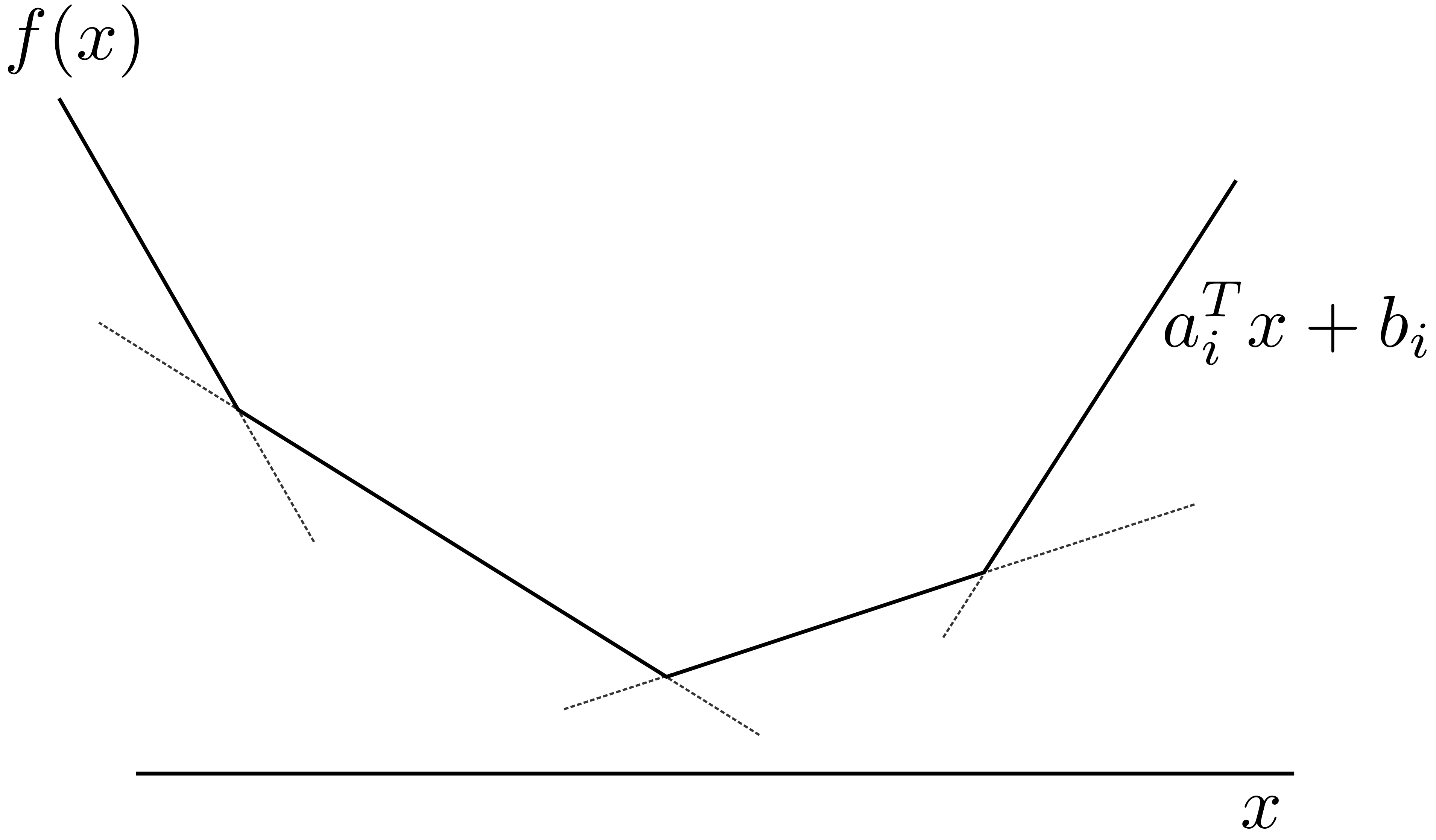
- Minimization
- Equality constraints
- Nonnegative variables

Useful to

- develop **algorithms**
- **algebraic** manipulations

Piecewise linear optimization

Piecewise-linear minimization



minimize $\quad f(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i)$

$$\downarrow$$

minimize $\quad t$

subject to $\quad a_i^T x + b_i \leq t, \quad i = 1, \ldots, m$

Matrix notation

minimize $\quad \tilde{c}^T \tilde{x}$

subject to $\quad \tilde{A}\tilde{x} \leq \tilde{b}$

$$\tilde{x} = \begin{bmatrix} x \\ t \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \tilde{A} = \begin{bmatrix} a_1^T & -1 \\ \vdots & \vdots \\ a_m^T & -1 \end{bmatrix}, \quad \tilde{b} = \begin{bmatrix} -b_1 \\ \vdots \\ -b_m \end{bmatrix}$$

1 and infinity norms reformulations

1-norm minimization:

minimize $\|Ax - b\|_1 = \sum_i |(Ax - b)_i|$

Equivalent to:

minimize $\mathbf{1}^T u$

subject to $-u \leq Ax - b \leq u$

Absolute value of every element $(Ax - b)_i$ is bounded by a component of the **vector** u

∞-norm minimization:

minimize $\|Ax - b\|_\infty = \max_i |(Ax - b_i)_i|$

Equivalent to:

minimize t

subject to $-t\mathbf{1} \leq Ax - b \leq t\mathbf{1}$

Absolute value of every element $(Ax - b)_i$ is bounded by the same **scalar** t

Duality

Inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

Relax the constraint

$$g(y) = \underset{x}{\text{minimize}} \quad \boxed{c^T x + y^T (Ax - b)}$$

Lagrangian

$L(x, y)$

Lower bound

$$g(y) \leq c^T x^\star + y^T (Ax^\star - b) \leq c^T x^\star$$

we must have $y \geq 0$

Dual of LP with inequalities

Derivation

Dual function

$$g(y) = \underset{x}{\text{minimize}} \left(c^T x + y^T (Ax - b)\right)$$

$$- b^T y + \underset{x}{\text{minimize}} \left(c + A^T y\right)^T x$$

$$g(y) = \begin{cases} -b^T y & \text{if } c + A^T y = 0 \quad \text{(and } y \geq 0\text{)} \\ -\infty & \text{otherwise} \end{cases}$$

Dual problem (find the best bound)

$$\underset{y}{\text{maximize}} \quad g(y) \quad = \quad \begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c = 0 \\ & y \geq 0 \end{array}$$

General forms

Inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

$$\begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c = 0 \\ & y \geq 0 \end{array}$$

Standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

$$\begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c \geq 0 \end{array}$$

LP with inequalities and equalities

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \\ & Dx = f \end{array}$$

$$\begin{array}{ll} \text{maximize} & -b^T y - f^T z \\ \text{subject to} & A^T y + D^T z + c = 0 \\ & y \geq 0 \end{array}$$

Weak duality

Theorem

If x, y satisfy:

- x is a feasible solution to the primal problem
- y is a feasible solution to the dual problem

$\longrightarrow$ $-b^T y \leq c^T x$

Proof

We know that $Ax \leq b$, $A^T y + c = 0$ and $y \geq 0$. Therefore,

$$0 \leq y^T(b - Ax) = b^T y - y^T Ax = c^T x + b^T y \qquad \blacksquare$$

Remark

- Any dual feasible y gives a **lower bound** on the primal optimal value
- Any primal feasible x gives an **upper bound** on the dual optimal value
- $c^T x + b^T y$ is the **duality gap**

Weak duality

Corollaries

Unboundedness vs feasibility

- Primal unbounded ($p^\star = -\infty$) $\Rightarrow$ dual infeasible ($d^\star = -\infty$)
- Dual unbounded ($d^\star = +\infty$) $\Rightarrow$ primal infeasible ($p^\star = +\infty$)

Optimality condition

If x, y satisfy:

- x is a feasible solution to the primal problem
- y is a feasible solution to the dual problem
- The duality gap is zero, *i.e.*, $c^T x + b^T y = 0$

Then x and y are **optimal solutions** to the primal and dual problem respectively

Strong duality

Primal

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \end{array}$$

Dual

$$\begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c \geq 0 \end{array}$$

Theorem

If a linear optimization problem has an optimal solution, then

- so does its dual
- the optimal values of the primal and dual are equal

Relationship between primal and dual

	$p^\star = +\infty$	$p^\star$ finite	$p^\star = -\infty$
$d^\star = +\infty$	primal inf. dual unb.		
$d^\star$ finite		optimal values equal	
$d^\star = -\infty$	exception		primal unb. dual inf

Complementary slackness

Primal

minimize $\quad c^T x$

subject to $\quad Ax \leq b$

Dual

maximize $\quad -b^T y$

subject to $\quad A^T y + c = 0$

$\qquad\qquad\qquad y \geq 0$

Theorem

Primal,dual feasible x, y are optimal if and only if

$$y_i(b_i - a_i^T x) = 0, \quad i = 1, \ldots, m$$

i.e., at optimum, $b - Ax$ and y have a **complementary sparsity** pattern:

$$\begin{aligned} y_i > 0 \quad &\Rightarrow \quad a_i^T x = b_i \\ a_i^T x < b_i \quad &\Rightarrow \quad y_i = 0 \end{aligned}$$

Complementary slackness

Primal

minimize $c^T x$

subject to $Ax \leq b$

Dual

maximize $-b^T y$

subject to $A^T y + c = 0$

$y \geq 0$

Proof

The duality gap at primal feasible x and dual feasible y can be written as

$$c^T x + b^T y = (-A^T y)^T x + b^T y = (b - Ax)^T y = \sum_{i=1}^{m} y_i (b_i - a_i^T x) = 0$$

Since all the elements of the sum are nonnegative, they must all be 0 ■

For **feasible** x and y **complementary slackness** = **zero duality gap**

Farkas lemma

Theorem

Given A and b, exactly one of the following statements is true:

1. There exists an x with $Ax = b$, $x \geq 0$
2. There exists a y with $A^T y \geq 0$, $b^T y < 0$

Farkas lemma

Proof

1 and 2 cannot be both true (easy)

$x \geq 0$, $Ax = b$ and $y^T A \geq 0$ $\longrightarrow$ $y^T b = y^T A x \geq 0$

---

1 and 2 cannot be both false (duality)

Primal		**Dual**	
minimize	0	maximize	$-b^T y$
subject to	$Ax = b$	subject to	$A^T y \geq 0$
	$x \geq 0$		

$\longleftarrow$ $y = 0$ always feasible

$d^\star \neq -\infty, \quad p^\star = d^\star$

Alternative 1: primal feasible $p^\star = d^\star = 0$

$b^T y \geq 0$ for all y such that $A^T y \geq 0$

Alternative 2: primal infeasible $p^\star = d^\star = +\infty$

There exists y such that $A^T y \geq 0$ and $b^T y < 0$

y is an **infeasibility certificate**

Sensitivity analysis

Changes in problem data

Goal: extract information from $x^\star, y^\star$ about their sensitivity with respect to changes in problem data

Modified LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b + u \\ & x \geq 0 \end{array}$$

Optimal value function

$$p^\star(u) = \min\{c^T x \mid Ax = b + u,\ x \geq 0\}$$

Assumption: $p^\star(0)$ is finite

Properties

- $p^\star(u) > -\infty$ everywhere (from global lower bound)
- $p^\star(u)$ is piecewise-linear on its domain

Global sensitivity

Dual of modified LP

$$\begin{array}{ll} \text{maximize} & -(b+u)^T y \\ \text{subject to} & A^T y + c \geq 0 \end{array}$$

Global lower bound

Given $y^\star$ a dual optimal solution for $u = 0$, then

$$\begin{aligned} p^\star(u) &\geq -(b+u)^T y^\star \\ &= p^\star(0) - u^T y^\star \end{aligned}$$

(from weak duality and dual feasibility)

It holds for any u

Local sensitivity

u in neighborhood of the origin

Original LP

minimize $\quad c^T x$

subject to $\quad Ax = b$

$\quad x \geq 0$

$\longrightarrow$

Optimal solution

Primal $\quad x_i = 0, \quad i \notin B$

$\quad x_B^\star = A_B^{-1} b$

Dual $\quad y^\star = -A_B^{-T} c_B$

Modified LP

minimize $\quad c^T x$

subject to $\quad Ax = b + u$

$\quad x \geq 0$

Modified dual

maximize $\quad -(b+u)^T y$

subject to $\quad A^T y + c \geq 0$

Optimal basis does not change

Modified optimal solution

$$x_B^\star(u) = A_B^{-1}(b+u) = x_B^\star + A_B^{-1} u$$
$$y^\star(u) = y^\star$$

Derivative of the optimal value function

Modified optimal solution

$$x_B^\star(u) = A_B^{-1}(b+u) = x_B^\star + A_B^{-1}u$$
$$y^\star(u) = y^\star$$

Optimal value function

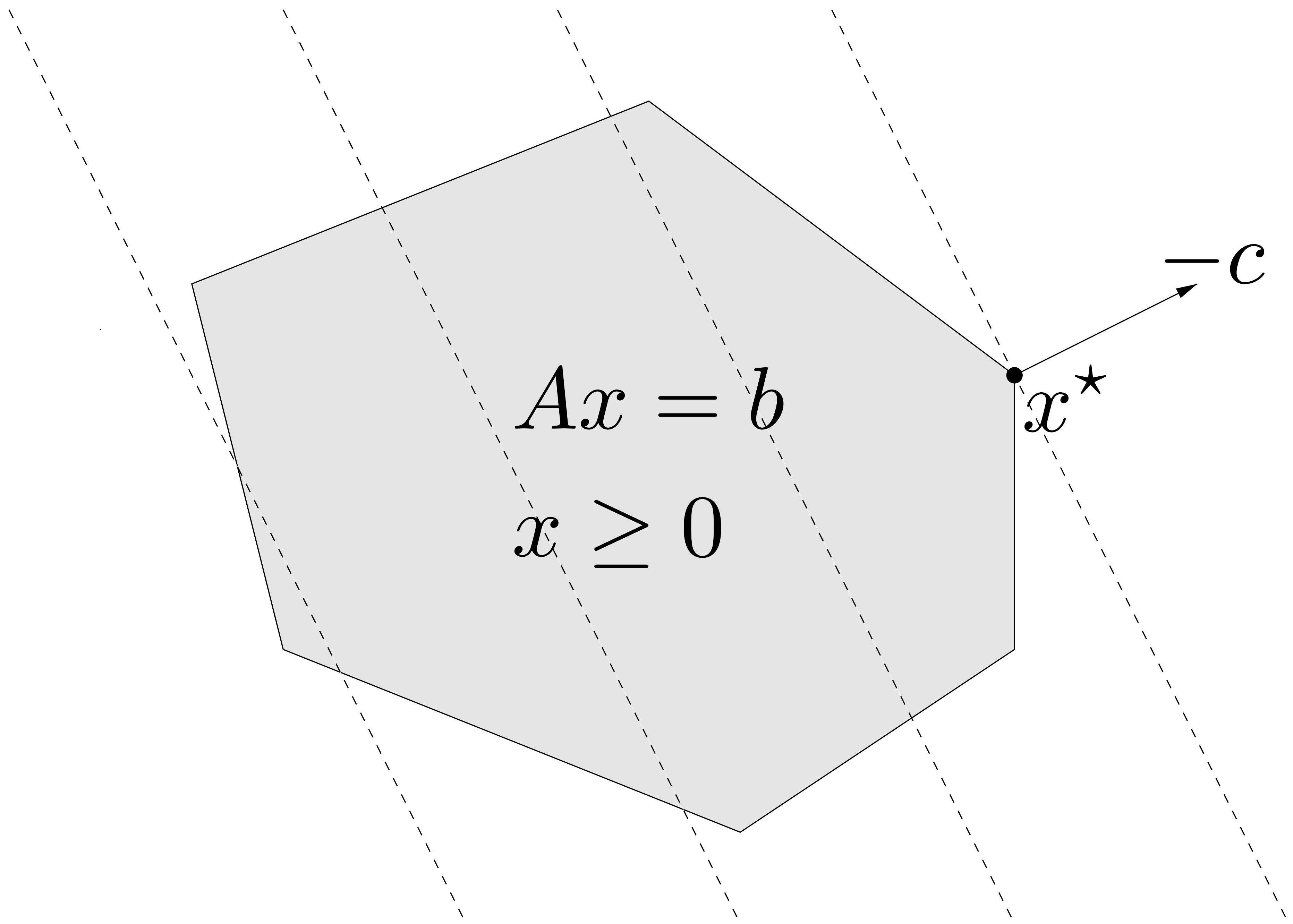
$$\begin{aligned} p^\star(u) &= c^T x^\star(u) \\ &= c^T x^\star + c_B^T A_B^{-1} u \\ &= p^\star(0) - y^{\star T} u \end{aligned}$$ (affine for small u)

Local derivative

$$\nabla p^\star(u) = -y^\star$$ ($y^\star$ are the **shadow prices**)

Simplex method

Optimality of extreme points



$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
& x \geq 0
\end{array}
$$

If
- P has at least one extreme point
- There exists an optimal solution $x^\star$

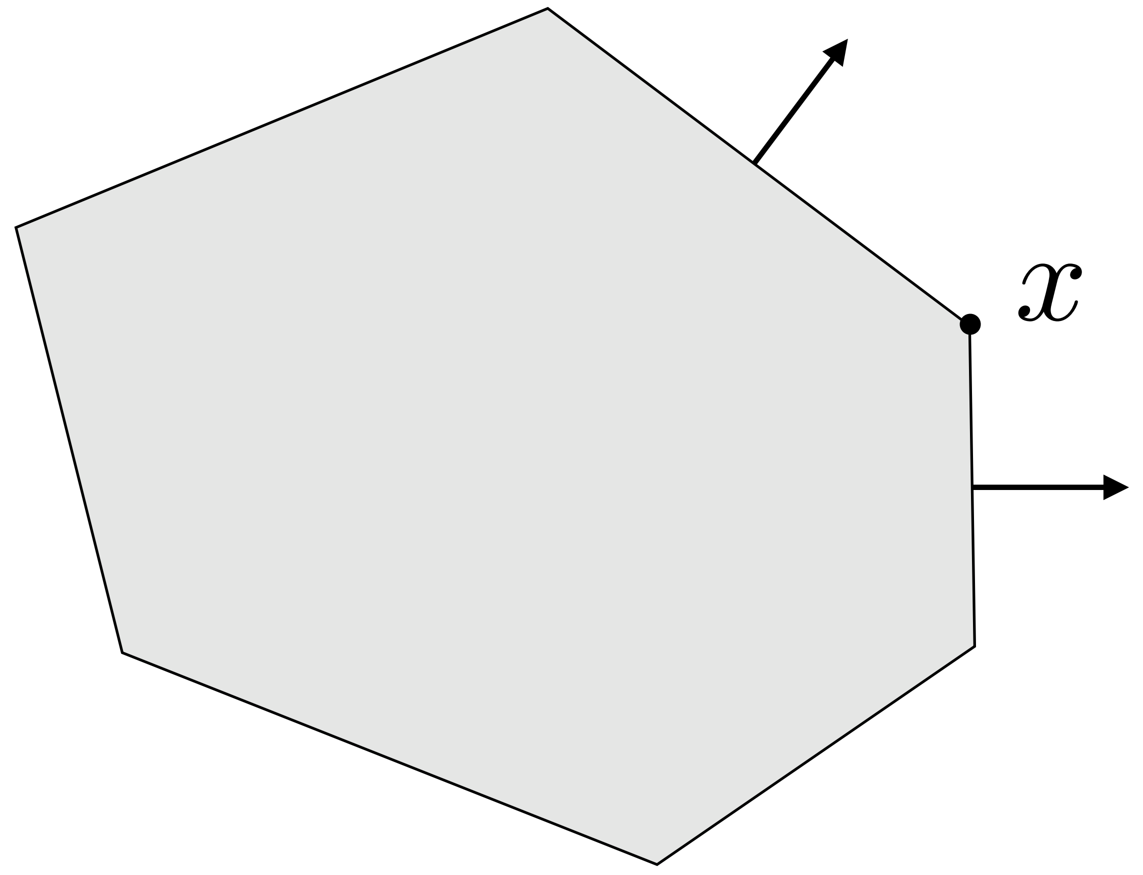
Then, there exists an optimal solution which is an **extreme point** of P

We only need to search between **extreme points**

Equivalence

Theorem

Given a nonempty polyhedron $P = \{x \mid Ax = b,\ x \geq 0\}$



Let $x \in P$

x is a **vertex** $\iff$ x is an **extreme point** $\iff$ x is a **basic feasible solution**

Constructing basic solution

1. Choose any m independent columns of A: $A_{B(1)}, \ldots, A_{B(m)}$
2. Let $x_i = 0$ for all $i \neq B(1), \ldots, B(m)$
3. Solve $Ax = b$ for the remaining $x_{B(1)}, \ldots, x_{B(m)}$

Basis matrix

Basis columns

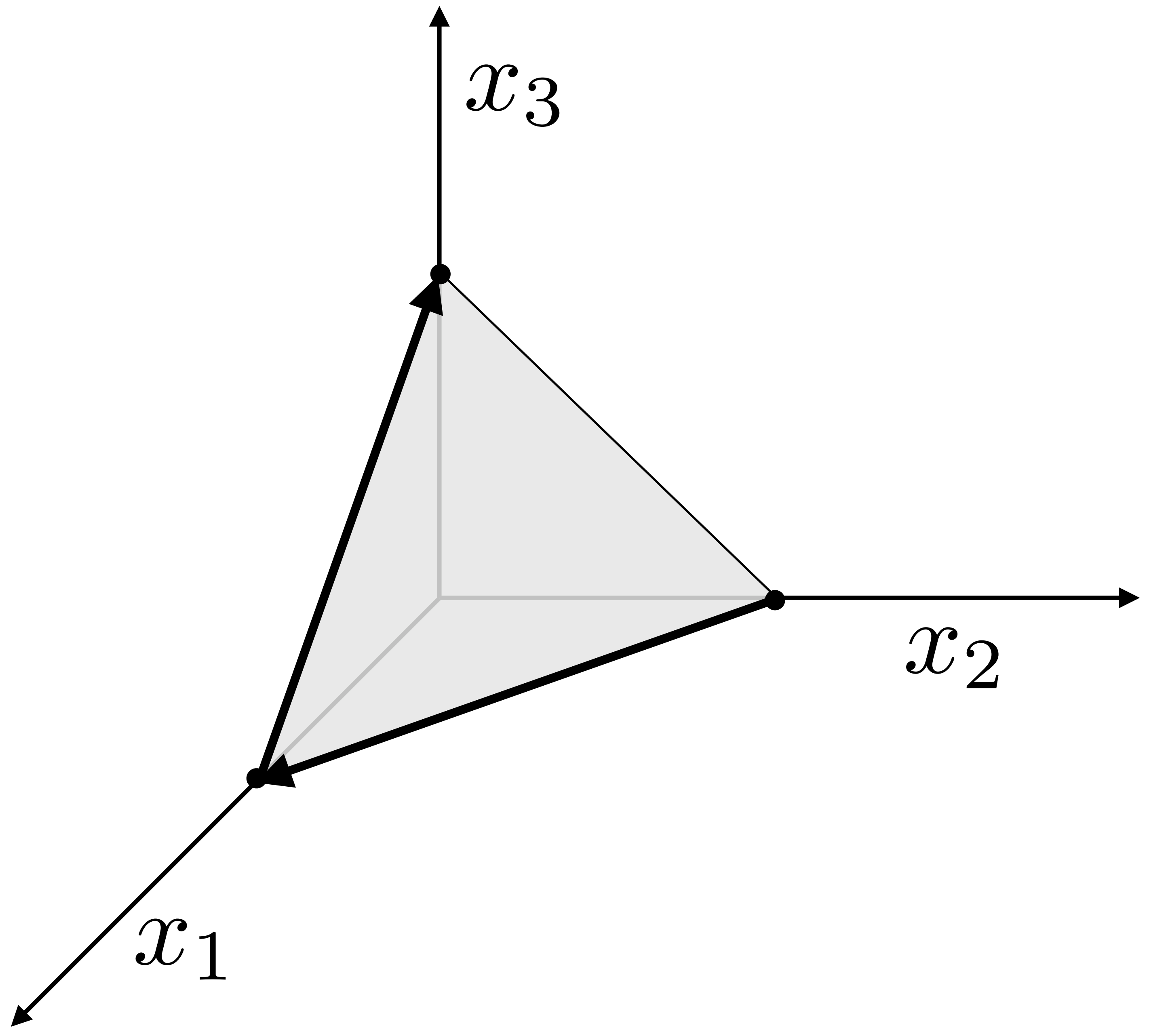
Basic variables

$$A_B = \begin{bmatrix} | & | & & | \\ A_{B(1)} & A_{B(2)} & \ldots & A_{B(m)} \\ | & | & & | \end{bmatrix}, \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix} \longrightarrow \text{Solve} \quad A_B x_B = b$$

If $x_B \geq 0$, then x is a **basic feasible solution**

Conceptual algorithm

- Start at corner
- Visit neighboring corner that improves the objective



How does the cost change?

Cost improvement

$$c^T(x+\theta d) - c^T x = \theta c^T d$$

New cost **Old cost**

We call $\bar{c}_j$ the **reduced cost** of (introducing) variable x_j in the basis

$$\bar{c}_j = c^T d = \sum_{i=1}^{n} c_j d_j = c_j + c_B^T d_B = c_j - c_B^T A_B^{-1} A_j$$

Reduced costs

Interpretation

Change in objective/marginal cost of adding x_j to the basis

$$\bar{c}_j = \boxed{c_j} \boxed{- c_B^T A_B^{-1} A_j}$$

Cost per-unit increase of variable x_j

Cost to change other variables compensating for x_j to enforce $Ax = b$

- $\bar{c}_j > 0$: adding x_j will increase the objective (bad)
- $\bar{c}_j < 0$: adding x_j will decrease the objective (good)

Reduced costs for basic variables is 0

$$\begin{aligned}\bar{c}_{B(i)} &= c_{B(i)} - c_B^T A_B^{-1} A_{B(i)} = c_{B(i)} - c_B^T (A_B^{-1} A_B) e_i \\ &= c_{B(i)} - c_B^T e_i = c_{B(i)} - c_{B(i)} = 0\end{aligned}$$

Optimality conditions

Theorem

Let x be a basic feasible solution associated with basis B

Let $\bar{c}$ be the vector of reduced costs.

If $\bar{c} \geq 0$, then x is **optimal**

Remark

This is a **stopping criterion** for the simplex algorithm.

If the **neighboring solutions** do not improve the cost, we are done

Single simplex iteration

1. Compute the reduced costs $\bar{c}$
   - Solve $A_B^T p = c_B$
   - $\bar{c} = c - A^T p$
2. If $\bar{c} \geq 0$, x **optimal. break**
3. Choose j such that $\bar{c}_j < 0$
4. Compute search direction d with $d_j = 1$ and $A_B d_B = -A_j$
5. If $d_B \geq 0$, the problem is **unbounded** and the optimal value is $-\infty$. **break**
6. Compute step length $\theta^\star = \min_{\{i \in B | d_i < 0\}} \left(-\frac{x_i}{d_i}\right)$
7. Define y such that $y = x + \theta^\star d$
8. Get new basis $\bar{B}$ (i exits and j enters)

Bottleneck
Two linear systems

→

Matrix inversion lemma trick
$\approx n^2$ per iteration
(very cheap)

How many iterations do we need?

Complexity of the simplex method

We do **not know any polynomial version of the simplex method**, no matter which pivoting rule we pick. → Still **open research question!**

Worst-case

There are problem instances where the simplex method will run an **exponential number of iterations** in terms of the dimensions, e.g. 2^n

Good news: average-case
Practical performance is very good. On average, it stops in n iterations.

Interior point method

Optimality conditions

Primal

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax + s = b \\
& s \geq 0
\end{array}
$$

Dual

$$
\begin{array}{ll}
\text{maximize} & -b^T y \\
\text{subject to} & A^T y + c = 0 \\
& y \geq 0
\end{array}
$$

KKT conditions

$$
\begin{aligned}
Ax + s - b &= 0 \\
A^T y + c &= 0 \\
s_i y_i &= 0, \quad i = 1, \dots, m \\
s, y &\geq 0
\end{aligned}
$$

$$
S = \begin{bmatrix} s_1 & & & \\ & s_2 & & \\ & & \ddots & \\ & & & s_m \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 & & & \\ & y_2 & & \\ & & \ddots & \\ & & & y_m \end{bmatrix}
$$

$$
\implies \quad SY\mathbf{1} = 0
$$

Main idea

$$h(x, s, y) = \begin{bmatrix} Ax + s - b \\ A^T y + c \\ SY\mathbf{1} \end{bmatrix} = 0$$

$$s, y \geq 0$$

$$S = \mathbf{diag}(s)$$
$$Y = \mathbf{diag}(y)$$

- Apply variants of Newton's method to solve $h(x, s, y) = 0$
- Enforce $s, y > 0$ (strictly) at every iteration
- **Motivation** avoid getting stuck in "corners"

Issue

Pure **Newton's step** does not allow significant progress towards $h(x, s, y) = 0$ **and** $x, y \geq 0$.

Smoothed optimality conditions

Optimality conditions

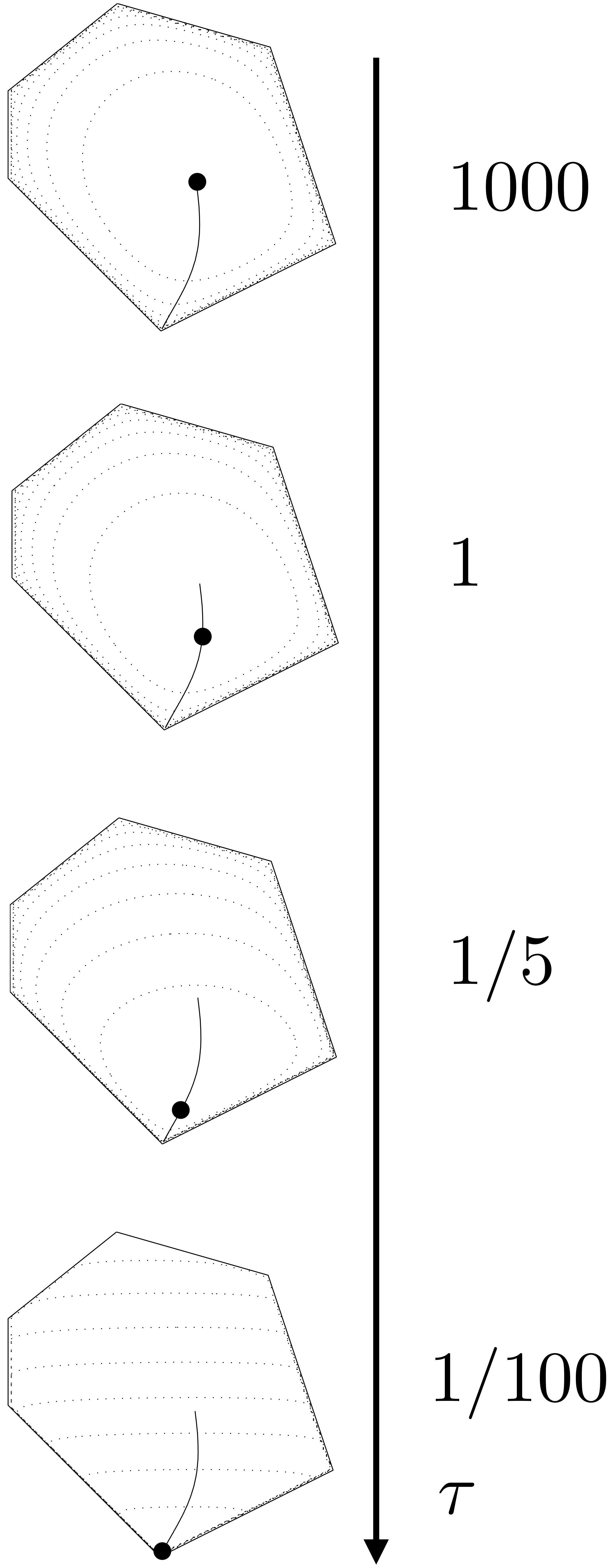
$$\begin{aligned} Ax + s - b &= 0 \\ A^T y + c &= 0 \\ s_i y_i &= \tau \\ s, y &\geq 0 \end{aligned}$$

$\longleftarrow$ Same τ for every pair

Same optimality conditions for a "smoothed" version of our problem

Central path

Analytic Center
$\tau \to \infty$



$$\begin{array}{ll} \text{minimize} & c^T x - \tau \sum_{i=1}^m \log(s_i) \\ \text{subject to} & Ax + s = b \end{array}$$

Set of points $(x^\star(\tau), s^\star(\tau), y^\star(\tau))$
with $\tau > 0$ such that

$$\begin{aligned} Ax + s - b &= 0 \\ A^T y + c &= 0 \\ s_i y_i &= \tau \\ s, y &\geq 0 \end{aligned}$$

Main idea

Follow central path as $\tau \to 0$

Newton's method for smoothed optimality conditions

Smoothed optimality conditions

$$h_\tau(x, s, y) = \begin{bmatrix} Ax + s - b \\ A^T y + c \\ SY\mathbf{1} - \tau\mathbf{1} \end{bmatrix} = 0$$

$$s, y \geq 0$$

Linear system

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Y \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_p \\ -r_d \\ -SY + \tau\mathbf{1} \end{bmatrix}$$

Line search to enforce $x, s > 0$

$$(x, s, y) \leftarrow (x, s, y) + \alpha(\Delta x, \Delta s, \Delta y)$$

Algorithm step

Linear system

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Y \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_p \\ -r_d \\ -SY\mathbf{1} + \sigma\mu\mathbf{1} \end{bmatrix}$$

Duality measure

$$\mu = \frac{s^T y}{m}$$

Centering parameter

$\sigma \in [0, 1]$

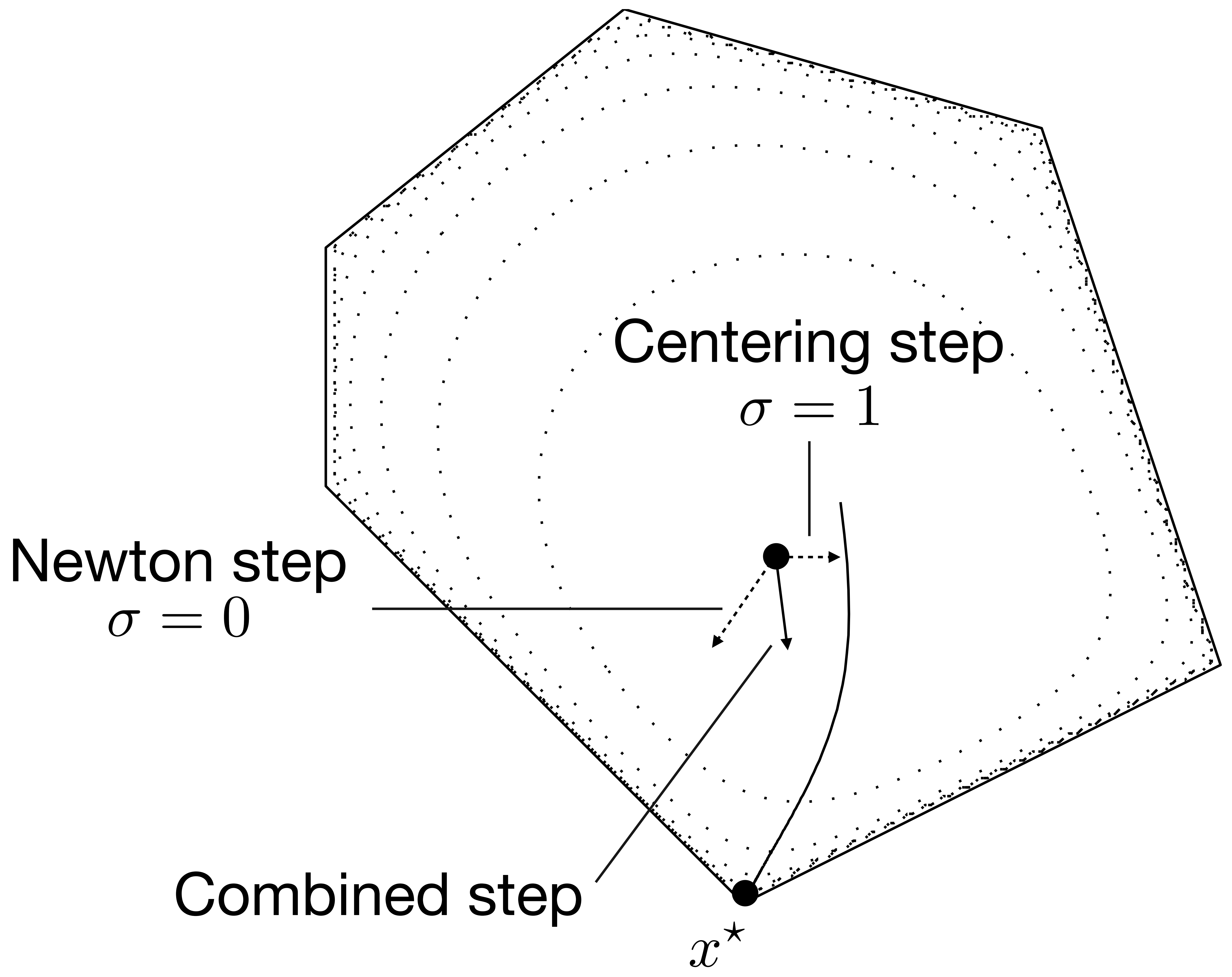
$\sigma = 0 \quad \Rightarrow \quad$ Newton step

$\sigma = 1 \quad \Rightarrow \quad$ Centering step towards $(x^\star(\mu), s^\star(\mu), y^\star(\mu))$

Line search to enforce $x, s > 0$

$$(x, s, y) \leftarrow (x, s, y) + \alpha(\Delta x, \Delta s, \Delta y)$$

Path-following algorithm idea



Centering step

Moves towards the **central path** and is usually biased towards $s, y > 0$. **No progress** on duality measure μ

Newton step

Moves towards the **zero duality measure** μ. Quickly violates $s, y > 0$.

Combined step

Best of both, with longer steps.

Choosing the centering parameter

Newton direction

$$(\Delta x_a, \Delta s_a, \Delta y_a)$$

- The Newton step might quickly hit nonnegativity.
- The centering step might not reduce complementary condition.

Maximum step-size

$$\alpha_p = \max\{\alpha \in [0,1] \mid s + \alpha \Delta s_a \geq 0\}$$

$$\alpha_d = \max\{\alpha \in [0,1] \mid y + \alpha \Delta y_a \geq 0\}$$

Centering parameter heuristic (after a Newton step)

$$\mu_a = \frac{(s + \alpha_p \Delta s_a)^T (y + \alpha_d \Delta y_a)}{m} \qquad \Longrightarrow \qquad \sigma = \left(\frac{\mu_a}{\mu}\right)^3$$

Mehrotra predictor-corrector algorithm

Initialization

Given (x, s, y) such that $s, y > 0$

1. **Termination conditions**

   $$r_p = Ax + s - b, \quad r_d = A^T y + c, \quad \mu = (s^T y)/m$$

   If $\|r_p\|, \|r_d\|, \mu$ are small, **break** Optimal solution $(x^\star, s^\star, y^\star)$

2. **Newton step** (affine scaling)

   $$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Y \end{bmatrix} \begin{bmatrix} \Delta y_a \\ \Delta x_a \\ \Delta s_a \end{bmatrix} = \begin{bmatrix} -r_p \\ -r_d \\ -SY\mathbf{1} \end{bmatrix}$$

Mehrotra predictor-corrector algorithm

3. **Centering parameter**

$$\alpha_p = \max\{\alpha \in [0,1] \mid s + \alpha \Delta s_a \geq 0\}$$

$$\alpha_d = \max\{\alpha \in [0,1] \mid y + \alpha \Delta y_a \geq 0\}$$

$$\mu_a = \frac{(s + \alpha_p \Delta s_a)^T (y + \alpha_d \Delta y_a)}{m}$$

$$\sigma = \left(\frac{\mu_a}{\mu}\right)^3$$

4. **Corrected direction**

$$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Y \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} -r_p \\ -r_d \\ -SY\mathbf{1} - \Delta S_a \Delta Y_a \mathbf{1} + \sigma\mu\mathbf{1} \end{bmatrix}$$

Linearization correction

Centering heuristic

Mehrotra predictor-corrector algorithm

5. **Update iterates**

$$\alpha_p = \max\{\alpha \geq 0 \mid s + \alpha \Delta s \geq 0\}$$
$$\alpha_d = \max\{\alpha \geq 0 \mid y + \alpha \Delta y \geq 0\}$$

$$(x, s) = (x, s) + \min\{1, \eta\alpha_p\}(\Delta x, \Delta s)$$
$$y = y + \min\{1, \eta\alpha_d\}\Delta y$$

Avoid corners

$$\eta = 1 - \epsilon \approx 0.99$$

Solving the search equations

Step 2 (**Newton**) and 4 (**Corrected direction**) solve equations of the form

(not symmetric) $\longrightarrow$ $$\begin{bmatrix} 0 & A & I \\ A^T & 0 & 0 \\ S & 0 & Y \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \\ \Delta s \end{bmatrix} = \begin{bmatrix} b_y \\ b_x \\ b_s \end{bmatrix}$$

Substitute last equation, $\Delta s = Y^{-1}(b_s - S\Delta y)$, into first

$$\begin{bmatrix} -Y^{-1}S & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} \Delta y \\ \Delta x \end{bmatrix} = \begin{bmatrix} b_y - Y^{-1}b_s \\ b_x \end{bmatrix}$$

Substitute first equation, $\Delta y = S^{-1}Y(A\Delta x - b_y + Y^{-1}b_s)$, into second

$$A^T S^{-1} Y A \Delta x = b_x + A^T S^{-1} Y b_y - A^T S^{-1} b_s$$

Reduced linear system

Coefficient matrix

$$B = A^T S^{-1} Y A$$

- B is **positive definite**
  if A has linearly independent columns
- Sparsity pattern of B is the **pattern** of $A^T A$
  (independent of $S^{-1}Y$)

Sparse cholesky factorization

$$B = PLL^T P^T$$

- Reorder only once to get P
- One numerical factorizaton per interior-point iteration $O(n^3)$
- Forward/backward substitution twice per iteration $O(n^2)$

$\longrightarrow$ **Per-iteration complexity** $O(n^3)$

Convergence

Mehrotra's algorithm

No convergence theory $\longrightarrow$ Examples where it **diverges** (rare!)

Fantastic convergence **in practice** $\longrightarrow$ Fewer than 30 iterations

Theoretical iteration complexity

Alternative versions (slower than Mehrotra) converge in $O(\sqrt{n})$ iterations $\longrightarrow$ **Operations** $O(n^{3.5})$

Average iteration complexity

Average iterations complexity is $O(\log n)$ $\longrightarrow$ $O(n^3 \log n)$

Interior-point vs simplex

Comparison between interior-point method and simplex

Primal simplex

- Primal feasibility

↓

- Zero duality gap
- Dual feasibility

Dual simplex

- Dual feasibility

↓

- Zero duality gap
- Primal feasibility

Primal-dual interior-point

- Interior condition

↓

- Primal feasibility
- Dual feasibility
- Zero duality gap

Simplex	Primal-dual interior-point
Exponential worst-case complexity	**Polynomial worst-case complexity**
Requires feasible point	**Allows infeasible start**
Can be warm-started	**Cannot be warm-started**

Which algorithm should I use?

Dual simplex

- Small-to-medium problems
- Repeated solves with varying constraints

Interior-point (barrier)

- Medium-to-large problems
- Sparse structured problems

How do solvers with multiple options decide?

<u>Concurrent Optimization</u>

Why not both? (crossover)

Interior-point $\longrightarrow$ Few simplex steps

Questions

Next lecture

- Integer optimization