S-Procedure

Chung-Yao Kao

Optimization and Systems Theory

Royal Institute of Technology

Sweden

Preliminary Material

- A set C in a linear vector space is *convex* if $\alpha x_1 + (1-\alpha)x_2 \in C$, for all $x_1, x_2 \in C$, and $\alpha \in [0,1]$.
- A set C is a *convex cone* if C is convex and in addition $\alpha x \in C$ for all $x \in C$ and all $\alpha \geq 0$.
- A *convex polytope with vertices* $\{x_1, \cdots, x_n\}$ is defined as the convex hull of these points: $C := \mathbf{CO}(\{x_1, \cdots, x_n\})$, where

$$\mathbf{CO}(\{x_1, \cdots, x_n\}) = \left\{ \sum_{i=1}^{n} \alpha_i x_i : \alpha_i \geq 0, \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

- An *ellipsoid* $\mathcal{E}$ with center x_c can be defined by $\mathcal{E} = \{x \mid (x-x_c)'Q(x-x_c) \leq 1\} = \{x \mid x'Qx + 2x'b + c \leq 0\} = \{By + u \mid \|y\| \leq 1\}$

- A *linear functional* on a vector space V (over real number $\mathbf{R}$) is a function $f : V \to \mathbf{R}$ which satisfies $f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f(x_1) + \alpha_2 f(x_2)$ for all $v_1, v_2 \in V$ and $\alpha_1, \alpha_2 \in \mathbf{R}$.
- The continuous linear functionals on V form a vector space which is called *the dual space* of V and is denoted as V^*.
- For any $v^* \in V^*$, $v^*(v)$ is also denoted as $\langle v, v^* \rangle$.
- For any finite dimensional Hilbert space V (over $\mathbf{R}$), any element in the dual space V^* can be identified by an element in V; i.e., $V^* = V$. The linear functional $\langle v, v^* \rangle$ is nothing but inner product.

Example 1. *The dual space of* $\mathbf{R}^n$ *is* $\mathbf{R}^n$. *The linear functionals* $\langle x, z \rangle$ *are defined as* $x'z$.

Example 2. *The dual space of the vector space* $S^{n\times n} := \{X \in \mathbf{R}^{n\times n} \mid X = X'\}$ *is* $S^{n\times n}$ *itself. The linear functionals* $\langle X, Z \rangle$ *are defined as* $\mathrm{tr}(X'Z)$.

- The *affine hull* of a set $S \subset \mathbf{R}^m$ (denoted as $\mathrm{aff}S$) is the set of all linear combination of the form $\{\sum_{i=1}^{n}, \alpha_i x_i \mid x_i \in S, \sum \alpha_i = 1\}$.
- The *relative interior* of a set $S \subset \mathbf{R}^m$, $\mathrm{ri}S$, is the interior relative to $\mathrm{aff}S$.
- A *hyperplane* is an affine subset of maximal dimension. Every hyperplane of V can be represented as $H = \{x \in V \mid \langle x, z \rangle = c\}$ for suitably chosen nonzero $z \in V^*$ and $c \in \mathbf{R}$.

Lemma 1 (Separation Lemma). *Let V be a finite dimensional space. Let $C \subset V$ be a relatively open convex set. Let W be a linear subspace of V such that $C \bigcap W = \emptyset$. Then, there exists a non-zero $z \in V^*$ such that*

$$\langle x, z \rangle > 0, \quad \forall x \in C,$$
$$\langle x, z \rangle \leq 0, \quad \forall x \in W.$$

Theorem 1. *Let V be a finite dimensional space. Let C_1 be an open convex cone and C_2 be a convex set. Furthermore, suppose that $C_1 \bigcap C_2 = \emptyset$. Then, there exists a non-zero $z \in V^*$ such that*

$$\langle x, z \rangle > 0, \quad \forall x \in C_1,$$
$$\langle x, z \rangle \leq 0, \quad \forall x \in C_2.$$

Proof. Let $C = C_1 - C_2$. Then C is open and does not contain 0.

Hence, by the separation lemma, there exists z such that $\langle x, z \rangle > 0$, for all $x \in C$. In other words, we have $\langle x_1, z \rangle > \langle x_2, z \rangle$ for all $x_1 \in C_1$ and $x_2 \in C_2$. Now let $\alpha = \inf_{x_1 \in C_1} \langle x_1, z \rangle$. Since C_1 is a cone, we conclude that $\alpha = 0$. We thus proof that

$$\langle x_1, z \rangle \geq 0, \quad \forall x_1 \in C_1,$$
$$\langle x_2, z \rangle \leq 0, \quad \forall x_2 \in C_2.$$

Finally, $\langle x_1, z \rangle > 0$ for all $x_1 \in C_1$ since C_1 is open. □

The S-Procedure

The basic idea: Let $\sigma_i : V \to \mathbf{R}$, $i = 0, 1, \cdots, n$, be functionals defined on a linear vector space V. Consider the following two statements

S_1: $\sigma_0(y) \leq 0$ for all $y \in V$ which satisfy $\sigma_i(y) \leq 0$, $i = 1, \cdots, n$.
That is, $\sigma_0(x) \leq 0$ for all $y \in \Omega := \{y \mid \sigma_i(y) \leq 0, \ i = 1, \cdots, n\}$.

S_2: There exist $\tau_i \geq 0$, $i = 1, \cdots, n$, such that

$$\sigma_0(y) - \sum_{i=1}^{n} \tau_i \sigma_i(y) \leq 0, \quad y \in V$$

A trivial fact: S_2 implies S_1.

The S-procedure is the method of checking S_1 using S_2.

Reason: checking S_2 is usually much simpler than checking S_1.

Example: Consider ellipsoids $E_i := \{y \mid y' Q_i y + 2y' s_i + r_i \leq 0\}$, $i = 0, \cdots, n$. Want to check whether E_0 contains the $\bigcap_{i=1}^{n} E_i$. This problem is NP-complete. A sufficient condition can be obtained using S_2 which reads:
Find $\tau_i \geq 0$, $i = 1, \cdots, n$, such that

$$\begin{bmatrix} y \\ 1 \end{bmatrix}' \left(\begin{bmatrix} Q_0 & s_0 \\ s_0' & r_0 \end{bmatrix} - \sum_{i=1}^{n} \tau_i \begin{bmatrix} Q_i & s_i \\ s_i' & r_i \end{bmatrix} \right) \begin{bmatrix} y \\ 1 \end{bmatrix} \leq 0, \quad \forall\ y$$

This is an LMI problem!

For those who do not know what NP-complete means (like me):
it means the problem is very difficult!

efinition 1. Let $\sigma_i : V \to R$. The constraints $\sigma_i(y) \geq 0$, $k = 1, \cdots, n$, are said to be regular if there exists $y^* \in V$ such that $\sigma_i(y^*) > 0$, $i = 1, \cdots, n$.

Theorem 2 (S-procedure lossless Theorem). *Let $\sigma_i : V \to R$, $k = 0, \cdots, n$, and assume that $\sigma_i(y) \geq 0$, $i = 1, \cdots, n$ are regular. Let sets*

$$\mathcal{M} = \{(\sigma_0(y), \sigma_1(y), \cdots, \sigma_n(y)) \mid y \in V\}$$
$$\mathcal{N} = \{(N_0, N_1, \cdots, N_n) \mid N_0 > 0,\ N_i < 0,\ i = 1, \cdots, n\}$$

If $\mathcal{M} \bigcap \mathcal{N} = \emptyset$ implies $\mathbf{CO}(\mathcal{M}) \bigcap \mathcal{N} = \emptyset$, then the S-procedure is lossless. That is S_1 is true **if and only if** *S_2 is true.*

Proof. S_1 implies that $\mathcal{M} \bigcap \mathcal{N} = \emptyset$, which in turn implies that $co(\mathcal{M}) \bigcap \mathcal{N} = \emptyset$. $\mathcal{N}$ is an open convex cone, and hence the separating theorem implies that there exists a non-zero $n+1$-tuple

$), -c_1, \cdots, -c_n)$ such that

$$c_0 N_0 - c_1 N_1 - \cdots - c_n N_n < 0, \quad \forall (N_0, N_1, \cdots, N_n) \in \mathcal{N} \tag{1}$$
$$c_0 m_0 - c_1 m_1 - \cdots - c_n m_n < 0, \quad \forall (m_0, m_1, \cdots, m_n) \in \mathcal{M} \tag{2}$$

For any $\epsilon < 0$, we have $(n_0, \epsilon, \cdots, \epsilon) \in \mathcal{N}$ for all $n_0 > 0$. Hence, (1) that $c_0 \leq 0$. By similar arguments, we can show that $c_i \leq 0$, $i = 1, \cdots, n$. By regularity assumption, there exists y^* such that $\sigma_i(y^*) < 0$ for $i = 1, \cdots, n$. This implies that one can find a point $(m_0, m_1, \cdots, m_n) \in \mathcal{M}$ such that $m_i < 0$, $i = 1, \cdots, n$. This and (2)in turn imply that $c_0 < 0$. Since for any $y \in V$, $(\sigma_0(y), \sigma_1(y), \ldots, \sigma_n(y)) \in co(\mathcal{M})$, we conclude that

$$\sigma_0(y) - \sum_{i=1}^{n} \frac{c_i}{c_0} \sigma_i(y) \leq 0, \forall y \in V.$$

This shows that S_2 holds for $\tau_i = \frac{c_i}{c_0} \geq 0$. □

emark 1. If $\mathcal{M}$ is convex then S-procedure is lossless.

Corollary 1 (S-procedure lossless for linear functionals). *Consider the case where $\sigma_i(y) = s_i' y + r_i$, $i = 0, \cdots, n$. If the constraints are regular then S-procedure is lossless.*

Proof. The statement follows since $\mathcal{M}$ is convex. Indeed, let m_1, $m_2 \in \mathcal{M}$, we have

$$m_1 = (s_0' y_1 + r_0, \cdots, s_n' y_1 + r_n), \quad \text{for some } y_1 \in V$$
$$m_2 = (s_0' y_2 + r_0, \cdots, s_n' y_2 + r_n), \quad \text{for some } y_2 \in V$$

Hence, for any $\lambda \in [0, 1]$

$$\lambda m_1 + (1 - \lambda) m_2 = (s_0' y_\lambda + r_0, \cdots, s_n' y_\lambda + r_n),$$

where $y_\lambda = \lambda y_1 + (1 - \lambda) y_2 \in V$. This implies that $\lambda m_1 + (1 - \lambda) m_2 \in \mathcal{M}$, and hence $\mathcal{M}$ is convex. □

heorem 3 (S-procedure lossless for quadratic functionals). *Consider the case where $\sigma_i(y) = y' Q_i y + 2 s_i' y + r_i$, $i = 0, 1$. If the constraint is regular then S-procedure is lossless. More precisely:*

- *The following two statements are equivalent*
  1. *$y' Q_0 y + 2 s_0' y + r_0 \leq 0$ for all $y \in \mathbf{R}^m$ such that $y' Q_1 y + 2 s_1' y + r_1 \leq 0$*
  2. *there exist $\tau \geq 0$ such that $y' Q_0 y + 2 s_0' y + r_0 - \tau(y' Q_1 y + 2 s_1' y + r_1) \leq 0$ for all $y \in \mathbf{R}^m$.*
- *The following two statements are equivalent*
  1. *$y' Q_0 y + 2 s_0' y + r_0 < 0$ for all $y \in \mathbf{R}^m$ such that $y' Q_1 y + 2 s_1' y + r_1 \leq 0$*
  2. *there exist $\tau \geq 0$ such that $y' Q_0 y + 2 s_0' y + r_0 - \tau(y' Q_1 y + 2 s_1' y + r_1) < 0$ for all $y \in \mathbf{R}^m$.*

Remark 2. In this theorem, the quadratic functionals are real-valued. $Q_i = Q_i' \in \mathbf{R}^{m\times m}$, $s_i \in \mathbf{R}^m$, and $r_i \in \mathbf{R}$. We have the following theorem for the complex-valued case. Note the number of constraints one can have without having a lossy S-procedure.

Theorem 4 (S-procedure lossless for quadratic functionals of complex valued). *Let $Q_i = Q_i^* \in \mathbf{C}^{m\times m}$, $s_i \in \mathbf{C}^m$, and $r_i \in \mathbf{R}$. We have the following two cases*

- *Let $\sigma_i(y) = y^*Q_iy + 2\mathbf{Re}(s_i^*y) + r_i$, $k = 0, 1, 2$. Assume that $\sigma_i(y) \le 0$, $k = 1, 2$ are regular. The followings are equivalent*
  1. *$\sigma_0(y) \le 0$ for all $y \in \mathbf{C}^m$ such that $\sigma_i(y) \le 0$, $i = 1, 2$.*
  2. *there exist $\tau_1 \ge 0$, $\tau_2 \ge 0$ such that*

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

- *Let $\sigma_i(y) = y^*Q_iy$, $k = 0, 1, 2$. Assume that $\sigma_i(y) \le 0$, $k = 1, 2$ are regular. The followings are equivalent*
  1. *$\sigma_0(y) \le 0$ for all $y \in \mathbf{C}^m$ such that $\sigma_i(y) \le 0$, $i = 1, 2$.*
  2. *there exist $\tau_1 \ge 0$, $\tau_2 \ge 0$ such that $Q_0 - \tau_1Q_1 - \tau_2Q_2 \le 0$.*

Proof of Theorems 3 and 4. The main steps are to show that closures of the sets

$$\{(\sigma_0(y), \sigma_1(y)) \mid y \in \mathbf{R}^n\},$$

(where σ_i are as defined in Theorem 3) and

$$\{(\sigma_0(y), \sigma_1(y), \sigma_2(y)) \mid y \in \mathbf{C}^n\}$$

(where σ_i are as defined in Theorem 4) are convex. Then the theorems follow Theorem 2. The proof of convexities are due to Yakubovich. □

Let $\mathcal{H}$ be a Hilbert space and define quadratic forms $\sigma_i : \mathcal{H} \to \mathbf{R}$ as

$$\sigma_i(y) = \langle \Psi_i y, y\rangle, \quad i = 0, 1, \cdots, n$$

where Ψ_i are linear bounded self-adjoint operators on $\mathcal{H}$. Let $S_\tau : \mathcal{H} \to \mathcal{H}$ be the shift operator defined as $(S_\tau y)(t) = y(t - \tau)$. Assume that the Hilbert space, its inner product, and the self-adjoint operators Ψ_i are such that the following properties hold

- if $y \in \mathcal{H}$ then $S_\tau y \in \mathcal{H}$ for all $\tau \ge 0$.
- $\langle \Psi_i S_\tau y_1, y_2\rangle \to 0$, as $\tau \to \infty$.
- $\langle \Psi_i y_1, S_\tau y_2\rangle \to 0$, as $\tau \to \infty$.
- $\sigma_i(S_\tau y) = \sigma_i(y)$ for all $\tau \ge 0$ and for all $y \in \mathcal{H}$.

Theorem 5. *Assume the quadratic forms and the Hilbert space which they are defined on satisfy the above mentioned properties. Furthermore assume that the constraints $\sigma_i(y) \le 0$, $i = 1, \cdots, n$, are regular. Then the following statements are equivalent*

1. *$\sigma_0(y) \le 0$ for all y such that $\sigma_i(y) \le 0$, $i = 1, \cdots, n$.*
2. *There exist $\tau_i \ge 0$, $i = 1, \cdots, n$, such that*

$$\sigma_0(y) - \sum_{i=1}^{n} \tau_i\sigma_i(y) \le 0, \quad \forall\ y \in \mathcal{H}$$

Proof. The main point is to show that the closure of

$$\mathcal{M} := \{(\sigma_0(y), \sigma_1(y), \cdots, \sigma_n(y)) \mid y \in \mathcal{H}\}$$

convex. Let y_1, y_2 belong to $\mathcal{H}$ and let

$$m_1 = (\sigma_0(y_1), \sigma_1(y_1), \cdots, \sigma_n(y_1)) \in \mathcal{M}$$
$$m_2 = (\sigma_0(y_2), \sigma_1(y_2), \cdots, \sigma_n(y_2)) \in \mathcal{M}$$

Let $\lambda \in [0, 1]$ and consider $\sigma_i(\sqrt{\lambda} y_1 + \sqrt{1-\lambda} S_\tau y_2)$. We have

$$\sigma_i(\sqrt{\lambda} y_1 + \sqrt{1-\lambda} S_\tau y_2) := \lambda \sigma_i(y_1) + (1-\lambda)\sigma_i(y_2) + \sqrt{\lambda(1-\lambda)})(\langle \Psi_i y_1, S_\tau y_2 \rangle + \langle \Psi_i S_\tau y_2, y_1 \rangle)$$

As $\tau \to \infty$, we have

$$\sigma_i(\sqrt{\lambda} y_1 + \sqrt{1-\lambda} S_\tau y_2) \to \lambda \sigma_i(y_1) + (1-\lambda)\sigma_i(y_2)$$

The implies that, as $\tau \to \infty$,

$$(\sigma_0(y_\tau), \cdots, \sigma_n(y_\tau)) \to \lambda m_1 + (1-\lambda) m_2.$$

Hence, $\lambda m_1 + (1-\lambda) m_2$ belong to the closure of $\mathcal{M}$. $\square$