Convex Optimization Problems

Convex Optimization, by Stephen Boyd

- Find an $\underline{x}$ that minimizes $f_0(\underline{x})$ among all $\underline{x}$ that satisfy the conditions $f_i(\underline{x}) \leq 0$, $i=1,\ldots,m$, and $h_i(\underline{x})=0$, $i=1,\ldots,p$.

$x \in \mathbb{R}^n$ = optimization variable
$f_0 : \mathbb{R}^n \to \mathbb{R}$ = objective function / cost function
$f_i : \mathbb{R}^n \to \mathbb{R}$ = inequality constraint functions
$h_i : \mathbb{R}^n \to \mathbb{R}$ = equality constraint functions

Can write this problem as:

(Optimization problem in standard form)

$$\begin{array}{ll} \text{minimize} & f_0(\underline{x}) \\ \text{subject to} & \left.\begin{array}{ll} f_i(\underline{x}) \leq 0, & i=1,\ldots,m \\ h_i(\underline{x}) = 0, & i=1,\ldots,p \end{array}\right\} \text{(constraints)} \end{array}$$

- A point $\underline{x}$ is *feasible* if it satisfies the constraints.
- The problem is *feasible* if $\exists$ at least one feasible point and *infeasible* otherwise.
- The *optimal value* p^* of the problem is:

$$p^* = \inf \{ f_0(\underline{x}) \mid \underset{i=1,\ldots,m}{f_i(\underline{x}) \leq 0}, \ \underset{i=1,\ldots,p}{h_i(\underline{x})=0} \}.$$

If problem is infeasible, then $p^* = \infty$.

- $\underline{x}^*$ is an *optimal point* if $\underline{x}^*$ is feasible and $f_0(\underline{x}^*) = p^*$.

Optimal set: $X_{opt} = \{ \underline{x} \mid \underset{i=1,\ldots,m}{f_i(\underline{x}) \leq 0}, \ \underset{i=1,\ldots,p}{h_i(\underline{x})=0}, \ f_0(\underline{x}) = p^* \}$.

- Feasibility problem: find $\underline{x}$
  subj. to $f_i(\underline{x}) \le 0$, $i=1,\dots,m$
  $h_i(\underline{x})=0$, $i=1,\dots,p$

$\boxed{\text{Ex}}$ Box constraints

minimize $f_0(\underline{x})$
subject to $l_i \le x_i \le u_i$, $i=1,\dots,n$

$\rightarrow$ Standard form: minimize $f_0(\underline{x})$
subject to:
$l_i - x_i \le 0$, $i=1,\dots,n$
$x_i - u_i \le 0$, $i=1,\dots,n$

- Maximization problems:
  maximize $f_0(\underline{x})$ = minimize $-f_0(\underline{x})$

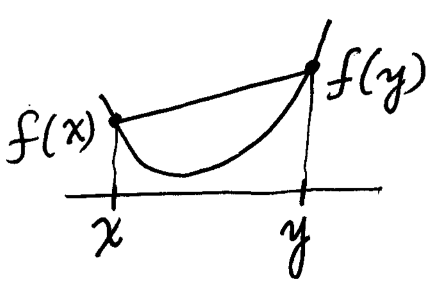
Convex optimization problems

minimize $f_0(\underline{x})$
subj. to $f_i(\underline{x}) \le 0$, $i=1,\dots,m$
$\underline{a}_i^T \underline{x} = b_i$, $i=1,\dots,p$
$\hookrightarrow h_i(\underline{x}) = \underline{a}_i^T \underline{x} - b_i = 0$ $\leftarrow$ h_i is an affine function

where $f_0, \dots, f_m$ are convex functions.

- A function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if the domain of f, $\text{dom}(f)$, is a convex set and if $\forall\, \underline{x}, \underline{y} \in \text{dom}(f)$, and $\theta \in [0,1]$:

$$f(\theta \underline{x} + (1-\theta)\underline{y}) \le \theta f(\underline{x}) + (1-\theta) f(\underline{y})$$

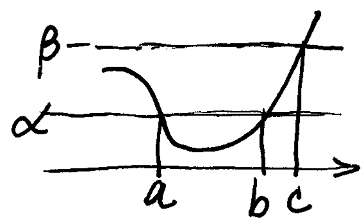


This means that the line segment between any 2 points on the graph lies above the graph.

- A function f is <u>concave</u> if $-f$ is convex.
- For an <u>affine</u> function, the $\leq$ sign is $=$
  (it is both convex and concave)

- A function f is <u>quasiconvex</u> if its domain and all its sublevel sets, $S_\alpha = \{x \in \text{dom}(f) \mid f(x) \leq \alpha\}$, $\alpha \in \mathbb{R}$, are convex.

ex)



For each α, the α-sublevel set S_α is convex (an interval).
$S_\alpha = [a, b]$ $\quad S_\beta = (-\infty, c]$

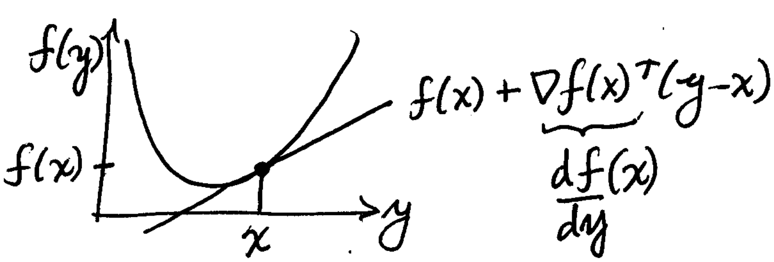
- A function is convex iff it is convex when restricted to any line that intersects its domain:

$$\forall\, \underline{x} \in \text{dom}(f) \text{ and } \forall\, \underline{v}, \quad g(t) = f(\underline{x} + t\underline{v}) \text{ is convex}$$
on its domain, $\{t \mid \underline{x} + t\underline{v} \in \text{dom}(f)\}$.

  - Allows you to check whether a function is convex

- First-order conditions:

If f is differentiable, then f is convex iff $\text{dom}(f)$ is convex and:
$$f(\underline{y}) \geq f(\underline{x}) + \nabla f(\underline{x})^T(\underline{y} - \underline{x}) \quad \forall\, \underline{x}, \underline{y} \in \text{dom}(f)$$
1st-order Taylor approximation



- If $\nabla f(\underline{x}) = 0$, then $\forall\, \underline{y} \in \text{dom}(f)$, $f(\underline{y}) \geq f(\underline{x})$
  $\Rightarrow \underline{x}$ is a <u>global minimizer</u> of f.

• Second-order conditions:

If f is twice differentiable, then f is convex iff $\text{dom}(f)$ is convex and $\forall\, \underline{x} \in \text{dom}(f)$:

$$\nabla^2 f(\underline{x}) \succeq 0 \leftarrow \text{Hessian of } f \text{ is positive semidefinite.} \quad \underline{H}_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

(Hessian of f = $\underline{H}$)

When $f: \mathbb{R} \to \mathbb{R}$, this condition is $f''(\underline{x}) \geq 0$ and $\text{dom}(f)$ is convex (an interval) $\Rightarrow$ the derivative is nondecreasing.

• Examples of convex functions:

e^{ax} is convex on $\mathbb{R}$ for any $a \in \mathbb{R}$

x^a is convex on $\mathbb{R}_{>0}$ when $a \geq 1$ or $a \leq 0$

$-\log x$ is convex on $\mathbb{R}_{>0}$

$f(x) = x \log x$ is convex on $\mathbb{R}_{>0}$ $\left(f'(x) = \log x + 1,\ f''(x) = \frac{1}{x} > 0 \text{ for } x > 0 \right)$

Every norm on $\mathbb{R}^n$ is convex.

$f(\underline{x}) = \max\{x_1, \ldots, x_n\}$ is convex on $\mathbb{R}^n$.

$f(x) = -\underbrace{\left(\prod_{i=1}^n x_i\right)^{1/n}}_{\text{geometric mean}}$ is convex on $\mathbb{R}^n_{>0}$

• Some operations that preserve convexity:

- If $w_i \geq 0$ and f_i are convex, $i = 1, \ldots, m$, then

$$f = w_1 f_1 + \cdots + w_m f_m \text{ is convex.}$$

Also, $f(\underline{x}) = \max\{f_1(\underline{x}), \ldots, f_m(\underline{x})\}$ is convex

• Showing convexity of the pointwise maximum function,

$$f(\underline{x}) = \max\{f_1(\underline{x}), \dots f_m(\underline{x})\},$$

where $f_i(\underline{x})$ are convex, for $m=2$:

$f(\underline{x}) = \max\{f_1(\underline{x}), f_2(\underline{x})\}$

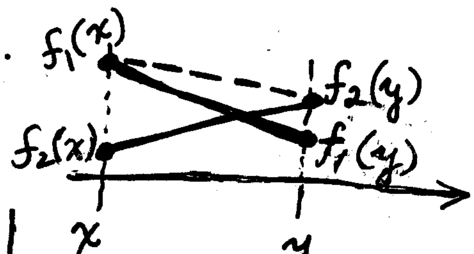
$\mathrm{dom}(f) = \mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$

$0 \le \theta \le 1 \qquad \underline{x}, \underline{y} \in \mathrm{dom}(f)$

$$f(\theta\underline{x} + (1-\theta)\underline{y}) = \max\{f_1(\theta\underline{x} + (1-\theta)\underline{y}), f_2(\theta\underline{x}+(1-\theta)\underline{y})\}$$

$$\le \max\{\theta f_1(\underline{x}) + (1-\theta) f_1(\underline{y}), \theta f_2(\underline{x}) + (1-\theta) f_2(\underline{y})\}$$ (due to the convexity of f_1, f_2)

For $\theta \in [0,1]$, $\theta f_i(\underline{x}) + (1-\theta) f_i(\underline{y})$ is the closed line segment between $f_i(\underline{x})$ and $f_i(\underline{y})$.



Note that these line segments will all lie below the line segment:

$\theta \max\{f_i(\underline{x}), i=1,\dots,m\} + (1-\theta)\max\{f_i(\underline{y}), i=1,\dots,m\}$.

Dashed line above is:
$\theta \max\{f_1(x), f_2(x)\} + (1-\theta)\max\{f_1(y), f_2(y)\}$

$$\Rightarrow \max\{\theta f_1(\underline{x}) + (1-\theta) f_1(\underline{y}), \theta f_2(\underline{x}) + (1-\theta) f_2(\underline{y})\}$$

$$\le \theta \max\{f_1(\underline{x}), f_2(\underline{x})\} + (1-\theta)\{f_1(\underline{y}), f_2(\underline{y})\}$$

$$= \theta f(\underline{x}) + (1-\theta) f(\underline{y}).$$

$\Rightarrow f(\theta\underline{x} + (1-\theta)\underline{y}) \le \theta f(\underline{x}) + (1-\theta) f(\underline{y}) \Rightarrow$ $\underline{f(\underline{x}) \text{ is convex}}$.

- Let $\underline{A} \in \mathbb{R}^{n \times n}$ be a symmetric matrix.

  Then it has real eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

$$\lambda_{max}(\underline{A}) = \lambda_1 \qquad \lambda_{min}(\underline{A}) = \lambda_n$$

$$\lambda_{max}(\underline{A}) = \sup_{\underline{x} \neq \underline{0}} \frac{\underline{x}^T \underline{A} \underline{x}}{\underline{x}^T \underline{x}}, \quad \lambda_{min}(\underline{A}) = \inf_{\underline{x} \neq \underline{0}} \frac{\underline{x}^T \underline{A} \underline{x}}{\underline{x}^T \underline{x}}$$

$f(\underline{A}) = \lambda_{max}(\underline{A})$ is convex. $\hookrightarrow = \sup\{\underline{x}^T \underline{A} \underline{x} \mid \|\underline{x}\|_2 = 1\}$

$f(\underline{A}) = \lambda_{min}(\underline{A})$ is concave.

- The distance of a point $\underline{x}$ to a set $S \subseteq \mathbb{R}^n$ is:

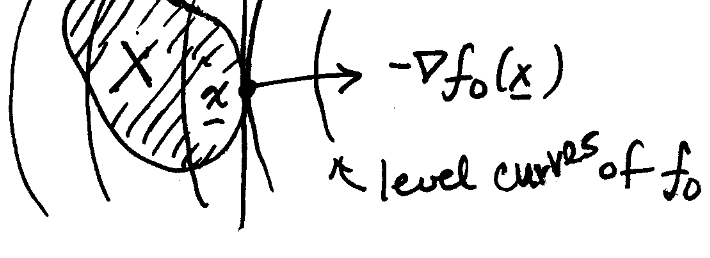
$$\text{dist}(\underline{x}, S) = \inf_{\underline{y} \in S} \|\underline{x} - \underline{y}\|.$$

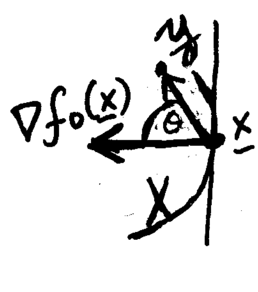
  $\|\underline{x} - \underline{y}\|$ is convex in $(\underline{x}, \underline{y})$, so if S is convex, $\text{dist}(\underline{x}, S)$ is a convex function of $\underline{x}$.

- The feasible set of a convex optimization problem is convex
- Any locally optimal point of a convex opt. prob. is also globally optimal.
- An optimality criterion for differentiable f_0: ← objective/cost function

  If f_0 is differentiable, and X is the feasible set, then $\underline{x}$ is optimal iff $\underline{x} \in X$ and:

$$\nabla f_0(\underline{x})^T (\underline{y} - \underline{x}) \geq 0 \quad \forall \underline{y} \in X.$$





$\nabla f_0(\underline{x})^T (\underline{y} - \underline{x}) = \|\nabla f_0(\underline{x})\| \|\underline{y} - \underline{x}\| \cdot \cos\theta \geq 0$ iff $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

Algebraic connectivity / Fiedler value, $\lambda_2(\underline{L})$

- Consider an undirected graph with Laplacian matrix $\underline{L} \in \mathbb{R}^{n \times n}$.

$\underline{L}$ is symmetric and positive semi-definite. Its eigenvalues are real and can be labeled as follows:

$\lambda_1 = 0 \le \lambda_2 \le \lambda_3 \le \dots \le \lambda_n$

- $\lambda_2 > 0$ iff the graph is connected.
- The number of times 0 appears as an eigenvalue of $\underline{L}$ = the number of connected components in the graph.

Define $\underline{1}^T = \{ \underline{x} \mid \underline{1}^T \underline{x} = 0 \}$.

The vector $\underline{1}$ is in the null space of $\underline{L}$:

$\underline{L}\,\underline{1} = \underline{0} \Rightarrow$ Since $\lambda_1 = 0$, $\underline{1}$ is the eigenvector corresponding to λ_1. $(\underline{L}\,\underline{1} = \lambda_1 \underline{1})$

$$\underline{L}\,\underline{x}_i = \lambda_i \underline{x}_i \Rightarrow \underline{x}_i^T \underline{L}\,\underline{x}_i = \underline{x}_i^T \lambda_i \underline{x}_i = \lambda_i \underline{x}_i^T \underline{x}_i$$

The Rayleigh-Ritz theorem can be used to define λ_2:

$$\lambda_2 = \inf_{\substack{\underline{x} \neq \underline{0}, \\ \underline{x} \in \underline{1}^T}} \left(\frac{\underline{x}^T \underline{L}\,\underline{x}}{\underline{x}^T \underline{x}} \right)$$

(See *Matrix Analysis* by Horn and Johnson, section 4.2 on variational characterizations of eigenvalues of Hermitian matrices)

λ_2 is a concave function of $\underline{L}$ in the space $\underline{1}^T$.

$$\max_{\underline{L} \in S^n} \lambda_2(\underline{L}) \Leftrightarrow \min_{\underline{L} \in S^n} (-\lambda_2(\underline{L})) \qquad S^n = \{ \underline{X} \in \mathbb{R}^{n \times n} \mid \underline{X} = \underline{X}^T \}$$

Linear optimization problems (Linear programs, LP)

Standard form: minimize $\underline{c}^T \underline{x}$

subj. to $\underline{A}\,\underline{x} = \underline{b}$

$\underline{x} \succeq \underline{0}$ ← componentwise inequality

- f_0 is affine, $f_1, \ldots, f_m$ are affine
- The feasible set is a polyhedron

Quadratic optimization problems (Quadratic programs, QP)

minimize $\frac{1}{2}\underline{x}^T \underline{P}\,\underline{x} + \underline{q}^T \underline{x} + r$

subj. to $\underline{G}\,\underline{x} \preceq \underline{h}$

$\underline{A}\,\underline{x} = \underline{b}$

- f_0 is convex quadratic, constraint fns. are affine.
- $\underline{P} \in \mathbb{R}^{n \times n}$ is a symmetric, positive semi-definite matrix

$\underline{G} \in \mathbb{R}^{m \times n}$ $\quad \underline{A} \in \mathbb{R}^{p \times n}$

- Linear programs are a special case

Semidefinite programs (SDP)

Can generalize the convex opt. problem to a problem with generalized inequality constraints:

minimize $f_0(x)$ $\quad (f_0 : \mathbb{R}^n \to \mathbb{R})$

subj. to $f_i(\underline{x}) \preceq_{K_i} 0, \quad i = 1, \ldots, m \quad (f_i : \mathbb{R}^n \to \mathbb{R}^{k_i})$

$\underline{A}\,\underline{x} = \underline{b}$

Here, $K_i \subseteq \mathbb{R}^{k_i}$ is a <u>proper cone</u>, which is a subset of $\mathbb{R}^{k_i}$ that is convex, closed, solid (has a nonempty interior) and pointed (contains no line, i.e., if $\underline{x} \in K_i$ and $-\underline{x} \in K_i$, then $\underline{x} = \underline{0}$).

Generalized inequality with proper cone K: $\underline{x} \preceq_K y \Leftrightarrow y - \underline{x} \in K$

$\underline{x} \prec_K y \Leftrightarrow y - x \in$ interior of K

<u>Examples</u>

- $K = \mathbb{R}^n_{\geq 0}$ is a proper cone.

  $\underline{x} \preceq_K y \Leftrightarrow x_i \leq y_i, \quad i = 1, \ldots, n$

- The set of positive semidefinite symmetric matrices,

  $S^n_{\geq 0} = \{ \underline{X} \in \mathbb{R}^{n \times n} \mid \underline{X} = \underline{X}^T, \underline{X} \succeq \underline{0} \}$, is a proper cone in the set of symm. matrices. It is also a convex set since:

  If $\theta_1, \theta_2 \geq 0$ and $\underline{A}, \underline{B} \in S^n_{\geq 0}$, then $\theta_1 \underline{A} + \theta_2 \underline{B} \in S^n_{\geq 0}$.

  $\underline{X} \preceq_K \underline{Y} \Leftrightarrow \underline{Y} - \underline{X} \succeq_K \underline{0} \Rightarrow (\underline{Y} - \underline{X})$ is a positive semidefinite matrix.

- Can drop the subscript K: $\underline{X} \preceq \underline{Y}$

Also, the $f_i(\underline{x})$ are K_i-convex, meaning that:

$$f_i(\theta \underline{x} + (1-\theta) y) \preceq_{K_i} \theta f_i(\underline{x}) + (1-\theta) f_i(y)$$

- Many results for ordinary convex opt. problems hold for problems with generalized inequalities.

- Conic form problems: (cone programs):

$$\begin{aligned} &\text{minimize} && \underline{c}^T \underline{x} && \leftarrow f_0 \text{ is linear} \\ &\text{subj. to} && \underline{F}\underline{x} + \underline{g} \preceq_K \underline{0} && \leftarrow \text{one affine } f_i \\ & && \underline{A}\underline{x} = \underline{b} \end{aligned}$$

- When $K = \mathbb{R}^n_{\geq 0}$, this reduces to a linear program.

- Semidefinite programs: (SDP)

$K = S^k_{\geq 0}$ (cone of positive semidef. $k \times k$ matrices)

$$\begin{aligned} &\text{minimize} && \underline{c}^T \underline{x} \\ &\text{subj. to} && x_1 \underline{F}_1 + \dots + x_n \underline{F}_n + \underline{G} \preceq_K 0 && \leftarrow \text{Linear Matrix inequality (LMI)} \\ & && \underline{A}\underline{x} = \underline{b} \quad (\underline{A} \in \mathbb{R}^{p \times n}) \end{aligned}$$

where $\underline{F}_1, \dots, \underline{F}_n$ and $\underline{G}$ are symmetric $k \times k$ matrices.
(If these matrices are all diagonal, then the SDP reduces to a linear program.)

In standard form:

$$\begin{aligned} &\text{minimize} && \operatorname{trace}(\underline{C}\underline{X}) \\ &\text{subj. to} && \operatorname{trace}(\underline{A}_i \underline{X}) = \underline{b}_i, \quad i = 1, \dots, p \\ & && \underline{X} \succeq \underline{0} \end{aligned}$$

where $\underline{C}, \underline{A}_1, \dots, \underline{A}_p$ are symm. $n \times n$ matrices.

- $\operatorname{trace}(\underline{C}\underline{X}) = \sum_{i,j=1}^n \underline{C}_{ij} \underline{X}_{ij}$ is a real-valued linear function on the set of symm. $n \times n$ matrices.

Epigraph problem form

- Standard problem:

  minimize $f_0(\underline{x})$

  subj. to $f_i(\underline{x}) \leq 0$, $i=1,\dots,m$

  $h_i(\underline{x}) = 0$, $i=1,\dots,p$

$=$

- Epigraph form:

  minimize t

  subj. to $f_0(\underline{x}) - t \leq 0$

  $f_i(\underline{x}) \leq 0$, $i=1,\dots,m$

  $h_i(\underline{x}) = 0$, $i=1,\dots,p$

  - objective function is <u>linear</u> in $\underline{x}, t$, and hence convex.
  - If $f_0(\underline{x})$ is convex, then $f_0(\underline{x}) - t$ is convex in $(\underline{x}, t)$.

"Epi" = "above"

the graph of $f: \mathbb{R}^n \to \mathbb{R}$ is: $\{(\underline{x}, f(\underline{x})) \mid \underline{x} \in \mathrm{dom}(f)\}$.

the epigraph of $f: \mathbb{R}^n \to \mathbb{R}$ is:

$$\mathrm{epi}(f) = \{(\underline{x}, t) \mid \underline{x} \in \mathrm{dom}(f),\ f(x) \leq t\}.$$

- f is convex iff its epigraph is a convex set.

If $f: \mathbb{R}^n \to \mathbb{R}^m$, and $K \subseteq \mathbb{R}^m$ is a proper convex cone with generalized ineq. $\preceq_K$, then:

$$\underset{(\text{wrt to } \preceq_K)}{\mathrm{epi}(f)} = \mathrm{epi}_K(f) = \{(\underline{x}, \underline{t}) \in \mathbb{R}^{n+m} \mid f(\underline{x}) \preceq_K \underline{t}\}$$

Graph-Theoretic Connectivity Control of Mobile Robot Networks

$\lambda_2(\underline{L})$ measures connectivity in robot networks.

- Can design connectivity controllers by maximizing $\lambda_2(\underline{L})$

$\boxed{P1}$ $\max_{L \in S^n} \lambda_2(\underline{L})$ $S^n = \{\underline{X} \in \mathbb{R}^{n\times n} \mid \underline{X} = \underline{X}^T\}$

- Centralized solution to $\boxed{P1}$

-Relate positive definiteness of $\lambda_2(\underline{L})$ to positive definiteness of a quadratic function of $\underline{L}$:

$$\underline{P} = [\underline{p}_1 \ \underline{p}_2 \ \cdots \ \underline{p}_{n-1}] \in \mathbb{R}^{n\times(n-1)}$$

$$\underline{p}_i^T \underline{1} = 0 \quad \forall i = 1, \ldots, n-1$$

$$\underline{p}_i^T \underline{p}_j = 0 \quad \forall i \neq j$$

$$\Rightarrow \quad \lambda_2(\underline{L}) > 0 \quad \text{iff} \quad \underline{P}^T \underline{L} \underline{P} \succ 0.$$

This result allows us to formulate an equivalent convex formulation for $\boxed{P1}$: (see next page)

$$\boxed{P2} \left\{ \begin{array}{l} \max\limits_{\underline{L} \in S^n} \gamma \\ \text{Subject to } \underline{P}^T \underline{L} \underline{P} \succ \gamma \underline{I}_{n-1} \end{array} \right.$$

Can solve it for the optimal $\underline{L}$.

- Inequality form SDP (no equality constraint $\underline{A}\underline{x} = \underline{b}$).

$\boxed{P2}$ is in epigraph form.

$\boxed{P1}$ $\underset{\underline{L} \in S^n}{\text{maximize}}\ \lambda_2(\underline{L}) \iff \underset{\underline{L} \in S^n}{\text{minimize}}\ -\lambda_2(\underline{L})$

$\iff$ minimize t
subject to $-\lambda_2(\underline{L}) - t < 0$

$\iff$ minimize t
subject to $\lambda_2(\underline{L}) > -t$

$\iff$ minimize t
subject to $\underline{P}^T \underline{L}\, \underline{P} \succ -t\, \underline{I}_{n-1}$

$\iff$ $\boxed{P2}$ $\begin{cases} \text{maximize } \gamma & (\gamma = -t) \\ \text{subject to } \underline{P}^T \underline{L}\, \underline{P} \succ \gamma\, \underline{I}_{n-1} \\ \qquad (\underline{L} \in S^n) \end{cases}$