Linear Programming VI : Interior-Point Method

Asst. Prof. Dr.-Ing. Sudchai Boonto

October 4, 2023

Department of Control System and Instrumentation Engineering
King Mongkut's Unniversity of Technology Thonburi
Thailand

Objective

- Understand the interior-Point Methods

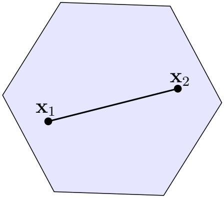
Convex Sets

Definition 6.1 Convex Sets

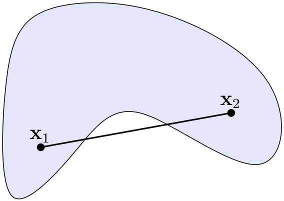
A sets $\mathbb{R}_c$ is said to be convex if for every pair of points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_c$ and for every real number $0 \leq \alpha \leq 1$, the point

$$\mathbf{x} = \alpha \mathbf{x}_1 + (1 - \alpha)\mathbf{x}_2$$

is located in $\mathbb{R}_c$



convex set



nonconvex set

Definition 6.2 Convex Functions

- A function $f(\mathbf{x})$ defined over a convex set $\mathbb{R}_c$ is said to be convex if for every pair of points $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}_c$ and every real number $0 \leq \alpha \leq 1$, the inequality

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)$$

  hold. If $x_1 \neq x_2$ and

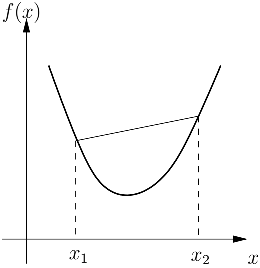
$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2)$$

  then $f(\mathbf{x})$ is said to be strictly convex.

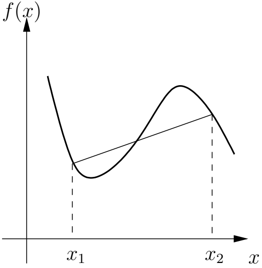
- If $\psi(\mathbf{x})$ is defined over a convex set $\mathbb{R}_c$ and $f(\mathbf{x}) = -\psi(\mathbf{x})$ is convex, then $\phi(\mathbf{x})$ is said to be concave. If $f(\mathbf{x})$ is strictly convex, $\psi(\mathbf{x})$ is strictly concave.

is located in $\mathbb{R}_c$

Convex Functions : Convexity



Convex



Nonconvex

Theorem 6.3: Convexify of linear combination of convex function

If

$$f(\mathbf{x}) = af_1(\mathbf{x}) + bf_2(\mathbf{x})$$

where $a, b \geq 0$ and $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ are convex functions on the convex set $\mathbb{R}_c$, then $f(\mathbf{x})$ is convex on the set $\mathbb{R}_c$.

Proof: Since $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are convex, and $a, b \geq 0$, then for $\mathbf{x} = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$, we have

$$af_1(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_1) \leq a(\alpha f_1(\mathbf{x}_1) + (1-\alpha)f_1(\mathbf{x}_2))$$
$$af_2(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_1) \leq b(\alpha f_1(\mathbf{x}_1) + (1-\alpha)f_1(\mathbf{x}_2))$$

Since,

$$\begin{aligned} f(\mathbf{x}) &= af_1(\mathbf{x}) + bf_2(\mathbf{x}) \\ f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &= af_1(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) + bf_2(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \\ &\leq \alpha(\underbrace{af_1(\mathbf{x}_1) + bf_2(\mathbf{x}_1)}_{f(\mathbf{x}_1)}) + (1-\alpha)(\underbrace{af_1(\mathbf{x}_2) + af_2(\mathbf{x}_2)}_{f(\mathbf{x}_2)}) \\ f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) &\leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2) \end{aligned}$$

That is $f(\mathbf{x})$ is convex.

Convex Functions : Convexity

Theorem 6.4: Relation between convex functions and convex sets.

If $f(x)$ is a convex function on a convex set $\mathbb{R}_c$, then the set

$$\mathcal{S}_c = \{\mathbf{x} : \mathbf{x} \in \mathbb{R}_c, f(\mathbf{x}) \leq K\}$$

is convex for every real number K.

Proof: If $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{S}_c$, then $f(\mathbf{x}_1) \leq K$ and $f(\mathbf{x}_2) \leq K$ from the definition of $\mathcal{S}_c$. Since $f(\mathbf{x})$ is convex

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \leq \alpha K + (1-\alpha)K$$
$$\text{or} \quad f(\mathbf{x}) \leq K \quad \text{for} \quad \mathbf{x} = \alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 \quad \text{and} \quad 0 < \alpha < 1$$

Therefore $\mathbf{x} \in \mathcal{S}_c$. That is, $\mathcal{S}_c$ is convex by virtue of the definition of convex set.

Convex Functions : Convexity

An alternative view of convexity can be generated by examining some theorems which involve the gradient and Hessian of $f(\mathbf{x})$.

Theorem 6.5: Property of convex functions relating to gradient

If $f(\mathbf{x}) \in \mathbb{C}^1$, then $f(\mathbf{x})$ is convex over a convex set $\mathbb{R}_c$ if and only if $f(\mathbf{x}_1) \geq f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x})$ for all $\mathbf{x}$ and $\mathbf{x}_1 \in \mathbb{R}_c$, where $\mathbf{g}(\mathbf{x})$ is the gradient of $f(\mathbf{x})$.

Proof:

- Show that if $f(\mathbf{x})$ is convex, then

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}) \leq \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x})$$
$$f(\mathbf{x} + \alpha(\mathbf{x}_1 - \mathbf{x})) - f(\mathbf{x}) \leq \alpha(f(\mathbf{x}_1) - f(\mathbf{x}))$$

As $\alpha \to 0$, the Taylor series of $f(\mathbf{x} + \alpha(\mathbf{x}_1 - \mathbf{x}))$ yields

$$f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T\alpha(\mathbf{x}_1 - \mathbf{x}) - f(\mathbf{x}) \leq \alpha(f(\mathbf{x}_1) - f(\mathbf{x}))$$
$$f(\mathbf{x}_1) \geq f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x})$$

Convex Functions : Convexity

- If the inequality holds at points $\mathbf{x}$ and $\mathbf{x}_2 \in \mathbb{R}_c$, then $f(\mathbf{x}_2) \geq f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_2 - \mathbf{x})$. Hence

$$\alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \geq \alpha f(\mathbf{x}) + \alpha \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x}) + (1-\alpha) f(\mathbf{x}) + (1-\alpha)\mathbf{g}(\mathbf{x})^T(\mathbf{x}_2 - \mathbf{x})$$

or

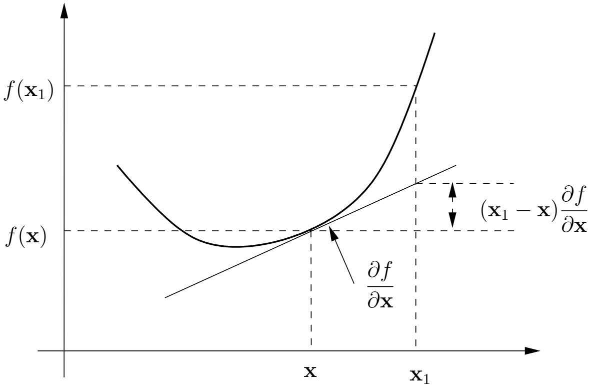
$$\alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) \geq f(\mathbf{x}) + \mathbf{g}^T(\mathbf{x})(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2 - \mathbf{x})$$

With the substitution $\mathbf{x} = \alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2$, we obtain

$$f(\alpha \mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2)$$

From the definition, $f(\mathbf{x})$ is convex.

The theorem 6.5 states the a linear approximation of $f(\mathbf{x})$ at point $\mathbf{x_1}$ based on the derivatives of $f(\mathbf{x})$ at $\mathbf{x}$ underestimates the value of the function.



Theorem 10.6: Property of convex functions relating to the Hessian

A function $f(\mathbf{x}) \in \mathbb{C}^2$ is convex over a convex set $\mathbb{R}_c$ if an only if the Hessian $\mathbf{H}(\mathbf{x})$ of $f(\mathbf{x})$ is positive semi-definite for $\mathbf{x} \in \mathbb{R}_c$

Proof: If $\mathbf{x}_1 = \mathbf{x} + \mathbf{d}$ where $\mathbf{x}_1$ and $\mathbf{x}$ are arbitrary points in $\mathbb{R}_c$, then the Taylor series yields

$$f(\mathbf{x}_1) = f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x}) + \frac{1}{2}\mathbf{d}^T\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})\mathbf{d}$$

If $\mathbf{H}(\mathbf{x})$ is positive semidefinite everywhere in $\mathbb{R}_c$, then

$$\frac{1}{2}\mathbf{d}^T\mathbf{H}(\mathbf{x} + \alpha\mathbf{d})\mathbf{d} \geq 0 \text{ and so } f(\mathbf{x}_1) \geq f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x})$$

Then $f(\mathbf{x})$ is convex.

If $\mathbf{H}(\mathbf{x})$ is not positive semidefinite everywhere in $\mathbb{R}_c$, then a point $\mathbf{x}$ and at least a $\mathbf{d}$ exist such that

$$\mathbf{d}^T\mathbf{H}(\mathbf{x}+\alpha\mathbf{d})\mathbf{d} < 0$$

$$f(\mathbf{x}_1) < f(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x})$$

and $f(\mathbf{x})$ is nonconvex from Theorem 6.5. Therefore, $f(\mathbf{x})$ is convex if and only if $\mathbf{H}(\mathbf{x})$ is positive semi-definite everywhere in $\mathbb{R}_c$.

Example: Check the following functions for convexity

(a) $f(\mathbf{x}) = e^{x_1} + x_2^2 + 5$ (b) $f(\mathbf{x}) = 3x_1^2 - 5x_1x_2 + x_2^2$ (c) $f(\mathbf{x}) = \frac{1}{4}x_1^4 - x_1^2 + x_2^2$
(d) $f(\mathbf{x}) = 50 + 10x_1 + x_2 - 6x_1^2 - 3x_2^2$

(a) The Hessian can be obtained as

$$\mathbf{H} = \begin{bmatrix} e^{x_1} & 0 \\ 0 & 2 \end{bmatrix}$$

For $-\infty < x_1 < \infty$, $\mathbf{H}$ is positive definite and $f(\mathbf{x})$ is strictly convex.

(b) We have

$$\mathbf{H} = \begin{bmatrix} 6 & -5 \\ -5 & 2 \end{bmatrix}$$

Since $\mathbf{H}$ is indefinte, $f(\mathbf{x})$ is neither convex nor concave.

(c) We get

$$\mathbf{H} = \begin{bmatrix} 3x_1^2 - 2 & 0 \\ 0 & 2 \end{bmatrix}$$

For $x_1 \leq -\sqrt{2/3}$ and $x_1 \geq \sqrt{2/3}$, $\mathbf{H}$ is positive semi-definite and $f(\mathbf{x})$ is convex; for $x_1 < -\sqrt{2/3}$ and $x_1 > \sqrt{2/3}$. $\mathbf{H}$ is positive definite and $f(x)$ is strictly convex; for $-\sqrt{2/3} < x_1 < \sqrt{2/3}$, $\mathbf{H}$ is indefinite, and $f(\mathbf{x})$ is neither convex nor concave.

Convex Functions : Optimization

Theorem 6.7: Relation between local and global minimizers in convex functions

If $f(\mathbf{x})$ is a convex function defined on a convex set $\mathbb{R}_c$, then

(a) the set of points $\mathbb{S}_c$ where $f(\mathbf{x})$ is minimum is convex;

(b) any local minimizer of $f(\mathbf{x})$ is a global minimizer

Proof: (a) If F^* is a minimum of $f(\mathbf{x})$, then $\mathbb{S}_c = \{\mathbf{x} : f(\mathbf{x}) \leq F^*, \mathbf{x} \in \mathbb{R}_c\}$ is convex by virtue of Theorem 6.4

(b) If $\mathbf{x}^* \in \mathbb{R}_c$ is a local minimizer but there is another point $\mathbf{x}^{**} \in \mathbb{R}_c$ which is a global minimizer such that $f(\mathbf{x}^{**}) < f(\mathbf{x}^*)$ the one the line $\mathbf{x} = \alpha\mathbf{x}^{**} + (1-\alpha)\mathbf{x}^*$

$$f(\alpha\mathbf{x}^{**} + (1-\alpha)\mathbf{x}^*) \leq \alpha f(\mathbf{x}^{**}) + (1-\alpha)f(\mathbf{x}^*) < \alpha f(\mathbf{x}^*) + (1-\alpha)f(\mathbf{x}^*)$$

or $f(\mathbf{x}) < f(\mathbf{x}^*)$ for all α

This is contradicts the fact that $\mathbf{x}^*$ is a local minimizer and so $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ for all $\mathbf{x} \in \mathbb{R}_c$. Therefore, any local minimizers are located in a convex set, and all are global minimizers.

Theorem 6.8: Existence of a global minimizer in convex functions

If $f(\mathbf{x}) \in \mathbb{C}^1$ is a convex function on a convex set $\mathbb{R}_c$ and there is a point $\mathbf{x}^*$ such that

$$\mathbf{g}(\mathbf{x}^*)^T\mathbf{d} \geq 0 \text{ where } \mathbf{d} = \mathbf{x}_1 - \mathbf{x}^*$$

for all $\mathbf{x}_1 \in \mathbb{R}_c$, then $\mathbf{x}^*$ is a global minimizer of $f(\mathbf{x})$.

Proof: From Theorem 6.6, we have $f(\mathbf{x}_1) \geq f(\mathbf{x}^*) + \mathbf{g}(\mathbf{x}^*)^T(\mathbf{x}_1 - \mathbf{x}^*)$ where $\mathbf{g}(\mathbf{x}^*)$ is the gradient of $f(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}^*$. Since $\mathbf{g}(\mathbf{x}^*)^T(\mathbf{x}_1 - \mathbf{x}^*) \geq 0$, we have

$$f(\mathbf{x}_1) \geq f(\mathbf{x}^*)$$

and so $\mathbf{x}^*$ is a local minimizer. By virtue of Theorem 6.7, $\mathbf{x}^*$ is also a global minimizer. Similarly, if $f(\mathbf{x})$ is a strictly convex function and $\mathbf{g}(\mathbf{x}^*)^T\mathbf{d} > 0$, then $\mathbf{x}^*$ is a strong global minimizer.

Duality : The Lagrangian

Consider an optimization problem in the standard form:

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & a_i^T \mathbf{x} - b = 0, \quad i = 1, \ldots, p \\ & \mathbf{c}_j(\mathbf{x}) \leq 0, \quad j = 1, \ldots q \end{aligned}$$

with variable $\mathbf{x} \in \mathbb{R}^n$, domain $\mathcal{D}$, and optimal value p^*.
The basic idea in Lagrangian duality is to take the constraints about into account by augmenting the objective function with a weighted sum of the constraint functions.

Lagrangian:

$\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$, with dom $\mathcal{L} = \mathcal{D} \times \mathbb{R}^p \times \mathbb{R}^q$,

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i (\mathbf{a}_i^T \mathbf{x} - b_i) + \sum_{j=1}^{q} \mu_j c_j(\mathbf{x})$$

- weighted sum of objective and constraint functions
- λ_i is Lagrange multiplier associated with $a_i^T \mathbf{x} - b_i = 0$
- μ_j is Lagrange multiplier associated with $c_j(\mathbf{x}) \leq 0$.

Duality: The Lagrange dual function

The Lagrange dual function:

$\mathbf{g} : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$ as the minimum value of the Lagrangian over $\mathbf{x}$: for $\boldsymbol{\lambda} \in \mathbb{R}^p$, $\boldsymbol{\mu} \in \mathbb{R}^q$,

$$q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{x \in \mathcal{D}} \mathcal{L}(\mathbf{x}, \lambda, \mu) = \inf_{\mathbf{x} \in \mathcal{D}} \left(f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i (\mathbf{a}_i^T \mathbf{x} - b_i) + \sum_{j=1}^{q} \mu_j c_j(x) \right)$$

$\mathbf{g}$ is concave, it can be $-\infty$ for some $\boldsymbol{\lambda}, \boldsymbol{\mu}$.

Lagrange dual problem

The Lagrange dual problem with respect to the convex problem is defined as

$$\begin{aligned} \underset{\boldsymbol{\lambda}, \boldsymbol{\mu}}{\text{maximize}} \quad & q(\boldsymbol{\lambda}, \boldsymbol{\mu}) \\ \text{subject to} \quad & \boldsymbol{\mu} \geq 0 \end{aligned}$$

Duality: The Lagrange dual function

- For any feasible $\mathbf{x}$ and any feasible $\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}$ of the above maximize problem, we have

$$f(\mathbf{x}) \geq q(\boldsymbol{\lambda}, \boldsymbol{\mu})$$

Because

$$\begin{aligned}\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) &= f(\mathbf{x}) + \sum_{i=1}^{p} \lambda_i (\mathbf{a}_i^T \mathbf{x} - b_i) + \sum_{j=1}^{q} \mu_j c_j(\mathbf{x}) \\ &= f(\mathbf{x}) + \sum_{j=1}^{q} \mu_j c_j(\mathbf{x}) \leq f(\mathbf{x}) \quad \text{since } j(\mathbf{x}) \leq 0\end{aligned}$$

Thus

$$q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \leq f(\mathbf{x})$$

Duality : Standard form LP

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq 0 \end{aligned} \tag{1}$$

Dual function

- Lagrangian is ($\mathbf{x} \geq 0 \quad \Rightarrow -\mathbf{x} \leq 0$)

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{c}^T\mathbf{x} + \boldsymbol{\lambda}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) - \boldsymbol{\mu}^T\mathbf{x} = -\mathbf{b}^T\boldsymbol{\lambda} + (\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} - \boldsymbol{\mu})^T\mathbf{x}$$

- $\mathcal{L}$ is affine in $\mathbf{x}$, hence (The linear function is bounded from below only when it is identically zero. Then $\mathbf{g}(\boldsymbol{\lambda}, \boldsymbol{\mu}) = -\infty$ except when $\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} - \boldsymbol{\mu} = 0$)

$$q(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \inf_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \begin{cases} -\mathbf{b}^T\boldsymbol{\lambda}, & \mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} - \boldsymbol{\mu} = 0 \\ -\infty, & \text{otherwise} \end{cases}$$

q is linear on affine domain $\{(\boldsymbol{\lambda}, \boldsymbol{\mu}) | \mathbf{A}^T\boldsymbol{\mu} - \boldsymbol{\lambda} + \mathbf{c} = 0\}$, hence concave

Lower bound property: $p^* \geq -\mathbf{b}^T\boldsymbol{\lambda}$ if $\mathbf{A}^T\boldsymbol{\lambda} + \mathbf{c} \geq 0$.

Duality : Standard form LP

The Lagrange dual problem

The lagrange dual problem is defined as

$$\begin{aligned} \underset{\boldsymbol{\lambda},\boldsymbol{\mu}}{\text{maximize}} \quad & q(\boldsymbol{\lambda},\boldsymbol{\mu}) \\ \text{subject to} \quad & \boldsymbol{\mu} \geq 0 \end{aligned}$$

For the standard form LP, we have

$$\begin{aligned} \underset{\boldsymbol{\lambda},\boldsymbol{\mu}}{\text{maximize}} \quad & -\mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject to} \quad & \boldsymbol{\mu} \geq 0 \end{aligned}$$

Since $\mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda} - \boldsymbol{\mu} = 0$ and $\boldsymbol{\mu} = \mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda}$, the above problem becomes

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & -\mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject to} \quad & -\mathbf{c} - \mathbf{A}^T\boldsymbol{\lambda} \leq 0 \end{aligned}$$

$$\begin{aligned} \underset{\boldsymbol{\lambda}}{\text{minimize}} \quad & \mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject to} \quad & (-\mathbf{A}^T)\boldsymbol{\lambda} \leq \mathbf{c} \end{aligned}$$

Primal-Dual Solutions and Central Path : Primal-Dual Solutions

The standard-form LP problem

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} \\
\text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geq 0
\end{aligned} \tag{2}
$$

The dual problem is

$$
\begin{aligned}
\underset{\boldsymbol{\lambda}}{\text{maximize}} \quad & \mathbf{h}(\boldsymbol{\lambda}) = -\mathbf{b}^T\boldsymbol{\lambda} \\
\text{subject to} \quad & -\mathbf{A}^T\boldsymbol{\lambda} + \boldsymbol{\mu} = \mathbf{c}, \quad \boldsymbol{\lambda} \geq 0 \text{ (or) } \mathbf{A}^T\boldsymbol{\lambda} + \mathbf{c} \geq 0
\end{aligned} \tag{3}
$$

- Under what conditions will the solutions of these problems exist?
- How are the feasible points and solutions of the primal and dual related?
- $\boldsymbol{\mu} \geq 0$

- An LP problem is said to be **feasible** if its feasible region is not empty. The problem in (2) is said to be **strictly feasible** if there exists and $\mathbf{x}$ that satisfies $-\boldsymbol{\lambda}^T\mathbf{A} + \boldsymbol{\mu} = \mathbf{c}$ with $\mathbf{x} \geq 0$
- The LP problem in (3) is said to be strictly feasible if there exist $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ that satisfy $-\boldsymbol{\lambda}^T\mathbf{A} + \boldsymbol{\mu} = \mathbf{c}$ with $\boldsymbol{\mu} \geq 0$.
- It is known that $\mathbf{x}^*$ is a minimizer of the problem in (2) if and only if there exist $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^* \geq 0$ such that

$$\begin{aligned} -\mathbf{A}^T\boldsymbol{\lambda}^* + \boldsymbol{\mu}^* &= \mathbf{c} \\ \mathbf{A}\mathbf{x}^* &= \mathbf{b} \\ x_i^*\mu_i^* &= 0 \text{ for } 1 \leq i \leq n \\ \mathbf{x}^* \geq 0, \quad \boldsymbol{\mu}^* &\geq 0 \end{aligned} \tag{4}$$

- A set $\{\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ satisfying (4) is called a **primal-dual solution**. The set $\{\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ is a primal-dual solution if and only if $\mathbf{x}^*$ solves the primal and $\{\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ solves the dual.

Theorem: 10.9 Existence of a primal-dual solution

A primal-dual solution exists if the primal and dual problems are both feasible.

Proof: If point $\mathbf{x}$ is feasible for the LP problem and $\{\boldsymbol{\lambda}, \boldsymbol{\mu}\}$ is feasible for the LP problem, then set

$$
\begin{aligned}
-\boldsymbol{\lambda}^T \mathbf{b} \leq -\boldsymbol{\lambda}^T \mathbf{b} + \boldsymbol{\mu}^T \mathbf{x} &= -\boldsymbol{\lambda}^T \mathbf{A}\mathbf{x} + \boldsymbol{\mu}^T \mathbf{x} \\
&= (-\mathbf{A}^T \boldsymbol{\lambda} + \boldsymbol{\mu})^T \mathbf{x} = \mathbf{c}^T \mathbf{x}
\end{aligned}
$$

Since $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x}$ has a finite lower bound in the feasible region, there exists a set $\{\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ that satisfies (4). This $\mathbf{x}^*$ solves the problem in (2). From above condition $\mathbf{h}(\boldsymbol{\lambda})$ has a finite upper bound and $\{\boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ solves the problem in (3). Consequently, the set $\{\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*\}$ is a primal-dual solution.

Primal-Dual Solutions and Central Path: Primal-Dual Solutions

Theorem 6.10: Strict feasibility of primal-dual solutions

If the primal and dual problems are both feasible, then

1. solutions of the primal problem are bounded if the dual is strictly feasible:
2. solutions of the dual problem are bounded if the primal is strictly feasible:
3. primal-dual solutions are bounded if the primal and dual are both strictly feasible.

Proof: see reference 5.

Duality gap From (4), we observe that

$$\mathbf{c}^T\mathbf{x}^* = [(\boldsymbol{\mu}^*)^T - (\boldsymbol{\lambda}^*)^T\mathbf{A}]\mathbf{x}^* = -(\boldsymbol{\lambda}^*)^T\mathbf{A}\mathbf{x}^* = -(\boldsymbol{\lambda}^*)^T\mathbf{b} \quad \Rightarrow \quad f(\mathbf{x}^*) = \mathbf{h}(\boldsymbol{\lambda}^*)$$

If we define the **duality gap** as

$$\delta(\mathbf{x}, \boldsymbol{\lambda}) = \mathbf{c}^T\mathbf{x} + \mathbf{b}^T\boldsymbol{\lambda}$$

Then the above equations imply that $\delta(\mathbf{x}, \boldsymbol{\lambda})$ is always nonnegative with $\delta(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$. For any feasible x and $\boldsymbol{\lambda}$, we have

$$\mathbf{c}^T\mathbf{x} \geq \mathbf{c}^T\mathbf{x}^* \geq -\mathbf{b}^T\boldsymbol{\lambda}^* \geq -\mathbf{b}^T\boldsymbol{\lambda}$$
$$\mathbf{c}^T\mathbf{x} - \mathbf{c}^T\mathbf{x}^* \geq 0 \geq -\mathbf{b}^T\boldsymbol{\lambda}^* - \mathbf{c}^T\mathbf{x}^* \geq -\mathbf{b}^T\boldsymbol{\lambda} - \mathbf{c}^T\mathbf{x}^* \Longrightarrow \quad 0 \leq \mathbf{c}^T\mathbf{x} - \mathbf{c}^T\mathbf{x}^* \leq \delta(\mathbf{x}, \boldsymbol{\lambda})$$

It indicates that the duality gap can serve as a bound on the closeness of $f(\mathbf{x})$ to $f(\mathbf{x}^*)$.

Primal-Dual Solutions and Central Path: Central Path

One of the important concept related to the primal-dual solutions is central path. By using (4), set $\{\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}\}$ with $\mathbf{x} \in \mathbb{R}^n$, $\boldsymbol{\lambda} \in \mathbb{R}^p$, and $\boldsymbol{\mu} \in \mathbb{R}^n$ is a primal-dual solution if it satisfies the conditions

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \quad \text{with } \mathbf{x} \geq 0 \\ -\mathbf{A}^T\boldsymbol{\lambda} + \boldsymbol{\mu} &= \mathbf{c} \quad \text{with } \boldsymbol{\mu} \geq 0 \\ \mathbf{X}\boldsymbol{\mu} &= 0 \end{aligned} \tag{5}$$

where $\mathbf{X} = \text{diag}\{x_1, x_2, \ldots, x_n\}$ The centeral path for a standard form LP problem is defined as a set of vectors $\{\mathbf{x}(\tau), \boldsymbol{\lambda}(\tau), \boldsymbol{\mu}(\tau)\}$ that satisfy the conditions

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b} \quad \text{with } \mathbf{x} > 0 \\ -\mathbf{A}^T\boldsymbol{\lambda} + \boldsymbol{\mu} &= \mathbf{c} \quad \text{with } \boldsymbol{\mu} > 0 \\ \mathbf{X}\boldsymbol{\mu} &= \tau\mathbf{e} \end{aligned} \tag{6}$$

where τ is a strictly positive scalar parameter, and $\mathbf{e} = [1\ 1\ \cdots\ 1]^T$

Primal-Dual Solutions and Central Path

- For each fixed $\tau > 0$, the vectors in the set $\{\mathbf{x}(\tau), \boldsymbol{\lambda}(\tau), \boldsymbol{\mu}(\tau)\}$ satisfying (6) can be viewed as sets of points in $\mathbb{R}^n$, $\mathbb{R}^p$, and $\mathbb{R}^n$, respectively.
- When τ varies, the corresponding points form a set of trajectories called the **central path**.
- By comparing (6) with (4), it is obvious that the centeral path is closely related to the primal-dual solutions. Every point on the central path is strictly feasible.
- The central path lies in the interior of the feasible regions of the problems in (2) and (3) and it approaches a primal-dual solution as $\tau \to 0$.
- Given $\tau > 0$, let $\{\mathbf{x}(\tau), \boldsymbol{\lambda}(\tau), \boldsymbol{\mu}(\tau)\}$ be on the central path. From (6), the duality gap $\delta[\mathbf{x}(\tau, \boldsymbol{\lambda}(\tau)]$ is given by

$$\begin{aligned}\delta[\mathbf{x}(\tau), \boldsymbol{\lambda}(\tau)] &= \mathbf{c}^T\mathbf{x}(\tau) + \mathbf{b}^T\boldsymbol{\lambda}(\tau) = [-\boldsymbol{\lambda}^T(\tau)\mathbf{A} + \boldsymbol{\mu}^T(\tau)]\mathbf{x}(\tau) + \mathbf{b}^T\boldsymbol{\lambda}(\tau) \\ &= \boldsymbol{\mu}^T(\tau)\mathbf{x}(\tau) = n\tau\end{aligned}$$

The central path converges linearly to zero a $\tau \to 0$. The objective function $\mathbf{c}^T\mathbf{x}(\tau)$, and $\mathbf{b}^T\boldsymbol{\lambda}(\tau)$ approach the same optimal value.

Primal-Dual Solutions and Central Path

Sketch the central path of the LP problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = -2x_1 + x_2 - 3x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{aligned}$$

Solution: With $\mathbf{c} = [-2\ 1\ \ -3]^T$, $\mathbf{A} = [1\ 1\ 1]$, and $\mathbf{b} = 1$, (6) become

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ -\lambda + \mu_1 &= -2 \\ -\lambda + \mu_2 &= 1 \\ -\lambda + \mu_3 &= -3 \\ x_1\mu_1 = \tau,\ x_2\mu_2 &= \tau,\ x_3\mu_3 = \tau \end{aligned}$$

where $x_i > 0$ and $\mu_i > 0$ for $i = 1, 2, 3$.

Primal-Dual Solutions and Central Path

From above equations, we have

$$\mu_1 = -2 + \lambda \quad \mu_2 = 1 + \lambda \quad \mu_3 = -3 + \lambda$$

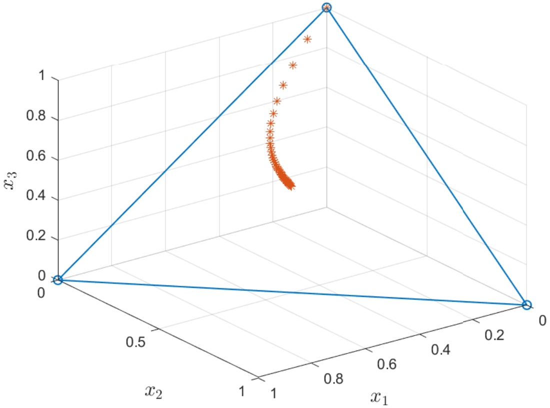
Hence $\mu_i > 0$ for $1 \leq i \leq 3$ if $\lambda > 3$. If we assume that $\lambda > 3$, then

$$\frac{1}{\lambda - 2} + \frac{1}{\lambda + 1} + \frac{1}{\lambda - 3} = \frac{1}{\tau}$$

i.e.,

$$\frac{1}{\tau}\lambda^3 - \left(\frac{4}{\tau} + 3\right)\lambda^2 + \left(\frac{1}{\tau} + 8\right)\lambda + \left(\frac{6}{\tau} - 1\right) = 0$$

The central path can be constructed by finding a root of above equation that satisfies $\lambda > 3$.


1
0.8
0.6
0.4
0.2
0
x_3
0
0.5
1
x_2
0
0.2
0.4
0.6
0.8
1
x_1

Primal-Dual Interior Methods: Path-Following Method

We need to find τ_k to make $\{\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k\}$ approach the minimizer vertex. In this lecture, we introduce the method that simultaneously solves the primal and dual LP problems and has emerged as *the modest efficient interior-point method for the LP problems.*

- Consider the standard form LP problem in (2) and its dual (3). Let $\mathbf{w}_k = \{\mathbf{x}_k, \boldsymbol{\lambda}_k, \boldsymbol{\mu}_k\}$ where $\mathbf{x}_k$ is strictly feasible for the primal and $\{\boldsymbol{\lambda}_k, \boldsymbol{\mu}_k\}$ is strictly feasible for the dual.
- We need to find the increment vector $\boldsymbol{\delta}_w = \{\boldsymbol{\delta}_x, \boldsymbol{\delta}_\lambda, \boldsymbol{\delta}_\mu\}$ such that the next iterate $\mathbf{w}_{k+1} = \{\mathbf{x}_{k+1}, \boldsymbol{\lambda}_{k+1}, \boldsymbol{\mu}_{k+1}\} = \{\mathbf{x}_k + \boldsymbol{\delta}_x, \boldsymbol{\lambda}_k + \boldsymbol{\delta}_\lambda, \boldsymbol{\mu}_k + \boldsymbol{\delta}_\mu\}$ remains strictly feasible and approaches the central path defined by (3) with $\tau = \tau_{k+1} > 0$.
- The path-following method, a suitable $\boldsymbol{\delta}_w$ is obtained as a first-order approximate solution of (6).

Primal-Dual Interior Methods: Path-Following Method

If $\mathbf{w}_{k+1}$ satisfies (6) with $\tau = \tau_{k+1}$, then

$$\begin{aligned} \mathbf{A}(\mathbf{x}_k + \boldsymbol{\delta}_x) &= \mathbf{b} \\ -\mathbf{A}^T(\boldsymbol{\lambda}_k + \boldsymbol{\delta}_\lambda) + (\boldsymbol{\mu}_k + \boldsymbol{\delta}_\mu) &= \mathbf{c} \\ \tilde{\mathbf{X}}(\boldsymbol{\mu}_k + \boldsymbol{\delta}_\mu) &= \tau_{k+1}\mathbf{e} \end{aligned}$$

$$\begin{aligned} \mathbf{A}\boldsymbol{\delta}_x &= 0 \\ -\mathbf{A}^T\boldsymbol{\delta}_\lambda + \boldsymbol{\delta}_\mu &= 0 \\ \Delta\mathbf{X}\boldsymbol{\mu}_k + X\boldsymbol{\delta}_\mu + \Delta\mathbf{X}\boldsymbol{\delta}_\mu &= \tau_{k+1}\mathbf{e} - \mathbf{X}\boldsymbol{\mu}_k \end{aligned}$$

where

$$\mathbf{X} = \mathrm{diag}\{x_1, x_2, \ldots, x_n\}, \tilde{\mathbf{X}} = \mathrm{diag}\{x_1 + (\boldsymbol{\delta}_x)_1, x_2 + (\boldsymbol{\delta}_x)_2, \ldots, x_n + (\boldsymbol{\delta}_x)_n\},$$
$$\Delta\mathbf{X} = \mathrm{diag}\{(\boldsymbol{\delta}_x)_1, (\boldsymbol{\delta}_x)_2, \ldots, (\boldsymbol{\delta}_x)_n\}, \quad \tilde{\mathbf{X}} = \mathbf{X} + \Delta\mathbf{X}$$

By neglecting the $\Delta\mathbf{X}\boldsymbol{\delta}_\mu$ and let $\Delta\mathbf{X}\boldsymbol{\mu}_k = \mathbf{M}\boldsymbol{\delta}_x$ (we need to find $\boldsymbol{\delta}_x$) , where $\mathbf{M} = \mathrm{diag}\{(\boldsymbol{\mu}_k)_1, (\boldsymbol{\mu}_k)_2, \ldots, (\boldsymbol{\mu}_k)_n\}$, we have

$$\mathbf{A}\boldsymbol{\delta}_x = 0, \quad -\mathbf{A}^T\boldsymbol{\delta}_\lambda + \boldsymbol{\delta}_\mu = 0, \quad \mathbf{M}\boldsymbol{\delta}_x + \mathbf{X}\boldsymbol{\delta}_\mu = \tau_{k+1}\mathbf{e} - \mathbf{X}\boldsymbol{\mu}_k \tag{7}$$

Primal-Dual Interior Methods: Path-Following Method

Solving (**??**) for $\boldsymbol{\delta}_w$, we obtain

$$
\begin{aligned}
\boldsymbol{\delta}_\lambda &= -\mathbf{YAy} \\
\boldsymbol{\delta}_\mu &= \mathbf{A}^T \boldsymbol{\delta}_\lambda \\
\boldsymbol{\delta}_x &= -\mathbf{y} - \mathbf{D}\boldsymbol{\delta}_\mu \\
\mathbf{D} = \mathbf{M}^{-1}\mathbf{X}, &\quad \mathbf{Y} = (\mathbf{ADA}^T)^{-1}, \quad \mathbf{y} = \mathbf{x}_k - \tau_{k+1}\mathbf{M}^{-1}\mathbf{e}
\end{aligned} \tag{8}
$$

where

$$
\begin{aligned}
\mathbf{M}^{-1}\mathbf{X}\boldsymbol{\mu}_k &= \begin{bmatrix} 1/(\boldsymbol{\mu}_k)_1 & & \\ & \ddots & \\ & & 1/(\boldsymbol{\mu}_k)_n \end{bmatrix} \begin{bmatrix} (\mathbf{x}_k)_1 & & \\ & \ddots & \\ & & (\mathbf{x}_k)_n \end{bmatrix} \begin{bmatrix} (\boldsymbol{\mu}_k)_1 \\ \vdots \\ (\boldsymbol{\mu}_k)_n \end{bmatrix} \\
&= \mathbf{x}_k
\end{aligned}
$$

Primal-Dual Interior Methods : Path-Following Method

Primal-dual path-following algorithm for the LP problem

1. Input $\mathbf{A}$ and a strictly feasible $\mathbf{w}_0 = \{\mathbf{x}_0, \boldsymbol{\lambda}_0, \boldsymbol{\mu}_0\}$. Set $k = 0$ and $\rho > \sqrt{n}$ (n is the dimension of $\mathbf{x}$), and initialize the tolerance ε for the duality gap.
2. If $\boldsymbol{\mu}_k^T \mathbf{x}_k \leq \varepsilon$, output solution $\mathbf{w}^* = \mathbf{w}_k$ and stop; otherwise, continue with Step 3
3. Set $\tau_{k+1} = \frac{\boldsymbol{\mu}_k^T \mathbf{x}_k}{n+\rho}$ and compute $\boldsymbol{\delta}_w = \{\boldsymbol{\delta}_x, \boldsymbol{\delta}_\lambda, \boldsymbol{\delta}_\mu\}$ using (**??**).
4. compute step size α_k as follow:

$$\alpha_k = (1 - 10^{-6})\alpha_{\max}, \quad \alpha_{\max} = \min(\alpha_p, \alpha_d)$$

where

$$\alpha_p = \min_{i \text{ with } (\boldsymbol{\delta}_x)_i < 0} \left[-\frac{(\mathbf{x}_k)_i}{(\boldsymbol{\delta}_x)_i} \right], \quad \alpha_d = \min_{i \text{ with } (\boldsymbol{\delta}_\mu)_i < 0} \left[-\frac{(\boldsymbol{\mu}_k)_i}{(\boldsymbol{\delta}_\mu)_i} \right]$$

Sketch the central path of the LP problem

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = -2x_1 + x_2 - 3x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{aligned}$$

Solution: Find an initial $\mathbf{w}_0$ on the central path by using the method described in the previous example with $\tau_0 = 5$. The vector $\mathbf{w}_0$ obtained is $\{\mathbf{x}_0, \boldsymbol{\lambda}_0, \boldsymbol{\mu}_0\}$ with

$$\mathbf{x}_0 = \begin{bmatrix} 0.344506 \\ 0.285494 \\ 0.370000 \end{bmatrix}, \quad \boldsymbol{\lambda}_0 = 16.513519, \quad \boldsymbol{\mu}_0 = \begin{bmatrix} 14.513519 \\ 17.513519 \\ 13.513519 \end{bmatrix}$$

With $\rho = 7\sqrt{n}$ and $\varepsilon = 10^{-6}$, the algorithm will converges after eight iterations to the solution $\mathbf{x}^* = \begin{bmatrix} 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$

A nonfeasible-Initialization: Path-Following Method

If $\mathbf{w}_{k+1}$ satisfies (6) with $\tau = \tau_{k+1}$, then

$$\begin{aligned} \mathbf{A}(\mathbf{x}_k + \boldsymbol{\delta}_x) &= \mathbf{b} \\ -\mathbf{A}^T(\boldsymbol{\lambda}_k + \boldsymbol{\delta}_\lambda) + (\boldsymbol{\mu}_k + \boldsymbol{\delta}_\mu) &= \mathbf{c} \\ \tilde{\mathbf{X}}(\boldsymbol{\mu}_k + \boldsymbol{\delta}_\mu) &= \tau_{k+1} e \end{aligned}$$

$$\begin{aligned} \mathbf{A}\boldsymbol{\delta}_x &= \mathbf{r}_p \\ -\mathbf{A}^T\boldsymbol{\delta}_\lambda + \boldsymbol{\delta}_\mu &= \mathbf{r}_d \\ \mathbf{M}\boldsymbol{\delta}_x + \mathbf{X}\boldsymbol{\delta}_\mu &= \tau_{k+1}\mathbf{e} - \mathbf{X}\boldsymbol{\mu}_k \end{aligned}$$

where $\mathbf{r}_p = \mathbf{b} - \mathbf{A}\mathbf{x}_k$ and $\mathbf{r}_d = \mathbf{c} + \mathbf{A}^T\boldsymbol{\lambda}_k - \boldsymbol{\mu}_k$ are the residuals for the primal and dual constraints, respectively.

$$\begin{aligned} \boldsymbol{\delta}_\lambda &= -\mathbf{Y}(\mathbf{A}\mathbf{y} + \mathbf{A}\mathbf{D}\mathbf{r}_d + \mathbf{r}_p) \\ \boldsymbol{\delta}_\mu &= \mathbf{A}^T\boldsymbol{\delta}_\lambda + r_d \\ \boldsymbol{\delta}_x &= -\mathbf{y} - \mathbf{D}\boldsymbol{\delta}_\mu \end{aligned} \tag{9}$$

Nonfeasible-Initialization Primal-Dual Path-Following Method

Nonfeasible-initialization Primal-dual path-following algorithm for the LP problem

1. Input $\mathbf{A}, \mathbf{b}, \mathbf{c}$, and $w_0 = \{x_0, \lambda_0, \mu_0\}$. Set $k = 0$ and $\rho > \sqrt{n}$ (n is a dimension of x), and initialize the tolerance ε for the duality gap.
2. If $\boldsymbol{\mu}_k^T \mathbf{x}_k \leq \varepsilon$, output solution $\mathbf{w}^* = \mathbf{w}_k$ and stop; otherwise, continue with Step 3
3. Set $\tau_{k+1} = \frac{\boldsymbol{\mu}_k^T \mathbf{x}_k}{n+\rho}$ and compute $\boldsymbol{\delta}_w = \{\boldsymbol{\delta}_x, \boldsymbol{\delta}_\lambda, \boldsymbol{\delta}_\mu\}$ using (7).
4. compute step size α_k as follow:

$$\alpha_k = (1 - 10^{-6})\alpha_{\max} \quad \alpha_{\max} = \min(\alpha_p, \alpha_d)$$

where

$$\alpha_p = \min_{i \text{ with } (\boldsymbol{\delta}_x)_i < 0} \left[-\frac{(\mathbf{x}_k)_i}{(\boldsymbol{\delta}_x)_i} \right], \quad \alpha_d = \min_{i \text{ with } (\boldsymbol{\delta}_\mu)_i < 0} \left[-\frac{(\boldsymbol{\mu}_k)_i}{(\boldsymbol{\delta}_\mu)_i} \right]$$

Nonfeasible-Initialization Primal-Dual Path-Following Method Example

Sketch the central path of the LP problem

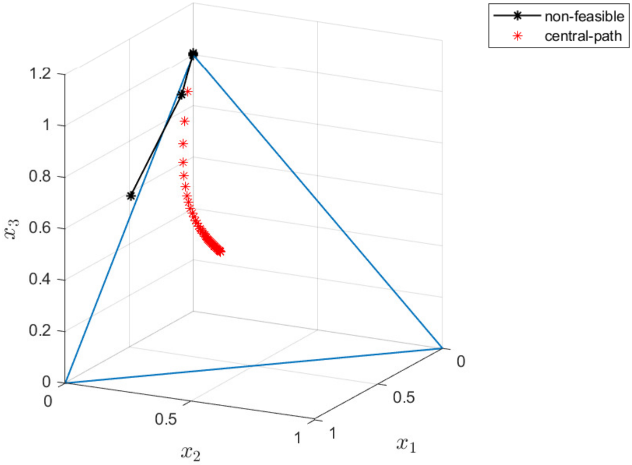
$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) = -2x_1 + x_2 - 3x_3 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0 \end{aligned}$$

Solution: The vector $\mathbf{w}_0$, which is not feasible, is $\{\mathbf{x}_0, \boldsymbol{\lambda}_0, \boldsymbol{\mu}_0\}$ with

$$\mathbf{x}_0 = \begin{bmatrix} 0.4 \\ 0.3 \\ 0.4 \end{bmatrix}, \quad \boldsymbol{\lambda}_0 = -0.5, \quad \boldsymbol{\mu}_0 = \begin{bmatrix} 1.0 \\ 0.5 \\ 1.0 \end{bmatrix}$$

With $\rho = 7\sqrt{n}$ and $\varepsilon = 10^{-6}$, the algorithm will converges after eight iterations to the solution $\mathbf{x}^* = \begin{bmatrix} 0.000000 & 0.000000 & 1.000000 \end{bmatrix}$

Nonfeasible-Initialization Primal-Dual Path-Following Method Example



Reference

1. Joaquim R. R. A. Martins, Andrew Ning, ”*Engineering Design Optimization*,” Cambridge University Press, 2021.
2. Mykel J. kochenderfer, and Tim A. Wheeler, ”*Algorithms for Optimization*,” The MIT Press, 2019.
3. Ashok D. Belegundu, Tirupathi R. Chandrupatla, ”*Optimization Concepts and Applications in Engineering*,” Cambridge University Press, 2019.
4. Laurent Lessarn , ”*Introduction to Optimization*,” Lecture Note, University of Wisconsin–Madison.
5. Andreas Antoniou and Wu-Sheng Lu, ”*Practical Optimization: Algorithms and Engineering Applications*,” 2^{nd} edition, Springer, 2021.
6. Stephen Boyd and Lieven Vandenberghe ”*Convex Optimization*,” Cambridge University Press, 2004.