A Lecture on the S-Procedure

Ulf T. Jönsson
Division of Optimization and Systems Theory
Royal Institute of Technology
10044 Stockholm, Sweden
Email: ulfj@math.kth.se

May 21, 2001
Revised 2006

1 Introduction

The purpose of this lecture is to give an introduction to the the S-procedure and its use in systems and control applications. The notion S-procedure was coined by researches in the former Soviet union who used it in stability analysis of nonlinear systems. Several important results in systems analysis, such as the circle criterion, the Popov criterion, and the Kalman-Yakubovich-Popov Lemma, are closely related to the S-procedure.

The S-procedure is really nothing but a Lagrange relaxation technique, which mostly has been used in problems with quadratic constraints. One of the main results that will be presented in this note is a condition for exactness (losslessness) of the relaxation. One of the few known, yet extremely important, cases when the S-procedure is lossless is when there is one quadratic objective subject to one quadratic constraint. We will use this to derive the circle criterion.

The S-procedure is useful also in cases when it is not exact (lossy). The reason is that it can be used to obtain sufficient sufficient conditions in terms of Linear Matrix Inequalities (LMI) for a large number of nonconvex problems in systems analysis. It has numerous applications in analysis of uncertain differential equations.

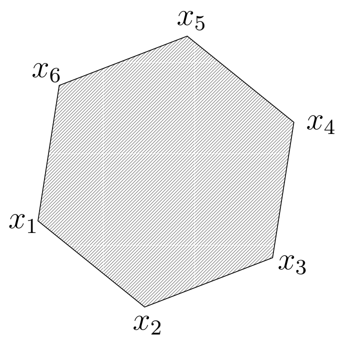

Figure 1: The polytope $C = \mathrm{co}\,\{x_1, \ldots, x_6\}$

.

The main results in Section 4 can be found in [10]. The book [1] contains a large number of applications of the S-procedure and all our examples in Section 5 can be derived from it. Other useful references with many applications of the S-procedure are [5, 3].

The note also contains some other material and examples related to the application of LMIs in system theory.

2 Preliminary material

A set $\mathcal{C}$ in a linear vector space is *convex* if $\alpha x_1 + (1-\alpha)x_2 \in \mathcal{C}$, for all $x_1, x_2 \in \mathcal{C}$ and $\alpha \in [0, 1]$. Furthermore, $\mathcal{C}$ is a *convex cone* if it is convex and in addition $\alpha x \in \mathcal{C}$ for all $x \in \mathcal{C}$ and all $\alpha > 0$.

Convex Polytope

The *convex polytope*, $\mathcal{C}$, with vertices at $x_1, \ldots, x_n \in \mathbf{R}^m$ is defined as the convex hull of these points, i.e.,

$$\mathcal{C} = \mathrm{co}\,\{x_1, \ldots, x_n\} = \left\{ \sum_{i=1}^{n} \alpha_i x_i : \alpha_i \geq 0,\ \sum_{i=1}^{n} \alpha_i = 1 \right\}.$$

Ellipsoids

An ellipsoid $\mathcal{E}$ with center m can be defined by

$$\begin{aligned}\mathcal{E} &= \{x : (x-m)^T Q^{-1}(x-m) \leq 1\} \\ &= \{x : x^T P x + 2x^T b + c \leq 0\}\end{aligned}$$

where Q is positive definite and symmetric. In the second characterization we have $P = Q^{-1}$, $b = -Q^{-1}m$, and $c = m^T Q^{-1} m - 1$. The size, shape, and location of the ellipsoid is determined by m and Q. The center is as we already noted at m. The orientation of $\mathcal{E}$ is determined by the eigenvectors of Q, and the lengths of the semi-axes of $\mathcal{E}$ are defined by the eigenvalues of Q. For example, consider the case when (this example is taken from [9])

$$m = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \text{and} \quad Q = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

Then Q has the eigenvalue decomposition

$$Q = \begin{bmatrix} e_1 & e_2 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} e_1 & e_2 \end{bmatrix}^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}^T.$$

Hence, we get the ellipsoid in Figure 2. The volume of an ellipsoid is given as

$$V(\mathcal{E}) = k(n)\det(Q)^{1/2} = k(n)\det((b^T P^{-1} b - c)P^{-1})^{1/2},$$

where $k(n)$ is a dimension dependent constraint. A linear transformation of the matrix $\mathcal{E}$ changes the center, shape and, location of the ellipsoid as follows

$$\begin{aligned}\widetilde{\mathcal{E}} &= A\mathcal{E} + b = \{y : y = Ax + b : x \in \mathcal{E}\} \\ &= \{y : (y - b - Am)^T (AQA^T)^{-1} (y - b - Am) \leq 1\}.\end{aligned}$$

As an example, the ellipsoid in Figure 2 is projected onto the x_2-axis by a linear transformation with $b = 0$ and $A = \begin{bmatrix} 0 & 1 \end{bmatrix}$, i.e., we get

$$\mathcal{E}_{x_2} = \{x_2 : Ax; x \in \mathcal{E}\} = \{x_2 : |x_2 - 4|^2 \leq 5\}.$$

See Figure 2 for an illustration.

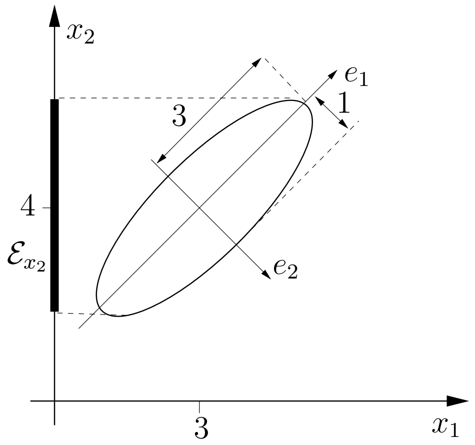

Figure 2: The ellipsoid $\mathcal{E}$ with its semi-axes. We also see the projection of $\mathcal{E}$ onto the x_2-axis in which gives the ellipsoid $\mathcal{E}_{x_2}$ in thick solid line.

2.1 The Schur Complement Formula

The Schur complement formula is a very useful tool for manipulating matrix inequalities. It is often use to inequlaities that are nonlinear in the parameters into an linear matrix inequality.

Proposition 1 (Nonstrict Schur Complement Formula). *Let Q and R be symmetric matrices. Then the following are equivalent.*

(*i*) $\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \geq 0$

(*ii*) $R \geq 0$, $Q - SR^\dagger S^T \geq 0$, $S(I - RR^\dagger) = 0$

where $R^\dagger$ denotes the pseudo inverse of R.

Remark 1. Let R have the singular value decomposition

$$R = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} U^T \tag{1}$$

where U is an orthogonal matrix ($U^T U = I$), and Σ is diagonal. Then

$$R^\dagger = U \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^T$$

Proof. We follow the proof in [1]. It is clear that the singular value decomposition of R in (1) will have $\Sigma > 0$. The inequality in (i) holds if and only if

$$\begin{bmatrix} I & 0 \\ 0 & U^T \end{bmatrix} \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} = \begin{bmatrix} Q & S_1 & S_2 \\ S_1^T & \Sigma & 0 \\ S_2^T & 0 & 0 \end{bmatrix} \geq 0 \tag{2}$$

where $\begin{bmatrix} S_1 & S_2 \end{bmatrix} := SU$ is consistently partitioned. We thus need

(a) $S_2 = 0$ which is equivalent to $S(I - RR^\dagger) = 0$, since $S(I - RR^\dagger) = \begin{bmatrix} 0 & S_2 \end{bmatrix} U^T$.

(b) The LMI

$$0 \leq \begin{bmatrix} Q & S_1 \\ S_1^T & \Sigma \end{bmatrix} = \begin{bmatrix} I & S_1\Sigma^{-1/2} \\ 0 & \Sigma^{1/2} \end{bmatrix} \begin{bmatrix} Q - S_1\Sigma^{-1}S_1^T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ \Sigma^{-1/2}S_1^T & \Sigma^{1/2} \end{bmatrix}$$

must hold. This implies that $Q - S_1\Sigma^{-1}S_1^T \geq 0$ or equivalently $Q - SR^\dagger S^T \geq 0$.

□

The strict version of this proposition will be used frequently in this chapter.

Corollary 1 (Strict Schur Complement Formula). *Let Q and R be symmetric matrices. Then the following are equivalent.*

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

(ii) $R > 0$, $Q - SR^{-1}S^T > 0$.

Remark 2. Note that the equality $S(I - RR^\dagger) = 0$ in (1) is redundant since $R^\dagger = R^{-1}$.

Example 1. Let $P = P^T$ be a matrix variable and let $Q = Q^T$, $R = R^T$ and S be given. Then the LMI

$$\begin{bmatrix} PA + A^TP + Q & PB + S \\ B^TP + S^T & R \end{bmatrix} > 0$$

is equivalent to the LMIs $R > 0$ and $PA + A^TP + Q - (PB + S)R^{-1}(PB + S)^T > 0$. The last LMI is called a Riccati inequality. The LMIs in this example appear in connection to LQ optimal control and the Kalman-Yakubovich-Popov Lemma.

3 Separating Hyperplane Theorem

The separating hyperplane theorem is one of the most useful results in convex analysis. We will here discuss a finite-dimensional version of this theorem.

We first need to introduce some notation. Let V be a linear vector space. A linear functional on V is a function $f : V \to \mathbf{R}$, which is linear, i.e. $f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2)$ for all $v_1, v_2 \in V$ and $\alpha_1, \alpha_2 \in \mathbf{R}$ (we assume V is a vector space over the real numbers). It is easy to see that the (continuous) linear functionals on V also make up a vector space, which is called the dual of V and is denoted V^*. For each $v^* \in V^*$ the linear functional is also denoted $v^*(v) =: \langle v, v^* \rangle$. In our applications we will only consider finite dimensional Hilbert spaces (over $\mathbf{R}$). Then the dual space can be identified with V itself, i.e. $V^* = V$, and the linear functional $\langle v, v^* \rangle$ is nothing but the inner product. We will often denote the element from the primal space V by x and the element from the dual space by z.

Example 2. The dual of the Euclidean space $\mathbf{R}^n$ is $\mathbf{R}^n$ itself and the linear functionals are defined as $\langle x, z \rangle = x^T z$, where $x, z \in \mathbf{R}^n$.

Example 3. The dual of the vector space of symmetric matrices $S^{m\times m} = \{X \in \mathbf{R}^m : X = X^T\}$ is $S^{m\times m}$ itself and the linear functionals are defined as $\langle X, Z \rangle = \mathrm{tr}(XZ)$, where $X, Z \in S^{m\times m}$.

We will need the notion relative interior of a convex set

Relative interior

The relative interior of a set $S \subset \mathbf{R}^m$ is denoted $\mathrm{ri}S$. It is defined in terms of the affine hull of S, which is denoted $\mathrm{aff}\, S$. The affine hull is defined as the set of all linear combinations on the form $\sum \alpha_i x_i$, where $x_i \in S$, and $\sum \alpha_i = 1$. We can now define $\mathrm{ri}S$ as the set of points of S, which are interior relative $\mathrm{aff}\, S$. This means that for any $x \in \mathrm{ri}S$, there exists $\varepsilon > 0$ such that all $y \in \mathrm{aff}\, S$ with $|x - y| < \varepsilon$ are also members of S. The definition is illustrated in Figure 3.

Discuss basic ideas behind quotient spaces, codimension, etc.

Definition 1 (Hyperplane). A *hyperplane* is an affine proper subsset of maximal dimension. In other words, if V has dimension n then every hyperplane is a translation of an $(n-1)$ dimensional subspace. Every hyperplane

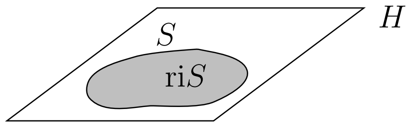

Figure 3: Illustration of the relative interior of a set $S \subset \mathbf{R}^3$. The affine hull of S is the hyperplane, H, illustrated as a plane in the figure. The relative interior of S consists of the grey area in the figure. It does not contain the solid contour.

can be represented as $H = \{x \in V : \langle v, z \rangle = c\}$ for suitably chosen nonzero $z \in V^*$ and $c \in \mathbf{R}$.

The sets $\{x \in V : \langle v, z \rangle \geq c\}$ and $\{x \in V : \langle v, z \rangle \leq c\}$ are (closed) poitive/negative halfspaces.

In the next three results we assume that all sets involved are nonempty.

Lemma 1 (Separation Lemma). *Let $C \subset V$ be a relatively open*[1] *convex set in a finite dimensional vector space V and let V_1 be a linear subspace of V such that $C \cap V_1 = \emptyset$. Then there exists a hyperplane containing V_1 and such one of the open half-spaces associated with the hyperplane contains C, i.e., there exists nonzero $z \in V^*$ such that*

$$\begin{aligned} \langle x, z \rangle &> 0, \quad \forall x \in C \\ \langle x, z \rangle &\leq 0, \quad \forall x \in V_1. \end{aligned}$$

Proof. See Rockafellars book [8]. □

We can specialize the above lemma to the following result, which will be used in the next section.

Theorem 1. *Let C_1 be an open convex cone and C_2 be a convex set, which are disjoint, i.e., $C_1 \cap C_2 = \emptyset$. Then there exists a separating hyperplane, i.e., there exists nonzero $z \in V^*$ such that*

$$\begin{aligned} \langle x_1, z \rangle &> 0, \quad \forall x_1 \in C_1 \\ \langle x_2, z \rangle &\leq 0, \quad \forall x_2 \in C_2. \end{aligned}$$

[1] C is relatively open if $\mathrm{ri}(C) = C$

Proof. The set $C = C_1 - C_2$ is open and convex and since $C_1 \cap C_2 = \emptyset$ it follows that $0 \notin C$. Hence, by the previous lemma there exists a nonzero $z \in V^*$ such that $\langle x, z \rangle > 0, \quad \forall x \in C$ (the claim follows since C is open and thus relatively open, and since $V_1 = \{0\}$ is a zero-dimensional subspace with $C \cap V_1 = \emptyset$). This can equivalently be written $\langle x_1, z \rangle > \langle x_2, z \rangle$, $\forall x_1 \in C_1$ and $\forall x_2 \in C_2$.

Now let $\alpha = \inf_{x_1 \in C_1} \langle x_1, z \rangle$. Since C_1 is a cone we have

$$\lambda \langle x_1, z \rangle = \langle \lambda x_1, z \rangle \geq \alpha, \quad \forall x \in C_1, \lambda > 0.$$

This implies that $\alpha \leq 0$, and since $\alpha = \inf_{x_1 \in C_1} \langle x_1, z \rangle$ we need $\alpha = 0$. We have thus proven that $\langle x_1, z \rangle \geq 0$, $x_1 \in C_1$ and $\langle x_2, z \rangle \leq 0$, $x_2 \in C_2$. Finally, note that the first inequality $\langle x_1, z \rangle > 0$ must be strict since C_1 is open. $\square$

4 The S-procedure

The S-procedure is frequently used in system theory to derive stability and performance results for nonlinear and uncertain systems. In fact, the idea has been used (often implicitly) in the former soviet union since the work of Lure and Postnikov, [6]. The idea has since then been developed by many researchers. The most notable early results are due to Yakubovich, see, for example, [10, 11, 4]. The S-procedure became popular in the robust control community during the 1990s, largely due to new developments by Megretski and Treil [7]. We will mainly discuss the use of the S-procedure in finite dimensional spaces.

The basic idea behind the S-procedure is trivial. Let $\sigma_k : V \to \mathbf{R}$, $k = 0, 1, \ldots, N$ be real valued functionals defined on a linear vector space V (e.g., $V = \mathbf{R}^m$) and consider the following two conditions

$S_1 :$ $\sigma_0(y) \geq 0$ for all $y \in V$ such that $\sigma_k(y) \geq 0$, $k = 1, \ldots, N$.

$S_2 :$ There exists $\tau_k \geq 0$, $k = 1, \ldots, N$ such that

$$\sigma_0(y) - \sum_{k=1}^{N} \tau_k \sigma_k(y) \geq 0, \quad \forall y \in V.$$

It is a obvious fact that S_2 implies S_1. Indeed,

$$S_2 \Rightarrow \sigma_0(y) \geq \sum_{k=1}^{N} \tau_k \sigma_k(y) \Rightarrow S_1$$

since the $\tau_k \geq 0$. The S-procedure is the method of verifying S_1 using S_2. This is useful since S_2 generally is much simpler to verify than S_1. We will consider the case with quadratic functionals defined over $V = \mathbf{R}^m$

$$\sigma_k(y) = y^T Q_k y + 2s_k^T y + r_k, \quad k = 0, 1, \ldots, N \tag{3}$$

where $Q_k = Q_k^T \in \mathbf{R}^{m\times m}$, $s_k \in \mathbf{R}^m$, $r_k \in \mathbf{R}$. The problem with S_1 is then that

1. σ_0 is not a convex function in general.

2. The constraint set

   $$\Omega = \{y \in \mathbf{R}^m : \sigma_k(y) \geq 0,\ k = 1, \ldots, N\}$$

   is not convex in general.

This means that condition S_1 in general corresponds to verifying that the minimum of a nonconvex function over a nonconvex set is positive, i.e., $S_1 \Leftrightarrow \min_{y\in\Omega} \sigma_0(y) \geq 0$. This is a NP hard problem.

Condition S_2 on the other hand correpsonds to a linear matrix inequality (LMI). In fact,

$$\begin{aligned}
S_2 \ &\Leftrightarrow\ \exists \tau_k \geq 0 \text{ s.t. } \sigma_0(y) + \textstyle\sum_{k=1}^N \tau_k \sigma_k(y) \geq 0,\ \forall y \in \mathbf{R}^m \\
&\Leftrightarrow\ \exists \tau_k \geq 0 \text{ s.t. } \begin{bmatrix} y \\ 1 \end{bmatrix}^T \begin{bmatrix} Q_0 + \sum \tau_k Q_k & s_0 + \sum \tau_k s_k \\ s_k^T + \sum \tau_k s_k^T & r_0 + \sum \tau_k r_k \end{bmatrix} \begin{bmatrix} y \\ 1 \end{bmatrix} \geq 0,\ \forall y \in \mathbf{R}^m \\
&\Leftrightarrow\ \exists \tau_k \geq 0 \text{ s.t. } \begin{bmatrix} Q_0 & s_0 \\ s_0^T & r_0 \end{bmatrix} + \sum_{k=1}^N \tau_k \begin{bmatrix} Q_k & s_k \\ s_k^T & r_k \end{bmatrix} \geq 0
\end{aligned}$$

The two conditions S_1 and S_2 are in general not equivalent. However, there are some special cases when $S_1 \Leftrightarrow S_2$ and the S-procedure is then called lossless. Before stating a number of important losslessness results for the S-procedure we supply some additional remarks.

1. We often use the S-procedure in applications where it can be lossy. This will in an example of control system stability mean that we obtain sufficient but not necessary conditions for stability. However, the computational advantage discussed above justifies the potential conservatism.

2. It often happens (for example in stability analysis) that we require strict inequality for the quadratic function σ_0. We then have that $S_2' \Rightarrow S_1'$, where

   S_1' : $\sigma_0(y) > 0$ for all $y \neq 0$ such that $\sigma_k(y) \geq 0$, $k = 1, \dots, N$.

   S_2' : There exists $\tau_k \geq 0$, $k = 1, \dots, N$ such that

$$\sigma_0(y) - \sum_{k=1}^{N} \tau_k \sigma_k(y) > 0, \quad \forall y \neq 0.$$

   Theorem 3 below shows that $S_1' \Leftrightarrow S_2'$ when $k = 1$ and $\sigma_k(y) = y^T Q_k y$, $k = 0, 1$, i.e. when we consider two quadratic forms.

We will next give some results due to Yakubovich on S-procedure losslessness. We need the following regularity assumption.

Definition 2. Let $\sigma_k 1 : V \to \mathbf{R}$ (for example a quadratic function). The constraint $\sigma_k(y) \geq 0$ for $k = 1, \dots, N$ is said to be regular if there exists $y^* \in V$ such that $\sigma(y^*) > 0$.

Theorem 2 (Yakubovich). *Let $\sigma_k : V \to \boldsymbol{R}$, $k = 0, \dots, N$ and assume the constraint $\sigma_k(y) \geq 0$, $k = 1, \dots, N$ is regular. Finally, let sets*

$$\mathcal{K} = \{(\sigma_0(y), \sigma_1(y), \dots, \sigma_N(y)) : y \in V\},$$
$$\mathcal{N} = \{(n_0, n_1, \dots, n_N) : n_0 < 0, n_k > 0\}.$$

If $\mathcal{K} \cap \mathcal{N} = \emptyset \Rightarrow \mathrm{co}(\mathcal{K}) \cap \mathcal{N} = \emptyset$ then the S-procedure is lossless, i.e., $S_1 \Leftrightarrow S_2$.

Remark 3. In particular, if $\mathcal{K}$ is a convex set then the S-procedure is lossless.

Proof. We have $S_1 \Rightarrow \mathcal{K} \cap \mathcal{N} = \emptyset \Rightarrow \mathrm{co}(\mathcal{K}) \cap \mathcal{N} = \emptyset$. Since $\mathcal{N}$ is an open convex cone we can use Theorem 1, i.e., there exists a nonzero $(N+1)$-tuple $(c_0, -c_1, \dots, -c_N)$ such that

$$c_0 n_0 - c_1 n_1 - \dots - c_N n_N < 0, \quad \forall (n_0, n_1, \dots, n_N) \in \mathcal{N} \tag{4}$$
$$c_0 \eta_0 - c_1 \eta_1 - \dots - c_N \eta_N \geq 0, \quad \forall (\eta_0, \eta_1, \dots, \eta_N) \in \mathrm{co}(\mathcal{K}) \tag{5}$$

For any given $\varepsilon > 0$, we have $(n_0, \varepsilon, \dots, \varepsilon) \in \mathcal{N}$, for all $n_0 < 0$. Hence, (4) implies that $c_0 \geq 0$. We can in a similar way show that $c_k \geq 0$, $k = 1, \dots, N$. By the regularity assumption there exists $y^* \in V$ such that $\eta_k = \sigma_k(y^*)$ for some $\eta_1, \ldots, \eta_N > 0$. Using this in (5) shows that $c_0 > 0$. For any $y \in V$ we have $(\sigma_0(y), \sigma_1(y), \ldots, \sigma_N(y)) \in \text{co}(\mathcal{K})$. Hence, (5) and the fact that $c_0 > 0$ gives

$$\sigma_0(y) - \sum_{k=1}^{N} \frac{c_k}{c_0} \sigma_k(y) \leq 0, \quad \forall y \in V$$

This shows that S_2 holds with $\tau_k = c_k/c_0$, for $k = 1, \ldots, N$. □

We will from now on consider the case when $V = \mathbf{R}^m$.

Corollary 2 (S-procedure Losslessness in linear case). *Consider the linear case, i.e., when $\sigma_k(y) = s_k^T y + r_k$, where $s_k \in \boldsymbol{R}^m$ and $r_k \in \boldsymbol{R}$, $k = 0, 1, \ldots, N$. If the cosntraint is regular then the S-procedure is lossless for any finite number of constraints N.*

Proof. This follows from Theorem 2 since $\mathcal{K}$ in this case is a convex set. □

Remark 4. The proposition is just a version of Farkas lemma. If we let

$$A = \begin{bmatrix} s_1^T \\ \vdots \\ s_N^T \end{bmatrix}, \quad B = -\begin{bmatrix} r_1 \\ \vdots \\ r_N \end{bmatrix}, \quad c = s_0^T, \quad d = -r_0$$

then the result shows that the following are equivalent

$S_1'':$ $cy \geq d$ for all y such that $Ay \geq B$

$S_2'':$ $\exists \tau = \begin{bmatrix} \tau_1 \ldots, \tau_N \end{bmatrix}$, $\tau_k \geq 0$, such that $\tau A = c$ and $\tau B + d \leq 0$.

The following important result can be derived from [10, 11].

Theorem 3. *The S-procedure is lossless in the case of one quadratic constraint. More precisely, we have the following two results*

(*i*) *Assume $\sigma_1(y) = y^T Q_1 y + 2s_1^T y + r_1 \geq 0$ is regular. Then the following are equivalent*

$S_1^{(3)}:$ *$y^T Q_0 y + 2s_0^T y + r_0 \geq 0$, for all $y \in \boldsymbol{R}^m$ such that $y^T Q_1 y + 2s_1^T y + r_1 \geq 0$*

$S_2^{(3)}:$ *there exists $\tau \geq 0$ such that the following LMI is feasible*

$$\begin{bmatrix} Q_0 & s_0 \\ s_0^T & r_0 \end{bmatrix} - \tau \begin{bmatrix} Q_1 & s_1 \\ s_1^T & r_1 \end{bmatrix} \geq 0$$

(*ii*) *Assume* $\sigma_1(y) = y^T Q_1 y \geq 0$ *is regular. Then the following are equivalent*

$S_1^{(4)}$: $y^T Q_0 y > 0$, *for all* $y \neq 0$ *such that* $y^T Q_1 y \geq 0$

$S_2^{(4)}$: *there exists* $\tau \geq 0$ *such that* $Q_0 - \tau Q_1 > 0$

Proof. For part (*i*) let $x = \begin{bmatrix} y^T & z \end{bmatrix}^T$ and $\sigma_k(x) = y^T Q_k y + 2 z s_k^T y + r_k z^2$, $k = 0, 1$. Then $\mathcal{K} = \{(\sigma_0(x), \sigma_1(x)) : x \in \mathbf{R}^{m+1}\}$ is convex by a result that follows from e.g. [2, 11]. Hence, by Theorem 2

$$\begin{aligned} S_1^{(3)} &\Leftrightarrow \sigma_0(x) \geq 0, \quad \forall x \in \mathbf{R}^{m+1} \quad \text{s.t} \quad \sigma_1(x) \geq 0 \\ &\Leftrightarrow \exists \tau \geq 0 \text{ s.t. } \sigma_0(x) - \tau \sigma_k(x) \geq 0, \ \forall x \in \mathbf{R}^{m+1} \end{aligned}$$

The last condition is equivalent to the LMI in $S_2^{(3)}$. The first equivalence follows by scaling both quadratic forms in $S_1^{(3)}$ by an arbitrary nonnegative number ρ. By defining $x = \begin{bmatrix} \sqrt{\rho} y^T & \sqrt{\rho} \end{bmatrix}^T$ we get the equivalence. This finishes the proof.

To prove (*ii*) we first notice that $S_2^{(4)} \Rightarrow S_1^{(4)}$ is trivial. For the opposite direction, note that $S_1^{(4)}$ implies

$$\min_{\{y : y^T Q_1 y \geq 0\}} \frac{y^T Q_0 y}{|y|^2} = \varepsilon$$

where $\varepsilon > 0$. Hence $y^T Q_0 y - \varepsilon |y|^2 \geq 0$, $\forall y \in \mathbf{R}^m$ such that $y^T Q_1 y \geq 0$. By (*i*) we know that there exists $\tau \geq 0$ such that $Q_0 - \tau Q_1 \geq \varepsilon$. $\square$

Example 4 (Circle criterion). We will here derive a necessary and sufficient condition for quadratic stability of the system

$$\begin{aligned} \dot{x} &= Ax + Bw, \quad x(0) = x_0 \\ v &= Cx \end{aligned}$$

where the input and output satisfies the sector constraint

$$\sigma(v, w) = (\beta v - w)(w - \alpha v) \geq 0,$$

where $\alpha < \beta$ are real numbers, see figure 4. In order to have quadratic stability it is necessary and sufficient that there exists $P = P^T > 0$ such that

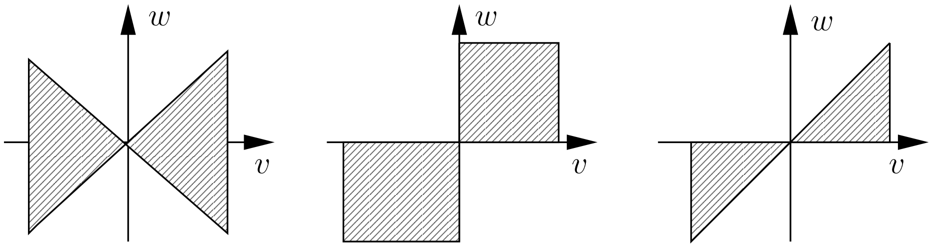

Figure 4: Illustration of the sector condition.

$V(x) = x^T P x$ is a Lyapunov function for the closed loop system. Hence, we require that

$$\dot{V}(x) = 2x^T P(Ax + Bw) < 0,\ \forall (x, w) \neq 0 \text{ s.t. } \sigma(cx, w) \geq 0. \tag{6}$$

The above condition tells us that the quadratic Lyapunov function $V(x) = x^T P x$ is strictly decreasing along every possible system trajectory. Let us now define the quadratic forms

$$\sigma_0(x, w) = \begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$$
$$\sigma_1(x, w) = 2\sigma(Cx, w) = \begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} -2\beta\alpha C^T C & (\beta + \alpha)C^T \\ (\beta + \alpha)C & -2 \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$$

Then condition (6) can then be rewritten as

$$\sigma_0(x, w) < 0 \quad \forall (x, w) \neq (0, 0) \text{ s.t. } \sigma_1(x, w) \geq 0$$

It follows from *(ii)* Theorem 3 that the S-procedure is lossless for this case of two quadratic forms (note that the constraint $\sigma_1(x, w) \geq 0$ is regular since $\alpha < \beta$). Hence, the above criterion is equivalent to the existence of $\tau \geq 0$ such that $\sigma_0(x, w) + \tau\sigma_1(x, w) < 0$ for all $(x, w) \neq 0$. It is easily seen that we need $\tau > 0$ for this to hold. We can then normalize such that $\tau = 1$ (and $P/\tau \to P$). We have thus shown that quadratic stability of a linear system with sector bounded uncertainty is equivalent to feasibility of the LMI: $\exists P = P^T > 0$ such that

$$\begin{bmatrix} A^T P + PA - 2\beta\alpha C^T C & PB + (\beta + \alpha)C^T \\ B^T P + (\beta + \alpha)C & -2 \end{bmatrix} < 0.$$

This LMI corresponds to the circle criterion.

S-Procedure Loslessness in Complex Space

We will here consider the S-procedure for the case of real valued functionals defined over the complex space $\mathbf{C}^m$. Let $\sigma_k : \mathbf{C}^m \to \mathbf{R}$, $k = 0, 1, \ldots, N$ be defined as

$$\sigma_k(y) = y^* Q_k y + 2\text{Re } s_k^* y + r_k$$

where $Q_k = Q_k^* \in \mathbf{C}^{m \times m}$, $s_k \in \mathbf{C}^m$, and $r_k \in \mathbf{R}$. Again, consider the conditions

$S_1 :$ $\sigma_0(y) \geq 0$ for all $y \in \mathbf{C}^m$ such that $\sigma_k(y) \geq 0$, $k = 1, \ldots, N$.

$S_2 :$ There exists $\tau_k \geq 0$, $k = 1, \ldots, N$ such that

$$\sigma_0(y) - \sum_{k=1}^{N} \tau_k \sigma_k(y) \geq 0, \quad \forall y \in \mathbf{C}^m.$$

It is obvious that $S_2 \Rightarrow S_1$, which is useful since S_2 corresponds to a complex valued LMI while S_1 in general is a NP hard problem. The next result due to Fradkov and Yakubovich [4] shows that S-procedure in complex linear space is lossless for the case of two (or one) constraints.

Theorem 4. *We have*

(*i*) *Assume $\sigma_k \geq 0$, $k = 1, 2$, is regular, i.e., there exists $\hat{y} \in \boldsymbol{C}^m$ such that $\sigma_1(\hat{y}) > 0$ and $\sigma_2(\hat{y}) > 0$. Then the following are equivalent*

$S_1 :$ $\sigma_0(y) \geq 0$ for all $y \in \boldsymbol{C}^m$ such that $\sigma_k(y) \geq 0$, $k = 1, 2$

$S_2 :$ there exists $\tau_1, \tau_2 \geq 0$ such that the following LMI is feasible

$$\begin{bmatrix} Q_0 & s_0 \\ s_0^* & r_0 \end{bmatrix} - \tau_1 \begin{bmatrix} Q_1 & s_1 \\ s_1^* & r_1 \end{bmatrix} - \tau_2 \begin{bmatrix} Q_2 & s_2 \\ s_2^* & r_2 \end{bmatrix} \geq 0$$

(*ii*) *Assume $\sigma_k(y) = y^* Q_k y \geq 0$, $k = 1, 2$, is regular. Then the following are equivalent*

$S_1 :$ $y^* Q_0 y > 0$, for all $y \neq 0$ such that $y^* Q_k y \geq 0$, $k = 1, 2$.

$S_2 :$ there exists $\tau_1, \tau_2 \geq 0$ such that $Q_0 - \tau_1 Q_1 - \tau_2 Q_2 > 0$

Proof. The key part of the proof is to show that $\mathcal{K} = \{(\sigma_0(y), \sigma_1(y), \sigma_2(y)) : y \in \mathbf{C}^m\}$ is convex. This is done in [4]. Then everything else can be proven in same way as the proof of Theorem 3. $\square$

Application to Quadratic Programming

We will next apply the S-procedure to a general class of quadratic programming problems

$$\begin{aligned} &\min\ y^T Q_0 y + 2s_0^T y + r_0 \\ &\text{subject to } y^T Q_k y + 2s_k^T y + r_k \geq 0,\ k = 1, \dots, N \end{aligned} \tag{7}$$

This class of optimization problem includes also integer valued constraints, since $x_i^2 - x_i = 0$ would be a valid constraint. It is thus a huge class of problems, which generally is NP hard unless the quadratic functions are convex. Lower bounds on the optimization problem in (7) can be obtained using the following semidefinite relaxation

$$\begin{aligned} &\min\ \operatorname{tr}(Q_0 Y) + 2s_0^T y + r_0 \\ &\text{subject to } \begin{cases} \operatorname{tr}(Q_k Y) + 2s_k^T y + r_k \geq 0,\ k = 1, \dots, N \\ \begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix} \geq 0 \end{cases} \end{aligned} \tag{8}$$

To see that (8) gives an lower bound of (7) we notice that the last constraint, by the nonstrict Schur complement formula is equivalent to $Y \geq yy^T$. Since $\operatorname{tr}(Q_k yy^T) = y^T Q_k y$ we see that (8) must have a lower (or equal) value than (7) since we optimize over a larger set in (8).

For the case of one constraint, we have the following interesting result.

Proposition 2. *Under the assumption that $\sigma_1(y) = y^T Q_1 y + 2s_1^T y + r_1 \geq 0$ is regular and that the last two optimization problems are strictly feasible then the following optimization problems are equivalent, i.e., they have the same objective values (and the optima are obtained)*

1. *Quadratic program*

$$\begin{aligned} &\min\ y^T Q_0 y + 2s_0^T y + r_0 \\ &\text{subject to } y^T Q_1 y + 2s_1^T y + r_1 \geq 0 \end{aligned} \tag{9}$$

2. *Eigenvalue problem for LMI*

$$\max \gamma \quad \text{subject to } \begin{cases} \begin{bmatrix} Q_0 - \lambda Q_1 & s_0 - \lambda s_1 \\ s_0^T - \lambda s_1^T & r_0 - \lambda r_1 - \gamma \end{bmatrix} \geq 0 \\ \lambda \geq 0,\ \gamma \in \boldsymbol{R} \end{cases} \tag{10}$$

3. Semidefinite relaxation of (9) (and Dual of (10))

$$\min \ tr(Q_0Y) + 2s_0^T y + r_0$$
$$\text{subject to } \begin{cases} tr(Q_1Y) + 2s_1^T y + r_1 \geq 0 \\ \begin{bmatrix} Y & y \\ y^T & 1 \end{bmatrix} \geq 0 \end{cases} \tag{11}$$

A survey with emphasis on engineering applications of general quadratic programs of the form (7) is given in [3]

5 Additional Examples

We here give some additional examples pertaining to the material in this lecture note.

Example 5. In this example we consider outer and inner approximation of polytopes by ellipsoids. First consider outer approximation. For simplicity we only consider ellipsoids and polytopes centered around the origin.

The ellipsoid $\mathcal{E} = \{x : x^T Q^{-1} x \leq 1\}$ covers the polytope $\mathcal{P} = \text{co}\{v_1, \dots, v_n\}$ if and only if

$$v_k^T Q^{-1} v_k \leq 1, \quad k = 1, \dots, N$$

The volume of the ellipsoid is proportional to $\det(Q)^{1/2}$. Hence, if we use $P = Q^{-1}$ then it follows that the smallest volume ellipsoid that covers the polyope is given by the optimization problem

$$\min \ \log\det(P^{-1}) \quad \text{subject to} \quad \begin{cases} v_k^T P v_k \leq 1, \quad k = 1, \dots N \\ P > 0 \end{cases}$$

The objective function $\log\det(P^{-1})$ can be proven to be convex since $P > 0$.

The polytope $\mathcal{P} = \text{co}\{v_1, \dots, v_n\}$ can equivalently be represented as an intersection of halfspaces, i.e., $\mathcal{P} = \{x : a_k^T x \leq 1, \ k = 1, \dots, K\}$ for appropriately chosen vectors a_k. With this representation it is easy to see that the polytope contains an ellipsoid $\mathcal{E} = \{x : x^T Q^{-1} x \leq 1\}$ if and only if

$$a_k^T x \leq 1, \quad \forall x \in \mathcal{E}, \text{ for } k = 1, \dots K$$

In order to obtain a more convenient condition we introduce the weighted inner product $\langle x, z \rangle_{Q^{-1}} = x^T Q^{-1} z$ and the corresponding norm $|x|_{Q^{-1}} =$

$(x^T Q^{-1} x)^{1/2}$. Then the condition $a_k^T x \leq 1$ for all $x \in \mathcal{E}$ can equivalently be written $\max_{|x|_{Q^{-1}} \leq 1} \langle Qa_k, x\rangle_{Q^{-1}} \leq 1$. Since, $\max_{|x|_{Q^{-1}} \leq 1} \langle Qa_k, x\rangle_{Q^{-1}} = |Qa_k|_{Q^{-1}}$, this is equivalently written $a_k^T Q a_k \leq 1$. The maximal volume ellipsoid included in the polytope can thus be obtained from the optimization problem

$$\min\ \log\det(Q^{-1}) \quad \text{subject to} \quad \begin{cases} a_k^T Q a_k \leq 1 \\ Q > 0 \end{cases}$$

Example 6. A polytopic linear differential inclusion (LDI) is a linear differential equation where the system matrix takes values in a convex polytope

$$\dot{x}(t) = A(t)x(t), \quad A(t) \in \mathrm{co}\{A_1, \ldots, A_N\},\ \forall t \geq 0.$$

The only restriction on the time-variation for the matrix function A is that the solution must be well-defined. The polytopic LDI can be used to model simple classes of switched hybrid systems and uncertain time-varying systems. The advantage of the LDI model is that it is simple to obtain conditions for quadratic stability, it is simple to compute invariant sets and other system properties.

A quadratic Lyapunov function $V(x) = x^T P x$ proves global assymptotic stability for the LDI if and only if

$$A_k^T P + P A_k < 0, \quad k = 1, \ldots, N$$

An ellipsoid $\mathcal{E} = \{x : x^T Q^{-1} x \leq 1\}$ is invariant for the LDI dynamics if every trajectory that starts in $\mathcal{E}$ also stays in this set for all future times, i.e., if $x(0) \in \mathcal{E}$ then $x(t) \in \mathcal{E}$ for $t \geq 0$. Let $V(x) = x^T Q^{-1} x$ and assume $x(0) \in \mathcal{E}$ and $\dfrac{dV(x)}{dt} = x^T(A^T Q^{-1} + Q^{-1} A)x \leq$ for all $x \in \mathbf{R}^n$. Then clearly $V(x) = x(t)^T Q^{-1} x(t) \leq 1$ for $t \geq 0$, i.e. $x(t) \in \mathcal{E}$ for $t \geq 0$.

Example 7. Consider the transfer function $G(s) = C(sI - A)^{-1}B$, where A is a Hurwitz matrix. We will use LMIs to find an upper bound on its induced $\mathbf{L}_\infty$-norm (for the case when the Euclidean norm is used as spatial norm of the input and output vectors. This is not so important in its own right, the idea can also be applied when there is uncertainty and then this is more useful.

In other words, we need to show that $|y(t)| \leq 1$, for all $y = Gu$, where $|u(t)| \leq 1$, for all $t \geq 0$. We address the problem by separating it into two subproblems. First we introduce the state space realization

$$\dot{x} = Ax + Bu, \quad x(0) = 0 \qquad\qquad y = Cx \tag{12}$$

and derive an ellipsoid that contains the reachable set

$$R = \{x(T) : x(T) = \int_0^T e^{A(t-\tau)} Bu(\tau)d\tau,\ x(0) = 0; |u(t)| \leq 1; T \geq 0\}.$$

Then we transform this ellipsoid with the output map $y = Cx$.

Consider the Lyapunov function $V(x) = x^T Px$, where $P = P^T > 0$. If we have

$$dV(x)/dt \leq 0, \text{ for all } (x, u) \text{ such that } V(x) \geq 1 \text{ and } |u| \leq 1, \tag{13}$$

then we know that the ellipsoid $\mathcal{E} = \{x : x^T Px \leq 1\}$ is an invariant set for the state equation (12). The condition (13) can equivalently be stated as

$$2x^T P(Ax + Bu) \leq 0, \text{ for all } (x, u) \text{ with } x^T Px \geq 1,\ |u| \leq 1,$$

which by the S-procedure is true if there exists $\tau_1, \tau_2 \geq 0$ such that

$$2x^T P(Ax + Bu) + \tau_1(x^T Px - 1) + \tau_2(1 - u^T u) \leq 0$$

holds for all (x, u). It is clear that $\tau_1 \geq \tau_2$, since otherwise the inequality would be violated for $(x, u) = 0$. It is also clear that if the inequality holds for some $\tau_1^0 > \tau_2$ then it also holds for $\tau_1 = \tau_2$. Hence, it is nonrestrictive to use $\tau_1 = \tau_2 =: \tau$. The last inequality is thus equivalent to the LMI

$$\begin{bmatrix} A^T P + PA + \tau P & PB \\ B^T P & -\tau I \end{bmatrix} \leq 0. \tag{14}$$

We have thus shown that if (14) holds then $y(t) \in C\mathcal{E} = \{y : y^T (CP^{-1}C^T)^{-1} y \leq 1\}$. Hence, we have $|y(t)| \leq \gamma := \lambda_{\max}(CP^{-1}C^T)$.

Putting everything together we can formulate the result as follows. If there exist $Q > 0$ and $\tau > 0$ such that

$$\begin{bmatrix} QA^T + AQ + \tau Q & B \\ B^T & -\tau I \end{bmatrix} \leq 0 \quad \text{and} \quad \begin{bmatrix} Q & QC^T \\ CQ & \gamma I \end{bmatrix} \geq 0$$

then $|y(t)| \leq \gamma$, for all $y = Gu$, where $|u(t)| \leq 1$. We obtained the first LMI by replacing by using $Q = P^{-1}$ and multiplying in (14) with

$$\begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix}.$$

The second LMI is by the Schur complements formula equivalent to $\gamma \geq \lambda_{\max}(CP^{-1}C^T)$.

Acknowledgement

We are grateful to Ugur Yildiran who pointed out an error in an earlier version of the manuscript.

References

[1] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory.* SIAM Studies in Applied Mathematics. SIAM, Philadelphia, 1994.

[2] Lloyd. L. Dines. On the mapping of quadratic forms. *Bull.Amer.Math.Soc*, 47:494–498, 1941.

[3] E. Feron. *Advances in Systems Theory*, chapter Nonconvex Quadratic Programming, Semidefinite Relaxations and Randomization Algorithms in Information and Decision Systems, pages 255–274. Kluwer Academic Publishers, 1999.

[4] A. L. Fradkov and V. A. Yakubovich. The s-procedure and a duality realations in nonconvex problems of quadratic programming. *Vestnik Leningrad Univ. Math.*, 5(1):101–109, 1979. In Russian 1973.

[5] L. El Ghaoui and S-I Niculescu, editors. *Advances in Matrix Inequality Methods in Control.* Advances in Design and Control. SIAM, 2000.

[6] A.I. Lure and V.N. Postnikov. On the theory of stability of control systems. *Prikl. Mat. i Mekh*, 8:3–13, 1944.

[7] A. Megretski and S. Treil. Power distribution inequalities in optimization and robustness of uncertain systems. *Journal of Mathematical Systems, Estimation, and Control*, 3(3):301–319, 1993.

[8] R.T. Rockafellar. *Convex Analysis*. Princeton, New Jersey, 1970.

[9] F. C. Schweppe. *Uncertain Dynamic Systems*. Prentice-Hall, Englewood Cliffs, N.J., 1973.

[10] V. A. Yakubovich. S-procedure in nonlinear control theory. *Vestnik Leningrad University*, pages 62–77, 1971. (English translation in Vestnik Leningrad Univ. 4:73-93, 1977).

[11] V. A Yakubovich. Minimization of quadratic functionals under quadratic constraints and the necessity of a frequency condition in the quadratic criterion for absolute stability of nonlinear control systems. *Soviet Math. Dokl.*, 14(2):593–597, 1973.