Part I

Mathematical Background and Linear Matrix Inequalities in Control Theory

Lecture 1

Mathematical Background

The fundamental characteristics of quadratic forms are addressed in the first lecture. The positive definitiveness of partitioned matrices is investigated using Schur's complement lemma. Finsler's lemma is provided, as well as the so-called S - method, which deals with extra restricting quadratic forms.

1.1 Quadratic forms

1.1.1 Nonnegative and positive definite matrices

Consider the Riemann space $\mathbb{R}^n$ of vectors and the quadratic form $S(x) = x^\intercal \bar{S} x$ with a matrix $\bar{S} \in \mathbb{R}^{n \times n}$. Since

$$S(x) = x^\intercal \bar{S} x = \left(x, \bar{S}x\right) = \left(\bar{S}^\intercal x, x\right) = \left(x, \bar{S}^\intercal x\right) = \left(x, \frac{\bar{S} + \bar{S}^\intercal}{2} x\right) = (x, Sx),$$
$$S = \frac{\bar{S} + \bar{S}^\intercal}{2} = S^\intercal$$

we may suppose hereafter that in any quadratic form $S(x) = (x, Sx)$ the matrix S is *symmetric*.

Nonnegative definiteness

Definition 1.1 *A* ***symmetric*** *matrix $S \in \mathbb{R}^{n \times n}$ is said to be* ***nonnegative definite*** *if*

$$\boxed{x^\intercal S x \geq 0} \tag{1.1}$$

for all $x \in \mathbb{R}^n$.

The next simple lemma (given without proof) holds.

Lemma 1.1 *The following statements are equivalent:*

1. *S is nonnegative definite;*

2. *S may be represented as*
$$\boxed{S = HH^{\intercal}} \tag{1.2}$$
for some matrix H;

3. *The eigenvalues of S are nonnegative, that is, for all $i = 1, ..., n$*
$$\boxed{\lambda_i(S) \geq 0\,;} \tag{1.3}$$

4. *There is a symmetric matrix $R \in \mathbb{R}^{n\times n}$ such that*
$$\boxed{S = R^2\,;} \tag{1.4}$$
R is called ***the square root*** *of S, and is denoted by the symbol $S^{1/2} = R$.*

Positive definiteness

Definition 1.2 *If S is nonnegative and nonsingular $(\det S \neq 0)$, it is said to be* ***positive definite****.*

Remark 1.1 *In the case when S is positive definite we have*
$$x^{\intercal} Sx > 0 \text{ for all } x \neq 0$$
and $S^{1/2}$ is also positive definite so that for all $x \neq 0$
$$x^{\intercal} S^{1/2} x > 0 \text{ for all } x \neq 0.$$

The statement "S is nonnegative definite" is abbreviated
$$S \geq 0,$$
and, similarly,
$$S > 0$$
means "S is positive definite".

Remark 1.2 *The abbreviation*

$$\boxed{A \geq B \ (or\ A > B)} \tag{1.5}$$

may be applied only to two symmetric matrices of the same size and means that

$$A - B \geq 0 \ (or\ A - B > 0).$$

Remark 1.3 *Evidently, if $A > 0$, then for any quadratic nonsingular matrix T $(\det T \neq 0)$ it follows*

$$TAT^{\intercal} > 0$$

and inverse, if $TAT^{\intercal} > 0$ for some nonsingular matrix T, then $A > 0$.

Remark 1.4 *If $A \geq B$ (or $A > B$), then for any quadratic nonsingular T $(\det T \neq 0)$*

$$TAT^{\intercal} \geq TBT^{\intercal} \ (or\ TAT^{\intercal} > TBT^{\intercal})$$

and, inverse, if $TAT^{\intercal} \geq TBT^{\intercal}$ (or $TAT^{\intercal} > TBT^{\intercal}$) for some nonsingular T, then $A \geq B$ (or $A > B$). In the special case $A > B$ there exists an orthogonal matrix T ($T^{\intercal} = T^{-1}$) transforming $A - B$ to a diagonal matrix with positive elements.

Proposition 1.1 *If*

$$\boxed{A > B > 0,}$$

then

$$\boxed{B^{-1} > A^{-1} > 0\,.}$$

Proof. Let T_A be an orthogonal transformation which transforms A to a diagonal matrix $\Lambda_A := \operatorname{diag}(\lambda_1(A), ..., \lambda_n(A))$ and

$$\Lambda_A = \Lambda_A^{1/2}\Lambda_A^{1/2}, \ \Lambda_A^{1/2} = \operatorname{diag}\left(\sqrt{\lambda_1(A)}, ..., \sqrt{\lambda_n(A)}\right) > 0\,.$$

Then, by the previous remark,

$$T_A A T_A^{\intercal} = \Lambda_A > T_A B T_A^{\intercal}$$

and, as the result,

$$I_{n\times n} > \Lambda_A^{-1/2} T_A B T_A^{\intercal} \Lambda_A^{-1/2}\,.$$

Denoting by T an orthogonal transformation which transforms the right-hand side of of the last inequality to a diagonal matrix Λ, we obtain

$$I_{n\times n} = TT^{\intercal} > T\left(\Lambda_A^{-1/2} T_A B T_A^{\intercal} \Lambda_A^{-1/2}\right) T^{\intercal} = \Lambda = \operatorname{diag}(\lambda_1, ..., \lambda_n)\,.$$

Inverting this inequality by components, one has

$$I_{n\times n} < \Lambda^{-1} = \left[T\left(\Lambda_A^{-1/2}T_A B T_A^{\intercal}\Lambda_A^{-1/2}\right)T^{\intercal}\right]^{-1} = \\ T\left(\Lambda_A^{1/2}T_A B^{-1} T_A^{\intercal}\Lambda_A^{1/2}\right)T^{\intercal},$$

that implies

$$I_{n\times n} < \Lambda_A^{1/2}T_A B^{-1} T_A^{\intercal}\Lambda_A^{1/2}$$

and

$$\Lambda_A^{-1} < T_A B^{-1} T_A^{\intercal}.$$

Hence,

$$T_A^{\intercal}\Lambda_A^{-1}T_A = A^{-1} < B^{-1}.$$

Proposition is proven. ■

Proposition 1.2 *If $S \geq 0$ and $T \geq 0$, then*

$$S + T \geq 0$$

with strict inequality holding if and only if

$$\mathcal{N}(S) \cap \mathcal{N}(T) = \varnothing$$

where

$$\mathcal{N}(S) = \ker S := \{x \mid Sx = 0\}$$

is the ***kernel*** *(or, nule-space) of the matrix S.*

The proofs of these statements are evident.

1.1.2 Positive definiteness of a partitioned matrix: Schur's complement

Theorem 1.1 (Schur's complement) *Let S be a square matrix partitioned as*

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^{\intercal} & S_{22} \end{bmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}$$

where $S_{11} \in \mathbb{R}^{n\times n}$ is a symmetric $n \times n$ matrix and $S_{22} \in \mathbb{R}^{m\times m}$ is a symmetric $m \times m$ matrix. Then $S > 0$ if and only if

$$\boxed{\begin{array}{c} S_{11} > 0, \\ S_{22} > 0, \\ S_{11} - S_{12}S_{22}^{-1}S_{12}^{\intercal} > 0, \\ S_{22} - S_{12}^{\intercal}S_{11}^{-1}S_{12} > 0, \end{array}} \tag{1.6}$$

moreover,

$$\boxed{\begin{aligned} \det S &= \det S_{11} \det\left(S_{22} - S_{12}^{\intercal} S_{11}^{-1} S_{12}\right) \\ &= \det S_{22} \det\left(S_{11} - S_{12} S_{22}^{-1} S_{12}^{\intercal}\right) . \end{aligned}} \tag{1.7}$$

Proof. *Necessity.* Suppose that $S > 0$. Then for any $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ ($x_1 \in \mathbb{R}^n$, $x_2 \in \mathbb{R}^m$) non equal to zero we have

$$0 < x^{\intercal} S x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{\intercal} \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^{\intercal} & S_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= x_1^{\intercal} S_{11} x_1 + x_1^{\intercal} S_{12} x_2 + x_2^{\intercal} S_{12}^{\intercal} x_1 + x_2^{\intercal} S_{22} x_2 .$$

Taking $x_2 = 0$ it follows $x_1^{\intercal} S_{11} x_1 > 0$, or, equivalently, $S_{11} > 0$. Analogously, Taking $x_1 = 0$ we get $x_2^{\intercal} S_{11} x_2 > 0$, or, equivalently, $S_{22} > 0$. Putting

$$x_2 = -S_{22}^{-1} S_{12} x_1$$

we get

$$0 < x_1^{\intercal} S_{11} x_1 - x_1^{\intercal} S_{12} S_{22}^{-1} S_{12}^{\intercal} x_1 - x_1^{\intercal} S_{12} S_{22}^{-1} S_{12}^{\intercal} x_1 +$$

$$x_1^{\intercal} S_{12} S_{22}^{-1} S_{22} S_{22}^{-1} S_{12}^{\intercal} x_1 = x_1^{\intercal} \left(S_{11} - S_{12} S_{22}^{-1} S_{12}^{\intercal}\right) x_1 ,$$

or, equivalently, $S_{11} - S_{12} S_{22}^{-1} S_{12}^{\intercal} > 0$. Analogously, taking

$$x_1 = -S_{11}^{-1} S_{12} x_2$$

we get

$$0 < x_2^{\intercal} S_{12}^{\intercal} S_{11}^{-1} S_{11} S_{11}^{-1} S_{12} x_2 - x_2^{\intercal} S_{12}^{\intercal} S_{11}^{-1} S_{12} x_2 - x_2^{\intercal} S_{12}^{\intercal} S_{11}^{-1} S_{12} x_2 + x_2^{\intercal} S_{22} x_2$$

$$= x_2^{\intercal} \left(S_{22} - S_{12}^{\intercal} S_{11}^{-1} S_{12}\right) x_1 ,$$

or, equivalently, $S_{22} - S_{12}^{\intercal} S_{11}^{-1} S_{12} > 0$. Analogousely, we can prove that $S_{11} - S_{12} S_{22}^{-1} S_{12}^{\intercal} > 0$.

This proves necessity of (1.6).

Sufficiency. Suppose that (1.6) holds. Define matrices

$$A := \left(S_{11} - S_{12} S_{22}^{-1} S_{12}^{\intercal}\right)^{-1} ,$$

$$C := \left(S_{22} - S_{12}^{\intercal} S_{11}^{-1} S_{12}\right)^{-1} ,$$

$$B := -S_{11}^{-1} S_{12} C .$$

It is easy to show that

$$B = -S_{11}^{-1}S_{12}C = -A\left[A^{-1}S_{11}^{-1}S_{12}C\right] =$$

$$-A\left[\left(S_{11} - S_{12}S_{22}^{-1}S_{12}^{\intercal}\right)S_{11}^{-1}S_{12}\left(S_{22} - S_{12}^{\intercal}S_{11}^{-1}S_{12}\right)^{-1}\right] = -$$

$$A\left[\left(S_{11}S_{11}^{-1}S_{12} - S_{12}S_{22}^{-1}S_{12}^{\intercal}S_{11}^{-1}S_{12}\right)\left(I_{m\times m} - S_{22}^{-1}S_{12}^{\intercal}S_{11}^{-1}S_{12}\right)^{-1}S_{22}^{-1}\right]$$

$$= -A\left[S_{12}\left(I_{m\times m} - S_{22}^{-1}S_{12}^{\intercal}S_{11}^{-1}S_{12}\right)\left(I_{m\times m} - S_{22}^{-1}S_{12}^{\intercal}S_{11}^{-1}S_{12}\right)^{-1}S_{22}^{-1}\right]$$

$$= -AS_{12}S_{22}^{-1}.$$

Then routine calculations verify that

$$S\begin{bmatrix} A & B \\ B^{\intercal} & C \end{bmatrix} = I_{(n+m)\times(n+m)}$$

So, $S^{-1} = \begin{bmatrix} A & B \\ B^{\intercal} & C \end{bmatrix}$ exists and, hence, S is nonsingular and $S > 0$.

The formula (1.7) results from the matrix identity

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{12}^{\intercal} & S_{22} \end{bmatrix} =$$

$$\begin{bmatrix} I_{n\times n} & 0 \\ S_{12}^{\intercal}S_{11}^{-1} & I_{m\times m} \end{bmatrix}\begin{bmatrix} S_{11} & 0 \\ 0 & S_{22} - S_{12}^{\intercal}S_{11}^{-1}S_{12} \end{bmatrix}\begin{bmatrix} I_{n\times n} & S_{11}^{-1}S_{12} \\ 0 & I_{m\times m} \end{bmatrix}$$

and the following determinat's properties

$$\begin{gathered} \det(FG) = \det(F)\det(G) \\ \det\left(\begin{bmatrix} F & 0 \\ H & G \end{bmatrix}\right) = \det(F)\det(G). \end{gathered}$$

■

Corollary 1.1 *A matrix S is negative definite ($S < 0$) if and only if*

$$\boxed{\begin{gathered} S_{11} < 0, \\ S_{22} < 0, \\ S_{11} - S_{12}S_{22}^{-1}S_{12}^{\intercal} < 0, \\ S_{22} - S_{12}^{\intercal}S_{11}^{-1}S_{12} < 0. \end{gathered}} \tag{1.8}$$

Corollary 1.2 *To guarantee that the matrix $S > 0$ it is necessary that all diagonal ellements would be positive.*

Corollary 1.3 *Suppose that in the previous theorem $m = 1$, that is, the following representation holds*

$$\boxed{\begin{array}{c} S_{n+1} = \begin{bmatrix} S_n & s_n \\ s_n^\intercal & \sigma_{n+1} \end{bmatrix}, \\ \\ s_n \in \mathbb{R}^n, \ \sigma_{n+1} \in \mathbb{R}, \end{array}} \tag{1.9}$$

where $0 < S_n \in \mathbb{R}^{n\times n}$. Then $S_{n+1} > 0$ if and only if

$$\boxed{\alpha_n = \sigma_{n+1} - s_n^\intercal S_n^{-1} s_n > 0} \tag{1.10}$$

and, as the result,

$$\boxed{S_{n+1}^{-1} = \begin{bmatrix} S_n^{-1} + \left[S_n^{-1} s_n s_n^\intercal S_n^{-1}\right] \alpha_n^{-1} & -\left(S_n^{-1} s_n\right) \alpha_n^{-1} \\ -\left(S_n^{-1} s_n\right)^\intercal \alpha_n^{-1} & \alpha_n^{-1} \end{bmatrix}.} \tag{1.11}$$

1.1.3 Sylvester's criterion

Here we present a simple proof of the known criterion which gives a power instrument for numerical test of positive definiteness.

Theorem 1.2 (Sylvester's criterion) *A symmetric matrix $S \in \mathbb{R}^{n\times n}$ is positive definite if and only if* ***all leading principle minors***

$$A\begin{pmatrix} 1 & 2 & \cdots & p \\ 1 & 2 & \cdots & p \end{pmatrix} := \begin{vmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{vmatrix}, \ p = 1, 2, ..., n$$

are ***strictly positive****, that is, for all $p = 1, 2, ..., n$*

$$\boxed{A\begin{pmatrix} 1 & 2 & \cdots & p \\ 1 & 2 & \cdots & p \end{pmatrix} > 0\,.} \tag{1.12}$$

Proof. Let us prove this result by the induction method implementation. For $n = 2$ the result is evident. Indeed, for

$$S = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}$$

under the assumption that $a_{11} \neq 0$, we have

$$x^\intercal S x = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 =$$
$$a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}\right)^2 + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 .$$

from which it follows that $x^\intercal S x > 0$ ($x \neq 0$), or equivalently, $S > 0$ if and only if

$$a_{11} > 0,\ a_{22} - \frac{a_{12}^2}{a_{11}} = \det S > 0 .$$

Let us represent $S \in \mathbb{R}^{n\times n}$ in the form (1.9)

$$S_n = \begin{bmatrix} S_{n-1} & s_{n-1} \\ s_{n-1}^\intercal & \sigma_n \end{bmatrix},$$
$$s_{n-1} \in \mathbb{R}^{n-1},\ \sigma_n \in \mathbb{R}$$

and suppose that $S_{n-1} > 0$. This implies that $\det S_{n-1} > 0$. Then by (1.3) $S_n > 0$ if and only if the condition (1.10) holds, that is, when

$$\alpha_{n-1} = \sigma_n - s_{n-1}^\intercal S_{n-1}^{-1} s_{n-1} > 0 .$$

But by the Schur's formulas (1.6) and (1.7) we have

$$A\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix} = \det S = \det\left(\sigma_n - s_{n-1}^\intercal S_{n-1}^{-1} s_{n-1}\right)(\det S_{n-1}) =$$
$$\left(\sigma_n - s_{n-1}^\intercal S_{n-1}^{-1} s_{n-1}\right)(\det S_{n-1}) = \alpha_{n-1}(\det S_{n-1}) > 0$$

if and only if (1.10) holds, that proves the result. ∎

1.1.4 Nonnegative definiteness of a partitioned matrix

Theorem 1.3 *Let S be a square matrix partitioned as*

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^\intercal & S_{22} \end{bmatrix},$$

where S_{11} is a symmetric $n\times n$ matrix and S_{22} is a symmetric $m\times m$ matrix. Then $S \geq 0$ if and only if

$$\boxed{\begin{array}{c} S_{11} \geq 0, \\ S_{22} \geq 0, \\ S_{11}S_{11}^{+}S_{12} = S_{12}, \\ S_{22}S_{22}^{+}S_{12}^\intercal = S_{12}^\intercal, \\ S_{22} - S_{12}^\intercal S_{11}^{+} S_{12} \geq 0, \\ S_{11} - S_{12}S_{22}^{+}S_{12}^\intercal \geq 0. \end{array}} \tag{1.13}$$

Here the H^+ is the matrix, pseudoinversed (in the Moore-Penrouse sence) to H, satisfying the identities

$$\boxed{\begin{array}{c} HH^+H = H,\ H^+HH^+ = H^+, \\ (HH^+)^\intercal = HH^+,\ H^+ = H^\intercal (HH^\intercal)^+ . \end{array}} \tag{1.14}$$

Proof. *Necessity.* Suppose that $S \geq 0$. Then there exist a matrix H with $(n+m)$ rows such that $S = HH^\intercal$. Let us write H as a partitioned matrix

$$H = \begin{bmatrix} X \\ Y \end{bmatrix},\ X \in \mathbb{R}^{n\times n}, Y \in \mathbb{R}^{m\times m}$$

Then

$$S = HH^\intercal = \begin{bmatrix} XX^\intercal & XY^\intercal \\ YX^\intercal & YY^\intercal \end{bmatrix},$$

so that

$$S_{11} = XX^\intercal \geq 0,\ S_{12} = XY^\intercal$$

By (1.14)

$$S_{11}S_{11}^+ = (XX^\intercal)(XX^\intercal)^+ = X\left[X^\intercal (XX^\intercal)^+\right] = XX^+,$$

so that

$$S_{11}S_{11}^+S_{12} = XX^+(XY^\intercal) = \left(XX^+X\right)Y^\intercal = XY^\intercal = S_{12}$$

Finally, if we let

$$U := Y - S_{12}^\intercal S_{11}^+ X$$

then

$$0 \leq UU^\intercal = S_{22} - S_{12}^\intercal S_{11}^+ S_{12}.$$

Analogously, the other two relations follows by changing X to Y.

Sufficiency. Let (1.13) holds. Define

$$U := \begin{bmatrix} I_{n\times n} & \vdots & \mathrm{O}_{n\times m} \end{bmatrix},\ V := \begin{bmatrix} \mathrm{O}_{m\times n} & \vdots & I_{m\times m} \end{bmatrix}$$
$$X := S_{11}^{1/2}U$$
$$Y := S_{12}^\intercal S_{11}^+ S_{11}^{1/2} U + \left(S_{22} - S_{12}^\intercal S_{11}^+ S_{12}\right)^{1/2} V$$

Here $\mathrm{O}_{n\times m}$ and $\mathrm{O}_{m\times n}$ are the matrices with null-elements. Since

$$UV^\intercal = \mathrm{O}_{n\times m}$$

one can see that

$$0 \leq \begin{bmatrix} X \\ Y \end{bmatrix}\begin{bmatrix} X \\ Y \end{bmatrix}^\intercal = \begin{bmatrix} XX^\intercal & XY^\intercal \\ YX^\intercal & YY^\intercal \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^\intercal & S_{22} \end{bmatrix} = S.$$

■

Corollary 1.4 *Suppose that in the previous theorem $m = 1$, that is, the following representation holds*

$$\boxed{\begin{array}{c} S_{n+1} = \begin{bmatrix} S_n & s_n \\ s_n^\intercal & \sigma_{n+1} \end{bmatrix}, \\ s_n \in \mathbb{R}^n, \ \sigma_{n+1} \in \mathbb{R}, \end{array}} \tag{1.15}$$

where $0 \le S_n \in \mathbb{R}^{n\times n}$. Let

$$\begin{array}{c} t_n := S_n^+ s_n, \ \alpha_n := \sigma_{n+1} - s_n^\intercal S_n^+ s_n, \\ \beta_n := 1 + \|t_k\|^2, \ T_n - t_n t_n^\intercal / \beta_n. \end{array}$$

Then $S_{n+1} \ge 0$ if and only if

$$\boxed{S_n t_n = s_n \ \textit{and} \ \ \alpha_n \ge 0} \tag{1.16}$$

and

$$\boxed{S_{n+1}^+ = \left\{ \begin{array}{cc} \begin{bmatrix} S_n^+ + t_n t_n^\intercal \alpha_n^{-1} & -t_n \alpha_n^{-1} \\ -t_n^\intercal \alpha_n^{-1} & \alpha_n^{-1} \end{bmatrix} & \textit{if} \quad \alpha_n > 0 \\ \begin{bmatrix} T_n S_n^+ T_n & T_n S_n^+ t_n \beta_n^{-1} \\ \left(T_n S_n^+ t_n \beta_n^{-1}\right)^\intercal & \left(t_n^\intercal S_n^+ t_n\right) \beta_n^{-2} \end{bmatrix} & \textit{if} \quad \alpha_n = 0 \end{array} \right.} \tag{1.17}$$

The proof of this corollary follows directly from the previous theorem and the application of the Cline's formula (1.18):

$$\begin{array}{c} \left[U \vdots V \right]^+ = \begin{bmatrix} U^+ - U^+ V J \\ J \end{bmatrix}, \\ J = C^+ + (I - C^+ C) K V^\intercal (U^+)^\intercal U^+ (I - V C^+), \\ C = (I - U U^+) V, \\ K = \left(I + [U^+ V (I - C^+ C)]^\intercal [U^+ V (I - C^+ C)]\right)^{-1}. \end{array} \tag{1.18}$$

1.2 Finsler's lemma and S - procedure

1.2.1 Finsler's lemma

Finsler's lemma [1] is also known as D´ebreu's lemma. Here we present its slightly modified version.

Lemma 1.2 (the modified version of Finsler's lemma) *For any two matrices $Q = Q^\intercal \in \mathbb{R}^{n\times n}$ and $F \in \mathbb{R}^{m\times n}$ (rank$F < n$) the following statements are equivalent:*

- $$\boxed{x^{\intercal}Qx < 0 \textit{ for any } x \in \{x \in \mathbb{R}^n \mid Fx = 0\}} \tag{1.19}$$

- *there exists $\mu > 0$ such that*

$$\boxed{Q - \mu F^{\intercal}F < 0} \tag{1.20}$$

for all $x \neq 0$.

Proof.

1) *Necessity.* Show that (1.19) implies (1.20). Represent $\mathbb{R}^n$ as the direct sum

$$\mathbb{R}^n = \mathcal{N}(F) \oplus \mathcal{R}(F)$$

where

$-\mathrm{Ker}F = \mathcal{N}(F) := \{x \in \mathbb{R}^n : Fx = 0\}$ is the kernel (or null space) of the linear transformation $F : \mathbb{R}^n \longmapsto \mathbb{R}^m$,

$-\,\mathrm{Im}\,F = \mathcal{R}(F) := \{y \in \mathbb{R}^m : y = Fx,\ x \in \mathbb{R}^n\}$ - the image (or range) of the linear transformation $F : \mathbb{R}^n \longmapsto \mathbb{R}^m$.

Let $\mathbb{R}$ is the set of real numbers. Select the special basis such that the matrices Q and $F^{\intercal}F$ can be represented as

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^{\intercal} & Q_{22} \end{bmatrix} \text{ and } F^{\intercal}F = \begin{bmatrix} 0 & 0 \\ 0 & D^{\intercal}D \end{bmatrix}$$

where

$$0 > Q_{11} \in \mathbb{R}^{(n-m)\times(n-m)},\ 0 > Q_{22} \in \mathbb{R}^{m\times m},$$

$$0 < D^{\intercal}D \in \mathbb{R}^{m\times m},\ \mathrm{rank}D = m\,.$$

Then for any $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $x_1 \in \mathbb{R}^{n-m}$, $x_2 \in \mathbb{R}^m$ we have

$$x^{\intercal}\left(Q - \mu F^{\intercal}F\right)x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^{\intercal} \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^{\intercal} & Q_{22} - \mu D^{\intercal}D \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$

$$x_1^{\intercal}Q_{11}x_1 + 2x_1^{\intercal}Q_{12}x_2 + x_2^{\intercal}\left(Q_{22} - \mu D^{\intercal}D\right)x_2 =$$

$$\left(x_1 + Q_{11}^{-1}Q_{12}x_2\right)^{\intercal} Q_{11}\left(x_1 + Q_{11}^{-1}Q_{12}x_2\right) +$$

$$x_2^{\intercal}\left(Q_{22} - Q_{12}^{\intercal}Q_{11}^{-1}Q_{12} - \mu DD^{\intercal}\right)x_2 \leq$$

$$x_2^{\intercal}\left(Q_{22} - Q_{12}^{\intercal}Q_{11}^{-1}Q_{12} - \mu DD^{\intercal}\right)x_2 < 0$$

if μ satisfies the inequality

$$Q_{22} - Q_{12}^{\mathsf{T}} Q_{11}^{-1} Q_{12} - \mu D^{\mathsf{T}} D < 0\,.$$

Hence it exists satisfying the inequality

$$\mu > \lambda_{\max}\left(D^{-1}\left[Q_{22} - Q_{12}^{\mathsf{T}} Q_{11}^{-1} Q_{12}\right](D^{\mathsf{T}})^{-1}\right)\,.$$

2) *Sufficiency.* Show now that (1.20) implies (1.19). Evidently that

$$x^{\mathsf{T}}\left(Q - \mu F^{\mathsf{T}} F\right) x = x^{\mathsf{T}} Q x < 0$$

for all x such that $Fx = 0$. ■

Corollary 1.5 *Two statements are equivalent:*

- *For some nonsingular X*

$$\boxed{X^{\mathsf{T}} Q X < 0 \text{ and } FX = 0\,.} \tag{1.21}$$

- *There exists $\mu > 0$ such that*

$$\boxed{Q - \mu F^{\mathsf{T}} F < 0\,.} \tag{1.22}$$

Proof. Defining $z := Xx$, the relations (1.21) equivalently can be represented as

$$z^{\mathsf{T}} Q z < 0,\; Fz = 0$$

which by Finsler's lemma 1.2 leads to (1.22). ■

1.2.2 S - Procedure (lemma)

S - Procedure deals with nonnegativity of a quadratic form under quadratic constrains as inequalities. In 1971, Yakubovich [2] proved S - Lemma which became very popular in control theory. There exist several methods to prove it (see for example [3], [4]) but we want to give here a proof that uses Dines' theorem to emphasize the link between convexity and S - Lemma which is a separation theorem for convex sets. Putinar's Positivstellensatz [12] can be viewed as a broad generalization of the S - procedure.

The case of two quadratic forms

Theorem 1.4 (S - Lemma: neccessary and sufficent condition) *Let A, B be symmetric $n \times n$ matrices, and assume that the quadratic inequality*

$$\mathcal{A}(x) = x^{\intercal} A x \geq 0 \tag{1.23}$$

is strictly feasible for some set of argument $x \in \mathbb{R}^n$, that is, there exists $\mathring{x}$ such that $\mathring{x}^{\intercal} A \mathring{x} > 0$. Then the quadratic inequality:

$$\mathcal{B}(x) = x^{\intercal} B x \geq 0 \tag{1.24}$$

is a consequence of (1.23), i.e.,

$$x^{\intercal} A x \geq 0 \Rightarrow x^{\intercal} B x \geq 0$$

(in other word, (1.24) holds for all $x \in \mathbb{R}^n$, satisfying (1.23)), if and only if there exists a nonnegative $\tau \geq 0$ such that for all $x \in \mathbb{R}^n$

$$\mathcal{B}(x) - \tau \mathcal{A}(x) \geq 0, \tag{1.25}$$

or equivalently

$$\boxed{B - \tau A \geq 0.}$$

Proof. *a) Sufficiency.* The sufficiency part immediately follows. Indeed,

$$x^{\intercal} B x \geq \tau x^{\intercal} A x \geq 0\,.$$

b) Neccessity. To prove neccesity let us assume that $x^{\intercal} B x \geq 0$ is a consequence of $x^{\intercal} A x \geq 0$. Define the sets

$$S := \{(x^{\intercal} A x, x^{\intercal} B x) : x \in \mathbb{R}^n\}$$

and

$$U = \{(u_1, u_2) : u_1 \in \mathbb{R}_+ = \{u \in \mathbb{R} : u > 0\}\,, u_2 \in \mathbb{R}_- = \{u \in \mathbb{R} : u < 0\}\}.$$

S is a convex set by Dines' theorem while U is a convex cone. Since their intersection is empty, a separating hyperplane exists, i.e., there exists nonzero $c = (c_1, c_2) \in \mathbb{R}^2$, such that $(c, s) \leq 0, \forall s \in S$ and $(c, u) \geq 0, \forall u \in U$. For $(0, -1) \in U$ we have $c_2 \leq 0$. For $(1, -\alpha) \in U$ where α is a small positive number arbitrarily chosen, we obtain $c_1 \geq \alpha c_2$. Letting α tend to zero, we get $c_1 \geq 0$. Since there exists $\mathring{x}$ such that $\mathring{x}^{\intercal} A \mathring{x} > 0$ and by the separation argument we have

$$c_1 x^{\intercal} A x + c_2 x^{\intercal} B x \leq 0$$

for all $x \in \mathbb{R}^n$ implying

$$c_1 \mathring{x}^\intercal A \mathring{x} + c_2 \mathring{x}^\intercal B \mathring{x} \leq 0.$$

Since $c_1 \geq 0$ and by hypothesis $\mathring{x}^\intercal A \mathring{x} > 0$ and $\mathring{x}^\intercal B \mathring{x} \geq 0$, and taking into account that c_1 and c_2 cannot both be zero, the last inequality implies that $c_2 < 0$. Indeed,

$$c_2 \mathring{x}^\intercal B \mathring{x} \leq -c_1 \mathring{x}^\intercal A \mathring{x} < 0$$

Therefore, we obtain:

$$x^\intercal B x \leq -\frac{c_1}{c_2} x^\intercal A x$$

for all $x \in \mathbb{R}^n$, which is equivalent to $B - \tau A \geq 0$ after defining $\tau = -\dfrac{c_1}{c_2}$. This completes the proof of the necessity part. Hence, the result is proven. ■

Remark 1.5 *If Theorem 1.4 is valid for some positive parameter τ, then it can be taken equal to 1, and the condition (1.25) may be always considered in the simplified form as*

$$\boxed{\mathcal{B}(x) - \mathcal{A}(x) \geq 0} \tag{1.26}$$

for all $x \in \mathbb{R}^n$. Indeed, if (1.25) holds for $\tau > 0$, then the inequality (1.23) can be equivalently represented as

$$x^\intercal \left(\frac{1}{\tau} A \right) x \geq 0$$

that transforms (1.25) directly to (1.26).

The case of m ≥ 2 quadratic forms

Consider the collection of quadratic forms

$$\mathcal{A}_i(x) := x^\intercal A_i x \ (i = 1, 2, ..., m), \ \mathcal{B}(x) := x^\intercal B x$$

where $x \in \mathbb{R}^n$ and A, B_i are symmetric $n \times n$ matrices.

Definition 1.3 *S - Procedure consists in the creation of the quadratic form*

$$\boxed{S(x) = \mathcal{B}(x) - \tau_1 \mathcal{A}_1(x) - ... - \tau_m \mathcal{A}_m(x)} \tag{1.27}$$

for some nonnegative scalars $\tau_i \geq 0$ $(i = 1, 2, ..., m)$.

In fact, the nonnegative parameters τ_i play the role of *Lagrange multipliers* for quadratic forms.

Lemma 1.3 (sufficient condition) *Let the inequalities*

$$\mathcal{A}_i(x) = x^\intercal A_i x \geq 0 \; (i = 1, 2, ..., m) \tag{1.28}$$

hold. Then the quadratic inequality:

$$\mathcal{B}(x) := x^\intercal B x \geq 0$$

is consequences of (1.28), i.e.,

$$\bigcap_{i=1}^{m} (x^\intercal A_i x \geq 0) \Rightarrow x^\intercal B x \geq 0$$

if there exists nonnegative τ_i such that

$$S(x) = \mathcal{B}(x) - \tau_1 \mathcal{A}_1(x) - ... - \tau_m \mathcal{A}_m(x) \geq 0 \tag{1.29}$$

or equivalently

$$B - \sum_{i=1}^{m} \tau_i A_i \geq 0$$

Proof. Indeed, suppose that (1.29) is met. Then

$$\mathcal{B}(x) \geq \sum_{i=1}^{m} \tau_i \mathcal{A}_i(x) \geq 0$$

Lemma is proven. ■

Remark 1.6 *It is important to note that for $m \geq 2$ the analogue of the neccessary condition in Theorem 1.4, is not valid, that is, S - procedure becomes to be the* ***flawed.***

Extension of Theorem 1.4

Theorem 1.5 *Let inequalities*

$$\boxed{\mathcal{A}_i(x) := x^\intercal A_i x \leq \alpha_i \; (i = 1, ..., m)} \tag{1.30}$$

imply

$$\boxed{\mathcal{B}(x) := x^\intercal B x \leq \alpha_0} \tag{1.31}$$

where α_i $(i = 0, 1, ..., m)$ *are some real numbers. If there exist* $\tau_i \geq 0$ $(i = 1, ..., m)$ *such that*

$$\boxed{B \leq \sum_{i=1}^{m} \tau_i A_i,\ \alpha_0 \geq \sum_{i=1}^{m} \tau_i \alpha_i,} \tag{1.32}$$

then (1.30) implies (1.31). Inversely, if (1.30) implies (1.31) and, additionally, one of the following conditions fulfilled:

1. $$\boxed{m = 1}$$

2. $$\boxed{m = 2,\ n \leq 3}$$

   and there exists a vector $x^{(0)}$, μ_1, μ_2 *such that*

   $$\boxed{\begin{array}{c} \mathcal{G}_i(x^{(0)}) < \alpha_i \ \ (i = 1, 2)\,, \\ \mu_1 A_1 + \mu_2 A_2 > 0, \end{array}}$$

   then there exist $\tau_i \geq 0$ $(i = 1, ..., m)$ *such that (1.32) holds.*

For $m > 2$ ***the analogue results is not true.***

Proof. *Sufficiency* is trivial. Indeed, the proof of this theorem can be converted to the conditions of lemma 1.3: the constraint

$$x^\intercal A_i x \leq \alpha_i$$

can be rewritten as

$$\begin{pmatrix} x \\ 1 \end{pmatrix}^\intercal \begin{bmatrix} A_i & 0 \\ 0 & -\alpha_i \end{bmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} \leq 0.$$

Therefore, by Lemma 1.3, we get

$$\begin{bmatrix} B & 0 \\ 0 & -\alpha_0 \end{bmatrix} - \sum_{i=1}^{m} \tau_i \begin{bmatrix} A_i & 0 \\ 0 & -\alpha_i \end{bmatrix} = \begin{bmatrix} B - \sum\limits_{i=1}^{m} \tau_i A_i & 0 \\ 0 & -\alpha_0 + \sum\limits_{i=1}^{m} \tau_i \alpha_i \end{bmatrix} \leq 0$$

which corresponds to (1.32).

Necessity follows from previous Theorem (1.4). Then simple counterexample may show that this theorem is not valid for $m > 2$. ■

1.3 Examples

Example 1.1 *Using Schur's lemma find for which α the matrix*

$$S = \begin{bmatrix} -1 & 1 & \alpha \\ 1 & -2 & 0 \\ \alpha & 0 & -1 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{12}^{\intercal} & S_{22} \end{bmatrix},$$

$$S_{11} = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix}, \; S_{12} = \begin{bmatrix} \alpha \\ 0 \end{bmatrix}, \; S_{22} = -1$$

is strictly negative, i.e.,

$$S < 0$$

Solution. By Schur's formula (1.8) $S < 0$ if and only if

$$S_{11} - S_{12}S_{22}^{-1}S_{12}^{\intercal} = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} - \begin{bmatrix} \alpha \\ 0 \end{bmatrix} [-1] \begin{bmatrix} \alpha & 0 \end{bmatrix} =$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} + \begin{bmatrix} \alpha^2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \alpha^2 - 1 & 1 \\ 1 & -2 \end{bmatrix} < 0\,,$$

or equivalently,

$$\begin{bmatrix} 1-\alpha^2 & -1 \\ -1 & 2 \end{bmatrix} > 0\,,$$

implying

$$1-\alpha^2 > 0, \; 2\left(1-\alpha^2\right) - 1 > 0$$

and

$$|\alpha| < 1 \text{ and } 2\alpha^2 < 1\,.$$

As the result we have that $S < 0$ for all α satisfying

$$\boxed{-\frac{1}{\sqrt{2}} < \alpha < \frac{1}{\sqrt{2}}\,.}$$

Example 1.2 *For which c the set $\mathcal{X}$ defined as*

$$\mathcal{X} = \left\{x \in \mathbb{R}^2 \mid x^{\intercal}Qx < 0 \textit{ for any } x \in \left\{x \in \mathbb{R}^2 \mid Fx = 0\right\}\right\}$$
$$n = 2, \; Q = \begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix}, \; F = \begin{pmatrix} 1 & c \end{pmatrix}$$

is non empty ($\mathcal{X} \neq \varnothing$).

Solution. By Finsler's lemma $\mathcal{X} \neq \varnothing$ if and only if there exists $\mu > 0$ such that

$$Q - \mu F^{\mathsf{T}} F < 0$$

for all $x \neq 0$ which is equivalent to the following condition

$$\begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix} - \mu \begin{bmatrix} 1 \\ c \end{bmatrix} \begin{bmatrix} 1 & c \end{bmatrix} =$$

$$\begin{bmatrix} 1 & -1 \\ -1 & -2 \end{bmatrix} - \mu \begin{bmatrix} 1 & c \\ c & 0 \end{bmatrix} = \begin{bmatrix} 1-\mu & -1-\mu c \\ -1-\mu c & -2 \end{bmatrix} < 0$$

which is met if

$$\begin{bmatrix} \mu - 1 & 1+\mu c \\ 1+\mu c & 2 \end{bmatrix} > 0\,,$$

or equivalently

$$\mu > 1,\ 2(\mu - 1) - (1+\mu c)^2 > 0\,.$$

The last inequality holds if

$$\phi(\mu) := -3 + 2\mu(c-1) - \mu^2 c^2 > 0$$

which is fulfilled together with $\mu > 1$ if

$$\max_{\mu>1} \phi(\mu) > 0\,.$$

This takes place if $0 = \phi'(\mu) = 2(c-1) - 2\mu c^2$ in the point $\mu = \mu^* > 1$:

$$\mu^* = \frac{1-c}{c^2} > 1. \tag{1.33}$$

The condition (1.33) is possible for all parameters c satisfying

$$1 - c > c^2 \Leftrightarrow c^2 + c - 1 < 0\,,$$

which corresponds to the interval (c_1, c_2) between the roots of the polynomial $p(c) := c^2 + c - 1$:

$$c_{1,2} = \frac{-1 \pm \sqrt{5}}{2}\,.$$

So, the following open interval

$$\boxed{c \in \left(\frac{-1-\sqrt{5}}{2}, \frac{-1+\sqrt{5}}{2} \right)}$$

of parameters c provides the property $\mathcal{X} \neq \varnothing$.

Example 1.3 *Using S - procedure find multipliers $\tau > 0$ such that*

$$S(x) = B\,(x) - \tau A(x) \geq 0 \tag{1.34}$$

for all $x \in \mathbb{R}^3$, satisfying $A(x) \geq 0$ with

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \geq 0\,,$$

implies $B\,(x) := x^\intercal Bx \geq 0$ with

$$B = \begin{bmatrix} 1 & -1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}.$$

Solution. By S - *procedure for (1.34) we have*

$$S(x) = B\,(x) - \tau A(x) =$$
$$x^\intercal \begin{bmatrix} 1 & -1 & 0 \\ -1 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix} x - \tau x^\intercal \begin{bmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} x =$$
$$x^\intercal \begin{bmatrix} 1-2\tau & -1+\tau & 0 \\ -1+\tau & -2-\tau & 0 \\ 0 & 0 & -3+\tau \end{bmatrix} x \geq 0.$$

This means that the last inequality holds for all $x \in \mathbb{R}^3$

$$\begin{bmatrix} 1-2\tau & -1+\tau & 0 \\ -1+\tau & -2-\tau & 0 \\ 0 & 0 & -3+\tau \end{bmatrix} \geq 0$$

By the Silvester criterion this is true if and only if

$$1-2\tau \geq 0,\ \det \begin{bmatrix} 1-2\tau & \tau-1 \\ \tau-1 & -2-\tau \end{bmatrix} \geq 0\,,$$
$$\det \begin{bmatrix} 1-2\tau & \tau-1 & 0 \\ \tau-1 & -2-\tau & 0 \\ 0 & 0 & -3+\tau \end{bmatrix} \geq 0\,,$$

which leads to

$$\tau \leq \frac{1}{2},\ \ (2\tau-1)\,(2+\tau) - (1-\tau)^2 = \tau^2 + 5\tau - 3 \geq 0,$$
$$\left[(2\tau-1)\,(2+\tau) - (1-\tau)^2\right](-3+\tau) \geq 0\,.$$

The set of the last constraints is as follows

$$\tau \leq \frac{1}{2}, \begin{cases} \tau \geq 3, & -3 + 5\tau + \tau^2 \geq 0 \\ \tau \leq 3, & -3 + 5\tau + \tau^2 \leq 0 \end{cases}$$

$$\Updownarrow$$

$$\tau \leq 0.5,\ \tau \geq 3,\ \tau \geq \frac{\sqrt{37} - 5}{2} = 0.54138 \Longleftrightarrow \tau \text{ no exists},$$

$$\tau \leq 0.5,\ \tau \leq 3,\ 0 \leq \tau \leq \frac{\sqrt{37} - 5}{2} = 0.54138 \Longleftrightarrow 0 < \tau \leq 0.5$$

So, the property (1.34) takes place for all

$$\boxed{0 < \tau \leq 0.5.}$$

1.4 Exercises

Exercise 1.1 *Find for which* β

$$S = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & -\beta \\ -1 & -\beta & 1 \end{bmatrix} > 0\,.$$

Exercise 1.2 *Suppose that for all* $x \in R^3$, *satisfying*

$$x^\intercal A x = x^\intercal \begin{bmatrix} 2 & -1 & 0 \\ -1 & \alpha & 0 \\ 0 & 0 & -1 \end{bmatrix} x \geq 0\,,$$

the inequality $B(x) := x^\intercal B x \geq 0$ *with*

$$\begin{bmatrix} 1 & -0.5 & 0 \\ -0.5 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \geq 0$$

is met. For which α *is it possible?*