Linear Programming

Chongnan Li
16271221@bjtu.edu.cn

Lecture 6 Interior Point Method

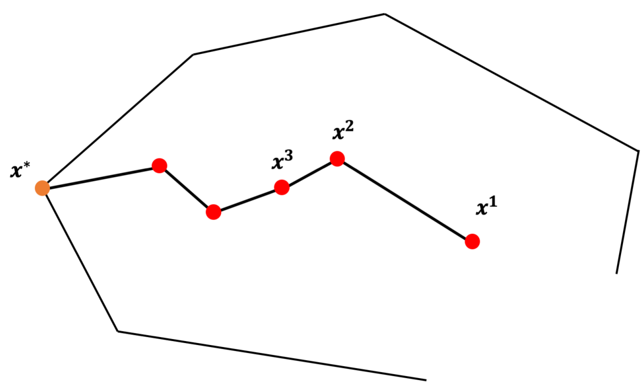
Outline:
1.Motivation
2.Basic concepts
3.Primal affine scaling algorithm
4.Dual affine scaling algorithm

● 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
= **# of iterations** × **# of operations in each iteration**

● Simplex method
-Simple operations: only check adjacent extreme points.
-May take many iterations: Klee-Minty example.

Question: **any fix?**

● 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 $\boldsymbol{x}^k$, we have

$$\boldsymbol{A}\boldsymbol{x}^k = \boldsymbol{b}, \boldsymbol{x}^k > \boldsymbol{0}.$$

-An interior movement has a **general** format

$$\boldsymbol{x}^{k+1} := \boldsymbol{x}^k + \alpha \boldsymbol{d}_x^k$$

$$\begin{cases} \alpha \geq 0: & \text{step-length} \\ \boldsymbol{d}_x^k \in R^n: & \text{moving direction} \end{cases}$$

Key knowledge

1. **Who is in the interior?**
   -Inital solution.

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

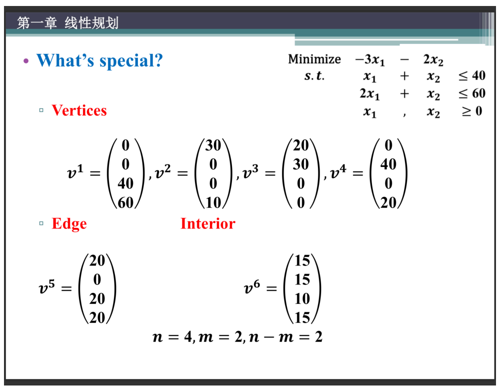
3. How to move to a new solution?
   -Which direction to move?(**good feasible direction**)
   -How far to go?(**step-length**)

Q1 - Who is in the interior?

Standard form LP

$$(\text{LP}) \quad \begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \boldsymbol{0} \end{array}$$

Who is at the vertex?
Who is on the edge?
Who is on the boundary?
Who is in the interior?



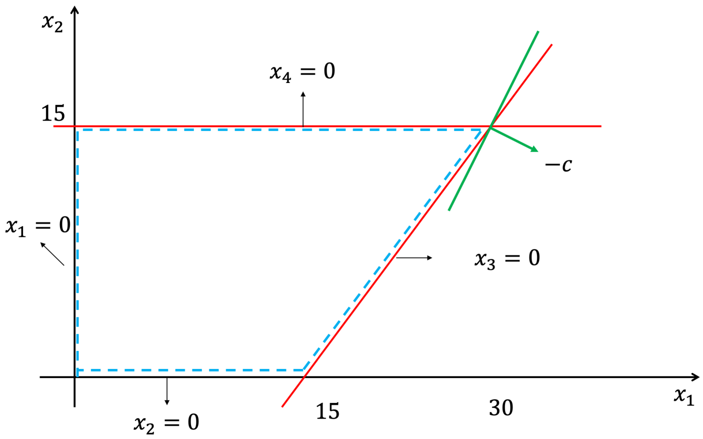
Two criteria for a point $\boldsymbol{x}$ to be an interior feasible solution:
1. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ (every linear constraint is satisfied)
2. $\boldsymbol{x} > \boldsymbol{0}$ (every component is positive)

Comments:

1. On a hyperplane $H = \{\boldsymbol{x} \in R^n \mid \boldsymbol{a}^T\boldsymbol{x} = \beta\}$. (**every point is interior relative to H since H does not have the concept of "boundary"!**)
2. For the first orthant $K = \{\boldsymbol{x} \in R^n \mid \boldsymbol{x} \geq \boldsymbol{0}\}$. (**only those $\boldsymbol{x} > \boldsymbol{0}$ are interior relative to K.**)

Example:

$$\begin{aligned} \min \quad & -2x_1 + x_2 \\ \text{s.t.} \quad & x_1 - x_2 \leq 15 \\ & x_2 \leq 15 \\ & x_1, x_2 \geq 0 \end{aligned}$$



How to find an initial interior solution?
Like the simplex method, we have
-**Big M** method
-**Two-phase** method
(to be discussed later in "Part Two")

● 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:

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha \boldsymbol{d}_x^k, \quad \boldsymbol{A}\boldsymbol{x}^k = \boldsymbol{b}, \boldsymbol{x}^k > \boldsymbol{0}.$$

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

$$\boldsymbol{A}\boldsymbol{x}^{k+1} = \boldsymbol{b} \Longleftrightarrow \boldsymbol{A}\boldsymbol{x}^k + \alpha \boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{b} \Rightarrow \boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{0},$$

i.e.
$\boldsymbol{d}_x^k \in \mathcal{N}(A)$: null space of A.

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

Observations:

$$\boldsymbol{c}^T\boldsymbol{x}^{k+1} \le \boldsymbol{c}^T\boldsymbol{x}^k \iff \boldsymbol{c}^T\boldsymbol{x}^k + \alpha\boldsymbol{c}^T\boldsymbol{d}_x^k \le \boldsymbol{c}^T\boldsymbol{x}^k \Rightarrow \boldsymbol{c}^T\boldsymbol{d}_x^k \le 0.$$

Optimality check
Principle: “no feasible direction at this point is a good direction.”
At a current solution, we check that:

$$\text{No } \boldsymbol{d}_x^k \in R^n \text{ with } \boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{0} \text{ can make } \boldsymbol{c}^T\boldsymbol{d}_x^k < 0.$$

● Q3 - How to move to a new solution?
1.Which direction to move?
- a good, feasible direction.

“Good” requires $\boldsymbol{c}^T\boldsymbol{d}_x^k \le 0$.
“Feasible” requires $\boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{0}$, $\boldsymbol{d}_x^k \in \mathcal{N}(A)$: null space of A.

Question: any suggestion?

A good direction
-Reduce the objective value

$\boldsymbol{c}^T\boldsymbol{d}_x^k \le 0$ Candidate: $\boldsymbol{d}_x^k = -\boldsymbol{c}$ **(negative gradient) (steepest descent)**

-Maintain feasibility

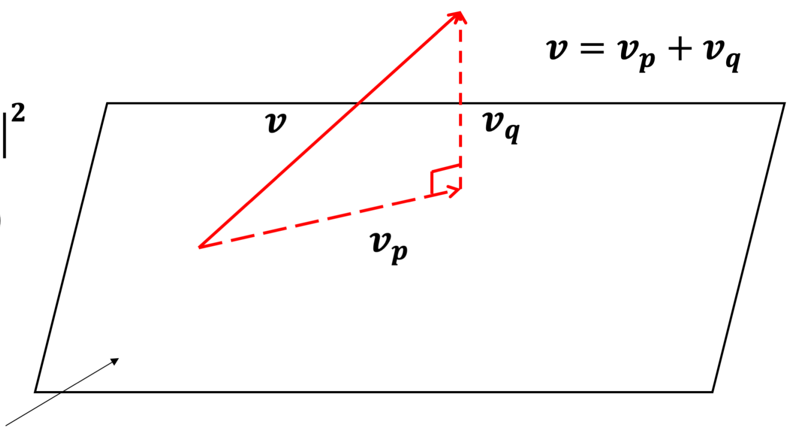
$\boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{0}$ Candidate: **projected** negative gradient

$$\boldsymbol{d}_x^k = [\boldsymbol{I} - \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{A}^T)^{-1}\boldsymbol{A}](-\boldsymbol{c})$$

Convex QP

$$\min_{\boldsymbol{v_p}} \|\boldsymbol{v_p} - \boldsymbol{v}\|^2$$

$$\text{s.t. } \boldsymbol{A}\boldsymbol{v_p} = \boldsymbol{0}$$



$\mathcal{N}(\boldsymbol{M})$ =Null space of matrix $\boldsymbol{M}$

$\mathcal{N}(\boldsymbol{M}) = \{\boldsymbol{x} \mid \boldsymbol{M}\boldsymbol{x} = \boldsymbol{0}\}$
$\boldsymbol{v}_p = [\boldsymbol{I} - \boldsymbol{M}^T(\boldsymbol{M}\boldsymbol{M}^T)^{-1}\boldsymbol{M}]\boldsymbol{v}$
$\boldsymbol{v}_q = \boldsymbol{M}^T(\boldsymbol{M}\boldsymbol{M}^T)^{-1}\boldsymbol{M}\boldsymbol{v}$

$d_x^k = [I - A^T(AA^T)^{-1}A](-c) = P_x^k(-c)$,

$P_x^k \triangleq I - A^T(AA^T)^{-1}A$: null space projection matrix

2. How far to go?
-To **satisfy every linear constraint**.
Since $Ad_x^k = 0$, $d_x^k \in \mathcal{N}(A)$: null space of A, $Ax^{k+1} = Ax^k + \alpha A d_x^{k+1} = b$, the step-length can be real number.

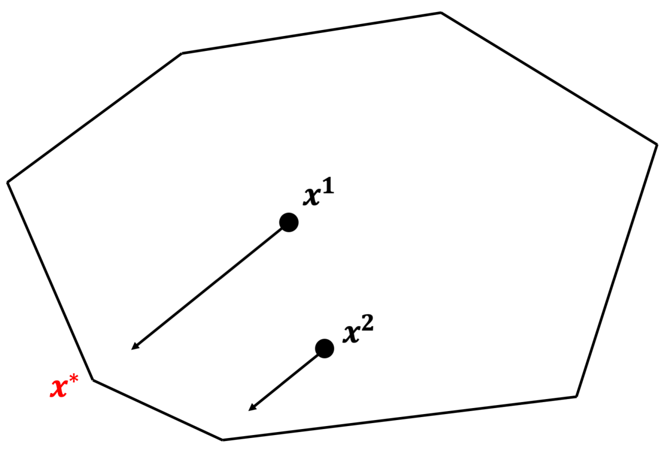
-To **stay to be an interior solution,** we need $x^{k+1} > 0$.

How to choose step-length?
One easy approach: in order to keep $x^{k+1} = x^k + \alpha d_x^k > 0$, 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?**

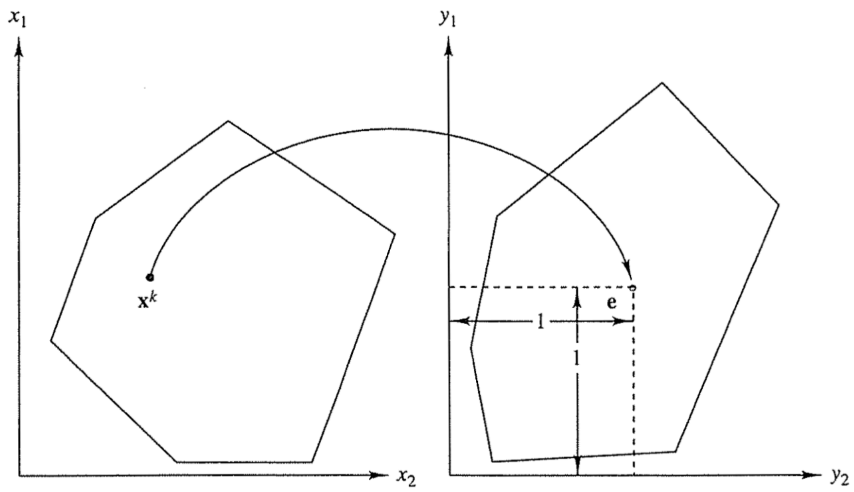
Obervations
If a current solution is **near the center** of the feasible domain (polyhedral set), it makes sense to move along the **steepest descent direction**.
If a current solution is not **near the center,** we need to **re-scale** its coordinates to transform it to become "near the center".



Question: **but how?**

● Where is the center?

We need to know where is the "**center**" of the **non-negative/first orthant**: $\{\boldsymbol{x} \in R^n \,|\, \boldsymbol{x} \geq \boldsymbol{0}\}$.



If $\boldsymbol{x}^k = \boldsymbol{e} = (1,1,\cdots,1)^T$, then

(1) $\boldsymbol{x}^k$ is one-unit away from the boundary.

(2) As long as $\alpha < 1$, $\boldsymbol{x}^{k+1} > 0$. **(if $|(\boldsymbol{d}_x^k)_i| \leq 1, \forall i$)**

Question: **If not, what to do?**

● Concept of scaling

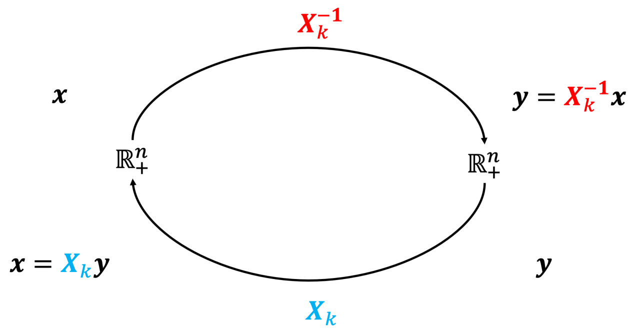
Scale $\boldsymbol{x}^k$ to be $\boldsymbol{e}$.

Define a diagonal scaling matrix:

$$\boldsymbol{X}_k = \mathrm{diag}(x_i^k) = \begin{pmatrix} x_1^k & 0 & \cdots & 0 \\ 0 & x_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n^k \end{pmatrix}$$

then $\boldsymbol{X}_k^{-1}\boldsymbol{x}^k = \boldsymbol{e}$.

● Transformation - affine scaling



The transformation is

1.one-to-one

2.onto

3.invertible

4.boundary to boundary

5.interior to interior

(7.1 ↓)

$$\begin{array}{lll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & A\boldsymbol{x}=\boldsymbol{b} \\ & \boldsymbol{x}\ge \boldsymbol{0} \end{array} \iff \begin{array}{ll} \min & \boldsymbol{c}^T X_k \boldsymbol{y} \\ \text{s.t.} & A X_k \boldsymbol{y}=\boldsymbol{b} \\ & \boldsymbol{y}\ge \boldsymbol{0} \end{array} \iff \begin{array}{ll} \min & (X_k^T\boldsymbol{c})^T\boldsymbol{y} \\ \text{s.t.} & (AX_k)\boldsymbol{y}=\boldsymbol{b} \\ & \boldsymbol{y}\ge \boldsymbol{0} \end{array} \iff \begin{array}{ll} \min & (X_k\boldsymbol{c})^T\boldsymbol{y} \\ \text{s.t.} & (AX_k)\boldsymbol{y}=\boldsymbol{b} \\ & \boldsymbol{y}\ge \boldsymbol{0} \end{array}$$

$\boldsymbol{x}^k > \boldsymbol{0}$ $\quad\quad$ $\boldsymbol{y}^k = \boldsymbol{e}$

$$\boldsymbol{d}_y^k = [\boldsymbol{I} - X_k A^T (A X_k^2 A^T)^{-1} A X_k](-X_k \boldsymbol{c})$$

$$\boldsymbol{y}^{k+1} = \boldsymbol{y}^k + \alpha_k \frac{\boldsymbol{d}_y^k}{\|\boldsymbol{d}_y^k\|}$$

$\alpha_k = 0.99$(say) $\quad 0 < \alpha_k < 1$

$$\boldsymbol{x}^{k+1} = X_k \boldsymbol{y}^{k+1} = X_k(\boldsymbol{y}^k + \alpha_k \frac{\boldsymbol{d}_y^k}{\|\boldsymbol{d}_y^k\|})$$

$$= X_k \boldsymbol{y}^k + \alpha_k X_k \frac{\boldsymbol{d}_y^k}{\|\boldsymbol{d}_y^k\|}$$

$$= \boldsymbol{x}^k + \frac{\alpha_k}{\|\boldsymbol{d}_y^k\|}\boldsymbol{d}_y^k$$

$$X_k \boldsymbol{d}_y^k = \boldsymbol{d}_x^k = -X_k[\boldsymbol{I} - X_k A^T (A X_k^2 A^T)^{-1} A X_k](X_k \boldsymbol{c})$$

● Step-length in the transformed space

Minimum ratio test in the y-space

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

$$\boldsymbol{y}^k + \alpha_k \boldsymbol{d}_y^k > \boldsymbol{0}$$

Case 1: $\boldsymbol{d}_k^y \ge \boldsymbol{0}$, then $\alpha_k \in (0, +\infty)$. Typically, $\boldsymbol{d}_y^k = \boldsymbol{0}$, then the objective value is a constant in its feasible domain.

Case 2: $(\boldsymbol{d}_y^k)_i < 0$ for some i.

Then

$$(\boldsymbol{y}^k)_i + \alpha_k(\boldsymbol{d}_y^k)_i > 0 \iff (\boldsymbol{y}^k)_i > -\alpha_k(\boldsymbol{d}_y^k)_i \iff \alpha_k < \frac{(\boldsymbol{y}^k)_i}{-(\boldsymbol{d}_y^k)_i} = \frac{1}{-(\boldsymbol{d}_y^k)_i}$$

$$\Rightarrow \alpha_k = \min_i \{ \frac{\alpha}{-(\boldsymbol{d}_y^k)_i} \mid (\boldsymbol{d}_y^k)_i < 0\} \text{ for some } \alpha \in (0,1).$$

● Property 1

Iteration in the x-space:

$$
\begin{aligned}
\boldsymbol{x}^{k+1} &= \boldsymbol{X}_k \boldsymbol{y}^{k+1} = \boldsymbol{X}_k(\boldsymbol{e} + \alpha_k \boldsymbol{d}_y^k) \\
&= \boldsymbol{x}^k + \alpha_k \boldsymbol{X}_k \boldsymbol{d}_y^k \\
&= \boldsymbol{x}^k + \alpha_k \boldsymbol{X}_k \boldsymbol{P}_k(-\boldsymbol{X}_k \boldsymbol{c}) \\
&= \boldsymbol{x}^k + \alpha_k \boldsymbol{X}_k[\boldsymbol{I} - \boldsymbol{X}_k \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k](-\boldsymbol{X}_k\boldsymbol{c}) \\
&= \boldsymbol{x}^k + \alpha_k[\boldsymbol{X}_k - \boldsymbol{X}_k^2 \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k](-\boldsymbol{X}_k\boldsymbol{c}) \\
&= \boldsymbol{x}^k - \alpha_k[\boldsymbol{X}_k^2\boldsymbol{c} - \boldsymbol{X}_k^2 \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}] \\
&= \boldsymbol{x}^k - \alpha_k\{\boldsymbol{X}_k^2[\boldsymbol{c} - \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}]\} \\
&= \boldsymbol{x}^k + \alpha_k\{-\boldsymbol{X}_k^2[\boldsymbol{c} - \boldsymbol{A}^T\underbrace{(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}}_{\boldsymbol{w}^k}]\} \\
&= \boldsymbol{x}^k + \alpha_k\{\underbrace{-\boldsymbol{X}_k^2[\boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{w}^k]}_{\boldsymbol{d}_x^k}\} \\
&= \boldsymbol{x}^k + \alpha_k \boldsymbol{d}_x^k
\end{aligned}
$$

● Property 2

Feasible direction in x-space:

$$
\begin{aligned}
\boldsymbol{x}^{k+1} = \boldsymbol{X}_k \boldsymbol{y}^{k+1} &= \boldsymbol{X}_k(\boldsymbol{y}^k + \alpha_k \frac{\boldsymbol{d}_y^k}{\|\boldsymbol{d}_y^k\|}) \\
&= \boldsymbol{X}_k \boldsymbol{y}^k + \alpha_k \boldsymbol{X}_k \frac{\boldsymbol{d}_y^k}{\|\boldsymbol{d}_y^k\|} \\
&= \boldsymbol{x}^k + \frac{\alpha_k}{\|\boldsymbol{d}_y^k\|}\boldsymbol{X}_k \boldsymbol{d}_y^k \\
&= \boldsymbol{x}^k + \frac{\alpha_k}{\|\boldsymbol{d}_y^k\|}\boldsymbol{d}_x^k
\end{aligned}
$$

Since $\boldsymbol{A}\boldsymbol{X}_k\boldsymbol{d}_y^k = \boldsymbol{0}$, $\boldsymbol{A}\boldsymbol{d}_x^k = \boldsymbol{0}$, i.e., $\boldsymbol{d}_x^k \in \mathcal{N}(A)$.

● Property 3

Good direction in x-space:

$$
\begin{aligned}
\boldsymbol{c}^T\boldsymbol{x}^{k+1} &= \boldsymbol{c}^T(\boldsymbol{x}^k + \alpha_k X_k \boldsymbol{d}_y^k) \\
&= \boldsymbol{c}^T\boldsymbol{x}^k + \alpha_k \boldsymbol{c}^T X_k \boldsymbol{d}_y^k \\
&= \boldsymbol{c}^T\boldsymbol{x}^k + \alpha_k \boldsymbol{c}^T X_k [\boldsymbol{P}_k(-X_k\boldsymbol{c})] \\
&= \boldsymbol{c}^T\boldsymbol{x}^k - \alpha_k (-X_k\boldsymbol{c})^T [\boldsymbol{P}_k(-X_k\boldsymbol{c})] \\
&= \boldsymbol{c}^T\boldsymbol{x}^k - \alpha_k \langle -X_k\boldsymbol{c}, \boldsymbol{P}_k(-X_k\boldsymbol{c})\rangle \\
&= \boldsymbol{c}^T\boldsymbol{x}^k - \alpha_k \|\boldsymbol{P}_k(-X_k\boldsymbol{c})\|^2 \\
&= \boldsymbol{c}^T\boldsymbol{x}^k - \alpha_k \|\boldsymbol{d}_y^k\|^2
\end{aligned}
$$

Hence, $\boldsymbol{c}^T\boldsymbol{x}^{k+1} \le \boldsymbol{c}^T\boldsymbol{x}^k$, and $\boldsymbol{c}^T\boldsymbol{x}^{k+1} < \boldsymbol{c}^T\boldsymbol{x}^k$ if $\boldsymbol{d}_y^k \neq \boldsymbol{0}$.

Lemma 7.1 If $\exists \boldsymbol{x}^k \in P$, $\boldsymbol{x}^k > \boldsymbol{0}$ with $\boldsymbol{d}_y^k > \boldsymbol{0}$, then the standard LP is unbounded below.

● Property 4

Optimality check

Lemma 7.2 If there exists an $\boldsymbol{x}^k \in P^0$ with $\boldsymbol{d}_y^k = \boldsymbol{0}$, then every feasible solution of the linear programming problem (7.1) is optimal. (the objective value is a constant)

For $\boldsymbol{x}^k \in P^0 = \{\boldsymbol{x} \in R^n \,|\, \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\}$, if $\boldsymbol{d}_y^k = -\boldsymbol{P}_k X_k \boldsymbol{c} = \boldsymbol{0}$, then $X_k\boldsymbol{c}$ falls in the orthogonal space of $\mathcal{N}(\boldsymbol{A}X_k)$, i.e., $X_k\boldsymbol{c} \in$ row space of $(\boldsymbol{A}X_k)$.

$\Rightarrow \exists \boldsymbol{u}^k$, s.t. $(\boldsymbol{A}X_k)^T\boldsymbol{u}^k = X_k\boldsymbol{c}$ or $(\boldsymbol{u}^k)^T\boldsymbol{A}X_k = \boldsymbol{c}^T X_k$

$\Rightarrow (\boldsymbol{u}^k)^T\boldsymbol{A} = \boldsymbol{c}^T$

For any feasible solution $\boldsymbol{x}$,

$$\boldsymbol{c}^T\boldsymbol{x} = (\boldsymbol{u}^k)^T\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{u}^k)^T\boldsymbol{b}$$

Since $(\boldsymbol{u}^k)^T\boldsymbol{b}$ does not depend on $\boldsymbol{x}$, the value of $\boldsymbol{c}^T\boldsymbol{x}$ remains constant over P.

● Property 5

Lemma 7.3 If the linear programming problem (7.1) is bounded below and its objective function is not constant, then the sequence $\{\boldsymbol{c}^T\boldsymbol{x}^k \,|\, k = 1,2,\cdots\}$ is well-defined and strictly decreasing.

● Property 6

We may define

$$\boldsymbol{w}^k \equiv (\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c} \qquad \text{(dual estimate)}$$
$$\boldsymbol{r}^k \equiv \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{w}^k \qquad \text{(reduced cost)}$$

If $\boldsymbol{r}^k \geq \boldsymbol{0}$, then $\boldsymbol{w}^k$ is dual feasible, and

$$(\boldsymbol{x}^k)^T\boldsymbol{r}^k = (\boldsymbol{X}_k\boldsymbol{y}^k)^T\boldsymbol{r}^k = (\boldsymbol{X}_k\boldsymbol{e})^T\boldsymbol{r}^k = \boldsymbol{e}^T\boldsymbol{X}_k^T\boldsymbol{r}^k = \boldsymbol{e}^T\boldsymbol{X}_k\boldsymbol{r}^k$$

becomes the duality gap, i.e., $\boldsymbol{c}^T\boldsymbol{x}^k - \boldsymbol{b}^T\boldsymbol{w}^k = \boldsymbol{e}^T\boldsymbol{X}_k\boldsymbol{r}^k$.

Therefore, if $\boldsymbol{r}^k \geq \boldsymbol{0}$ and $\boldsymbol{e}^T\boldsymbol{X}_k\boldsymbol{r}^k = 0$, then $\boldsymbol{x}^k \leftarrow \boldsymbol{x}^*$, $\boldsymbol{w}^k \leftarrow \boldsymbol{w}^*$.

● Property 7 Moving direction and reduced cost

$$\begin{aligned}
\boldsymbol{d}_y^k &= [\boldsymbol{I} - \boldsymbol{X}_k\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k](-\boldsymbol{X}_k\boldsymbol{c}) \\
&= -[\boldsymbol{X}_k\boldsymbol{c} - \boldsymbol{X}_k\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}] \\
&= -\boldsymbol{X}_k[\boldsymbol{c} - \boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}] \\
&= -\boldsymbol{X}_k[\boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{w}^k] \\
&= -\boldsymbol{X}_k\boldsymbol{r}^k
\end{aligned}$$

● Key steps of primal affine scaling algorithm

Step 1 Set $k \leftarrow 0$, $\epsilon > 0$, $0 < \alpha < 1$,
find $\boldsymbol{x}^0 > \boldsymbol{0}$ and $\boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{b}$.

Step 2 Compute
$\boldsymbol{w}^k = (\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{A}^T)^{-1}\boldsymbol{A}\boldsymbol{X}_k^2\boldsymbol{c}$
$\boldsymbol{r}^k = \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{w}^k$
If $\boldsymbol{r}^k \geq \boldsymbol{0}$ and $\boldsymbol{e}^T\boldsymbol{X}_k\boldsymbol{r}^k \leq \epsilon$,
then STOP! $\boldsymbol{x}^k \leftarrow \boldsymbol{x}^*$, $\boldsymbol{w}^k \leftarrow \boldsymbol{w}^*$.
Otherwise,

Step 3 Compute $\boldsymbol{d}_y^k = -\boldsymbol{X}_k\boldsymbol{r}^k$.
If $\boldsymbol{d}_y^k \geq \boldsymbol{0}$(but $\neq \boldsymbol{0}$), then STOP! Unbounded below.
If $\boldsymbol{d}_y^k = \boldsymbol{0}$, then STOP! $\boldsymbol{x}^k \leftarrow \boldsymbol{x}^*$ (the objective value is a constant).
Otherwise,

Step 4 Find $\alpha_k = \min_i\{\frac{\alpha}{-(\boldsymbol{d}_y^k)_i} \mid (\boldsymbol{d}_y^k)_i < 0\}$.
$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \alpha_k\boldsymbol{X}_k\boldsymbol{d}_y^k$
$k := k + 1$.
Go to Step 2.