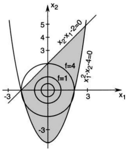
Chapter 22 Convex Optimization Problems

Constrained Optimization Problem

★ Because of the nonlinearity of objective function and constraint, optimization problems are usually very difficult to solve. Even if the objective function is simple, the properties of constraint may make the problem challenging.

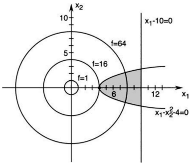
Example $(\min\ x_1^2 + x_2^2, \quad \text{s.t. } x_2 - x_1 - 2 \leq 0, \ \ x_1^2 - x_2 - 4 \leq 0.)$

Observe that all the constraints are inactive at the minimizer. Hence, we can solve this problem as an unconstrained optimization.



Example $(\min\ x_1^2 + x_2^2, \quad \text{s.t. } x_1 - 10 \leq 0, \ \ x_1 - x_2^2 - 4 \geq 0.)$

At the minimizer, only one constraint is active. If we had known about this we could have handled this problem as an equality constrained optimization.

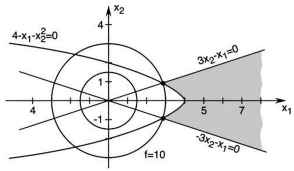


Constrained Optimization Problem

Example

$$\boxed{\begin{aligned} \min \quad x_1^2 + x_2^2, \quad \text{s.t.} \quad & 4 - x_1 - x_2^2 \le 0, \\ & 3x_2 - x_1 \le 0, \\ & -3x_2 - x_1 \le 0. \end{aligned}}$$

The optimization problem has one optimal value, but there are two local minimizers.



★ If the constraints of optimization problem are convex sets, then the difficulties in the solution process of the above three examples can be eliminated. The key to solving optimization problems is "whether it is a convex optimization".

Convex Function

Definition (graph and epigraph of function)

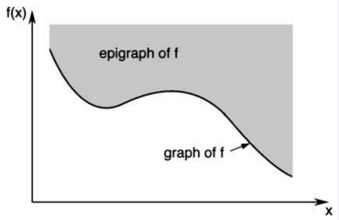
Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ be a function.

- The graph of f is the set of points in $\Omega \times \mathbb{R}$

$$\operatorname{gph}(f) = \left\{ \begin{bmatrix} x \\ \beta \end{bmatrix} \subset \mathbb{R}^{n+1} \;\middle|\; \beta = f(x),\ x \in \Omega \right\}.$$

- The epigraph of f is the set of points in $\Omega \times \mathbb{R}$

$$\operatorname{epi}(f) = \left\{ \begin{bmatrix} x \\ \beta \end{bmatrix} \subset \mathbb{R}^{n+1} \;\middle|\; \beta \geq f(\boldsymbol{x}),\ x \in \Omega \right\}.$$



Convex Function

Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ be a function.

Definition (convex function)

f is a convex function on $\Omega \iff \mathrm{epi}(f)$ is a convex set in $\mathbb{R}^{n+1}$.

Theorem

If $f : \Omega \to \mathbb{R}$ is a convex function, then Ω is a convex set.

proof. (By contradiction) Suppose that Ω is a nonconvex set.

Then, $\exists\ \boldsymbol{y}_1, \boldsymbol{y}_2 \in \Omega$ and $\alpha \in (0,1)$ such that $\boldsymbol{z} = \alpha \boldsymbol{y}_1 + (1-\alpha)\boldsymbol{y}_2 \notin \Omega$.

Let $\beta_1 = f(\boldsymbol{y}_1)$ and $\beta_2 = f(\boldsymbol{y}_2)$.

Then, $\begin{bmatrix} \boldsymbol{y}_1 \\ \beta_1 \end{bmatrix}$ and $\begin{bmatrix} \boldsymbol{y}_2 \\ \beta_2 \end{bmatrix}$ belong to $\mathrm{epi}(f)$.

Let $\boldsymbol{w} = \alpha \begin{bmatrix} \boldsymbol{y}_1 \\ \beta_1 \end{bmatrix} + (1-\alpha) \begin{bmatrix} \boldsymbol{y}_2 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{z} \\ \alpha\beta_1 + (1-\alpha)\beta_2 \end{bmatrix}$.

$\because\ \boldsymbol{z} \notin \Omega, \quad \therefore\ \boldsymbol{w} \notin \mathrm{epi}(f), \quad \therefore\ \mathrm{epi}(f)$ is not convex.

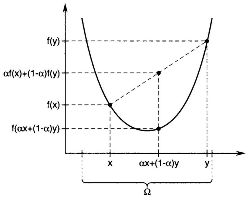
Convex Function

Definition (convex function)

Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}$ be a convex function
$\iff f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}), \ \forall \boldsymbol{x}, \boldsymbol{y} \in \Omega, \ \alpha \in (0,1).$

geometric interpretation of convex function

The segments connecting any two points on the graph of a convex function belong to its epigraph, i.e., epi(f) is convex set.



- f is strictly convex $\iff f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) < \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y})$.
- f is strongly convex $\iff$
  $f(\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) - \frac{\alpha(1-\alpha)c}{2} \|\boldsymbol{x} - \boldsymbol{y}\|^2$.

Convex Function

Example

$f(\boldsymbol{x}) = x_1 x_2$. Is f convex over the set $\Omega = \{\boldsymbol{x} \mid x_1 \geq 0, x_2 \geq 0\}$?

Ans: For $\boldsymbol{x} = [1, 2]^\top \in \Omega$ and $\boldsymbol{y} = [2, 1]^\top \in \Omega$, let $\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} = \begin{bmatrix} 2-\alpha \\ 1+\alpha \end{bmatrix}$.

$$\therefore \begin{cases} f(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}) = (2-\alpha)(1+\alpha) = 2 + \alpha - \alpha^2, \\ \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}) = 2. \end{cases}$$

By taking $\alpha = \frac{1}{2} \in (0, 1)$, then $f\left(\frac{1}{2}\boldsymbol{x} + \frac{1}{2}\boldsymbol{y}\right) = \frac{9}{4} > \frac{1}{2} f(\boldsymbol{x}) + \frac{1}{2} f(\boldsymbol{y})$,

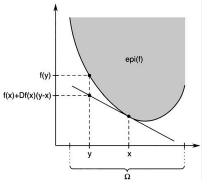
$\therefore$ f is not convex over Ω.

Theorem (differentiable convex function)

Let $\Omega \subset \mathbb{R}^n$ be an open convex set and $f : \Omega \to \mathbb{R}$ belongs to $\mathcal{C}^1$. Then, f is convex on Ω $\iff$

$$\boxed{f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}), \quad \forall \boldsymbol{x}, \boldsymbol{y} \in \Omega.}$$

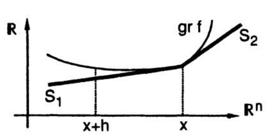
Namely, the graph of f always lies above its linear approximation at any point.



Differentiable Convex Function

Definition (generalization of the gradient)

Let $\Omega \subset \mathbb{R}^n$ be an open convex set and $f : \Omega \to \mathbb{R}$. A vector $\boldsymbol{g} \in \mathbb{R}^n$ is said to be a subgradient of f at $\boldsymbol{x} \in \Omega$ if
$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{g}^\top(\boldsymbol{y} - \boldsymbol{x}), \ \ \forall \boldsymbol{y} \in \Omega.$



Theorem

Let $\Omega \subset \mathbb{R}^n$ be an open convex set and $f : \Omega \to \mathbb{R}$ and ket $f \in \mathcal{C}^2$. Then, f is convex on $\Omega \iff \boldsymbol{F}(\boldsymbol{x}) \succeq 0, \ \forall \boldsymbol{x} \in \Omega$. [hereafter, $\boldsymbol{F}(\boldsymbol{x})$ is the Hessian of f at $\boldsymbol{x}$]

proof. $\Leftarrow$) Let $\boldsymbol{x}, \boldsymbol{y} \in \Omega$.

$\because \ f \in \mathcal{C}^2$, it follows Taylor's theorem that: $\exists \ \alpha \in (0,1)$ such that
$f(\boldsymbol{y}) = f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) + \frac{1}{2}(\boldsymbol{y} - \boldsymbol{x})^\top \boldsymbol{F}(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x}))(\boldsymbol{y} - \boldsymbol{x})$.
$\because \ \boldsymbol{F}(\boldsymbol{x} + \alpha(\boldsymbol{y} - \boldsymbol{x})) \succeq 0 \iff (\boldsymbol{y} - \boldsymbol{x})^\top \boldsymbol{F}(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) \geq 0$,
$\therefore \ f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x})$. $\quad \therefore \ f$ is convex.

Differentiable Convex Function

$\Rightarrow$) If $\exists\, \boldsymbol{x} \in \Omega$ such that $\boldsymbol{F}(\boldsymbol{x}) \nsucceq 0 \iff \exists\, \boldsymbol{d} \in \mathbb{R}^n$ such that $\boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x})\boldsymbol{d} < 0$.

$$\xRightarrow{\text{continuity of Hessian}} \boldsymbol{F}(\boldsymbol{x}+s\boldsymbol{d}) \nsucceq 0.$$

Let $\boldsymbol{y} = \boldsymbol{x} + s\boldsymbol{d}$. By Taylor's theorem, $\exists\, \alpha \in (0,1)$ such that

$$\begin{aligned} f(\boldsymbol{y}) &= f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x}) + \frac{1}{2}(\boldsymbol{y}-\boldsymbol{x})^\top \boldsymbol{F}(\boldsymbol{x}+\alpha(\boldsymbol{y}-\boldsymbol{x}))(\boldsymbol{y}-\boldsymbol{x}) \\ &= f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x}) + \frac{1}{2}s^2\boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}+\alpha s\boldsymbol{d})\boldsymbol{d}. \end{aligned}$$

$\because\ \boldsymbol{d}^\top \boldsymbol{F}(\boldsymbol{x}+\alpha s\boldsymbol{d})\boldsymbol{d} < 0 \implies f(\boldsymbol{y}) < f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y}-\boldsymbol{x})$. $\therefore\ f$ is not convex.

$f(\boldsymbol{x}) = 4x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + x_1x_3 - 3x_1 - 2x_2 + 15$

$\because$ The Hessian of f is $\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{bmatrix}$, $\therefore\ \forall\, \boldsymbol{x} \in \mathbb{R}^3$, $\boldsymbol{F}(\boldsymbol{x}) \succ 0$, $\therefore\ f$ is convex on $\mathbb{R}^3$.

$f(\boldsymbol{x}) = 2x_1x_2 - x_1^2 - x_2^2$

$\because$ The Hessian of f is $\boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$, $\therefore\ \forall\, \boldsymbol{x} \in \mathbb{R}^2$, $\boldsymbol{F}(\boldsymbol{x}) \preceq 0$, $\therefore\ f$ is concave on $\mathbb{R}^2$.

Operations that Preserve Convexity

Table 2.5.1. Main operations yielding convexity

Operations on functions: $f =$	Operations on sets: $\operatorname{epi} f$ or $\operatorname{epi}_s f =$	Closedness
$\sum_{j=1}^{m} t_j f_j$	nothing interesting	preserved
$\sup_{j \in J} f_j$	$\bigcap_{j \in J} \operatorname{epi} f_j$	preserved
$g \circ A$ (A affine)	$\overset{-1}{A'}(\operatorname{epi} g)$	preserved
$ug(x/u)$	$\mathbb{R}_*^+(\{1\} \times \operatorname{epi} g)$	must be forced
$f_1 \overset{+}{\vee} f_2$	$\operatorname{epi}_s f_1 + \operatorname{epi}_s f_2$	destroyed
$f_1 \bar{\vee} f_2$	$\operatorname{epi}_s f_1 \underline{*} \operatorname{epi}_s f_2$	preserved from f_1
Ag (A linear)	epigr. hull of $A'(\operatorname{epi} g)$	destroyed
$\inf_y g(\cdot, y)$	$\operatorname{Proj}_{\mathbb{R}^n \times \mathbb{R}} \operatorname{epi}_s g$	destroyed
$\operatorname{co} g$	epigr. hull of $\operatorname{co} \operatorname{epi} g$	can be forced

Theorem (local vs global minimizer)

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function. Then, a point is a global minimizer of f over Ω $\iff$ it is a local minimizer of f.

proof. $\Rightarrow$) Obviously.

$\Leftarrow$) Suppose that $\boldsymbol{x}^*$ is not a global minimizer of f over Ω.

Then, $\exists\ \boldsymbol{y} \in \Omega$, such that $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$.

$\because$ f is convex,

$\therefore$ $f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) \leq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}^*)$, $\forall\alpha \in (0,1)$.

$\because$ $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$,

$\therefore$ $\alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}^*) = \alpha (f(\boldsymbol{y}) - f(\boldsymbol{x}^*)) + f(\boldsymbol{x}^*) < f(\boldsymbol{x}^*)$.

$\therefore$ $\forall\alpha \in (0,1)$, $f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) < f(\boldsymbol{x}^*)$.

$\therefore$ $\boldsymbol{x}^*$ is not a local minimizer.

Lemma (sublevel set of convex function)

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function. Then, the sublevel set of f w.r.t. a constant c, defined by $S_f(c) = \{\boldsymbol{x} \in \Omega \mid f(\boldsymbol{x}) \leq c, \ \forall c \in \mathbb{R}\}$, is a convex set.

★ The above lemma is a necessary but not sufficient condition (why?).

Corollary

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function. Then, the set of all global minimizers of f is a convex set.

Lemma (global minimum condition)

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function and $f \in \mathcal{C}^1$. For any $\boldsymbol{x}^* \in \Omega$, if

$$\boxed{Df(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) \geq 0, \quad \forall \boldsymbol{x} \in \Omega, \ \boldsymbol{x} \neq \boldsymbol{x}^*,}$$

then $\boldsymbol{x}^*$ is a global minimizer of f over Ω.

Convex Optimization Problems

Theorem

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function and $f \in \mathcal{C}^1$. For an $\boldsymbol{x}^* \in \Omega$, if

$\boxed{\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0 \text{ for any feasible direction } \boldsymbol{d} \text{ at } \boldsymbol{x}^*,}$

then $\boldsymbol{x}^*$ is a global minimizer of f over Ω.

Corollary

$f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function and $f \in \mathcal{C}^1$. For any $\boldsymbol{x}^* \in int(\Omega)$, if

$\boxed{\nabla f(\boldsymbol{x}^*) = \boldsymbol{0},}$ *then $\boldsymbol{x}^*$ is a global minimizer of f over Ω.*

Corollary (equality constrained optimization)

Let $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}$, where $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{h} \in \mathcal{C}^1$, and Ω is convex.
Let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}$ is a convex function and $f \in \mathcal{C}^1$. If

$$\boxed{\exists\, \boldsymbol{x}^* \in \Omega,\ \boldsymbol{\lambda}^* \in \mathbb{R}^m \text{ such that } Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) = \boldsymbol{0}^\top, \qquad (\star)}$$

then $\boldsymbol{x}^*$ is a global minimizer of f over Ω.

★ This implies that the feasible set is convex, and the Lagrange condition is sufficient for a minimizer.

proof. $\because$ f is a convex function,

$\therefore$ $\forall \boldsymbol{x} \in \Omega, \quad f(\boldsymbol{x}) \geq f\left(\boldsymbol{x}^*\right) + Df\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x} - \boldsymbol{x}^*\right) \overset{(\star)}{=\!=} \left(\boldsymbol{x}^*\right) - \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x} - \boldsymbol{x}^*\right),$

$\because$ Ω is convex, $\therefore$ $\forall \alpha \in (0,1), \quad (1-\alpha)\boldsymbol{x}^* + \alpha\boldsymbol{x} \in \Omega.$

$\therefore$ $\boldsymbol{h}\left(\boldsymbol{x}^* + \alpha\left(\boldsymbol{x} - \boldsymbol{x}^*\right)\right) = \boldsymbol{h}\left((1-\alpha)\boldsymbol{x}^* + \alpha\boldsymbol{x}\right) = \mathbf{0}.$

Premultiplying by $\boldsymbol{\lambda}^{*\top}$, subtracting $\boldsymbol{\lambda}^{*\top}\boldsymbol{h}\left(\boldsymbol{x}^*\right) = 0$, and dividing by α, we have

$$\frac{\boldsymbol{\lambda}^{*\top}\boldsymbol{h}\left(\boldsymbol{x}^* + \alpha\left(\boldsymbol{x} - \boldsymbol{x}^*\right)\right) - \boldsymbol{\lambda}^{*\top}\boldsymbol{h}\left(\boldsymbol{x}^*\right)}{\alpha} = 0.$$

If we take the limit as $\alpha \to 0$, we get

$$\boldsymbol{\lambda}^{*\top} D\boldsymbol{h}\left(\boldsymbol{x}^*\right)\left(\boldsymbol{x} - \boldsymbol{x}^*\right) = 0.$$

$\therefore$ $f(\boldsymbol{x}) \geq f\left(\boldsymbol{x}^*\right).$

$\therefore$ $\boldsymbol{x}^*$ is a global minimizer of f over Ω.

Convex Optimization Problems

Theorem (inequality constrained optimization)

$f:\mathbb{R}^n \to \mathbb{R}$ and $f \in \mathcal{C}^1$ is a convex function on the set of feasible points

$$\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0},\ \boldsymbol{g}(\boldsymbol{x}) \le \mathbf{0}\},$$

where $\boldsymbol{h}:\mathbb{R}^n \to \mathbb{R}^m, \boldsymbol{g}:\mathbb{R}^n \to \mathbb{R}^p$ and $\boldsymbol{h}, \boldsymbol{g} \in \mathcal{C}^1$, and Ω is convex. If $\exists\ \boldsymbol{x}^* \in \Omega$, $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ and $\boldsymbol{\mu}^* \in \mathbb{R}^p$ such that

1. $\mu^* \ge 0$.
2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*\top} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^\top$.
3. $\boldsymbol{\mu}^{*\top}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.

Then, $\boldsymbol{x}^*$ is a global minimizer of f over Ω.

proof. Let $\boldsymbol{x} \in \Omega$. $\because$ f is convex $\implies$ $f(\boldsymbol{x}) \ge f(\boldsymbol{x}^*) + Df(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)$.

By condition 2 , we have

$$f(\boldsymbol{x}) \ge f(\boldsymbol{x}^*) - \boldsymbol{\lambda}^{*\top} D\boldsymbol{h}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) - \boldsymbol{\mu}^{*\top} D\boldsymbol{g}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*).$$

By Theorem 22.8, we have

$$\boldsymbol{\lambda}^{*\top} D\boldsymbol{h}\left(\boldsymbol{x}^{*}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right)=0.$$

We now claim that $\boldsymbol{\mu}^{*\top} D\boldsymbol{g}\left(\boldsymbol{x}^{*}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right) \leq 0$.
$\because$ Ω is convex, $\therefore$ $(1-\alpha)\boldsymbol{x}^{*}+\alpha\boldsymbol{x} \in \Omega, \quad \forall \alpha \in (0,1)$.
$\therefore$ $\boldsymbol{g}\left(\boldsymbol{x}^{*}+\alpha\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right)\right)=\boldsymbol{g}\left((1-\alpha)\boldsymbol{x}^{*}+\alpha\boldsymbol{x}\right) \leq \mathbf{0}$.
Premultiplying by $\boldsymbol{\mu}^{*\top} \geq \mathbf{0}^{\top}$ (by condition 1), subtracting $\boldsymbol{\mu}^{*\top}\boldsymbol{g}\left(\boldsymbol{x}^{*}\right)=0$ (by condition 3), and dividing by α, we have

$$\frac{\boldsymbol{\mu}^{*\top}\boldsymbol{g}\left(\boldsymbol{x}^{*}+\alpha\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right)\right)-\boldsymbol{\mu}^{*\top}\boldsymbol{g}\left(\boldsymbol{x}^{*}\right)}{\alpha} \leq 0.$$

By taking the limit as $\alpha \rightarrow 0$ to obtain $\boldsymbol{\mu}^{*\top} D\boldsymbol{g}\left(\boldsymbol{x}^{*}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right) \leq 0$.

$$\therefore\ f(\boldsymbol{x}) \geq f\left(\boldsymbol{x}^{*}\right)-\boldsymbol{\lambda}^{*\top} D\boldsymbol{h}\left(\boldsymbol{x}^{*}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right)-\boldsymbol{\mu}^{*\top} D\boldsymbol{g}\left(\boldsymbol{x}^{*}\right)\left(\boldsymbol{x}-\boldsymbol{x}^{*}\right) \geq f\left(\boldsymbol{x}^{*}\right).$$

Concave Optimization

Theorem

A concave function $f\mathbb{R}^n \to \mathbb{R}$, which attains its minimum over a convex set X at an $\boldsymbol{x}^* \in ri(X)$, must be constant over X.

proof. (By contradiction)

Let $\boldsymbol{x} \in X$ be such that $f(\boldsymbol{x}) > f(\boldsymbol{x}^*)$.

Prolong beyond $\boldsymbol{x}^*$ the line segment $\boldsymbol{x}$-to-$\boldsymbol{x}^*$ to a point $\boldsymbol{x} \in X$.

By concavity of f, we have for some $\alpha \in (0,1)$,

$$f(\boldsymbol{x}^*) \geq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{x}),$$

since $f(\boldsymbol{x}) > f(\boldsymbol{x}^*)$, we have $f(\boldsymbol{x}^*) > f(\boldsymbol{x})$ (contradiction).

Semi-definite Programming

Definition (semi-definite programming)

minimizing linear function subject to linear matrix inequalities

$$\min\ \boldsymbol{c}^\top \boldsymbol{x}, \quad \text{s.t. } \boldsymbol{F}(\boldsymbol{x}) \succeq 0,$$

where $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + x_1\boldsymbol{F}_1 + \cdots + x_n\boldsymbol{F}_n = \boldsymbol{F}_0 + \sum_{i=1}^{n} x_i\boldsymbol{F}_i$ is called an affine function of $\boldsymbol{x}$, and $\boldsymbol{F}_i = \boldsymbol{F}_i^\top \in \mathbb{R}^{m\times m}$ for $i = 0, 1, \ldots, n$.

Semi-definite Programming

★ The inequality constraint of the form (positive semidefinite):

$$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + x_1\boldsymbol{F}_1 + \cdots + x_n\boldsymbol{F}_n \succeq 0.$$

★ $\{\boldsymbol{x} \mid \boldsymbol{F}(\boldsymbol{x}) \succeq 0\}$ is convex.

Definition (systems of multiple linear matrix inequalities)

$$\boldsymbol{F}_1(\boldsymbol{x}) \succeq 0, \cdots, \boldsymbol{F}_k(\boldsymbol{x}) \succeq 0 \iff \boldsymbol{F}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{F}_1(\boldsymbol{x}) & & & \\ & \boldsymbol{F}_2(\boldsymbol{x}) & & \\ & & \ddots & \\ & & & \boldsymbol{F}_k(\boldsymbol{x}) \end{bmatrix} \succeq 0.$$

Definition (block matrix contract transformation)

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^\top & \boldsymbol{D} \end{bmatrix} \succeq 0 \iff \begin{bmatrix} \boldsymbol{D} & \boldsymbol{B}^\top \\ \boldsymbol{B} & \boldsymbol{A} \end{bmatrix} \succeq 0.$$

Definition (Schur complement)

A blocky matrix $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix}$, where $\boldsymbol{A}_{11}$ and $\boldsymbol{A}_{22}$ are square submatrices.

1. If $\boldsymbol{A}_{11}$ is invertible, then the Schur complement of $\boldsymbol{A}_{11}$ is $\boldsymbol{\Delta}_{11} = \boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12}$.
2. If $\boldsymbol{A}_{22}$ is invertible, then the Schur complement of $\boldsymbol{A}_{22}$ is $\boldsymbol{\Delta}_{22} = \boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21}$.

Theorem

$\boldsymbol{A} \succ 0 \iff \boldsymbol{A}_{11} \succ 0$ *and* $\boldsymbol{\Delta}_{11} \succ 0 \iff \boldsymbol{A}_{22} \succ 0$, *and* $\boldsymbol{\Delta}_{22} \succ 0$.

Example

Does $\boldsymbol{A} \in \mathbb{R}^{m\times m}$ have all its eigenvalues in the open left half-complex plane?

Ans: If this condition is true $\iff$ $\exists$ a symmetric positive definite matrix $\boldsymbol{P} \in \mathbb{R}^{m\times m}$ such that (Lyapunov inequality)

$$\boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} \prec 0 \iff \begin{bmatrix} \boldsymbol{P} & \boldsymbol{O} \\ \boldsymbol{O} & -\boldsymbol{A}^\top \boldsymbol{P} - \boldsymbol{P}\boldsymbol{A} \end{bmatrix} \succ 0.$$

$$\text{Let } \boldsymbol{P} = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ x_2 & x_{m+1} & \cdots & x_{2m-1} \\ \vdots & & & \vdots \\ x_m & x_{2m-1} & \cdots & x_n \end{bmatrix}, \quad n = \frac{m(m+1)}{2}.$$

Properties of Matrices

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, P_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \cdots, P_n = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

★ $\boldsymbol{P}_i$ has only nonzero elements corresponding to $\boldsymbol{x}_i$ in $\boldsymbol{P}$.

Let $\boldsymbol{F}_i = -\boldsymbol{A}^\top \boldsymbol{P}_i - \boldsymbol{P}_i \boldsymbol{A}, \quad i = 1, 2, \ldots, n.$

$$\begin{aligned} \Longrightarrow \ \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} =& x_1 \left(\boldsymbol{A}^\top \boldsymbol{P}_1 + \boldsymbol{P}_1 \boldsymbol{A}\right) + \cdots + x_n \left(\boldsymbol{A}^\top \boldsymbol{P}_n + \boldsymbol{P}_n \boldsymbol{A}\right) \\ =& - x_1 \boldsymbol{F}_1 - x_2 \boldsymbol{F}_2 - \cdots - x_n \boldsymbol{F}_n \prec 0. \end{aligned}$$

Let $\boldsymbol{F}(\boldsymbol{x}) = x_1 \boldsymbol{F}_1 + x_2 \boldsymbol{F}_2 + \cdots + x_n \boldsymbol{F}_n$.

Then, $\boldsymbol{P} = \boldsymbol{P}^\top \succ 0, \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} \prec 0 \iff \boldsymbol{F}(\boldsymbol{x}) \succ 0.$

Thus, the problem is equivalent to feasibility of the above matrix inequality.

Linear Matrix Inequality (LMI) Solvers

Definition (canonical representation of LMI)

$\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}_0 + x_1\boldsymbol{F}_1 + \cdots + x_n\boldsymbol{F}_n \succeq 0.$

Matlab Solver (I): Consider the form

$$\boldsymbol{N}^\top \mathcal{L}\left(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\right) \boldsymbol{N} \preceq \boldsymbol{M}^\top \boldsymbol{R}\left(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\right) \boldsymbol{M},$$

- $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$: matrix variables;
- $\boldsymbol{N}$: left outer factor; $\boldsymbol{M}$: right outer factor.
- $\mathcal{L}\left(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\right)$: left inner factor; $\boldsymbol{R}\left(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k\right)$: right one.

The matrices $\mathcal{L}(\cdot)$ and $\mathcal{R}(\cdot)$ are, in general, symmetric block matrices.

Example

$$\boldsymbol{A}_1 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} -2 & 0 \\ 1 & -1 \end{bmatrix}.$$

Find $\boldsymbol{P}$ such that $\boldsymbol{P} \succeq 0.5\boldsymbol{I}_2$, $\boldsymbol{A}_1^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A}_1 \preceq 0$ and $\boldsymbol{A}_2^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A}_2 \preceq 0$.

Linear Matrix Inequality (LMI) Solvers

Matlab Syntax:

```
A_1=[-1  0 ; 0  -1];
A_2=[-2  0 ; 1  -1];
setlmis([ ])
P=lmivar(1,[2,1])
lmiterm([1 1 1 P],A_1',1,'s')
lmiterm([2 1 1 P],A_2',1,'s')
lmiterm([3 1 1 0],.5)
lmiterm([-3 1 1 P],1,1)
lmis=getlmis;
[tmin, xfeas] = feasp(lmis);
P=dec2mat(lmis,xfeas,P)
```

Matlab Solver (II) for problem in the form: $\boxed{\min \ \ \boldsymbol{c}^\top \boldsymbol{x}, \quad \text{s. t. } \ \boldsymbol{A}(\boldsymbol{x}) \succeq \boldsymbol{B}(\boldsymbol{x}).}$

Example

$\min \ \ \boldsymbol{c}^\top \boldsymbol{x}, \quad \text{s. t. } \ \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}$, where $\boldsymbol{c} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}, \ \boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 2 & 1 \end{bmatrix}, \ \boldsymbol{b} = \begin{bmatrix} 8 \\ 18 \\ 14 \end{bmatrix}.$

We first use the `feasp` solver to find a feasible point. Then, using the `mincx` solver to solve the optimization problem.

Matlab Syntax:

```
A=[1 1;1 3;2 1];
b=[8 18 14]';
c=[-4 -5]';
setlmis([ ]);
X=lmivar(2,[2 1]);
lmiterm([1 1 1 X],A(1,:),1);
lmiterm([1 1 1 0],-b(1));
lmiterm([1 2 2 X],A(2,:),1);
lmiterm([1 2 2 0],-b(2));
lmiterm([1 3 3 X],A(3,:),1);
lmiterm([1 3 3 0],-b(3));
lmis=getlmis;
```

```
[tmin, xfeas] = feasp(lmis);
x_feasp=dec2mat(lmis,xfeas,X);
[objective,x_mincx]
=mincx(lmis,c,[0.0001 1000 0 0 1])
```

The feasp function produces

$$\boldsymbol{x}_{\mathsf{feasp}} = \begin{bmatrix} -64.3996 \\ -25.1712 \end{bmatrix},$$

The mincx function produces

$$\boldsymbol{x}_{\mathsf{mincx}} = \begin{bmatrix} 3.0000 \\ 5.0000 \end{bmatrix}.$$

Matlab Solver(II) for the form: $\min\ \text{trace}(\boldsymbol{P}),\quad \text{s.t. } \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} \preceq 0,\ \boldsymbol{P} \succeq 0$. To use mincx, we need a vector $\boldsymbol{c}$ such that

$$\boldsymbol{c}^\top \boldsymbol{x} = \text{trace}(\boldsymbol{P}).$$

We can obtain the desired $\boldsymbol{c}$ with the following Matlab Syntax:

```
q= decnbr(lmisys);
c = zeros(q,1);
for j = 1:q
Pj = defcx(lmisys,j,P);
c(j) = trace(Pj);
end
```

Having obtained the vector $\boldsymbol{c}$, we can use mincx to solve the optimization problem.

Linear Matrix Inequality (LMI) Solvers

Example (minimize generalized eigenvalue under LMI constraints)

$\min\ \lambda, \quad$ s.t. $\boldsymbol{C}(\boldsymbol{x}) \preceq \boldsymbol{D}(\boldsymbol{x}),\ 0 \preceq \boldsymbol{B}(\boldsymbol{x}),\ \boldsymbol{A}(\boldsymbol{x}) \preceq \lambda \boldsymbol{B}(\boldsymbol{x}).$

- $\boldsymbol{C}(\boldsymbol{x}) \preceq \boldsymbol{D}(\boldsymbol{x})$: standard LMI constraints.
- $\boldsymbol{A}(\boldsymbol{x}) \preceq \lambda \boldsymbol{B}(\boldsymbol{x})$: linear-fractional LMIs.

Matlab Solver (II): gevp

[lopt,xopt]=gep{lmisys,nflc}

- lopt: global minimum of the generalized eigenvalue.
- xopt: optimal decision vector variable.
- lmisys: system of LMIs

$$\boldsymbol{C}(\boldsymbol{x}) \preceq \boldsymbol{D}(\boldsymbol{x}), \boldsymbol{C}(\boldsymbol{x}) \preceq \boldsymbol{D}(\boldsymbol{x}), \text{ and } \boldsymbol{A}(\boldsymbol{x}) \preceq \lambda \boldsymbol{B}(\boldsymbol{x}) \text{ for } \lambda = 1.$$

- nflc: the number of linear-fractional constraints.

Example

Find the smallest α such that

$$\boldsymbol{P} \succ 0,\ \boldsymbol{A}^\top \boldsymbol{P} + \boldsymbol{P}\boldsymbol{A} \preceq -\alpha \boldsymbol{P}, \text{ where } \boldsymbol{A} = \begin{bmatrix} -1.1853 & 0.9134 & 0.2785 \\ 0.9058 & -1.3676 & 0.5469 \\ 0.1270 & 0.0975 & -3.0000 \end{bmatrix}.$$

Matlab Syntax

```
A = [-1.1853 0.9134 0.2785; 0.9058 -1.3676 0.5469; 0.1270 0.0975 -3.0000];
setlmis([ ]) ;
P = lmivar(1,[3 1])
lmiterm([-1 1 1 P], 1,1)     % P
lmiterm([1 1 1 0],.01)        % P >= 0.01*I
lmiterm([2 1 1 P],1,A,'s')   % linear fractional constraint --- LHS
lmiterm( [-2 1 1 P] ,1,1)     % linear fractional constraint --- RHS
lmis = getlmis; [gamma,P_opt]=gevp(lmis,1);
P=dec2mat(lmis,P_opt)
alpha=-gamma
```

Linear Matrix Inequality (LMI) Solvers

The result is $\alpha = 0.6561$, $\boldsymbol{P} = \begin{bmatrix} 0.6996 & -0.7466 & -0.0296 \\ -0.7466 & 0.8537 & -0.2488 \\ -0.0296 & -0.2488 & 3.2307 \end{bmatrix}$.