ME 7247: Advanced Control Systems **Fall 2021–22**

The S-lemma

Date: Wednesday December 1, 2021

1 The lossless S-lemma

The lossless S-lemma (or S-procedure) is a statement about one quadratic inequality implying another. It has applications in robust control and constrained optimization.

Where does the name *S-lemma* come from? The following explanation is an excerpt from a survey on the S-Lemma by Pólik and Terlaky[1]

> The term S-method was coined by Aizerman and Gantmacher in their book[2], but later it changed to S-procedure. The S-method tries to decide the stability of a system of linear differential equations by constructing a Lyapunov matrix. During the process an auxiliary matrix S (for stability) is introduced. This construction leads to a system of quadratic equations (the Lur'e resolving equations, 1944). If that quadratic system can be solved, then a suitable Lyapunov function can be constructed. The term S-lemma refers to results stating that such a system can be solved under certain conditions; the first such result is due to Yakubovich (1971).

Here is the result.

Theorem 1.1 (Lossless S-lemma)**.** *Suppose P_0 and P_1 are symmetric matrices of the same size. The following statements are equivalent.*

(i) If x satisfies $x^\mathsf{T} P_1 x \le 0$, then we have $x^\mathsf{T} P_0 x \le 0$.

(ii) There exists $\lambda \ge 0$ such that $P_0 \preceq \lambda P_1$.

Proof. Suppose that (ii) holds. Then there exists $\lambda \ge 0$ such that $P_0 \preceq \lambda P_1$. Therefore,

$$x^\mathsf{T} P_0 x \le \lambda x^\mathsf{T} P_1 x \qquad \text{for all } x. \tag{1}$$

If $x^\mathsf{T} P_1 x \le 0$, then from Eq. (1), we have $x^\mathsf{T} P_0 \le 0$, and therefore (i) holds. This proves (ii) $\implies$ (i).

Now we prove the difficult direction. Define the sets

$$S := \left\{ \begin{bmatrix} x^\mathsf{T} P_1 x \\ x^\mathsf{T} P_0 x \end{bmatrix} \;\middle|\; x \in \mathbb{R}^n \right\}, \qquad T := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^2 \;\middle|\; u \le 0 \text{ and } v > 0 \right\}.$$

Both S and T are subsets of $\mathbb{R}^2$. Now suppose that (i) holds. We prove two important properties:

S and T are disjoint. To see why, suppose $\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in S$. Then, we have $u = x^\mathsf{T} P_1 x$ and $v = x^\mathsf{T} P_0 x$ for some $x \in \mathbb{R}^n$. We know from (i) that if $u \le 0$, then we must have $v \le 0$. This means that $\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \notin T$, and therefore S and T are disjoint.

[1] I. Pólik and T. Terlaky, *A Survey of the S-Lemma*, SIAM Review, Volume 49, 2007, Pages 371–418.

[2] M. A. Aizerman and F. R. Gantmacher, Absolute Stability of Regulator Systems, Holden-Day Series in Information Systems, Holden-Day, San Francisco, 1964. Originally published as *Absolutnaya Ustoichivost' Reguliruyemykh Sistem* by The Academy of Sciences of the USSR, Moscow, 1963

S is a cone. In other words, if $z \in S$, then $\alpha z \in S$ for all $\alpha \geq 0$. To see why, suppose $\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in S$. Then, we have $u = x^\mathsf{T} P_1 x$ and $v = x^\mathsf{T} P_0 x$ for some $x \in \mathbb{R}^n$. Therefore:

$$\alpha \begin{bmatrix} u \\ v \end{bmatrix} = \alpha \begin{bmatrix} x^\mathsf{T} P_1 x \\ x^\mathsf{T} P_0 x \end{bmatrix} = \begin{bmatrix} (\sqrt{\alpha} x)^\mathsf{T} P_1 (\sqrt{\alpha} x) \\ (\sqrt{\alpha} x)^\mathsf{T} P_0 (\sqrt{\alpha} x) \end{bmatrix} \in S$$

S is convex. To see why, suppose $z_1 \in S$ and $z_2 =\in S$. We would like to show that $\alpha z_1 + (1-\alpha) z_2 \in S$ for all $\alpha \in [0,1]$. Since S is a cone, this is equivalent to proving that $\alpha z_1 + \beta z_2 \in S$ for all $\alpha, \beta \geq 0$. We consider two cases.

If z_1 and z_2 are linearly dependent, then $z_2 = c z_1$ for some $c \neq 0$. But since S is a cone, we have

$$\alpha z_1 + \beta z_2 = (\alpha + c\beta) z_1 = \tfrac{1}{c} (\alpha + c\beta) z_2$$

If $c > 0$, then we have $(\alpha + c\beta) z_1 \in S$ because of the cone property. If $c < 0$, then either $(\alpha + c\beta)$ or $\frac{1}{c}(\alpha + c\beta)$ will be positive, so we can use the cone property again to show that either $(\alpha + c\beta) z_1 \in S$ or $\frac{1}{c}(\alpha + c\beta) z_2 \in S$. In conclusion, we have $\alpha z_1 + \beta z_2 \in S$.

Suppose instead that z_1 and z_2 are linearly independent. Since $z_1, z_2 \in S$, we can let x and y be defined such that

$$z_1 = \begin{bmatrix} x^\mathsf{T} P_1 x \\ x^\mathsf{T} P_0 x \end{bmatrix}, \qquad \text{and} \qquad z_2 = \begin{bmatrix} y^\mathsf{T} P_1 y \\ y^\mathsf{T} P_0 y \end{bmatrix}.$$

Since z_1 and z_2 are linearly independent, then every vector in $\mathbb{R}^2$ can be expressed as a linear combination of z_1 and z_2. In particular, there must exist $a, b \in \mathbb{R}$ such that

$$\begin{bmatrix} x^\mathsf{T} P_1 y \\ x^\mathsf{T} P_0 y \end{bmatrix} = a z_1 + b z_2$$

In order to show that $\alpha z_1 + \beta z_2 \in S$, we must show that there exists some $w \in \mathbb{R}^n$ such that

$$\alpha z_1 + \beta z_2 = \begin{bmatrix} w^\mathsf{T} P_1 w \\ w^\mathsf{T} P_0 w \end{bmatrix}$$

We will look for w of the form $w = px + qy$. This leads to:

$$\begin{aligned} \alpha z_1 + \beta z_2 &= \begin{bmatrix} (px+qy)^\mathsf{T} P_1 (px+qy) \\ (px+qy)^\mathsf{T} P_0 (px+qy) \end{bmatrix} \\ &= \begin{bmatrix} p^2 x^\mathsf{T} P_1 x + 2pq x^\mathsf{T} P_1 y + q^2 y^\mathsf{T} P_1 y \\ p^2 x^\mathsf{T} P_0 x + 2pq x^\mathsf{T} P_0 y + q^2 y^\mathsf{T} P_0 y \end{bmatrix} \\ &= p^2 z_1 + 2pq(a z_1 + b z_2) + q^2 z_2 \\ &= (p^2 + 2pqa) z_1 + (q^2 + 2pqb) z_2. \end{aligned}$$

In other words:

$$\alpha = p^2 + 2pqa \quad \text{and} \quad \beta = q^2 + 2pqb. \tag{2}$$

Remember: α, β, a, b are fixed, and we are tasked with showing that we can always find a real pair (p, q) that satisfies the above two equations. Let's write $q = mp$ and eliminate q. This leads to:

$$\alpha = p^2(1 + 2am) \quad \text{and} \quad \beta = p^2(m^2 + 2bm).$$

Eliminating p, we obtain

$$\alpha(m^2 + 2bm) = \beta(1 + 2am)$$

The solutions are given by

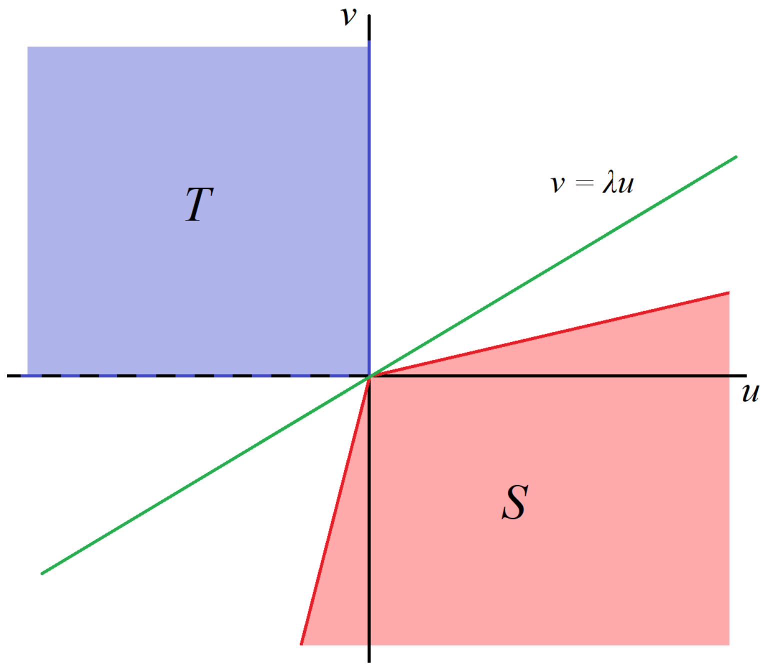
$$m = \frac{(2a\beta - 2b\alpha) \pm \sqrt{(2a\beta - 2b\alpha)^2 + 4\alpha\beta}}{2\alpha} \tag{3}$$

Since $\alpha > 0$ and $\beta > 0$, these solutions are real. One of them solutions is positive and the other is negative. Pick the one for which $am \geq 0$. Then we can solve for p and q and obtain

$$p = \frac{\sqrt{\alpha}}{\sqrt{1 + 2am}} \quad \text{and} \quad q = \frac{m\sqrt{\alpha}}{\sqrt{1 + 2am}}$$

It is straightforward to check that these choices together with (3) satisfy (2).

Putting everything together. We have established that S is a convex cone, and disjoint from T. Here is a diagram of what S and T might look like.



We now make use of the fact that disjoint convex sets can always be *separated*. In other words, we can find a hyperplane such that each set is on a different side of the hyperplane. Specifically, we can find a $\lambda \geq 0$ such that S lies below while T lies above, as shown in the figure above. The reason for λ being nonnegative is so that the line does not intersect T (note that the lower boundary of T is not included in T). Now S lies below, which means that: for all $\left[\begin{smallmatrix} u \\ v \end{smallmatrix}\right] \in S$, we have $v \leq \lambda u$. From the definition of S, this is the same as saying that for all x, we have $x^\mathsf{T} P_0 x \leq \lambda x^\mathsf{T} P_1 x$. In other words, we have $P_0 \preceq \lambda P_1$, as required. Therefore (i) $\Longrightarrow$ (ii) and the proof is complete. ■

2 The lossy S-lemma

When we have multiple quadratic constraints, only the easy direction holds.

Theorem 2.1 (Lossy S-lemma)**.** *Suppose $P_0, \dots, P_m$ are symmetric matrices of the same size. Consider the following statements.*

(i) If x satisfies $x^\mathsf{T} P_k x \le 0$ for $k = 1, \dots, m$, then $x^\mathsf{T} P_0 x \le 0$.

(ii) There exist $\lambda_1, \dots, \lambda_m \ge 0$ such that $P_0 \preceq \sum_{k=1}^m \lambda_k P_k$.

Then we have (ii) $\Longrightarrow$ (i).

.

The proof is the same as in Theorem 1.1. The reason the same approach cannot be used to prove the converse here is that the set

$$S = \left\{ (x^\mathsf{T} P_0 x, \, \dots, \, x^\mathsf{T} P_m x) \in \mathbb{R}^{m+1} \;\middle|\; x \in \mathbb{R}^m \right\}$$

is only guaranteed to be convex when $m = 1$.