Chapter 1

Introduction

Exercise 1.1

Let $A \in \mathbb{S}^m$. Show that for arbitrary $M \in \mathbb{R}^{m\times n}$, $A \leq 0$ implies $M^{\mathrm{T}}AM \leq 0$.

Solution I. Since $A \leq 0$, we have

$$y^{\mathrm{T}}Ay \leq 0, \ \forall y \in \mathbb{R}^m.$$

Therefore, for arbitrary $x \in \mathbb{R}^n$ and $M \in \mathbb{R}^{m\times n}$, there holds

$$x^{\mathrm{T}}M^{\mathrm{T}}AMx = (Mx)^{\mathrm{T}}AMx \leq 0.$$

This completes the proof.

Solution II. Since $A \leq 0$, there exists a matrix $T \in \mathbb{R}^{m\times m}$ such that

$$A = -T^{\mathrm{T}}T.$$

Therefore, for arbitrary $M \in \mathbb{R}^{m\times n}$, there holds

$$M^{\mathrm{T}}AM = -(TM)^{\mathrm{T}}(TM) \leq 0.$$

This completes the proof.

Remark. On the other side, if for arbitrary $M \in \mathbb{R}^{m\times n}$, there holds $M^{\mathrm{T}}AM \leq 0$, we can simply choose $M = I$, the identity, and obtain $A \leq 0$. Therefore, we actually have the conclusion that $A \leq 0$ if and only if $M^{\mathrm{T}}AM \leq 0$ for arbitrary $M \in \mathbb{R}^{m\times n}$.

Exercise 1.2 (Duan and Patton (1998), Zhang and Yang (2003), page 175)

Let $A \in \mathbb{C}^{n\times n}$. Show that A is Hurwitz stable if $A + A^{\mathrm{H}} < 0$.

Solution. First, we remark that, like the case for a real matrix, a complex square matrix is called Hurwitz stable if all its eigenvalues have negative real parts. Let λ be an eigenvalue of A, x be a corresponding eigenvector, then we have

$$Ax = \lambda x,$$

and further

$$x^{\mathrm{H}}(A^{\mathrm{H}} + A)x = (\lambda + \bar{\lambda})x^{\mathrm{H}}x.$$

Thus $A^{\mathrm{H}} + A < 0$ implies

$$\mathrm{Re}\,\lambda(A) = \frac{\lambda + \bar{\lambda}}{2} < 0.$$

This completes the proof.

Exercise 1.3 (Duan and Patton (1998))

Let $A \in \mathbb{R}^{n\times n}$. Show that A is Hurwitz stable if and only if

$$A = PQ, \tag{s1.1}$$

with $P > 0$ and Q being some matrix satisfying $Q + Q^{\mathrm{T}} < 0$.

Solution. Suppose $A = PQ$ holds with $P > 0$ and Q satisfying

$$Q + Q^{\mathrm{T}} < 0. \tag{s1.2}$$

Let

$$\hat{P} = P^{-1} > 0,$$

then it is easy to see

$$A^{\mathrm{T}}\hat{P} + \hat{P}A = Q + Q^{\mathrm{T}} < 0.$$

Therefore, the matrix A is Hurwitz stable.

Conversely, if A is Hurwitz stable, then there exists a matrix $\hat{P} > 0$, such that

$$A^{\mathrm{T}}\hat{P} + \hat{P}A < 0. \tag{s1.3}$$

Let

$$Q = \hat{P}A,$$

we can easily get (s1.1), with $P = \hat{P}^{-1} > 0$, and the matrix Q obviously satisfies (s1.2) because of (s1.3).

Exercise 1.4

Give an example to show that certain set of nonlinear inequalities can be converted into LMIs.

Solution. Let $Q(x) = Q^{\mathrm{T}}(x)$, $R(x) = R^{\mathrm{T}}(x)$ and $S(x)$ depend affinely on x. It is clearly that

$$Q(x) - S(x)R(x)^{-1}S^{\mathrm{T}}(x) > 0,\ R(x) > 0,$$

is quadratic with respect to $S(x)$. Using Schur completion lemma the above two relations can be equivalently converted into

$$\begin{bmatrix} Q(x) & S(x) \\ S^{\mathrm{T}}(x) & R(x) \end{bmatrix} > 0,$$

which is now linear in $S(x)$.

Exercise 1.5

Verify for which integer i the following inequality is true:

$$\begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} > \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Solution I. Let

$$\Theta = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & i-1 \\ i-1 & 1 \end{bmatrix}, \tag{s1.4}$$

then we know

$$sI - \Theta = \begin{bmatrix} s-1 & -i+1 \\ -i+1 & s-1 \end{bmatrix},$$

$$\det(sI - \Theta) = (s-1)^2 - (i-1)^2 = (s-i)(s+i-2).$$

Thus

$$\lambda(\Theta) = \{i, 2-i\},$$

which indicates $\Theta > 0$ if and only if $i = 1$. Therefore the conclusion holds if and only if $i = 1$.

Solution II. It follows from (s1.4) and the Schur complement lemma that $\Theta > 0$ if and only if

$$1 - (i-1)^2 > 0,$$

which is equivalent to

$$(i-1)^2 < 1.$$

Obviously, this holds if and only if $i = 1$.

Exercise 1.6

Consider the combined constraints (in the unknown x) of the form

$$\begin{cases} F(x) < 0 \\ Ax = a \end{cases}, \tag{s1.5}$$

where the affine function $F : \mathbb{R}^n \to \mathbb{S}^m$, matrix $A \in \mathbb{R}^{m\times n}$ and vector $a \in \mathbb{R}^m$ are given, and the equation $Ax = a$ has a solution. Show that (s1.5) can be converted into an LMI.

Solution. Suppose rank$A = r$, then it is well-known that all the solution vectors of the equation $Ax = a$ constitute a manifold, of dimension r, in $\mathbb{R}^n$, and a general form of all the solutions can be written as

$$x = x_0 + z_1 e_1 + z_2 e_2 + \cdots + z_r e_r,$$

where x_0 is a particular solution to the matrix equation $Ax = a$, while $e_1, e_2, \ldots, e_r$ are a set of linearly independent solutions to the homogeneous equation $Ax = 0$, and z_i, $i = 1, 2, \ldots, r$, are a series of arbitrary scalars.

Let

$$\begin{aligned} x &= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^{\mathrm{T}}, \\ x_0 &= \begin{bmatrix} x_1^0 & x_2^0 & \cdots & x_n^0 \end{bmatrix}^{\mathrm{T}}, \\ e_i &= \begin{bmatrix} e_{1i} & e_{2i} & \cdots & e_{ni} \end{bmatrix}^{\mathrm{T}}, \ i = 1, 2, \ldots, r, \end{aligned}$$

then the components of vector x can be written as

$$x_j = x_j^0 + z_1 e_{j1} + z_2 e_{j2} + \cdots + z_r e_{jr}, \ j = 1, 2, \ldots, n, \tag{s1.6}$$

and the affine function F can be expressed as

$$F(x) = F_0 + x_1 F_1 + x_2 F_2 + \cdots + x_n F_n. \tag{s1.7}$$

Substituting (s1.6) into (s1.7), yields,

$$\begin{aligned} F(x) &= F_0 + \left(x_1^0 + z_1 e_{11} + z_2 e_{12} + \cdots + z_r e_{1r}\right) F_1 \\ &\quad + \left(x_2^0 + z_1 e_{21} + z_2 e_{22} + \cdots + z_r e_{2r}\right) F_2 \\ &\quad + \cdots + \left(x_n^0 + z_1 e_{n1} + z_2 e_{n2} + \cdots + z_r e_{nr}\right) F_n \\ &= F_0 + x_1^0 F_1 + \cdots + x_n^0 F_n \\ &\quad + z_1 \left(e_{11} F_1 + e_{21} F_2 + \cdots + e_{n1} F_n\right) \\ &\quad + \cdots + z_r \left(e_{1r} F_1 + e_{2r} F_2 + \cdots + e_{nr} F_n\right). \end{aligned}$$

Put

$$\begin{aligned} &\widetilde{F}_0 = F_0 + x_1^0 F_1 + \cdots + x_n^0 F_n, \\ &\widetilde{F}_i = e_{1i} F_1 + e_{2i} F_2 + \cdots + e_{ni} F_n, \ i = 1, 2, \ldots, r, \\ &z = \begin{bmatrix} z_1 & z_2 & \cdots & z_r \end{bmatrix}^{\mathrm{T}}, \end{aligned}$$

we finally have

$$F(x) = \widetilde{F}_0 + z_1 \widetilde{F}_1 + \cdots + z_r \widetilde{F}_r \overset{\triangle}{=} \widetilde{F}(z).$$

This implies that $x \in \mathbb{R}^n$ satisfies (1.22) if and only if $\widetilde{F}(z) < 0$, $z \in \mathbb{R}^r$.

Exercise 1.7

Write the Hermite matrix $A \in \mathbb{C}^{n\times n}$ as $X + iY$ with real X and Y. Show that $A < 0$ only if $X < 0$.

Solution. Considering the conjugate symmetry of matrix A, we know that

$$X^{\mathrm{T}} = X, \quad Y^{\mathrm{T}} = -Y.$$

Therefore, for arbitrary $z \in \mathbb{R}^n$, we have

$$\left(z^{\mathrm{T}} Y z\right)^{\mathrm{T}} = z^{\mathrm{T}} Y^{\mathrm{T}} z = -\left(z^{\mathrm{T}} Y z\right),$$

which results in

$$z^{\mathrm{T}} Y z = 0, \ \forall z \in \mathbb{R}^n. \tag{s1.8}$$

Using the above relation, we further have

$$\begin{aligned} z^{\mathrm{T}}(X + iY)z &= z^{\mathrm{T}} X z + i z^{\mathrm{T}} Y z \\ &= z^{\mathrm{T}} X z. \end{aligned} \tag{s1.9}$$

When $A < 0$, we have

$$z^{\mathrm{T}}(X + iY)z < 0, \ \forall z \in \mathbb{R}^n, \ z \neq 0$$

this, together with (s1.9), implies

$$z^{\mathrm{T}} X z < 0, \ \forall z \in \mathbb{R}^n, \ z \neq 0.$$

This gives the negative definiteness of X.

Exercise 1.8

Let A, B be symmetric matrices of the same dimension. Show

1. $A > B$ implies $\lambda_{\max}(A) > \lambda_{\max}(B)$,
2. $\lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B)$.

Solution. Let $M \in \mathbb{S}^n$, $\lambda_{\max}(M)$ be the maximum eigenvalue of matrix M. We can easily show that

$$\lambda_{\max}(M) I \geq M.$$

Then, there holds

$$\lambda_{\max}(M) x^{\mathrm{T}} x \geq x^{\mathrm{T}} M x, \ \ \forall x \in \mathbb{R}^n. \tag{s1.10}$$

Proof of conclusion 1

Let x be the eigenvector of matrix B corresponding to the eigenvalue $\lambda_{\max}(B)$, then

$$Bx = \lambda_{\max}(B) x.$$

Considering $A > B$, we have

$$x^{\mathrm{T}}(A-B)x > 0,$$

which means

$$x^{\mathrm{T}}Ax > x^{\mathrm{T}}Bx = \lambda_{\max}(B)x^{\mathrm{T}}x. \tag{s1.11}$$

On the other hand, using (s1.10) and (s1.11), gives

$$\lambda_{\max}(A)x^{\mathrm{T}}x > \lambda_{\max}(B)x^{\mathrm{T}}x,$$

which implies $\lambda_{\max}(A) > \lambda_{\max}(B)$, in view of $x^{\mathrm{T}}x > 0$, $x \neq 0$.

Proof of conclusion 2

Let x be the eigenvector of matrix $A+B$ corresponding to the eigenvalue $\lambda_{\max}(A+B)$, then

$$\lambda_{\max}(A+B)x = (A+B)x,$$

from which we have

$$\lambda_{\max}(A+B)x^{\mathrm{T}}x = x^{\mathrm{T}}Ax + x^{\mathrm{T}}Bx. \tag{s1.12}$$

Using (s1.10) again, we obtain

$$x^{\mathrm{T}}Ax \leq \lambda_{\max}(A)x^{\mathrm{T}}x,\ x^{\mathrm{T}}Bx \leq \lambda_{\max}(B)x^{\mathrm{T}}x. \tag{s1.13}$$

Combining (s1.12) with (s1.13), yields

$$\lambda_{\max}(A+B)x^{\mathrm{T}}x \leq (\lambda_{\max}(A) + \lambda_{\max}(B))x^{\mathrm{T}}x,$$

which clear implies, in view of $x^{\mathrm{T}}x > 0$, $x \neq 0$, the relation to be proven.