Semidefinite Programming

Computational Intelligence, Lecture 9

by Sergei Savin

Spring 2024

Content

General form of a semidefinite program is:

$$
\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{c}^{\top}\mathbf{x}, \\
\text{subject to} \quad & \begin{cases} \mathbf{G} + \sum \mathbf{F}_i x_i \preceq 0, \\ \mathbf{A}\mathbf{x} = \mathbf{b}. \end{cases}
\end{aligned}
\tag{1}
$$

where $\mathbf{F}_i \succeq 0$ and $\mathbf{G} \succeq 0$ (meaning they are positive semidefinite).

Constraint $\mathbf{G} + \sum \mathbf{F}_i x_i \preceq 0$ is called *linear matrix inequality* or *LMI*.

SDP - Multiple LMI

SDP can have several LMIs. Assume you have:

$$\begin{cases} \mathbf{G} + \sum \mathbf{F}_i x_i \preceq 0 \\ \mathbf{D} + \sum \mathbf{H}_i x_i \preceq 0 \end{cases} \tag{2}$$

This is equivalent to:

$$\begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} + \sum \begin{bmatrix} \mathbf{F}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_i \end{bmatrix} x_i \preceq 0 \tag{3}$$

SDP decision variable

Sometimes it is easier to directly think of semidefinite matrices as of decision variables. This leads to programs with such formulation:

$$\begin{aligned} \underset{\mathbf{X}}{\text{minimize}} \quad & \operatorname{tr}(\mathbf{E}^\top \mathbf{X}), \\ \text{subject to} \quad & \begin{cases} \operatorname{tr}(\mathbf{A}_i^\top \mathbf{X}) = \mathbf{b}_i, \\ \mathbf{C}\mathbf{X} \preceq \mathbf{D}. \end{cases} \end{aligned} \tag{4}$$

where cost and constraints should adhere to SDP limitations.

Trace of a matrix product

Consider a matrices $\mathbf{E} = \begin{bmatrix} \mathbf{e}_1 & ... & \mathbf{e}_n \end{bmatrix}$ and $\mathbf{X} = \begin{bmatrix} \mathbf{x}_1 & ... & \mathbf{x}_n \end{bmatrix}$. Their product can be written as:

$$\mathbf{E}^\top \mathbf{X} = \begin{bmatrix} \mathbf{e}_1^\top \mathbf{x}_1 & \mathbf{e}_1^\top \mathbf{x}_2 & ... \\ \mathbf{e}_2^\top \mathbf{x}_1 & \mathbf{e}_2^\top \mathbf{x}_2 & ... \\ ... & ... & ... \end{bmatrix} \tag{5}$$

Thus, the trace of this product is given as:

$$\mathrm{tr}(\mathbf{E}\mathbf{X}) = \mathbf{e}_1^\top \mathbf{x}_1 + ... + \mathbf{e}_n^\top \mathbf{x}_n \tag{6}$$

We can see that this is equivalent to an element-wise dot product.

In a cost function, matrix $\mathbf{E}$ plays the role of weights, similar to $\mathbf{f}$ in the linear cost $\mathbf{f}^\top \mathbf{x}$. Quadratic cost can be expressed as $\mathbf{X}^\top \mathbf{X}$.

Continuous Lyapunov eq. as SDP/LMI (1)

In control theory, Lyapunov equation is a condition of whether or not a continuous LTI system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is stable:

$$\begin{cases} \mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} \preceq -\mathbf{Q} \\ \mathbf{P} \succeq 0 \end{cases} \tag{7}$$

where $\mathbf{Q} \succeq 0$ is a constant and decision variable is $\mathbf{P}$. This can be represented as an SDP:

$$\begin{aligned} & \underset{\mathbf{P}}{\text{minimize}} && 0, \\ & \text{subject to} && \begin{cases} \mathbf{P} \succeq 0, \\ \mathbf{A}^\top \mathbf{P} + \mathbf{P}\mathbf{A} + \mathbf{Q} \preceq 0. \end{cases} \end{aligned} \tag{8}$$

```
n = 7; A = randn(n, n) - 3*rand*eye(n);
Q = eye(n);

cvx_begin sdp
    variable P(n, n) symmetric
    minimize 0
    subject to
        P >= 0;
        A'*P + P*A + Q <= 0;
cvx_end

if strcmp(cvx_status, 'Solved')
    [eig(A), eig(A*P + P*A' + Q), eig(P)]
else
    eig(A)
end
```

Discrete Lyapunov eq. as SDP/LMI (1)

In control theory, Discrete Lyapunov equation is a condition of whether or not a discrete LTI system $\mathbf{x}_{i+1} = \mathbf{A}\mathbf{x}_i$ is stabilizable:

$$\begin{cases} \mathbf{A}^\top \mathbf{P} \mathbf{A} - \mathbf{P} + \mathbf{Q} \preceq 0 \\ \mathbf{P} \succeq 0 \end{cases} \tag{9}$$

where $\mathbf{Q} \succeq 0$ is a constant and decision variable is $\mathbf{P}$. This can be represented as an SDP:

$$\begin{aligned} \underset{\mathbf{P}}{\text{minimize}} \quad & 0, \\ \text{subject to} \quad & \begin{cases} \mathbf{P} \succeq 0, \\ \mathbf{A}^\top \mathbf{P} \mathbf{A} - \mathbf{P} + \mathbf{Q} \preceq 0. \end{cases} \end{aligned} \tag{10}$$

```
n = 7; A = 0.35*randn(n, n);
Q = eye(n);

cvx_begin sdp
    variable P(n, n) symmetric
    minimize 0
    subject to
        P >= 0;
        A'*P*A - P + Q <= 0;
cvx_end

if strcmp(cvx_status, 'Solved')
    [abs(eig(A)), eig(A'*P*A - P), eig(P)]
else
    abs(eig(A))
end
```

Schur compliment. Given $\mathbf{M}$

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \tag{11}$$

with full-rank $\mathbf{A}$, we can make the following statement:

- $\mathbf{M} \succ 0$ iff $\mathbf{A} \succ 0$ and $\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} \succ 0$

If $\mathbf{C}$ is full-rank, we can make the following statement:

- $\mathbf{M} \succ 0$ iff $\mathbf{C} \succ 0$ and $\mathbf{A} - \mathbf{B} \mathbf{C}^{-1} \mathbf{B}^\top \succ 0$

Let us prove that $\mathbf{M} \succ 0$ iff $\mathbf{A} \succ 0$ and $\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} \succ 0$.

First we prove that $\mathbf{A} \succ 0$ and $\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} \succ 0$ implies $\mathbf{M} \succ 0$. We need to prove that the following quadratic form is positive definite:

$$f = \begin{bmatrix} \mathbf{x}^\top & \mathbf{y}^\top \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^\top & \mathbf{C} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \tag{12}$$

$$= \mathbf{x}^\top \mathbf{A} \mathbf{x} + \mathbf{x}^\top \mathbf{B} \mathbf{y} + \mathbf{y}^\top \mathbf{B}^\top \mathbf{x} + \mathbf{y}^\top \mathbf{C} \mathbf{y} \tag{13}$$

Since $\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B} \succ 0$, the following quadratic form is positive-definite:

$$\mathbf{y}^\top (\mathbf{C} - \mathbf{B}^\top \mathbf{A}^{-1} \mathbf{B}) \mathbf{y} > 0 \tag{14}$$

We define a change of variables $\mathbf{x} = -\frac{1}{2}\mathbf{A}^{-1}\mathbf{B}\mathbf{y}$, giving us two equations:

$$\mathbf{y}^\top \mathbf{C}\mathbf{y} + 2\mathbf{x}^\top \mathbf{B}\mathbf{y} > 0 \tag{15}$$

$$\mathbf{y}^\top \mathbf{C}\mathbf{y} + 2\mathbf{y}^\top \mathbf{B}^\top \mathbf{x} > 0 \tag{16}$$

Their sum gives us:

$$2\mathbf{y}^\top \mathbf{C}\mathbf{y} + 2\mathbf{x}^\top \mathbf{B}\mathbf{y} + 2\mathbf{y}^\top \mathbf{B}^\top \mathbf{x} > 0 \tag{17}$$

$$\mathbf{y}^\top \mathbf{C}\mathbf{y} + \mathbf{x}^\top \mathbf{B}\mathbf{y} + \mathbf{y}^\top \mathbf{B}^\top \mathbf{x} > 0 \tag{18}$$

Since $\mathbf{A} \succ 0$ we conclude that:

$$\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{y}^\top \mathbf{C}\mathbf{y} + \mathbf{x}^\top \mathbf{B}\mathbf{y} + \mathbf{y}^\top \mathbf{B}^\top \mathbf{x} > 0 \tag{19}$$

This finishes first part of the proof.

SOCP as SDP

Let us prove that SOCP is a sub-set of SDP. SOC constraint is:

$$||\mathbf{A}\mathbf{x} + \mathbf{b}|| \leq \mathbf{c}^\top \mathbf{x} + d \tag{20}$$

where $\mathbf{c}^\top \mathbf{x} + d \geq 0$, and we can rewrite the SOC as:
$(\mathbf{A}\mathbf{x} + \mathbf{b})^\top (\mathbf{A}\mathbf{x} + \mathbf{b}) = (\mathbf{c}^\top \mathbf{x} + d)^2$, and assuming $\mathbf{c}^\top \mathbf{x} + d > 0$ we can write it as:

$$\frac{(\mathbf{A}\mathbf{x} + \mathbf{b})^\top (\mathbf{A}\mathbf{x} + \mathbf{b})}{\mathbf{c}^\top \mathbf{x} + d} \leq \mathbf{c}^\top \mathbf{x} + d \tag{21}$$

which is equivalent to:

$$-\frac{(\mathbf{A}\mathbf{x} + \mathbf{b})^\top (\mathbf{A}\mathbf{x} + \mathbf{b})}{-(\mathbf{c}^\top \mathbf{x} + d)} \leq \mathbf{c}^\top \mathbf{x} + d \tag{22}$$

SOCP as SDP

Note that $-\frac{(\mathbf{Ax}+\mathbf{b})^\top(\mathbf{Ax}+\mathbf{b})}{-(\mathbf{c}^\top\mathbf{x}+d)} \leq \mathbf{c}^\top\mathbf{x} + d$ is equivalent to:

$$-\frac{(\mathbf{Ax}+\mathbf{b})^\top(\mathbf{Ax}+\mathbf{b})}{(\mathbf{c}^\top\mathbf{x}+d)} + (\mathbf{c}^\top\mathbf{x}+d) \geq 0 \tag{23}$$

Using Schur we can re-write it as:

$$\begin{bmatrix} (\mathbf{c}^\top\mathbf{x}+d) & (\mathbf{Ax}+\mathbf{b}) \\ (\mathbf{Ax}+\mathbf{b})^\top & (\mathbf{c}^\top\mathbf{x}+d) \end{bmatrix} \succeq 0 \tag{24}$$

which is an LMI constraint.

Consider the following constraint, where $\mathbf{X} \succeq 0$:

$$||\mathbf{X}\mathbf{v} + \mathbf{b}|| \leq \mathbf{c}^\top \mathbf{x} + d \tag{25}$$

Can we re-write it as an LMI? Using the same process as before we get:

$$\begin{bmatrix} (\mathbf{c}^\top \mathbf{x} + d) & (\mathbf{X}\mathbf{v} + \mathbf{b}) \\ (\mathbf{X}\mathbf{v} + \mathbf{b})^\top & (\mathbf{c}^\top \mathbf{x} + d) \end{bmatrix} \succeq 0 \tag{26}$$

So, (25) is an admissible constraint in an SDP.

Consider the problem: minimize the largest eigenvalue of A. The solution is:

$$\begin{aligned} \underset{\mathbf{A},t}{\text{minimize}} \quad & t, \\ \text{subject to} \quad & \mathbf{A} \preceq t\mathbf{I} \end{aligned} \tag{27}$$

Proof. If λ is an eigenvalue of $\mathbf{A}$, then $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$, hence $(\mathbf{A} - t\mathbf{I})\mathbf{v} = (\lambda - t)\mathbf{v}$, meaning $\lambda - t$ is eigenvalue of $(\mathbf{A} - t\mathbf{I})$. Thus, if $(\mathbf{A} - t\mathbf{I})$ is negative semi-definite, then $\lambda - t \leq 0$ and $\lambda \leq t$. □

Let us first remember how we describe a unit sphere:

$$\mathcal{S} = \{\mathbf{x} : \ ||\mathbf{x}|| \leq 1\} \tag{28}$$

An ellipsoid can be seen as a linear transformation of a unit sphere:

$$\mathcal{E} = \{\mathbf{A}\mathbf{x} + \mathbf{b} : \ ||\mathbf{x}|| \leq 1\} \tag{29}$$

Let us introduce a change of variables $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{b}$. Assuming $\mathbf{A}$ is invertible, we get:

$$\mathbf{x} = \mathbf{A}^{-1}(\mathbf{z} - \mathbf{b}) \tag{30}$$

So, we can describe the exact same ellipsoid using an alternative formula:

$$\mathcal{E} = \{\mathbf{z} : \ ||\mathbf{B}\mathbf{z} + \mathbf{c}|| \leq 1\} \tag{31}$$

where $\mathbf{B} = \mathbf{A}^{-1}$ and $\mathbf{c} = -\mathbf{A}^{-1}\mathbf{b}$.

For an ellipsoid of the form

$$\mathcal{E} = \{\mathbf{A}\mathbf{x} + \mathbf{b} : \ ||\mathbf{x}|| \leq 1\} \tag{32}$$

the "bigger" the $\mathbf{A}$, the bigger the ellipsoid. This concept can be made concrete by talking about the determinant of $\mathbf{A}$.

Thus, maximizing the volume of this ellipsoid is the same as maximizing $\det(\mathbf{A})$. Or, it is the same as *minimizing* the $\det(\mathbf{A}^{-1})$, since $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.

Finally, note that $\log\det(\mathbf{A})$ is a concave function and $\log\det(\mathbf{A}^{-1})$ is a convex function.

Volume of an ellipsoid (2)

For an ellipsoid of the form

$$\mathcal{E} = \{\mathbf{z} : \ ||\mathbf{B}\mathbf{z} + \mathbf{c}|| \leq 1\} \tag{33}$$

the "bigger" the $\mathbf{B}$, the *smaller* the ellipsoid. We can make it obvious by thinking that increasing $\mathbf{B}$ leaves less room for valid $\mathbf{z}$, and it is the volume of valid $\mathbf{z}$ that makes the volume of the ellipsoid in this case.

This concept can be made concrete by talking about the determinant of $\mathbf{B}$. Thus, maximizing the volume of this ellipsoid is the same as *minimizing* $\det(\mathbf{B})$. Or, it is the same as *maximizing* the $\det(\mathbf{B}^{-1})$.

Consider the problem: given V-polytope, defined by its vertices $\mathbf{v}_i$, find minimum-volume ellipsoid $\mathcal{E}$ containing the polytope. We will start with defining the ellipsoid as $\mathcal{E} = \{\mathbf{z} : \ ||\mathbf{B}\mathbf{z} + \mathbf{c}|| \leq 1\}$. The ellipsoid is smaller when $||\mathbf{B}||$ is bigger, and thus we can write the minimization as minimizing $\det(\mathbf{B}^{-1})$.

$$\begin{aligned} \underset{\mathbf{B},\mathbf{c}}{\text{minimize}} \quad & \log(\det(\mathbf{B}^{-1})), \\ \text{subject to} \quad & \begin{cases} \mathbf{B} \succeq 0, \\ ||\mathbf{B}\mathbf{v}_i + \mathbf{c}|| \leq 1. \end{cases} \end{aligned} \tag{34}$$

The solution gives us Löwner-John ellipsoid.

Max volume inscribed ellipsoid (1)

Consider the problem: given H-polytope, defined by its half-spaces $\mathbf{a}_i^\top \mathbf{x} \leq b_i$, find maximum-volume ellipsoid $\mathcal{E}$ contained in the polytope. We will start with defining the ellipsoid as $\mathcal{E} = \{\mathbf{C}\mathbf{x} + \mathbf{d} : \; ||\mathbf{x}|| \leq 1\}$. The ellipsoid is larger when $||\mathbf{C}||$ is bigger, and thus we can write the minimization as minimizing $\det(\mathbf{C}^{-1})$.

Let us write down the constraint requiring that $\mathcal{E}$ lies in the polytope. We know that $\mathbf{a}_i^\top(\mathbf{C}\mathbf{x} + \mathbf{d}) \leq b_i$ holds for all $||\mathbf{x}|| \leq 1$. The worst-case scenario is when $\mathbf{x}$ aligned with $\mathbf{a}_i^\top \mathbf{C}$ and has length 1:

$$\mathbf{x} = \frac{\mathbf{a}_i^\top \mathbf{C}}{||\mathbf{a}_i^\top \mathbf{C}||} \tag{35}$$

Thus the constraint becomes

$$||\mathbf{a}_i^\top \mathbf{C}|| + \mathbf{a}_i^\top \mathbf{d} \leq b_i \tag{36}$$

Max volume inscribed ellipsoid (2)

Here is the resulting problem:

$$\begin{aligned} \underset{\mathbf{C},\mathbf{d}}{\text{minimize}} \quad & \log(\det(\mathbf{C}^{-1})), \\ \text{subject to} \quad & \begin{cases} \mathbf{C} \succeq 0, \\ ||\mathbf{a}_i^\top \mathbf{C}|| + \mathbf{a}_i^\top \mathbf{d} \leq b_i. \end{cases} \end{aligned} \tag{37}$$

The solution gives us inscribed (inner) Löwner-John ellipsoid.

Homework

Implement both examples from page 2 of the LMI CVX documents.

Lecture slides are available via Github, links are on Moodle:

github.com/SergeiSa/Computational-Intelligence-2024