SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The S-Procedure and the Kalman-Yakubovich-Popov Lemma

av

Anu Kokkarinen

2012 - No 30

The S-Procedure and the Kalman-Yakubovich-Popov Lemma

Anu Kokkarinen

Självständigt arbete i matematik 30 högskolepoäng, Avancerad nivå

Handledare: Yishao Zhou

2012

Abstract

In this paper we study two classical control theory topics: the S-procedure and the Kalman-Yakubovich-Popov Lemma. Using Fenchel duality one can show that the S-procedure is lossless for a class of quadratic functions. We apply this result to derive a convex dual problem for certain optimization problems. Fenchel duality is also used to prove an extended version of the Kalman-Yakubovich-Popov lemma.

Contents

1 Introduction

The *S-procedure* is a method of confirming that a hard-to-access inequality holds by showing that another, stronger result is true. The *losslessness of the S-procedure* refers to the equivalence of this "inequality" and the "stronger result". This equivalence lies behind Lagrangian duality, but it also has various applications in control theory. We shall demonstrate how the S-procedure can be used to derive a convex dual problem for a non-convex optimization problem.

The *Kalman-Yakubovich-Popov (KYP) lemma* has its origins in the stability analysis of non-linear control systems. There are various different formulations of the lemma, and not all of them are equivalent. The KYP-lemma is a more general version of the positive real lemma, and it is closely related to the bounded real lemma.

Generally speaking, the lemma states that the following assertions are equivalent:

(1) The frequency condition holds

(2) There exists a solution to the Lur'e equation

(3) There exists a solution to the corresponding LMI

In this paper we use a relatively uncommon optimization method to establish our results: Fenchel duality. Although Fenchel duality has many similarities to the more popular Lagrangian duality – in fact, they can be shown to be equivalent – it in some cases leads to more approachable and even more general results. The main focus of this paper is on S. V. Gusev's article *The Fenchel duality, S-procedure, and the Yakubovich-Kalman Lemma* [1], in which the author uses Fenchel duality to find conditions under which the S-procedure is lossless and to prove an extended version of the Kalman-Yakubovich-Popov lemma.

Chapters 2 and 3 concern convex optimization. We go through the basic definitions and state and prove both the Lagrangian and Fenchel's duality theorem. In Chapter 4 we introduce the S-procedure, and Chapter 5 is dedicated to the Kalman-Yakubovich-Popov lemma.

1.1 Notation

Scalars and scalar-valued functions are denoted by small letters: $x \in \mathbb{C}$, $f : \mathcal{X} \to \mathbb{R}$. Bold small letters are used for column vectors and vector-valued functions: $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^T \in \mathbb{R}^n$, $\boldsymbol{g} = (g_1, g_2, \ldots, g_m)^T : \mathcal{X} \to \mathbb{R}^m$. Matrices are denoted by capital letters: $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{C}^{k \times m}$.

2 Convex Optimization

2.1 Basic Definitions

This section works as a reminder of basic definitions in the field of convex optimization.

Definition 2.1 (Convex set)**.** *A set $C \subseteq \mathbb{R}^n$ is called* convex *if for all $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in C$ and $\lambda \in [0,1]$ the following holds:*

$$\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in C$$

In other words a set C is convex if the line segment joining two arbitrary points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in C is entirely contained in the set. This is illustrated in Figure 1.

Definition 2.2 (Convex and concave function)**.** *A function $f : C \to \mathbb{R}$ is convex if the following inequality holds for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ and $\lambda \in [0,1]$:*

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$$

If the above inequality is strict, the function is called strictly convex*. A function g is called (strictly)* concave *if $-g$ is (strictly) convex.*

Figure 2 gives an illustration of Definition 2.2. It is worth noting that the only functions that are both convex and concave are the affine functions. It is also good to remember that a function cannot be convex if its domain C is not convex (naturally, the opposite is not true; the convexity of C does not guarantee the convexity of f).

Definition 2.3 (Epigraph)**.** *The* epigraph *of a function f is a subset of $\mathbb{R}^{n+1}$ defined by:*

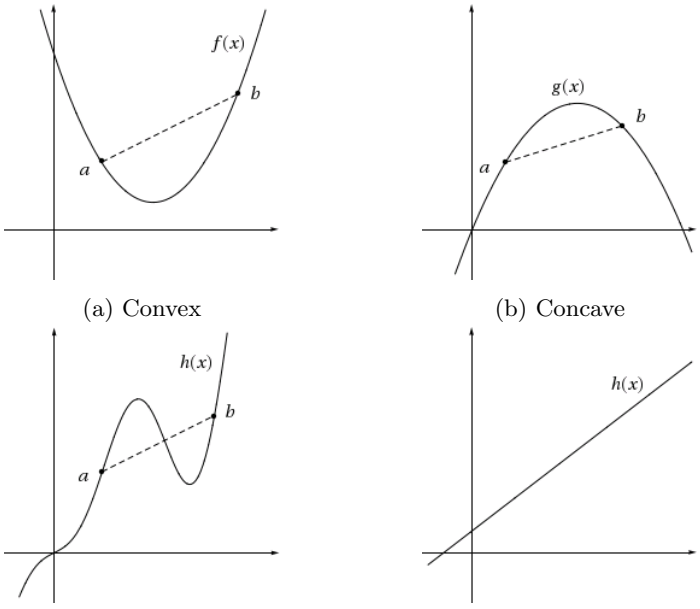
$$\operatorname{epi} f = \{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} \mid \boldsymbol{x} \in C,\ y \in \mathbb{R},\ y \geq f(\boldsymbol{x})\}$$

Similarly, the hypograph *of f is given by*

$$\operatorname{hyp} f = \{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} \mid \boldsymbol{x} \in C,\ y \in \mathbb{R},\ y \leq f(\boldsymbol{x})\}$$

(a) (b)

Figure 1: The set (a) is convex; regardless of how we pick $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ the line segment joining them lies in (a). As seen above, this is not true for the set (b).



(a) Convex (b) Concave

(c) Neither convex nor concave (d) Both convex and concave

Figure 2: The function f in (a) is convex: all points that lie on the line segment joining a and b lie above the graph of the function. Similarly, the function g in (b) is concave since the line segment always lies below the graph. h in (c) is neither convex nor concave, whereas the affine function in (d) is both.

In $\mathbb{R}^2$ and $\mathbb{R}^3$ the epigraph can be characterized as all points that lie above the graph of the function. Similarly, the hypograph consists of all points that lie below the function. One way to define a convex function f is to require that the line segment joining the images of two arbitrary points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ in C lies entirely in the epigraph of the function. Notice that in some literature the hypograph of a concave function is called the epigraph.

The next proposition establishes an important connection between the epigraph and convexity. It is so fundamental that it is sometimes used as the definition of a convex function.

Proposition 2.4. *A function f is convex if and only if its epigraph is convex. Similarly, f is concave if and only if its hypograph is convex.*

In order to be able to apply our theory more generally, we shall need to distinguish between different kinds of interiors.

Definition 2.5 (Interior and relative interior)**.** *Let $C \subseteq \mathbb{R}^n$. The* interior (int) *of C is given by all points that are surrounded by a sphere completely contained in C. The* relative interior (ri) *consists of all points that lie in the interior of C with respect to the smallest subspace containing the set C.*

We shall illustrate the difference between the regular and the relative interior with an exmple.

Example 2.6. Let C be the unit disk in $\mathbb{R}^2$, that is $C = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. Its interior points are the points $(x, y) \in \mathbb{R}^2$ that satisfy the strict inequality $x^2 + y^2 < 1$. The smallest subspace is containing C is $\mathbb{R}^2$ itself. Hence $\operatorname{ri} C$ is the same as $\operatorname{int} C$.

Consider now the unit disk in $\mathbb{R}^3$: $D = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq 1 \text{ and } z = 0\}$. This set is lies on the xy-plane. It has no interior points in $\mathbb{R}^3$. The smallest subspace containing D is $\mathbb{R}^2$. Hence the relative interior consists of all points (x, y, z) that satisfy $x^2 + y^2 < 1$ and $z = 0$. This can be considered the same as $\operatorname{int} C$. $\Diamond$

Lastly, by a *cone* we shall mean a cone with vertex at $\mathbf{0}$ as given in the next definition. Generally, a cone can have an arbitrary point as its vertex, but in order to simplify our calculations in the later chapters we shall restrict our attention to this special case.

Definition 2.7 (Cone)**.** *A set C is called a* cone *if $\boldsymbol{x} \in C$ implies $\lambda \boldsymbol{x} \in C$ for all $\lambda > 0$ and $\boldsymbol{x} \in C$. A cone that is convex is called a* convex cone*.*

2.2 Convex Optimization

Convexity is a very useful concept in optimization. There are many reasons for this, the most important of which being the fact that a local optimum under certain convexity assumptions becomes a global optimum. Optimization problems are usually expressed in the standard form:

$$\begin{array}{lll} \text{Minimize} & f(\boldsymbol{x}) & \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 & i = 1, 2, \ldots, m \\ & h_j(\boldsymbol{x}) = 0 & j = 1, 2, \ldots, l \\ & \boldsymbol{x} \in C & \end{array} \tag{2.1}$$

where f, g_i and h_j are real-valued functions defined (at least) on a subset C of $\mathbb{R}^n$.

f, the function we wish to minimize, is called the *objective function*, g_i, $i = 1, 2, \ldots, m$ are called the *inequality constraints*, h_j, $j = 1, 2, \ldots, l$ the *equality constraints* and the set restricted by the constraints, i.e.

$$S = \{x \in C \mid g_i(x) \leq 0,\ i = 1, 2, \ldots, m,\ h_j(\boldsymbol{x}) = 0,\ j = 1, 2, \ldots, l\}$$

is called the *feasible region*. A point that lies in the set S is called *a feasible point* and an optimization problem that can be solved is called *feasible*.

(2.1) is called convex if the objective function f and the g_i's are convex, h_j's are affine and C is convex.

In a way convex optimization problems are the simplest after linear programs. Various different methods have been developed to solve convex problems. However, merely assuming that the functions involved are convex or affine is not always enough. Some regularity condition is usually pressed on the set of constraints. One very common such is given in the following definition.

Definition 2.8 (Slater's condition)**.** *Consider the optimization problem given by* (2.1)*.* Slater's condition *is said to hold if there exists $\bar{\boldsymbol{x}} \in \operatorname{ri} C$ such that $g_i(\bar{\boldsymbol{x}}) < 0$ for $i = 1, 2, \ldots, m$ and $h_j(\bar{\boldsymbol{x}}) = 0$, $j = 1, 2, \ldots, m$.*

Remark. Constraints satisfying Slater's conditions are sometimes referred to as "regular constraints". $\Diamond$

Remark. Due to the vast amount of different applications in which Slater's condition appears there are many different versions of the above definition. For instance, a problem may only have inequality constraints and no equality constraints. Or the inequality constraints may be expressed in the form $\boldsymbol{g}(\boldsymbol{x}) \geq \mathbf{0}$ in which case the regularity condition becomes $\boldsymbol{g}(\bar{\boldsymbol{x}}) > \mathbf{0}$. We shall use the term Slater's condition even when referring to conditions not strictly speaking equivalent to the above definition. $\Diamond$

2.3 The Separating Hyperplane Theorem

Definition 2.9 (Hyperplane)**.** *A* hyperplane *is a set determined by an affine function as follows:*

$$H = \{\boldsymbol{x} \in \mathbb{R}^n \mid \langle \boldsymbol{x}, \boldsymbol{p} \rangle = \alpha\}$$

where $\boldsymbol{p} \in \mathbb{R}^n / \{\boldsymbol{0}\}$ *and* $\alpha \in \mathbb{R}$*.*

In $\mathbb{R}^2$ a hyperplane is a line, in $\mathbb{R}^3$ a plane (hence the word hyperplane). In general, a hyperplane in $\mathbb{R}^n$ has dimension $n-1$. What characterizes a hyperplane is that it divides the space into two separate subspaces.

Suppose C_1 and C_2 are sets in $\mathbb{R}^n$. A hyperplane H, as given in Definition 2.9, is said to *separate* C_1 and C_2 if $\langle \boldsymbol{x}_1, \boldsymbol{p} \rangle \geq \alpha$ for all $\boldsymbol{x}_1 \in C_1$ and $\langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \leq \alpha$ for all $\boldsymbol{x}_2 \in C_2$. This leads to the inequalities

$$\sup_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \leq \inf_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle \tag{2.2}$$

$$\inf_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \leq \sup_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle \tag{2.3}$$

It can be shown that the above conditions are fulfilled if and only if C_1 and C_2 can be separated.

If both C_1 and C_2 are contained in the hyperplane H, this separation is called improper, otherwise it is called proper. We make this distinction because otherwise we would be talking about "separation" even when the sets involved have inner points in common. If the separation is proper, then the inequality (2.3) must be strict. It is very straightforward to prove that the opposite implication also holds, and hence we omit the proof of the following proposition.

Proposition 2.10. [[2], Theorem 11.1] *Two nonempty sets* C_1 *and* C_2 *in* $\mathbb{R}^n$ *can be separated properly if and only if there exists a nonzero vector* $\boldsymbol{p} \in \mathbb{R}^n$ *such that*

(i) $\sup_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \leq \inf_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle$

(ii) $\inf_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle < \sup_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle$

What is special for disjoint convex sets is that they can always be separated by a hyperplane. This is not true in general, as Figure 3 illustrates.

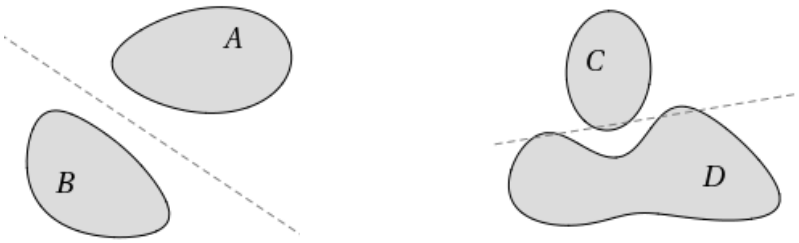
Another important property of convex sets is that the distance to a point outside of the set can always be minimized to a unique point in the set, as the following theorem states.

Proposition 2.11. [[3], 2.4.1 Theorem] *Let* C *be a nonempty closed convex set in* $\mathbb{R}^n$*, and let* $\bar{\boldsymbol{y}} \in \mathbb{R}^n$ *be a point outside of* C*. Then there exists a unique point* $\bar{\boldsymbol{x}} \in C$ *with minimum distance to* $\bar{\boldsymbol{y}}$*. Furthermore,* $\bar{\boldsymbol{x}}$ *is the minimizing point if and only if* $(\bar{\boldsymbol{y}} - \bar{\boldsymbol{x}})^T(\boldsymbol{x} - \bar{\boldsymbol{x}}) \leq 0$ *for all* $\boldsymbol{x} \in C$*.*

We omit the proof. This result does not hold for non-convex sets: although the distance itself is always uniquely determined, the point where the minimal distance is attained is not always unique. See Figure 4 for illustration. We can now prove that a closed convex set and a point can be separated by a hyperplane.

Proposition 2.12. [[3], 2.4.4 Theorem] *Let* C *be a nonempty closed convex set in* $\mathbb{R}^n$*, and let* $\bar{\boldsymbol{y}} \in \mathbb{R}^n$ *be a point outside* C*. Then there exists a hyperplane that separates* C *and* $\bar{\boldsymbol{y}}$*.*

Proof. We have to show that there exists a nonzero vector $\boldsymbol{p} \in \mathbb{R}^n$ and a scalar $\alpha \in \mathbb{R}$ such that $\langle \bar{\boldsymbol{y}}, \boldsymbol{p} \rangle \geq \alpha$ and $\langle \boldsymbol{x}, \boldsymbol{p} \rangle \leq \alpha$ for all $\boldsymbol{x} \in C$. This shall establish that the hyperplane $\{\boldsymbol{y} \in \mathbb{R}^n \mid \langle \boldsymbol{y}, \boldsymbol{p} \rangle = \alpha\}$ separates the point and the set.



(a) (b)

Figure 3: (a) The convex sets A and B can be separated. (b) No hyperplane separates C and D.

By Proposition 2.11 there exists a unique minimizing point $\bar{\boldsymbol{x}} \in C$ such that

$$(\bar{\boldsymbol{y}} - \bar{\boldsymbol{x}})^T(\boldsymbol{x} - \bar{\boldsymbol{x}}) \leq 0$$

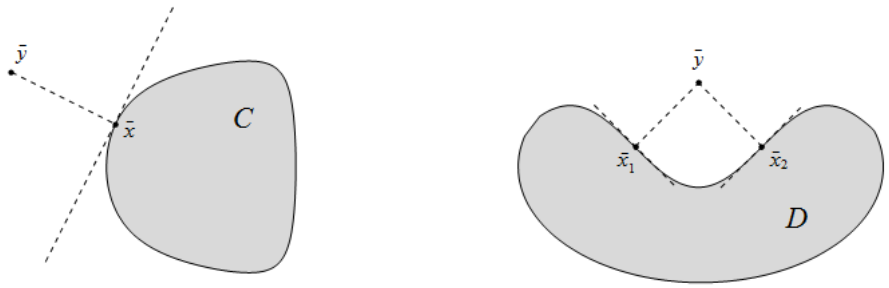
for all $\boldsymbol{x} \in C$. The result follows by setting $\boldsymbol{p} = \bar{\boldsymbol{y}} - \bar{\boldsymbol{x}} \neq \boldsymbol{0}$ and $\alpha = \bar{\boldsymbol{x}}^T(\bar{\boldsymbol{y}} - \bar{\boldsymbol{x}}) = \langle \boldsymbol{p}, \bar{\boldsymbol{x}} \rangle$. ■

If the point in question lies on the border of the set, it is more natural to use the word "support" than "separate". Hence the following proposition.

Proposition 2.13. [[3], 2.4.7 Theorem] *Let C be a nonempty convex set in $\mathbb{R}^n$, and let $\bar{\boldsymbol{y}} \in \mathbb{R}^n$ be a point on the border of C. Then there exists a hyperplane that supports C at $\bar{\boldsymbol{y}}$. In other words, there exists $\boldsymbol{p} \in \mathbb{R}^n$ such that $\boldsymbol{p}^T(\boldsymbol{x} - \bar{\boldsymbol{y}}) \leq 0$ for every $\boldsymbol{x} \in \operatorname{cl} C$.*

This result follows quite easily from Proposition 2.12. Note that although we say that the hyperplane supports C at $\bar{\boldsymbol{y}}$, it technically speaking separates C and $\bar{\boldsymbol{y}}$ and hence this situation comes under Proposition 2.10. We are now ready to present the *Separating Hyperplane Theorem.*

Theorem 2.14 (Separating Hyperplane Theorem)**.** [[2], Theorem 11.3] *Let C_1 and C_2 be nonempty convex subsets of $\mathbb{R}^n$. The sets C_1 and C_2 can be separated properly if and only if $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \varnothing$.*



(a) (b)

Figure 4: (a) The smallest distance between C and $\bar{y}$ is attained only at $\bar{x}$. (b) The smallest distance to the set D is attained both at $\bar{x}_1$ and at $\bar{x}_2$.

Proof. Let C_1 and C_2 be nonempty convex sets such that $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \varnothing$, and consider the set $C = C_1 - C_2 = \{\boldsymbol{x}_1 - \boldsymbol{x}_2 \mid \boldsymbol{x}_1 \in C_1,\ \boldsymbol{x}_2 \in C_2\}$. It is easily seen that C is convex and that

$$\operatorname{ri} C = \operatorname{ri} C_1 - \operatorname{ri} C_2 \tag{2.4}$$

from which it follows that $\mathbf{0} \notin \operatorname{ri} C$ by the assumption $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \varnothing$. If $\mathbf{0} \notin \operatorname{cl} C$, then by Proposition 2.12 the point $\mathbf{0}$ and $\operatorname{cl} C$ can be separated by a hyperplane. If $\mathbf{0} \in \operatorname{cl} C$ there exists a hyperplane that supports $\operatorname{cl} C$ at $\mathbf{0}$. Hence $\mathbf{0}$ and C can be separated properly and by Proposition 2.10 there exists a vector $\boldsymbol{p} \in \mathbb{R}^n$ such that

$$0 \leq \inf_{\boldsymbol{x} \in C} \langle \boldsymbol{x}, \boldsymbol{p} \rangle = \inf_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle - \sup_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \tag{2.5}$$

$$0 < \sup_{\boldsymbol{x} \in C} \langle \boldsymbol{x}, \boldsymbol{p} \rangle = \sup_{\boldsymbol{x}_1 \in C_1} \langle \boldsymbol{x}_1, \boldsymbol{p} \rangle - \inf_{\boldsymbol{x}_2 \in C_2} \langle \boldsymbol{x}_2, \boldsymbol{p} \rangle \tag{2.6}$$

Applying Proposition 2.10 on the above inequalities gives us the desired result: the sets C_1 and C_2 can be separated properly.

Proving the opposite implication is straightforward: if C_1 and C_2 can be separated properly, then the inequalities (2.5) and (2.6) hold and it follows that $\mathbf{0} \notin C$. By (2.4), we then get $\operatorname{ri} C_1 \cap \operatorname{ri} C_2 = \varnothing$. The proof is now complete. ■

2.4 Lagrangian duality

Although we in this paper mainly use Fenchel duality we also state and prove a much more common duality theory, namely Lagrangian duality. Or more precisely, we prove a weaker version of the Lagrangian duality theorem. Denote $\boldsymbol{g} = (g_1, g_2, \ldots, g_m)$ and $\boldsymbol{h} = (h_1, h_2, \ldots, h_l)$.

Theorem 2.15 (Lagrangian duality theorem)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$ be convex functions, let $h_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1, 2, \ldots, l$ be affine, and let C be a convex subset of $\mathbb{R}^n$. Suppose that Slater's condition holds (see Definition 2.8). Then the following equality holds:*

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0} \\ \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) = \sup_{\substack{(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^{m+l} \\ \boldsymbol{p} \geq \mathbf{0}}} \inf_{\boldsymbol{x} \in C} \{ f(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{q}^T \boldsymbol{h}(\boldsymbol{x}) \}$$

We shall prove the above theorem without the equality constraints. Before we proceed, let us note that the following always holds true:

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0} \\ \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) \geq \sup_{\substack{(\boldsymbol{p}, \boldsymbol{q}) \in \mathbb{R}^{m+l} \\ \boldsymbol{p} \geq \mathbf{0}}} \inf_{\boldsymbol{x} \in C} \{ f(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x}) + \boldsymbol{q}^T h(\boldsymbol{x}) \} \tag{2.7}$$

This relation is called *weak duality.*

Consider the following assertions:

(I) $\phi(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in C$ such that $g_i(\boldsymbol{x}) \leq 0$, $i = 1, \ldots, n$.

(II) There exists $p_i \geq 0$, $i = 1, \ldots, m$ such that for all $\boldsymbol{x} \in C$ we have

$$\phi(\boldsymbol{x}) + \sum_{k=1}^{m} p_i g_i(\boldsymbol{x}) \geq 0$$

We shall now attempt to construct conditions under which these statements are equivalent. The discussion here is a combination of the proof of Lemma 6.2.3 in [3] and [[4], page 391-392].

It is easily seen that the assertion (II) implies (I). To prove the converse, suppose that (I) holds. Denote

$$\Omega(\boldsymbol{x}) = \begin{pmatrix} \phi(\boldsymbol{x}) \\ g_1(\boldsymbol{x}) \\ g_2(\boldsymbol{x}) \\ \vdots \\ g_m(\boldsymbol{x}) \end{pmatrix} \tag{2.8}$$

and consider the following set:

$$D = \{(a, \boldsymbol{b}) \in \mathbb{R}^{m+1} \mid a < 0,\ \boldsymbol{b} \leq \boldsymbol{0}\}$$

It follows from the assumption (I) that the sets $\Omega(C)$ and D are disjoint. Let us claim that a hyperplane separates these sets. Then by Proposition 2.10 there exists a nonzero vector $(u, \boldsymbol{v}) \in \mathbb{R}^{m+1}$ such that

$$\inf_{(\phi(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x})) \in \Omega(C)} (u\phi(\boldsymbol{x}) + \boldsymbol{v}^T \boldsymbol{g}(\boldsymbol{x})) \geq \sup_{(a, \boldsymbol{b}) \in D} (ua + \boldsymbol{v}^T \boldsymbol{b})$$

Since a and $\boldsymbol{b}$ in D can be made arbitrarily small, this only makes sense when $(u, \boldsymbol{v}) \geq \boldsymbol{0}$. Hence we have $\sup_{(a, \boldsymbol{b}) \in D}(ua + \boldsymbol{v}^T \boldsymbol{b}) = 0$ and the following inequality holds for each $\boldsymbol{x} \in C$:

$$u\phi(\boldsymbol{x}) + \boldsymbol{v}^T \boldsymbol{g}(\boldsymbol{x}) \geq 0 \tag{2.9}$$

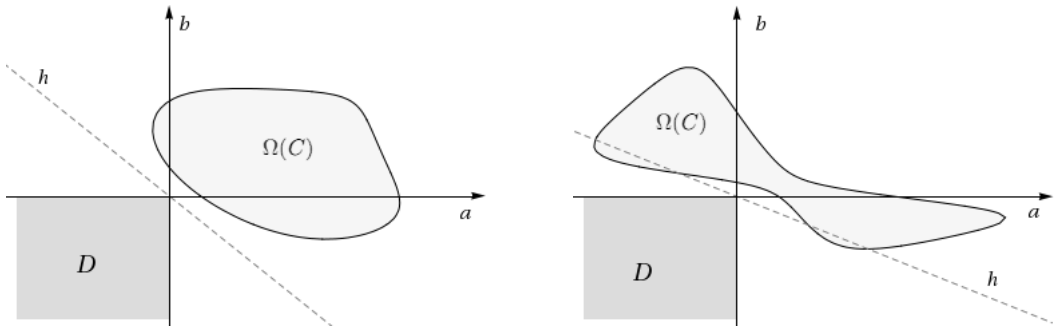
Now, (2.9) translates to assertion (II) whenever $u > 0$. The result follows by setting $\boldsymbol{p} = \boldsymbol{v}/u$.

So the case $u = 0$ must be impossible. Suppose, to get a contradiction, that $u = 0$. Now, if Slater's condition holds then there exists an $\bar{\boldsymbol{x}} \in \text{ri}(C)$ such that $\boldsymbol{g}(\bar{\boldsymbol{x}}) < 0$. From (2.9) we get

$$\boldsymbol{v}^T \boldsymbol{g}(\bar{\boldsymbol{x}}) \geq 0$$

Since $\boldsymbol{v} \geq \boldsymbol{0}$ and $\boldsymbol{g}(\bar{\boldsymbol{x}}) < 0$ this is only possible when $\boldsymbol{v} = \boldsymbol{0}$. But this contradicts the choice of $(u, \boldsymbol{v})$. Hence $u > 0$ under Slater's condition.

So, in order for (I) and (II) to be equivalent we need to consider whether the sets $\Omega(C)$ and D can be separated and if we can choose u to be nonzero. When the set $\Omega(C)$ is convex the



(a) $\Omega(C)$ is convex.

(b) $\Omega(C)$ is not convex.

Figure 5: We wish to separate D and $\Omega(C)$ by a hyperplane. A sufficient condition is the convexity of $\Omega(C)$.

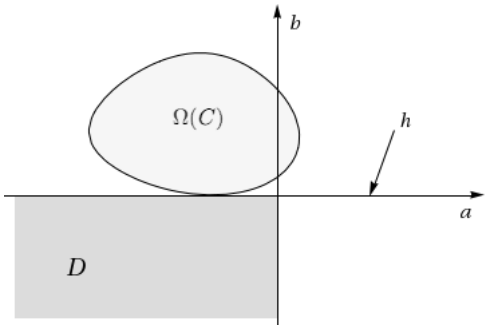


Figure 6: If $\Omega(C)$ takes values from the negative a-axis then the only separating hyperplane is the a-axis itself. If Slater's condition holds this is not possible: a convex set that contains points from the negative a-axis and the second quadrant must intersect with D. If $\Omega(C)$ is not convex it may contain points from the negative a-axis and from the second quadrant. However, this means that a hyperplane cannot separate $\Omega(C)$ from D.

existence of the separating hyperplane is self-evident (see Theorem 2.14). In the case when $\Omega(C)$ is not convex, this is not guaranteed as is illustrated in Figure 5. Notice though that $\Omega(C)$ does not *have* to be convex for a separating hyperplane to exist.

Now suppose that the only separating hyperplane is such that $u = 0$. As illustrated in Figure 6 this means that there exists an $\bar{\boldsymbol{x}}$ such that $f(\bar{\boldsymbol{x}}) < 0$ and $\boldsymbol{g}(\bar{\boldsymbol{x}}) = \boldsymbol{0}$. To ensure that this unfortunate situation does not occur it is sufficient to suppose that there exists $\bar{\boldsymbol{y}}$ such that $\boldsymbol{g}(\bar{\boldsymbol{y}}) < \boldsymbol{0}$; that way a hyperplane cannot separate $\Omega(C)$ and D which is a contradiction. Observe again that Slater's condition is not a necessary condition.

The convexity of $\Omega(C)$ is trivial in the case when all the functions and sets involved are convex. Hence the next lemma.

Lemma 2.16. [[3], Lemma 6.2.3] *Let $\phi : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, 2, \ldots, m$ be convex functions, and let C be a convex subset of $\mathbb{R}^n$. If Slater's condition is fulfilled, then (I) and (II) are equivalent.*

Using the above lemma, we can prove Lagrangian duality theorem.

Proof of Theorem 2.15 without equality constraints. Denote

$$\alpha = \inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \\ \boldsymbol{x} \in C}} f(\boldsymbol{x})$$

By Slater's condition the problem is feasible and hence $\alpha < \infty$. If $\alpha = -\infty$, the theorem follows from weak duality (2.7). Hence we can assume α to be finite. We shall now find a vector $\boldsymbol{p}$ such that $\inf_{\boldsymbol{x} \in C}\{f(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x})\} \geq \alpha$; this shall establish equality in (2.7). Let

$$\phi(\boldsymbol{x}) = f(\boldsymbol{x}) - \alpha$$

and consider the following statement:

$$\phi(\boldsymbol{x}) \geq 0 \text{ for all } \boldsymbol{x} \in C \text{ such that } g_i(\boldsymbol{x}) \leq 0,\ i = 1, \ldots, n.$$

By the choice of α, this is trivially true. Also, the function ϕ is convex. It follows from Lemma 2.16 that there exists $\boldsymbol{p} \geq \mathbf{0}$ such that

$$\phi(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x}) \geq 0$$

for all $\boldsymbol{x} \in C$. Going back to f, we get

$$f(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x}) \geq \alpha$$

Taking the infimum gives us the desired expression. The proof is now complete. ■

Notice that in the above proof we only use the convexity of f and $\boldsymbol{g}$ to show that the assertions (I) and (II) are equivalent. It should therefore not come as a surprise that Lagrangian duality can be applied to certain non-convex optimization problems.

Theorem 2.17. [[5], Theorem 3.1] *Let $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$, $f : \mathcal{X}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathcal{X}^n \to \mathbb{R}^m$. Let $\phi(\boldsymbol{x}) = f(\boldsymbol{x}) - \alpha$, where α is any real scalar. Suppose that the assertions (I) and (II) are equivalent regardless of how we pick α. Then the following duality result holds:*

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\boldsymbol{p} \geq \mathbf{0}} \inf_{\boldsymbol{x} \in \mathcal{X}^n} \{f(\boldsymbol{x}) + \boldsymbol{p}^T \boldsymbol{g}(\boldsymbol{x})\} \tag{2.10}$$

where the supremum is attained. Conversely, if the relation (2.10) holds and the supremum is attained, then (I) and (II) are equivalent regardless of how we pick α.

The implications of the above theorem shall be discussed in more detail in Chapter 4.

3 Conjugate Functions and Duality

Duality is not really a method of solving an optimization problem. Instead its main purpose is to convert the original problem (the primal problem) to another, hopefully more approachable optimization problem (the dual). It often happens that the dual does not have any constraints, or the constraints are significantly simpler than those of the primal problem. Solving unconstrained optimization problems is a lot easier; it often suffices to differentiate the objective function. Another application which we shall see later on is using duality to prove other results. Statements of the form "the following systems are equivalent" are especially approachable. Also, it is fairly common to come across a new problem that is easier to solve numerically.

3.1 Modifications and Generalizations

So far, and generally in optimization, we have used real-valued functions defined on a subset of $\mathbb{R}^n$. In Fenchel duality this leads to unnecessarily cumbersome notation. Luckily, there is an easy way to come around this restriction. In a regular optimization problem one wishes to minimize a function $f : C \to \mathbb{R}$, where C is a subset of $\mathbb{R}^n$. If we redefine f as

$$f_0(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in C \\ +\infty & \text{if } \boldsymbol{x} \notin C \end{cases}, \qquad \boldsymbol{x} \in \mathbb{R}^n$$

then we get an extended real-valued function defined on the entire $\mathbb{R}^n$. We use the notation $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ to denote the extended real line. Naturally, f_0 has the same minimum as f. Another important observation is that a function defined in this fashion is convex if and only if the original function is convex. This can be seen by considering the inequality a convex function must by definition fulfill:

$$f_0(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f_0(\boldsymbol{x}_1) + (1-\lambda) f_0(\boldsymbol{x}_2)$$

If $\boldsymbol{x}_1 \notin C$ or $\boldsymbol{x}_2 \notin C$, then the right-hand side equals infinity and since there is nothing greater than $+\infty$, the inequality must hold everywhere.

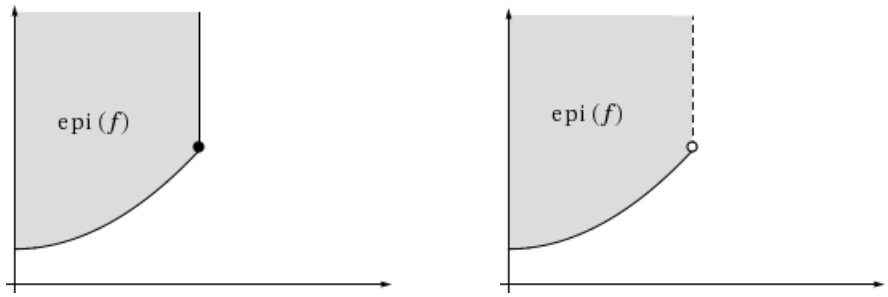
The extension is useful for convex functions, but it does not give desired results if f is concave. The inequality

$$f_0(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \geq \lambda f_0(\boldsymbol{x}_1) + (1-\lambda) f_0(\boldsymbol{x}_2)$$

immediately leads to problems: we would require the left-hand side to be $+\infty$, which does not have to be true. The problem can be solved by replacing $+\infty$ with $-\infty$. Hence a concave function is extended to the entire real space by setting $f_0(\boldsymbol{x}) = -\infty$ for $\boldsymbol{x} \notin C$.

The notation f_0 was only introduced to make the definition rigorous and shall not be used to distinguish between real-valued and extended real-valued functions.

Now that the domain is $\mathbb{R}^n$, it is natural to give a name to the original domain.



(a) (b)

Figure 7: Consider a convex function f which at a point makes a jump to $+\infty$. Whether or not f is closed depends on how it behaves at the point of discontinuity. The function in (a) is closed whereas the one in (b) is not.

Definition 3.1 (Effective domain)**.** *The* effective domain *of a convex function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is given by*

$$\operatorname{dom} f = \{\boldsymbol{x} \in \mathbb{R}^n \mid f(\boldsymbol{x}) < +\infty\}$$

The effective domain of a concave function $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined as the effective domain of $-g$, that is

$$\operatorname{dom} g = \{\boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) > -\infty\}$$

Notice that if f never takes the value $-\infty$, then the effective domain is the largest domain where the function is real-valued. The same goes for $-g$. Hence the next definition.

Definition 3.2 (Proper function)**.** *A convex function f is called* proper *if it satisfies the following two conditions:*

(i) The effective domain of f is nonempty

(ii) $f(\boldsymbol{x}) > -\infty$ for all $\boldsymbol{x} \in \mathbb{R}^n$

Similarly, a concave function g is proper if it fulfills the following conditions:

(i) The effective domain of g is nonempty

(ii) $g(\boldsymbol{x}) < +\infty$ for all $\boldsymbol{x} \in \mathbb{R}^n$

A function that is not proper is called improper*.*

The functions we have derived are discontinuous at the borders of their effective domains. This is not really a problem but these functions must behave in a certain way at points of discontinuity, as will be seen in the next section.

Definition 3.3 (Closed function)**.** *A convex function f is called* closed *if its epigraph is a closed set. A concave function g is closed if its hypograph is closed.*

The above definition is illustrated in Figure 7. A little loosely we may say that a closed function attains the "lesser value". In some publications the term *lower semi-continuous* is used instead of closed – although the definitions look quite different it can be shown that these two concepts are equivalent for proper convex functions.

Notice that all continuous functions are closed.

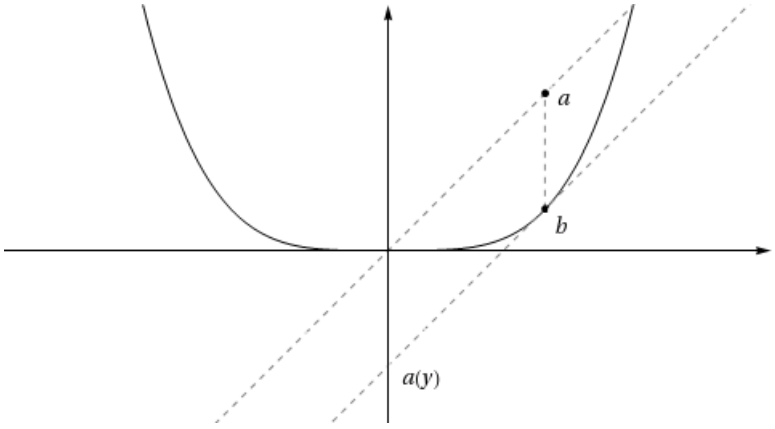


Figure 8: The largest difference between the line xy and the function f is $a - b$. This is equal to $0 - a(y) = -a(y)$.

3.2 Conjugate functions

Before we can formulate the Fenchel dual problem we must introduce the concept of conjugate functions. We illustrate the conjugate functions in the univariate case in order to get a better intuition on the subject.

So, let us consider a differentiable convex function f of one variable. Fix $y \in \mathbb{R}$ and choose a real number $a(y)$ so that the line $xy + a(y)$ is tangential to f. Notice that a tangent of slope y does not always exist, but when it does exist it is uniquelly determined because f is a convex function. The distance between f and the tangent is trivially 0 at a point of intersection x_0:

$$x_0 y + a(y) - f(x_0) = 0$$

or equivalently

$$-a(y) = x_0 y - f(x_0)$$

Now, say that we begin with the expression $xy - f(x)$ without knowing x_0 and wish to determine how large $a(y)$ is. Not so surprisingly, there is a very straightforward way to recover $a(y)$: as illustrated in Figure 8, the maximal value of $xy - f(x)$ must equal $-a(y)$.

Hence we may determine $a(y)$ by taking the supremum over all x:

$$-a(y) = \sup_{x \in \mathbb{R}} \{xy - f(x)\}$$

The above expression is so important it deserves a name of its own.

Definition 3.4 (Conjugate functions)**.** *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$. The* conjugate convex function *$f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ of f is defined by*

$$f^*(\boldsymbol{y}) = \sup_{\boldsymbol{x} \in \mathbb{R}^n} \{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - f(\boldsymbol{x})\}$$

Similarly, the conjugate concave function *$f_* : \mathbb{R}^n \to \overline{\mathbb{R}}$ of f is given by*

$$f_*(\boldsymbol{y}) = \inf_{\boldsymbol{x} \in \mathbb{R}^n} \{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - f(\boldsymbol{x})\}$$

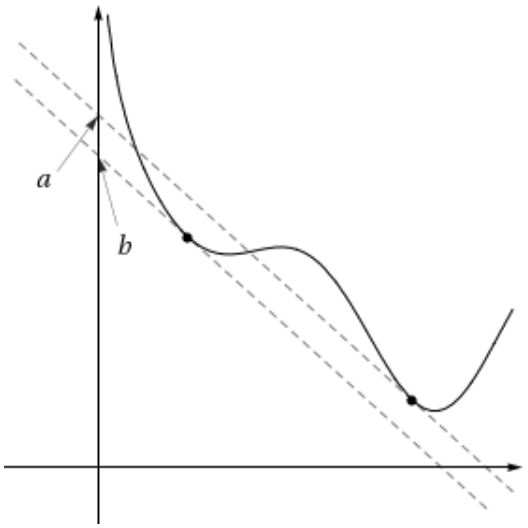


Figure 9: If the function is not convex, there may be multiple points where the tangent has the same slope. We choose the greater value and hence loose all information about the other value. In this picture we have $f_*(y) = -b > -a$.

The convex conjugate function sometimes goes under the name Legendre-Fenchel transformation because it generalizes the Legendre transformation. We shall in Section 3.4 give a more general form of Definition 3.4. Right now we stick to the real case to avoid confusion.

Since most functions we shall consider later on are defined on the entire $\mathbb{R}^n$, the terms "$\inf_{\boldsymbol{x}}$" and "$\sup_{\boldsymbol{x}}$" shall from now on refer to $\inf_{\boldsymbol{x}\in\mathbb{R}^n}$ and $\sup_{\boldsymbol{x}\in\mathbb{R}^n}$, respectively. We also use the convenient notations $\sup\{\varnothing\} = -\infty$ and $\inf\{\varnothing\} = +\infty$.

The function

$$f^{**}(\boldsymbol{x}) = (f^*)^*(\boldsymbol{x}) = \sup_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - f^*(\boldsymbol{x})\}$$

is called the *biconjugate* of f. f_{**} is defined in the same manner, and if it is clear from the context, it shall also be referred to as the biconjugate.

The aim of formulating a dual problem is to find another problem with the same optimal solution. If one is lucky, this other problem is more approachable than the original problem. With this in mind, the next step is to ask when does $f^{**} = f$ hold?

Let us look a little closer at Definition 3.4. Notice that although we required f to be differentiable and convex in the introduction, we do not mention these things in the definition. As we saw earlier, when f is convex and differentiable, the conjugate transformation returns information about the tangent with slope y. But even if a tangent with slope y does not exists, the conjugate convex function is well-defined.

In the case when f is not convex, it may happen that speaking of *the* tangent with slope y is not possible; there might be several tangents with the same slope. In this case, only one of the tangents is used, as illustrated in Figure 9. The conjugate convex transformation is well-defined for non-convex functions, but it does not preserve all information. It seems that the relation $f = f^{**}$ cannot hold if f is not convex. Let us now apply the conjugate transformations on a convex function.

Example 3.5. Consider the function $f(x) = x^2$. It is easy to verify that f is a convex function.

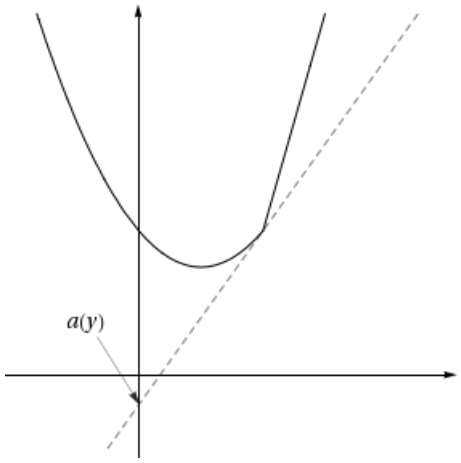


Figure 10: The conjugate convex function exists although the tangent does not.

The conjugate convex function can be calculated by finding the zero of the first derivative:

$$f^*(y) = \sup_x \left\{xy - x^2\right\} = \frac{1}{2}y \cdot y - \left(\frac{1}{2}y\right)^2 = \frac{y^2}{4}$$

It is hard to say anything about this function by just looking at it, so let us proceed to calculating the biconjugate of this function:

$$f^{**}(x) = \sup_y \left\{xy - \frac{1}{4}y^2\right\} = x \cdot 2x - \frac{1}{4}(2x)^2 = x^2$$

We have come back to the original function and hence the relation $f = f^{**}$ holds true. It is natural to wonder what the conjugate concave function looks like. Some simple calculations yield

$$f_*(y) = \inf_x \left\{xy - x^2\right\} = -\infty$$

The conjugate concave transformation completely destroys the function and consequently, it is not possible to recover the original function by applying the conjugate concave transformation on the above expression. $\Diamond$

A little loosely we may say that the conjugate convex transformation is suitable for convex functions, and the concave transformation for concave functions. In fact, f^* is always convex, regardless of f, and similarly f_* is always concave.

We required f to be differentiable in the introduction in order to justify the use of the term "tangent". However, non-differentiable points are not a problem, as illustrated in Figure 10. Interestingly though, the slope y is a subgradient to f at the points where f and the "tangent" intersect. More information about subgradients can be found for instance in [[3], Section 3.2].

Example 3.6. Let us redefine the function from the previous example as

$$f(x) = \begin{cases} x^2 & \text{if } x > 0 \\ +\infty & \text{if } x \leq 0 \end{cases}$$

We observe that f is convex but its epigraph is not closed. Is the conjugate transformation useful for functions that are not closed? We have

$$f^*(y) = \sup_x\{xy - f(x)\} = \sup_{x>0}\left\{xy - x^2\right\} = \begin{cases} \frac{1}{4}y^2 & \text{if } y \geq 0 \\ 0 & \text{if } y < 0 \end{cases}$$

After some calculations we get the biconjugate:

$$f^{**}(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ +\infty & \text{if } x < 0 \end{cases}$$

The function f and its biconjugate take different values at $x = 0$. Hence the relation $f = f^{**}$ does not hold. $\Diamond$

Notice that the biconjugate in the previous example is closed. It seems that we must require f to be closed if we wish to have $f^{**} = f$. Why is that? Well, consider Figure 11. At the point where the function makes a jump, we choose the "lower" point, because it gives a greater $-a(y)$-value. In fact, both the convex and the concave conjugate function are always closed.

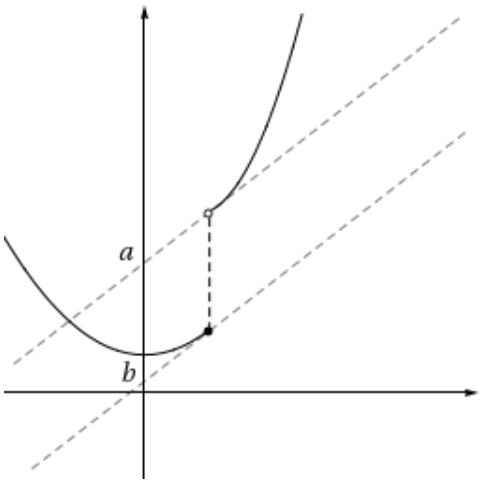


Figure 11: At a point of discontinuity there are two different "tangents". Since the conjugate convex function is defined by means of supremum, we choose the greater value. Hence we pick $f_*(y) = -b > -a$ in the picture.

Finally, a convex function is improper only when it is identically equal to $-\infty$ in its effective domain. Such functions lead to problems but are luckily of no interest. Based on the previous examples, the reader should be confident that the following theorem holds.

Theorem 3.7 (Conjugacy Theorem)**.** [[6], Proposition 7.1.1 c)] *If f is a closed proper convex function, then $f = f^{**}$. If g is a closed proper concave function, then $g = g_{**}$.*

For all $\boldsymbol{x}$ and $\boldsymbol{y}$ we have

$$f^*(\boldsymbol{y}) = \sup\{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - f(\boldsymbol{x})\} \geq \langle \boldsymbol{x}, \boldsymbol{y}\rangle - f(\boldsymbol{x})$$

from which it follows that the inequality

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle \leq f(\boldsymbol{x}) + f^*(\boldsymbol{y}) \tag{3.1}$$

holds for all $\boldsymbol{x}$ and $\boldsymbol{y}$. Similarly we get

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle \geq f(\boldsymbol{x}) + f_*(\boldsymbol{y}) \tag{3.2}$$

The identities (3.1) and (3.2) are called *Fenchel's inequalities* and they will come handy later on.

Finally, let us apply the conjugate transformation to a couple of "real-world"-examples.

Example 3.8 (The Indicator Function)**.** [[6], Example 7.1.2] Consider the indicator function of the set C:

$$I_C(\boldsymbol{x}) := \begin{cases} 0 & \text{if } \boldsymbol{x} \in C \\ +\infty & \text{if } \boldsymbol{x} \notin C \end{cases}$$

Interestingly, the indicator function is a proper convex function if and only if C is a nonempty convex set. Furthermore, I_C is closed if and only if C is. Its conjugate convex function is given by

$$I_C^*(\boldsymbol{y}) = \sup_{\boldsymbol{x}} \{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - I_C(\boldsymbol{x})\} = \sup_{\boldsymbol{x} \in C} \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \Psi_C(\boldsymbol{y})$$

This function is called the *support function* of the set C. The special case in which C is a convex cone, $\Psi_C(\boldsymbol{y})$ becomes the indicator function of the polar cone of C.

Notice that there is nothing that dictates that the indicator function should be defined by means of $+\infty$; if we wish the indicator function to be concave we may simply let it take the value $-\infty$ outside C. Then the support function would be defined by means of infimum. We shall use the words "indicator function" and "support function" in both cases if there is no risk for confusion. $\Diamond$

Example 3.9 (Differentiable functions)**.** [7] In the special case in which f is a differentiable convex function, calculating the conjugate is straightforward. The expression $\langle \boldsymbol{x}, \boldsymbol{y} \rangle - f(\boldsymbol{x})$ is concave in $\boldsymbol{x}$ and therefore attains its maximum at the zero of the gradient. So if for every $\boldsymbol{y} \in \mathbb{R}^n$ the equation $\boldsymbol{y} - \nabla f(\boldsymbol{x}) = \boldsymbol{0}$ has a solution $\boldsymbol{x} = s(\boldsymbol{y})$ then we simply get

$$f^*(\boldsymbol{y}) = \sup_{\boldsymbol{x}} \{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - f(\boldsymbol{x})\} = \langle s(\boldsymbol{y}), \boldsymbol{y} \rangle - f(s(\boldsymbol{y}))$$

In this case the conjugate transformation and the Legendre transformation coincide. $\Diamond$

Remark. One way to interpret the conjugate transformation is as a collection of non-vertical half-spaces (i.e. sets of the form $\{\boldsymbol{x} \in \mathbb{R}^n \mid \langle \boldsymbol{p}, \boldsymbol{x} \rangle \leq \alpha\}$) containing $\operatorname{epi} f$. It is generally true that a closed set is convex if and only if is an intersection of closed half-spaces. As is seen in the figures of this section, the conjugate function determines for every $\boldsymbol{y}$ a half-space that contains the epigraph. The epigraph is the intersection of all such half-spaces if and only if f is a proper closed convex function. $\Diamond$

3.3 Fenchel's Duality Theorem

The Fenchel dual problem was originally formulated by Werner Fenchel and therefore carries his name. Since the dual problem is defined by means of the conjugate transformations the term "conjugate duality" is also frequently used.

The primal problem for Fenchel duality looks slightly different from (2.1):

$$\begin{array}{ll} \text{Minimize} & f(\boldsymbol{x}) - g(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in C \end{array}$$

Here $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a convex and $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a concave function, and C is a subset of $\mathbb{R}^n$. Actually we often drop the constraint: we may simply define f or $-g$ to be infinitely large outside of C. The problem above is a convex optimization problem: the sum of two convex functions is convex.

The main idea behind Fenchel duality is the observation that minimizing the difference between f and g is equal to maximizing the difference between tangents of same slope. In other words, minimizing $f - g$ is under certain restrictions equivalent to maximizing $g_* - f^*$.

So let us consider the problem of maximizing $g_* - f^*$. By the definition of the convex conjugate, we get:

$$\begin{aligned}\sup_{\boldsymbol{y}}\{g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})\} &= \sup_{\boldsymbol{y}}\left\{\inf_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - g(\boldsymbol{x})\} - \sup_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - f(\boldsymbol{x})\}\right\} \\ &= \sup_{\boldsymbol{y}}\left\{\inf_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y}\rangle - g(\boldsymbol{x})\} + \inf_{\boldsymbol{x}}\{f(\boldsymbol{x}) - \langle \boldsymbol{x}, \boldsymbol{y}\rangle\}\right\}\end{aligned}$$

As is commonly known, the relation

$$\inf_{\boldsymbol{x}} \phi(\boldsymbol{x}) + \inf_{\boldsymbol{x}} \psi(\boldsymbol{x}) = \inf_{\boldsymbol{x}}\{\phi(\boldsymbol{x}) + \psi(\boldsymbol{x})\} \tag{3.3}$$

does not hold in general. But when it does hold, the expression $\sup_{\boldsymbol{y}}\{g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})\}$ becomes $\inf_{\boldsymbol{x}}\{f(\boldsymbol{x}) - g(\boldsymbol{x})\}$ as desired. Fenchel duality essentially gives us conditions under which the equation (3.3) is true.

There are many different versions of Fenchel's Duality Theorem. Below we give a simple real version and then reformulate it in the following section.

Theorem 3.10 (Fenchel's Duality Theorem)**.** [[2], Theorem 31.1] *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper convex function and $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ a proper concave function. Then we have*

$$\inf_{\boldsymbol{x}}\{f(\boldsymbol{x}) - g(\boldsymbol{x})\} = \sup_{\boldsymbol{y}}\{g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})\}$$

if one of the following conditions hold:

(i) $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\operatorname{dom} g) \neq \varnothing$

(ii) f and g are closed, and $\operatorname{ri}(\operatorname{dom} f^*) \cap \operatorname{ri}(\operatorname{dom} g_*) \neq \varnothing$

Under (i) the supremum is attained, under (ii) the infimum is attained. If both (i) and (ii) hold, both the supremum and infimum are finite.

Proof of Theorem 3.10. Pick any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. By Fenchel's inequalities (relations (3.1) and (3.2)), we have for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$ that

$$f(\boldsymbol{x}) + f^*(\boldsymbol{y}) \geq \langle \boldsymbol{x}, \boldsymbol{y}\rangle \geq g(\boldsymbol{x}) + g_*(\boldsymbol{y})$$

By rearranging the terms we gain

$$f(\boldsymbol{x}) - g(\boldsymbol{x}) \geq g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})$$

and hence

$$\inf_{\boldsymbol{x}}\{f(\boldsymbol{x}) - g(\boldsymbol{x})\} \geq \sup_{\boldsymbol{y}}\{g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})\} \tag{3.4}$$

This is called weak duality. Set $\alpha = \inf\{f(\boldsymbol{x}) - g(\boldsymbol{x})\}$. If $\alpha = -\infty$, then (3.4) forces the supremum to be $-\infty$ too and hence the theorem holds. Let us now assume that $\alpha > -\infty$.

Suppose that the condition (i) is true. Then $\alpha < \infty$ and hence α is finite. We shall now find a vector $\boldsymbol{y}$ such that $g_*(\boldsymbol{y}) - f^*(\boldsymbol{y}) \geq \alpha$; this shall establish equality in (3.4). By Proposition 2.4 the sets $A = \text{epi}(f)$ and $B = \text{hyp}(g(x) + \alpha)$ are convex. On the other hand,

$$\text{ri}(\text{epi}(f)) = \{(\boldsymbol{x}, z) \in \mathbb{R}^{n+1} \mid \boldsymbol{x} \in \text{ri}(\text{dom}\, f),\ f(\boldsymbol{x}) < z < \infty\} \tag{3.5}$$

as can easily be verified (for the proof, see [[2], Lemma 7.3]). Since

$$\alpha = \inf_{\boldsymbol{x}}\{f(\boldsymbol{x}) - g(\boldsymbol{x})\} \leq f(\boldsymbol{x}) - g(\boldsymbol{x})$$

holds for every $\boldsymbol{x} \in \mathbb{R}^n$, we have $f \geq g + \alpha$. Hence the set (3.5) and $\text{hyp}(g(x) + \alpha)$ are disjoint and by Theorem 2.14 we can separate them properly with a hyperplane, call it H.

If H were vertical, then its projection on $\mathbb{R}^n$ would separate the sets A and B properly, which would contradict (i). Hence H is not vertical and we can characterize it by an affine function

$$h(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{y} \rangle - \beta$$

Since H separates A and B, we have

$$f(x) \geq \langle \boldsymbol{x}, \boldsymbol{y} \rangle - \beta \geq g(\boldsymbol{x}) + \alpha$$

for all $\boldsymbol{x} \in \mathbb{R}^n$. From these inequalities we can deduce that

$$\beta \geq \sup_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - f(\boldsymbol{x})\} = f^*(\boldsymbol{y})$$

and also

$$\alpha + \beta \leq \inf_{\boldsymbol{x}}\{\langle \boldsymbol{x}, \boldsymbol{y} \rangle - g(\boldsymbol{x})\} = g_*(\boldsymbol{y})$$

and hence $\alpha = \alpha + \beta - \beta \leq g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})$. This is the desired expression.

If (ii) holds, the result follows from Theorem 3.7. ■

Remark. The reason we assume that f and g are proper functions is to avoid undefined situations like $\infty - \infty$. ◊

Remark. The fact that the infimum is attained under the condition (ii) and supremum is attained under (i) may seem a little unintuitive. As an example, take the following functions:

$$f(x) = e^x$$
$$g(x) = -e^x$$

As seen in Figure 12, the difference $f - g$ gets smaller and smaller as x tends to $-\infty$.

The solution to the primal problem is hence 0 but the infimum is not attained. This happens although the condition (i) holds. It is easily seen that the only common point for $\text{dom}\, f$ and $\text{dom}\, g$ is 0, which is not an interior point of either set. Hence (ii) does not hold. ◊

Remark. Although the Fenchel dual and Lagrangian dual problems look very different they can in fact shown to be equivalent. For the proof, see [8]. This does not mean that one of the problems is redundant. There are applications in which Fenchel duality is more suitable than Lagrangian duality, and vice versa. ◊

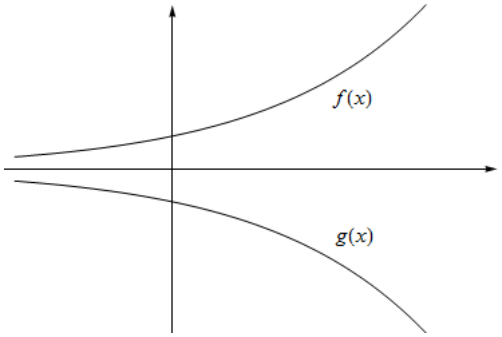


Figure 12: We have $\inf_x\{f(x)-g(x)\}=0$. Although $\operatorname{dom}(f)\cap\operatorname{dom}(g)\neq\varnothing$ holds true, the infimum is not attained.

3.4 Different Forms of the Duality Theorem

In the previous section we introduced the Fenchel primal and the associated dual problem. There are, however, other ways to define the primal problem. We dedicate this section to talking about possible modifications and generalizations.

The first thing we note is that a regular optimization problem consists of more than merely the objective function. The problem often requires that some constraints are fulfilled. There is an easy way to smuggle in linear constraints to Fenchel duality. Consider the following optimization problem:

$$\begin{aligned} &\text{Minimize} && f(\boldsymbol{x}) \\ &\text{subject to} && A\boldsymbol{x}\leq 0 \end{aligned}$$

where $f:\mathbb{R}^n\to\overline{R}$ is convex and A is a real $m\times n$ matrix. If we let $g:\mathbb{R}^m\to\overline{\mathbb{R}}$ be defined by

$$g(A\boldsymbol{x})=\begin{cases}0 & \text{if } A\boldsymbol{x}\leq 0\\ -\infty & \text{if } A\boldsymbol{x}>0\end{cases} \tag{3.6}$$

then we can write

$$\inf_{A\boldsymbol{x}\leq 0} f(\boldsymbol{x})=\inf_{\boldsymbol{x}}\{f(\boldsymbol{x})-g(A\boldsymbol{x})\}$$

This looks like a problem suitable for Fenchel duality. In fact, we may assume that g is any proper concave function. The above is a good illustration of how duality is used; a complicated problem is converted into a couple of unconstrained problems that are either trivial or can be solved by simple differentiation.

Theorem 3.11. [[2], Corollary 31.2.1] *Let $f:\mathbb{R}^n\to\overline{\mathbb{R}}$ be a closed proper convex function and $g:\mathbb{R}^m\to\overline{\mathbb{R}}$ a closed proper concave function. Suppose that A is a real $m\times n$ matrix. Then we have*

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n}\{f(\boldsymbol{x})-g(A\boldsymbol{x})\}=\sup_{\boldsymbol{y}\in\mathbb{R}^m}\{g_*(\boldsymbol{y})-f^*(A^T\boldsymbol{y})\}$$

if one of the following conditions hold:

(i) There exists an $\boldsymbol{x}\in\operatorname{ri}(\operatorname{dom} f)$ such that $A\boldsymbol{x}\in\operatorname{ri}(\operatorname{dom} g)$.

(ii) There exists a $\boldsymbol{y}\in\operatorname{ri}(\operatorname{dom} g_*)$ such that $A^T\boldsymbol{y}\in\operatorname{ri}(\operatorname{dom} f^*)$.

Under (i) the supremum is attained, under (ii) the infimum is attained. If both (i) and (ii) hold, both the supremum and infimum are finite.

Remark. In the special case (3.6) the condition (i) translates to Slater's condition (see Definition 2.8). ◊

The above theorem can be taken one step further.

Theorem 3.12 (Extended Fenchel Duality)**.** [[9], Theorem 2] *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a closed proper convex function and $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ a closed proper concave function. Suppose that A is a real $m \times n$ matrix and let $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{c} \in \mathbb{R}^m$. Then we have*

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n}\{f(\boldsymbol{x}) - g(\boldsymbol{c} - A\boldsymbol{x}) + \boldsymbol{b}^T\boldsymbol{x}\} = \sup_{\boldsymbol{y}\in\mathbb{R}^m}\{h_*(A^T\boldsymbol{y} - \boldsymbol{b}) - f^*(\boldsymbol{y}) + \boldsymbol{c}^T\boldsymbol{y}\}$$

if one of the following conditions hold:

(i) There exists an $\boldsymbol{x} \in \operatorname{ri}(\operatorname{dom} f)$ such that $\boldsymbol{c} - A\boldsymbol{x} \in \operatorname{ri}(\operatorname{dom} g)$.

(ii) There exists a $\boldsymbol{y} \in \operatorname{ri}(\operatorname{dom} g_*)$ such that $A^T\boldsymbol{y} - \boldsymbol{b} \in \operatorname{ri}(\operatorname{dom} f^*)$.

Under (i) the supremum is attained, unless the common value is $-\infty$. Under (ii) the infimum is attained, unless the common value is ∞.

The second thing that can be made more general is the domain of definition. There is nothing that really dictates that f and g should be defined on $\mathbb{R}^n$. It is more common to assume that $f, g : \mathcal{X} \to \overline{\mathbb{R}}$, where $\mathcal{X}$ is a finite-dimensional real vector-space. The problem that arises from changing the domain of definition for f and g is that the domains of definition for their conjugates are also affected. The reason is that the definition of the conjugate function includes scalar product. In order for us to be able to define $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ properly, the notion of *dual space* must be introduced.

Definition 3.13 (Dual Space)**.** *Let $\mathcal{X}$ be a finite-dimensional real vector space. The dual space of $\mathcal{X}$ is denoted by $\mathcal{X}'$ and it consists of all linear maps from $\mathcal{X}$ to $\mathbb{R}$.*

Example 3.14.

(a) Consider the case $\mathcal{X} = \mathbb{R}^n$. A linear functional from $\mathbb{R}^n$ to $\mathbb{R}$ is given by an $1 \times n$ matrix C:

$$C\boldsymbol{x} \in \mathbb{R} \qquad \text{for all } \boldsymbol{x} \in \mathbb{R}^n$$

Hence the dual space is $\mathbb{R}^{1\times n}$. However, the scalar product of two n-vectors $\boldsymbol{x}$ and $\boldsymbol{c}$ is usually defined by means of transpose: $\langle \boldsymbol{c}, \boldsymbol{x}\rangle = \boldsymbol{c}^T\boldsymbol{x}$. Therefore, we generally identify $\mathbb{R}^{1\times n}$ with $\mathbb{R}^n$ and say that $\mathbb{R}^n$ is the dual of itself.

(b) Consider the case $\mathcal{X} = \mathbb{R}^{n\times m}$. A linear functional can be expressed by means of the scalar product of two real matrices: trace. Given an $m \times n$ matrix C we can write:

$$\langle C, X\rangle = \operatorname{tr}(CX) \in \mathbb{R} \qquad \text{for all } X \in \mathbb{R}^{n\times m}$$

Hence the dual space of $\mathbb{R}^{n\times m}$ is $\mathbb{R}^{m\times n}$. ◊

Theorem 3.15. *Let $\mathcal{X}$ be a finite-dimensional vector space and let $f : \mathcal{X} \to \overline{\mathbb{R}}$ be a proper convex function and $g : \mathcal{X} \to \overline{\mathbb{R}}$ a proper concave function. Then we have*

$$\inf_{\boldsymbol{x}\in\mathcal{X}}\{f(\boldsymbol{x}) - g(\boldsymbol{x})\} = \sup_{\boldsymbol{y}\in\mathcal{X}'}\{g_*(\boldsymbol{y}) - f^*(\boldsymbol{y})\}$$

if one of the following conditions hold:

(i) $\operatorname{ri}(\operatorname{dom} f) \cap \operatorname{ri}(\operatorname{dom} g) \neq \varnothing$

(ii) *f and g are closed, and* $\operatorname{ri}(\operatorname{dom} f^*) \cap \operatorname{ri}(\operatorname{dom} g_*) \neq \varnothing$

Under (i) the supremum is attained, under (ii) the infimum is attained. If both (i) and (ii) hold, both the supremum and infimum are finite.

In addition to the theorems in this section, there are plenty of other conditions under which Fenchel duality holds. Banach spaces are discussed in [10]. There is even a discrete version of the theorem involving discrete convex functions and sets that are not only suitable for the cause but also arise in practical applications. See [[11], Theorem 8.21] for the discrete Fenchel-Type duality result.

4 The S-Procedure

We start by defining the S-procedure. Let $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$ and $f, g_1 \ldots, g_m : \mathcal{X}^n \to \mathbb{R}$. Consider the following two assertions:

S_1: $f(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathcal{X}^n$ such that $g_i(\boldsymbol{x}) \geq 0$, $i = 1, \ldots, n$.

S_2: There exists $p_i \geq 0$, $i = 1, \ldots, m$ such that for all $\boldsymbol{x} \in \mathcal{X}^n$ we have

$$f(\boldsymbol{x}) - \sum_{k=1}^{m} p_i g_i(\boldsymbol{x}) \geq 0$$

Regardless of how we choose our functions, S_2 always implies S_1. The opposite implication, on the other hand, is not always true. In some cases it is easier to show that S_2 holds, and using S_2 to verify S_1 is called the S-procedure. The equivalence of S_1 and S_2 is called the losslessness of the S-procedure.

We have already seen one special case in which the S-procedure is lossless. According to Lemma 2.16, S_1 and S_2 are equivalent when f and $-\boldsymbol{g}$ are convex. We shall in this chapter establish other conditions under which these assertions are equivalent. Although these conditions shall involve certain convexity assumptions we shall not directly require f and $-\boldsymbol{g}$ to be convex. By Theorem 2.17 this gives rise to a class of non-convex optimization problems that can be solved using regular methods.

It is quite common to express the S-procedure by means of equality constraints:

S_1': $f(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathcal{X}^n$ such that $g_i(\boldsymbol{x}) = 0$, $i = 1, \ldots, n$.

S_2': There exists $p_i \in \mathbb{R}$, $i = 1, \ldots, m$ such that for all $\boldsymbol{x} \in \mathcal{X}^n$ we have

$$f(\boldsymbol{x}) - \sum_{k=1}^{m} p_i g_i(\boldsymbol{x}) \geq 0$$

Also, strict inequalities come up frequently:

S_1'': $f(\boldsymbol{x}) > 0$ for all nonzero $\boldsymbol{x} \in \mathcal{X}^n$ such that $g_i(\boldsymbol{x}) \geq 0$, $i = 1, \ldots, n$.

S_2'': There exists $p_i \geq 0$, $i = 1, \ldots, m$ such that for all $\boldsymbol{x} \in \mathcal{X}^n$ we have

$$f(\boldsymbol{x}) - \sum_{k=1}^{m} p_i g_i(\boldsymbol{x}) > 0$$

Remark. In Lagrangian duality the condition $\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$ is more common, whereas in papers concerning the S-procedure the form $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$ is more frequently used. As mentioned before, this creates an ambiguous situation concerning the term "Slater's condition". By Slater's condition we shall mean the existence of an $\bar{\boldsymbol{x}}$ such that $\boldsymbol{g}(\bar{\boldsymbol{x}}) > 0$, $\boldsymbol{g}(\bar{\boldsymbol{x}}) < 0$ or $\boldsymbol{g}(\bar{\boldsymbol{x}}) = 0$, depending on the constraints we use. $\Diamond$

4.1 Preliminaries

So far we have mainly considered functions defined on a real set. From now on our theory will also be applicable on the complex case. However, the notion of symmetric matrices and matrix transpose do not give desired results. Instead we need to generalize these terms to suit our purpose. Let $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$.

Definition 4.1 (Hermitian matrix)**.** *A complex-valued $n \times n$ matrix A is called Hermitian if each entry a_{ij} is the complex conjugate of the entry a_{ji}, that is if*

$$a_{ij} = \overline{a_{ji}}$$

for all $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$.

It follows from the above definition that the diagonal entries of a Hermitian matrix are real.

Definition 4.2 (Conjugate transpose)**.** *The* conjugate transpose *of a complex-valued $n \times m$ matrix A is defined as the complex conjugate of the transpose of the matrix: $A^* = \overline{A^T}$*

In the real case, the conjugate and the regular transpose coincide, just like Hermitian and symmetric matrices do. Also, the conjugate of a Hermitian matrix is the matrix itself. Most rules that apply for the regular transpose work for the conjugate transpose as well. An interesting consequence of the above definitions is that for any Hermitian $n \times n$ matrix A the expression $\boldsymbol{x}^* A \boldsymbol{x}$ is real for all $\boldsymbol{x} \in \mathbb{C}^n$. This is essentially the reason we do not use regular transpose for complex-valued matrices. This also means that the following definition makes sense.

Definition 4.3. *A Hermitian matrix $A \in \mathcal{X}^{n \times n}$ is called positive semidefinite if for every $\boldsymbol{x} \in \mathcal{X}^n$ the following inequality holds:*

$$\boldsymbol{x}^* A \boldsymbol{x} \geq 0$$

We denote $A \geq 0$. If the equality is taken only at $\boldsymbol{x} = \boldsymbol{0}$, then A is called positive definite, and we write $A > 0$. A Hermitian matrix B is called negative (semi)definite if $-B$ is positive (semi)definite.

We shall use the following notations:

$\mathcal{S}^n$	The set of symmetric real matrices of size n
$\mathcal{S}^n_+$	The set of positive semidefinite symmetric real matrices of size n
$\mathcal{S}^n_{++}$	Positive definite symmetric real matrices of size n
$\mathcal{H}^n$	Hermitian matrices of size n
$\mathcal{H}^n_+$	Positive semidefinite Hermitian matrices of size n

Functions defined by means of Hermitian matrices have a special name.

Definition 4.4 (Quadratic Form)**.** *A function $f : \mathcal{X}^n \to \mathbb{R}$ is called a* quadratic form *if there exists a Hermitian matrix $F \in \mathcal{X}^{n \times n}$, a vector $\boldsymbol{u} \in \mathcal{X}^n$ and a real scalar $v \in \mathbb{R}$ such that*

$$f(\boldsymbol{x}) = \boldsymbol{x}^* F \boldsymbol{x} + 2 \operatorname{Re} \boldsymbol{u}^* \boldsymbol{x} + v$$

for all $\boldsymbol{x} \in \mathcal{X}^n$. If $\boldsymbol{u} = \boldsymbol{0}$ and $v = 0$, then f is called a homogeneous quadratic form*.*

Quadratic forms possess certain highly desirable qualities. For instance, it is very easy to determine whether a homogeneous quadratic form is convex. We simply have

$$f(\boldsymbol{x}) = \boldsymbol{x}^* F \boldsymbol{x} \text{ is convex} \quad \Leftrightarrow \quad F \text{ is positive semidefinite}$$

Furthermore, f is strictly convex if and only if F is positive definite.

In some publications quadratic forms are always assumed to be homogeneous. The reason is simple: one can always derive a homogeneous quadratic form from a regular quadratic form. It is easily seen that

$$f(\boldsymbol{x}) = \boldsymbol{x}^* F \boldsymbol{x} + 2\,\mathrm{Re}(\boldsymbol{u}^* \boldsymbol{x}) + v = \mathrm{Re}\left(\begin{pmatrix} \boldsymbol{x}^* & 1 \end{pmatrix} \begin{pmatrix} F & \boldsymbol{u} \\ \boldsymbol{u}^* & v \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix} \right)$$

Furthermore, it follows from the fact that $\mathrm{Re}(\boldsymbol{u}^* \boldsymbol{x})$ is linear that

$$f(\boldsymbol{x}) = \boldsymbol{x}^* F \boldsymbol{x} + 2\,\mathrm{Re}(\boldsymbol{u}^* \boldsymbol{x}) + v \text{ is convex} \quad \Leftrightarrow \quad F \text{ is positive semidefinite}$$

A linear matrix inequality (LMI) is an expression of the form

$$A = F + \sum_{i=1}^{m} x_i G_i > 0 \tag{4.1}$$

where the matrices F, $G_i \in \mathcal{X}^{n \times n}$ are Hermitian and $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathcal{X}^n$. Here the inequality means that A is positive definite. Naturally, other inequalities rise in applications, not only "$<$".

Linear matrix inequalities are common in both optimization and control theory. Various problems that arise in these fields can be transformed into standard statements involving LMI's, see [12] for a more thorough discussion. That is why there are various different, seemingly useless results concerning LMI's.

Notice a strong connection between LMI's and quadratic forms: the inequality (4.1) holds if and only if the homogeneous quadratic form $f(\boldsymbol{x}) = \boldsymbol{x}^* A \boldsymbol{x}$ is strictly convex.

4.2 Dynamical Systems

A dynamical system is usually expressed as a differential equation. The state vector $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_n(t))$ describes the state in which the system is in at a given time t. Its derivative gives the rate of change at different points in time:

$$\dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}(t) = \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{x}(t) = f(t, \boldsymbol{x}(t), \boldsymbol{u}(t))$$

The variable t is often absorbed from the calculations. It is common to study an initial value problem with $x(0)$ given, often $x(0) = 0$. The input vector $\boldsymbol{u}(t) = (u_1(t), u_2(t), \ldots, u_m(t))$ tells us what we can do at a given point in time to get desired results. A dynamical system is called linear if f is linear in $\boldsymbol{x}$ and $\boldsymbol{u}$.

Stability analysis is one of the most fundamental areas in control theory. A dynamical system is called *Lyapunov stable* at a critical point $\dot{\boldsymbol{x}} = \boldsymbol{0}$ if all the trajectories are bounded. It is called *Lyapunov asymptotically stable* if the trajectories converge to zero. A system can be shown to be asymptotically stable if there exists a Lyapunov function $V(\boldsymbol{x}, \boldsymbol{u})$ such that $V(\boldsymbol{x}, \boldsymbol{u}) > 0$ and $\dot{V}(\boldsymbol{x}, \boldsymbol{u}) < 0$ for all $(\boldsymbol{x}, \boldsymbol{u}) \neq 0$. What the Lyapunov function looks like depends on the problem at hand, so a general form including all cases is hard to construct. Quite often, however, one comes to ask if the function can be chosen to be homogeneous quadratic.

We shall soon investigate the most basic linear system there is: $\dot{\boldsymbol{x}} = A\boldsymbol{x}$. Before proceeding we need a couple of definitions.

Definition 4.5 (Kronecker product)**.** *Let $A \in \mathbb{C}^{n\times p}$ and $B \in \mathbb{C}^{m\times q}$. The* Kronecker product *of these matrices is given by*

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1m}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \dots & a_{nm}B \end{pmatrix}$$

If $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{m\times m}$ we can define their *Kronecker sum*:

$$A \oplus B = (I_m \otimes A) + (B \otimes I_n)$$

If the eigenvalues of A and B are λ_i, $i = 1, 2, \dots, n$ and μ_i, $i = 1, 2, \dots, m$, respectively, then the eigenvalues of their Kronecker sum are simply

$$\lambda_1 + \mu_1, \dots, \lambda_1 + \mu_m, \quad \lambda_2 + \mu_1, \dots, \lambda_2 + \mu_m, \quad \dots \quad \lambda_n + \mu_1, \dots, \lambda_n + \mu_m \tag{4.2}$$

For the proof, see [[14], Theorem 13.16]. Note that neither the Kronecker product nor the sum is commutative.

Definition 4.6 (Spectrum)**.** *The set of eigenvalues of a matrix A is called its* spectrum *and is denoted by* $\operatorname{Sp} A$.

Example 4.7. [15] Consider the following dynamical system:

$$\dot{\boldsymbol{x}} = A\boldsymbol{x}$$

where $A \in \mathbb{R}^{n\times n}$ and $\boldsymbol{x} \in \mathbb{R}^n$. A Lyapunov function for the above system can be assumed to be of the form

$$V(\boldsymbol{x}) = \boldsymbol{x}^T P \boldsymbol{x} > 0$$

As mentioned before, the above inequality holds when $P \in \mathcal{S}^n_{++}$. To determine whether $\dot{V}$ is negative definite, consider:

$$\dot{V}(\boldsymbol{x}) = \dot{\boldsymbol{x}}^T P\boldsymbol{x} + \boldsymbol{x}^T P\dot{\boldsymbol{x}} = \boldsymbol{x}^T A^T P\boldsymbol{x} + \boldsymbol{x}^T PA\boldsymbol{x} = \boldsymbol{x}^T(A^T P + PA)\boldsymbol{x} < 0$$

Hence a Lyapunov function exists if and only if the system

$$A^T P + PA < 0 \qquad P \in \mathcal{S}^n_{++}$$

is feasible. Indeed, we can pick any $Q > 0$ and solve the equation

$$A^T P + PA = -Q \tag{4.3}$$

If a solution $P \in \mathcal{S}^n_{++}$ exists then V must be a Lyapunov function.

Let us define the operator "vec" by stacking the columns of a matrix on top of each other:

$$\operatorname{vec}(P) = \begin{pmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{1n} \\ p_{21} \\ \vdots \\ p_{nn} \end{pmatrix}$$

Then (4.3) takes the form

$$(A \oplus A^T)\,\mathrm{vec}(P) = [(I_n \otimes A) + (A^T \otimes I_n)]\,\mathrm{vec}(P) = \mathrm{vec}(Q)$$

It follows that a unique solution $P \in \mathcal{S}^n$ to (4.3) exists if and only if the Kronecker sum $A \oplus A^T$ is nonsingular. This happens when all the eigenvalues of the sum are nonzero. Since A is a square matrix, it has the same eigenvalues as A^T. By (4.2) a solution exists if and only if $\lambda + \mu \neq 0$ for all $\lambda, \mu \in \mathrm{Sp}\, A$.

We know that the eigenvalues of a real matrix come in complex conjugate pairs, meaning that if $\lambda = a + ib \in \mathrm{Sp}\, A$ then $\lambda^* = a - ib \in \mathrm{Sp}\, A$. Hence we only have to assume that the conjugate pairs are never equal. Therefore, a solution $P \in \mathcal{S}^n$ exists if and only if $\lambda + \lambda^* = 2\,\mathrm{Re}\,\lambda \neq 0$ for all $\lambda \in \mathrm{Sp}\, A$. This argument is valid even if some eigenvalues happened to be real. Furthermore, it can be shown that P is positive definite exactly when $\mathrm{Re}(\lambda) < 0$ for each $\lambda \in \mathrm{Sp}\, A$. $\Diamond$

Remark. The function A^TP+PA is called the Lyapunov operator, the expression $A^TP+PA < 0$ Lyapunov's inequality and $A^TP + PA = -Q$ Lyapunov's equality. $\Diamond$

The fact that the system in the above example is stable when all eigenvalues of A have negative real parts is so important that such matrices have earned a name.

Definition 4.8 (Hurwitz)**.** *A matrix $A \in \mathbb{C}^{n\times n}$ is called* Hurwitz *if all its eigenvalues have negative real part. It is called* antihurwitz *if all the eigenvalues have positive real part.*

Multiplying Lyapunov's inequality by -1 and setting $F = -A$ gives us the system

$$F^TP + PF > 0 \qquad P > 0$$

Since the eigenvalues of $-A$ are the negatives of the eigenvalues of A it follows that the above is solvable if and only if F is antihurwitz.

$\dot{\boldsymbol{x}}$ is enough to determine a dynamical system, but in some cases an output vector is included in the calculations:

$$\boldsymbol{y} = \boldsymbol{y}(t) = h(t, \boldsymbol{x}(t), \boldsymbol{u}(t))$$

We shall consider systems of the following form:

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{u} \\ \boldsymbol{y} &= C\boldsymbol{x} + D\boldsymbol{u} \end{aligned} \tag{4.4}$$

Here $\boldsymbol{x} \in \mathcal{X}^n$, $\boldsymbol{u}, \boldsymbol{y} \in \mathcal{X}^m$, $A \in \mathcal{X}^{n\times n}$, $B \in \mathcal{X}^{n\times m}$, $C \in \mathcal{X}^{m\times m}$ and $D \in \mathcal{X}^{m\times n}$. To study stabilization of a system, one often requires it to be *controllable*. By this we mean that using the input vector we can, given any initial state, always force the system into another state under a finite amount of time. We say that the pair $\{A, B\}$ is controllable.

One way to show that a pair is controllable is by showing that the *controllability matrix*

$$\begin{pmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{pmatrix}$$

has full row rank.

Another important assumption is that the system is *observable*. We are quite handicapped if the output vector does not tell us everything we need to know about the state of the system. One can show that the pair $\{A, C\}$ is observable by showing that the *observability matrix*

$$\begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

has full row rank.

Example 4.9 (Sector constraint)**.** [17] Consider the dynamical system (4.4) with $D = 0$. Let us assume that the system is real. Suppose that the so-called sector constraint holds:

$$\sigma(\boldsymbol{y}, \boldsymbol{u}) = (\beta \boldsymbol{y} - \boldsymbol{u})^T (\boldsymbol{u} - \alpha \boldsymbol{y}) \geq 0$$

where α and β are real numbers that satisfy $\alpha < \beta$.

In order to find out whether the critical point $(\boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{0}$ is stable let us seek for a quadratic Lyapunov function of the form $V(\boldsymbol{x}, \boldsymbol{u}) = V(\boldsymbol{x}) = \boldsymbol{x}^T P \boldsymbol{x} > 0$. The derivative of V becomes:

$$\begin{aligned}
\dot{V}(\boldsymbol{x}, \boldsymbol{u}) &= \dot{\boldsymbol{x}}^T P \boldsymbol{x} + \boldsymbol{x}^T P \dot{\boldsymbol{x}} = (A\boldsymbol{x} + B\boldsymbol{u})^T P \boldsymbol{x} + \boldsymbol{x}^T P (A\boldsymbol{x} + B\boldsymbol{u}) = \\
&= \boldsymbol{x}^T (A^T P + PA) \boldsymbol{x} + \boldsymbol{u}^T B^T P \boldsymbol{x} + \boldsymbol{x}^T P B \boldsymbol{u} + \boldsymbol{u}^T \cdot 0 \cdot \boldsymbol{u} = \\
&= \begin{pmatrix} \boldsymbol{x}^T & \boldsymbol{u}^T \end{pmatrix} \begin{pmatrix} A^T P + PA & PB \\ B^T P & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{u} \end{pmatrix} < 0
\end{aligned}$$

The inequality must be satisfied for all $(\boldsymbol{x}, \boldsymbol{u}) \neq 0$ that satisfy the sector constraint: $\sigma(\boldsymbol{y}, \boldsymbol{u}) = \sigma(C\boldsymbol{x}, \boldsymbol{u}) \geq 0$. Let us define $f(\boldsymbol{x}, \boldsymbol{u}) = -\dot{V}(\boldsymbol{x}, \boldsymbol{u})$ and

$$\begin{aligned}
g(\boldsymbol{x}, \boldsymbol{u}) &= 2\sigma(C\boldsymbol{x}, \boldsymbol{u}) = 2(\beta C\boldsymbol{x} - \boldsymbol{u})^T (\boldsymbol{u} - \alpha C\boldsymbol{x}) = \\
&= -2\alpha\beta \boldsymbol{x}^T C^T C \boldsymbol{x} + 2(\alpha + \beta) \boldsymbol{x}^T C^T \boldsymbol{u} - 2\boldsymbol{u}^T \boldsymbol{u} = \\
&= \begin{pmatrix} \boldsymbol{x}^T & \boldsymbol{u}^T \end{pmatrix} \begin{pmatrix} -2\beta\alpha C^T C & (\beta + \alpha) C^T \\ (\beta + \alpha) C & -2 \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{u} \end{pmatrix}
\end{aligned}$$

We can now reformulate the existence of a quadratic Lyapunov function as

$$f(\boldsymbol{x}, \boldsymbol{u}) > 0 \text{ for all } (\boldsymbol{x}, \boldsymbol{u}) \neq \boldsymbol{0} \text{ such that } g(\boldsymbol{x}, \boldsymbol{u}) \geq 0$$

This is equivalent to S_1''. We shall later see that in this case the S-procedure is lossless since f and g are quadratic. Hence a quadratic Lyapunov function exists if and only if there exists $p > 0$ such that $f + pg > 0$ for all $(\boldsymbol{x}, \boldsymbol{u}) \neq \boldsymbol{0}$. $\Diamond$

4.3 Special Case: Farkas Lemma

Farka's Lemma is a very well-known result from convex optimization and it is closely related to the S-procedure. Farka's Lemma states that if A is a real $m \times n$ matrix and $\boldsymbol{b}$ a real m-vector, then only one of the following systems has a solution:

(i) There exists $\boldsymbol{x} \in \mathbb{R}^n$ such that $A\boldsymbol{x} \geq \boldsymbol{0}$ and $\boldsymbol{b}^T \boldsymbol{x} < 0$.

(ii) There exists $\boldsymbol{y} \in \mathbb{R}^m$ such that $A^T \boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{y} \geq \boldsymbol{0}$.

What Farka's Lemma says about convex cones is not that interesting. Instead it is very useful when proving other results in convex optimization. Proving Farka's Lemma using Fenchel duality is very straightforward. This is not surprising since both of them follow from the Separating Hyperplane Theorem.

Let us begin by rewriting the lemma. Farka's lemma is true when the following assertions are equivalent:

(i)* $A\boldsymbol{x} \geq \boldsymbol{0}$ implies $\boldsymbol{b}^T \boldsymbol{x} \geq 0$.

(ii)* There exists $\boldsymbol{y} \in \mathbb{R}^m$ such that $A^T \boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{y} \geq \boldsymbol{0}$.

The assertion (i)* is equivalent to requiring that

$$\inf_{A\boldsymbol{x}\geq 0} \boldsymbol{b}^T\boldsymbol{x} \geq 0$$

It is easily seen that the equality holds by noting that $\boldsymbol{x} = \boldsymbol{0}$ is included in the set of points fulfilling (i)*. In fact, since the feasible set defines a convex cone it follows that the infimum is either 0 or it is $-\infty$; no other case is possible.

Let $f(\boldsymbol{x}) = \boldsymbol{b}^T\boldsymbol{x}$ and

$$g(A\boldsymbol{x}) = \begin{cases} 0 & \text{if } A\boldsymbol{x} \geq \boldsymbol{0} \\ -\infty & \text{if } A\boldsymbol{x} < \boldsymbol{0} \end{cases}$$

Then f is a proper convex and g a proper concave function. Hence we can apply Theorem 3.11 on the following optimization problem:

$$\inf_{A\boldsymbol{x}\geq 0} \boldsymbol{b}^T\boldsymbol{x} = \inf_{\boldsymbol{x}\in\mathbb{R}^n} \{f(\boldsymbol{x}) - g(A\boldsymbol{x})\}$$

We get

$$f^*(A^T\boldsymbol{y}) = \sup_{\boldsymbol{x}} \left\{ \left(A^T\boldsymbol{y}\right)^T \boldsymbol{x} - \boldsymbol{b}^T\boldsymbol{x} \right\} = \begin{cases} 0 & \text{if } A^T\boldsymbol{y} = \boldsymbol{b} \\ \infty & \text{otherwise} \end{cases}$$

and

$$g_*(\boldsymbol{y}) = \inf_{A\boldsymbol{x}} \left\{ \boldsymbol{y}^T A\boldsymbol{x} - g(A\boldsymbol{x}) \right\} = \inf_{A\boldsymbol{x}\geq\boldsymbol{0}} \left\{ \boldsymbol{y}^T A\boldsymbol{x} \right\} = \begin{cases} 0 & \text{if } \boldsymbol{y} \geq \boldsymbol{0} \\ -\infty & \text{otherwise} \end{cases}$$

Keeping in mind that $\sup\{\varnothing\} = -\infty$ we obtain

$$\inf_{\boldsymbol{x}\in\mathbb{R}^n} \{f(\boldsymbol{x}) - g(A\boldsymbol{x})\} = \sup_{\boldsymbol{y}\in\mathbb{R}^m} \{g_*(\boldsymbol{y}) - f^*(A^T\boldsymbol{y})\} = \sup_{\substack{A^T\boldsymbol{y}=\boldsymbol{b} \\ \boldsymbol{y}\geq\boldsymbol{0}}} \{0\}$$

If (ii)* holds, then the supremum is attained and equals 0. This in turn implies that (i)* holds. If (ii)* does not hold, then the supremum is $-\infty$ and hence the infimum is also $-\infty$. From this it follows that (i)* is not true. Hence (i)* and (ii)* are equivalent.

4.4 Relation to Fenchel duality

The S-procedure is connected to duality in a very similar way as Farka's lemma: The statement S_1 is equivalent to

$$\inf\{f(\boldsymbol{x}) \mid g_i(\boldsymbol{x}) \geq 0,\ i = 1, 2, \ldots, m,\ \boldsymbol{x} \in \mathcal{X}^n\} \geq 0$$

In fact, the assumptions we shall press on f and $\boldsymbol{g}$ shall force the infimum to be either zero or $-\infty$ as in the proof of Farka's lemma. The objective of this section is to find a dual problem that can be associated with S_2. Notice that unlike with Farka's lemma, we cannot directly apply Fenchel duality to the above problem because the domain is not necessarily real.

Consider the following optimization problem:

$$\begin{array}{ll} \text{Minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{g}(\boldsymbol{x}) \in D \\ & \boldsymbol{x} \in C \end{array}$$

where $C \subseteq \mathcal{X}^n$ and $D \subseteq \mathbb{R}^m$. In order to connect the above problem to assertion S_2, let us consider all points $\boldsymbol{p} \in \mathbb{R}^m$ that satisfy the following inequality

$$p_1 g_1(\boldsymbol{x}) + p_2 g_2(\boldsymbol{x}) + \cdots + p_m g_m(\boldsymbol{x}) = \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x}) \rangle \leq f(\boldsymbol{x}) \tag{4.5}$$

for all $\boldsymbol{x} \in C$. Denote the set of such $\boldsymbol{p}$'s by P. Notice that unlike in S_2, we are not interested in the signs of p_i's at this stage. The infimum taken over the right-hand side must naturally be larger than the supremum over the left-hand side. Hence

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \in D \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in D} \left\{ \inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{y} \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) \right\} \geq \inf_{\boldsymbol{y} \in D} \sup_{\boldsymbol{p} \in P} \langle \boldsymbol{p}, \boldsymbol{y} \rangle$$

We also have

$$\inf_{\boldsymbol{y} \in D} \sup_{\boldsymbol{p} \in P} \langle \boldsymbol{p}, \boldsymbol{y} \rangle \geq \sup_{\boldsymbol{p} \in P} \inf_{\boldsymbol{y} \in D} \langle \boldsymbol{p}, \boldsymbol{y} \rangle$$

Keeping in mind how a support function is defined (see Example 3.8) we write $\Psi_D(\boldsymbol{p}) = \inf_{\boldsymbol{y} \in D} \langle \boldsymbol{p}, \boldsymbol{y} \rangle$ and get

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \in D \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) \geq \sup_{\substack{\boldsymbol{p} \in P \\ \boldsymbol{p} \in \operatorname{dom} \Psi_D}} \Psi_D(\boldsymbol{p}) \tag{4.6}$$

In the next theorem we establish conditions under which the equality holds.

Theorem 4.10. [[1], Theorem 1] *Let $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$. Suppose that C is a nonempty subset of $\mathcal{X}^n$, D a nonempty convex subset of $\mathbb{R}^m$ and let $f : \mathcal{X}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathcal{X}^n \to \mathbb{R}^m$. Denote $\Omega(\boldsymbol{x}) = (f(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x}))$. Suppose further that $\Omega(C)$ a convex cone, and that*

$$\operatorname{ri}(D) \cap \operatorname{ri}(\boldsymbol{g}(C)) \neq \varnothing \tag{4.7}$$

Then the following equality holds:

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \in D \\ \boldsymbol{x} \in C}} f(\boldsymbol{x}) = \sup_{\boldsymbol{p} \in P \cap \operatorname{dom} \Psi_D} \Psi_D(\boldsymbol{p}) \tag{4.8}$$

where $\Psi_D(\boldsymbol{p})$ is the support function of the set D and P is the set of vectors satisfying (4.5)*. If in addition we have*

$$P \cap \operatorname{dom}(\Psi_D) \neq \varnothing \tag{4.9}$$

then the supremum is finite and attained.

Remark. This theorem holds even in the case where the functions are defined on a set of matrices $\mathcal{X}^{n \times m}$. $\mathbb{R}^m$ in the definition of $\boldsymbol{g}$ may be replaced by any finite-dimensional real vector space $\mathcal{Y}$, in which case D is a subset of $\mathcal{Y}$. $\Diamond$

Proof. Define $\phi : \mathbb{R}^m \to \overline{\mathbb{R}}$ by

$$\phi(\boldsymbol{y}) = \inf\{ f(\boldsymbol{x}) \mid \boldsymbol{x} \in C,\ \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{y} \}$$

and $I_D : \mathbb{R}^m \to \{-\infty, 0\}$ by

$$I_D(\boldsymbol{y}) = \begin{cases} 0 & \text{if } \boldsymbol{y} \in D \\ -\infty & \text{it } \boldsymbol{y} \notin D \end{cases}$$

We notice that the set $\{\boldsymbol{x} \in C \mid \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{y}\}$ may be empty for some $\boldsymbol{y}$, in which case we use the convention $\phi(\boldsymbol{y}) = \inf \varnothing = +\infty$.

The function ϕ is convex since $\Omega(C)$ is convex. The indicator function I_D of the convex set D is a proper concave function. The effective domains of ϕ and I_D are easily calculated:

$$\begin{aligned}\operatorname{dom}(\phi) &= \boldsymbol{g}(C)\\ \operatorname{dom}(I_D) &= D\end{aligned}$$

from which it follows that

$$\operatorname{ri}(\operatorname{dom}(\phi)) \cap \operatorname{ri}(\operatorname{dom}(I_D)) = \operatorname{ri}(\boldsymbol{g}(C)) \cap \operatorname{ri}(D) \neq \varnothing \tag{4.10}$$

as assumed in the theorem. And most importantly, we get

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x})\in D\\ \boldsymbol{x}\in C}} f(\boldsymbol{x}) = \inf_{\boldsymbol{y}\in\mathbb{R}^m}\{\phi(\boldsymbol{y}) - I_D(\boldsymbol{y})\} \tag{4.11}$$

We wish to apply Fenchel's duality theorem on the right-hand side. However, we first have to consider the case in which ϕ is improper. If ϕ is an improper convex function, that is we have $\phi(\boldsymbol{y}) = -\infty$ for some $\boldsymbol{y} \in \mathbb{R}^m$, then the convexity of ϕ forces the function to attain the value $-\infty$ everywhere in the relative interior of the domain of ϕ. By (4.10) there exists a point that lies both in the effective domain of ϕ and in the effective domain of I_D. The first set only giving function value $-\infty$ and latter set 0, it follows from (4.11) that

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x})\in D\\ \boldsymbol{x}\in C}} f(\boldsymbol{x}) = \inf_{\boldsymbol{y}}\{\phi(\boldsymbol{y}) - I_D(\boldsymbol{y})\} = -\infty$$

The inequality (4.6) forces the supremum of $\Psi_D(\boldsymbol{p})$ to be $-\infty$, and hence the equality (4.8) holds.

Let us now assume that ϕ is a proper convex function. By (4.10) and Theorem 3.15 the following equality holds:

$$\inf_{\boldsymbol{y}\in\mathbb{R}^m}\{\phi(\boldsymbol{y}) - I_D(\boldsymbol{y})\} = \sup_{\boldsymbol{p}\in\mathbb{R}^m}\{I_{D*}(\boldsymbol{p}) - \phi^*(\boldsymbol{p})\}$$

By (4.10), the supremum is attained for some $\boldsymbol{p}$. Let us now calculate the conjugate functions. We have already seen in Example 3.8 that the conjugate of the indicator function is the support function: $I_{D*}(\boldsymbol{p}) = \Psi_D(\boldsymbol{p})$.

Computing ϕ^* is not as straightforward. We shall divide the calculations in two cases: (i) $\boldsymbol{p} \notin P$ and (ii) $\boldsymbol{p} \in P$.

(i) Suppose that $\boldsymbol{p} \notin P$. Then there exists an $\boldsymbol{x}_0 \in C$ such that

$$\langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x}_0)\rangle - f(\boldsymbol{x}_0) = \delta > 0$$

Since $\Omega(C)$ is a cone there exists $\boldsymbol{x}_\lambda \in C$ for any $\lambda > 0$ such that

$$\langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x}_\lambda)\rangle - f(\boldsymbol{x}_\lambda) = \langle \boldsymbol{p}, \lambda\boldsymbol{g}(\boldsymbol{x}_0)\rangle - \lambda f(\boldsymbol{x}_0) = \lambda\delta$$

We now get

$$\phi_*(\boldsymbol{p}) = \sup_{\boldsymbol{y}\in\mathbb{R}^m}\{\langle \boldsymbol{p}, \boldsymbol{y}\rangle - \phi(\boldsymbol{y})\} \geq \sup_{\lambda>0}\{\langle \boldsymbol{p}, \lambda\boldsymbol{g}(\boldsymbol{x}_0)\rangle - \lambda f(\boldsymbol{x}_0)\} = +\infty$$

and hence $\phi_*(\boldsymbol{p}) = +\infty$ for all $\boldsymbol{p} \notin P$.

(ii) Suppose now that $\boldsymbol{p} \in P$. Then

$$\phi_*(\boldsymbol{p}) = \sup_{\boldsymbol{y}\in\mathbb{R}^m} \left\{ \langle \boldsymbol{p}, \boldsymbol{y}\rangle - \inf_{\substack{\boldsymbol{x}\in C \\ \boldsymbol{g}(\boldsymbol{x})=\boldsymbol{y}}} f(\boldsymbol{x}) \right\}$$

Since $\phi(\boldsymbol{y}) = +\infty$ whenever an $\boldsymbol{x} \in C$ such that $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$ does not exist, we can ignore such points and simply write

$$\phi_*(\boldsymbol{p}) = \sup_{\boldsymbol{x}\in C} \left\{ \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x})\rangle - \inf_{\boldsymbol{z}\in C:\boldsymbol{g}(\boldsymbol{z})=\boldsymbol{g}(\boldsymbol{x})} f(\boldsymbol{z}) \right\}$$

By the choice of $\boldsymbol{p}$, we always have $f(\boldsymbol{z}) \geq \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{z})\rangle$ and hence

$$\sup_{\boldsymbol{x}\in C} \left\{ \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x})\rangle - \inf_{\boldsymbol{z}\in C:\boldsymbol{g}(\boldsymbol{z})=\boldsymbol{g}(\boldsymbol{x})} f(\boldsymbol{z}) \right\} \leq \sup_{\boldsymbol{x}\in C} \{ \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x})\rangle - \langle \boldsymbol{p}, \boldsymbol{g}(\boldsymbol{x})\rangle \} = 0$$

since C is nonempty. Since $\Omega(C)$ is a cone, we have $\mathbf{0} \in \operatorname{cl}\Omega(C)$. Hence the equality must hold and we get $\phi_*(\boldsymbol{p}) = 0$ for all $\boldsymbol{p} \in P$.

Combining (i) and (ii) we see that ϕ^* disappears from the calculations and we simply get

$$\sup_{\boldsymbol{p}\in\mathbb{R}^m} \{I_{D*}(\boldsymbol{p}) - \phi^*(\boldsymbol{p})\} = \sup_{\boldsymbol{p}\in P\cap\operatorname{dom}(\Psi_D)} \Psi_D(\boldsymbol{p})$$

Hence the equality (4.8) holds.

Furthermore, if the condition $P \cap \operatorname{dom}(\Psi_D) \neq \varnothing$ holds, then the supremum must be larger than $-\infty$, which, combined with the above discussion, means that ϕ cannot be improper. As we saw earlier, the supremum is attained when ϕ is proper. The proof is now complete. ■

4.5 S-Procedure for Homogeneous Quadratic Forms

One frequently used condition under which the assertions S_1 and S_2 are equivalent is that the functions involved are homogeneous quadratic (together with some additional requirements). This means that the assertion S_2 transforms into a linear matrix inequality. That is very convenient since various numerical and explicit methods have been constructed to solve problems of that form.

The earliest result of the losslessness of the S-procedure is due to Finsler. In [18] he proved that if $\boldsymbol{x}^T B\boldsymbol{x} = 0$ implies that $\boldsymbol{x}^T A\boldsymbol{x} > 0$, then $A + pB$ is positive definite for some real p. Yakubovich was the first to formulate the S-procedure in abstract terms. Like Finsler, he only proved the result for $m = 1$. The proof Yakubovich used can be found for example in [4]. The idea is very similar to the discussion is Section 2.4. Yakubovich uses a result that states that when f and g are homogeneous quadratic, then $\Omega(\mathbb{R}^n)$ is a convex set.

Using Theorem 4.10 we can easily show the following result.

Proposition 4.11. *Let $\mathcal{X} = \mathbb{R}$ or $\mathcal{X} = \mathbb{C}$. Let $f : \mathcal{X}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathcal{X}^n \to \mathbb{R}^m$ be quadratic functions determined by*

$$\begin{aligned} f(\boldsymbol{x}) &= \boldsymbol{x}^* F\boldsymbol{x} \\ g_i(\boldsymbol{x}) &= \boldsymbol{x}^* G_i\boldsymbol{x} \qquad i = 1, 2, \ldots, m \end{aligned}$$

where $F, G_1, \ldots, G_m \in \mathcal{X}^{n\times n}$ are Hermitian matrices. Denote $\Omega(\boldsymbol{x}) = (f(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x}))$ and suppose that $\Omega(\mathcal{X}^n)$ convex. Also, suppose that Slater's condition holds. Then the S-procedure is lossless, that is, S_1 and S_2 equivalent.

Proof. Let $D = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{y} \leq \boldsymbol{0}\}$. It is easily seen that D is a convex subset of $\mathbb{R}^m$. $\Omega(\mathcal{X}^n)$ is a cone because f and $\boldsymbol{g}$ are quadratic. We wish to apply Theorem 4.10. In order to do that, we must show that the condition

$$\operatorname{ri}(D) \cap \operatorname{ri} \boldsymbol{g}(\mathcal{X}^n) \neq \varnothing$$

holds. This is true when there exists a $\boldsymbol{x} \in \mathcal{X}^n$ such that $g(\boldsymbol{x}) < \boldsymbol{0}$, i.e. when Slater's condition holds.

We now get from Theorem 4.10 that

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\boldsymbol{p} \in P} \inf_{\boldsymbol{y} \geq \boldsymbol{0}} \langle \boldsymbol{p}, \boldsymbol{y} \rangle \tag{4.12}$$

The infimum is easily calculated:

$$\inf_{\boldsymbol{y} \geq \boldsymbol{0}} \langle \boldsymbol{p}, \boldsymbol{y} \rangle = \begin{cases} 0 & \text{if } \boldsymbol{p} \geq \boldsymbol{0} \\ -\infty & \text{otherwise} \end{cases}$$

And hence (4.12) gets the form

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \begin{cases} 0 & \text{if there exists } \boldsymbol{p} \geq \boldsymbol{0} \text{ in } P \\ -\infty & \text{otherwise} \end{cases}$$

If a suitable $\boldsymbol{p}$ exists (S_2) then all $f(\boldsymbol{x})$-values are greater than or equal to zero whenever $\boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{0}$ (S_1). On the other hand, if the infimum on the left-hand side equals zero (S_1), then a suitable $\boldsymbol{p}$ exists (S_2). ■

Considering the labor behind Theorem 4.10, the above proof cannot really be considered more elegant than the one found for instance in [5]. The latter is analogous to the proof of Lemma 2.16. Instead, the results of this section can be used in quadratic programming.

4.6 Quadratic Duality

Suppose that $f : \mathcal{X}^n \to \mathbb{R}$ and $\boldsymbol{g} : \mathcal{X}^n \to \mathbb{R}^m$ are homogeneous quadratic functions determined by

$$\begin{aligned} f(\boldsymbol{x}) &= \boldsymbol{x}^* F \boldsymbol{x} \\ g_i(\boldsymbol{x}) &= \boldsymbol{x}^* G_i \boldsymbol{x} \qquad i = 1, 2, \ldots, m \end{aligned}$$

where F and the G_i's are Hermitian matrices. Consider once again the following optimization problem:

$$\begin{aligned} &\text{Minimize} && f(\boldsymbol{x}) \\ &\text{subject to} && \boldsymbol{g}(\boldsymbol{x}) \in D \\ & && \boldsymbol{x} \in \mathcal{X}^n \end{aligned} \tag{4.13}$$

This is a special case of the class of a quadratically constrained quadratic programs (QCQP) with the exception that the functions involved are assumed to be homogeneous quadratic instead of merely quadratic.

The next proposition follows from Theorem 4.10.

Proposition 4.12 (Quadratic Duality)**.** *Let f and $\boldsymbol{g}$ be homogeneous quadratic and denote $\Omega(\boldsymbol{x}) = (f(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x}))$. Suppose that $\Omega(\mathcal{X}^n)$ convex. Let $D \subseteq \mathbb{R}^m$ be convex. If*

$$\operatorname{ri}(D) \cap \operatorname{ri}(\boldsymbol{g}(\mathcal{X}^n)) \neq \varnothing \tag{4.14}$$

then we have the following duality relation:

$$\inf_{\substack{\boldsymbol{g}(\boldsymbol{x}) \in D \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\boldsymbol{p} \in P} \inf_{\boldsymbol{y} \in D} \langle \boldsymbol{p}, \boldsymbol{y} \rangle$$

What we have done is that we have translated the original non-convex problem to a convex one; the set P can easily be shown to be convex. Notice that Slater's conditions is replaced by the more general regularity condition (4.14).

Now pick two real m-vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that $-\boldsymbol{\infty} < \boldsymbol{\alpha} < \boldsymbol{\beta} \leq \boldsymbol{\infty}$. If in the above proposition we set

$$D = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{\alpha} \leq \boldsymbol{y} \leq \boldsymbol{\beta}\}$$

then we get the following duality relation:

$$\inf_{\substack{\boldsymbol{\alpha} \leq \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{\beta} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\boldsymbol{p} \in P} \left\{ \inf_{\boldsymbol{\alpha} \leq \boldsymbol{y} \leq \boldsymbol{\beta}} \langle \boldsymbol{p}, \boldsymbol{y} \rangle \right\}$$

Consider the sum $\langle \boldsymbol{p}, \boldsymbol{y} \rangle = p_1 y_1 + p_2 y_2 + \cdots + p_m y_m$. When $p_i \geq 0$, then $p_i \alpha_i$ is smaller than $p_i \beta_i$. When $p_i < 0$, then the opposite holds. If $\beta_i = \infty$ for some i, then the infimum simply becomes $-\infty$ if $p_i \leq 0$. Since we are interested in the supremum, we may ignore such points and assume p_i to be nonnegative when $\beta_i = \infty$. We now get:

$$\inf_{\substack{\boldsymbol{\alpha} \leq \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{\beta} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\substack{\boldsymbol{p} \in P \\ p_i \geq 0 \text{ when } \beta_i = +\infty}} \left\{ \sum_{p_i \geq 0} p_i \alpha_i + \sum_{p_i < 0} p_i \beta_i \right\} \tag{4.15}$$

In the special case $\beta = \boldsymbol{\infty}$ this expression can be simplified to

$$\inf_{\substack{\boldsymbol{\alpha} \leq \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{\beta} \\ \boldsymbol{x} \in \mathcal{X}^n}} f(\boldsymbol{x}) = \sup_{\substack{\boldsymbol{p} \in P \\ \boldsymbol{p} \geq \boldsymbol{0}}} \langle \boldsymbol{p}, \boldsymbol{\alpha} \rangle$$

Hence the problem 4.13 for $D = \{\boldsymbol{y} \in \mathbb{R}^m \mid \boldsymbol{y} \geq \boldsymbol{\alpha}\}$ has the following dual problem:

$$\begin{aligned} \text{Maximize} \quad & \langle \boldsymbol{p}, \boldsymbol{\alpha} \rangle \\ \text{subject to} \quad & F - \sum_{i=1}^{m} p_i G_i \geq 0 \\ & \boldsymbol{p} \geq \boldsymbol{0} \end{aligned} \tag{4.16}$$

As mentioned before, this is a convex optimization problem, and it is a standard problem in the theory of LMI's. It is simply a question of maximizing a linear function over a convex set, and there are effective numerical methods for solving this kind of problems. The Lagrangian dual problem

$$\begin{aligned} \text{Maximize} \quad & \inf_{\boldsymbol{x} \in \mathcal{X}^n} \left\{ \boldsymbol{x}^* F \boldsymbol{x} - \sum_{i=1}^{m} p_i (\boldsymbol{x}^* G_i \boldsymbol{x} - \alpha_i) \right\} \geq 0 \\ \text{subject to} \quad & \boldsymbol{p} \geq \boldsymbol{0} \end{aligned} \tag{4.17}$$

is not quite as approachable. Also, it is not that straightforward to find the Lagrangian dual problem in the general case (4.15) where the entries of $\boldsymbol{\beta}$ are allowed be finite. Summing up we may say that the result derived using Fenchel duality has certain advantages compared to the Lagrangian version and also it applies to a wider circle of quadratic optimization problems. For further discussion on how to solve (4.15), see for instance [12] or [3].

We have so far seen the broad uses the convexity of $\Omega(C)$ implies. Unfortunately, finding and proving conditions under which this set is convex is quite laborious. In the real case, the easiest and perhaps the most useful result can be found in [19]. It is the following:

> If $m = 1$, then $\Omega(C)$ is convex for all symmetric matrices F and G_1.

This result cannot be generalized to $m > 1$. Instead, one has to construct special cases. The simplest condition for $\mathcal{X} = \mathbb{C}$ is the following:

> If $m = 2$, then $\Omega(C)$ is convex for arbitrary Hermitian matrices F, G_1 and G_2.

For the proof, see [20]. This result cannot be generalized, either.

In addition to the ones mentioned above, several other conditions for the convexity of $\Omega(C)$ have been derived. See [21] for more thorough discussion.

5 Kalman-Yakubovich-Popov Lemma

As mentioned in the introduction, the Kalman-Yakubovich-Popov (KYP) lemma states that certain assertions are equivalent. We start by giving an example from passivity analysis and then generalize the result thus gaining a standard form of the KYP-lemma. After that we generalize the result even further in order to state and prove the extended version of the lemma that Gusev established in [1].

We shall only consider the real case since the complex case leads to certain cumbersome definitions.

5.1 Standard Form of the Lemma

Example 5.1 (Passive systems)**.** [22] Consider the following real dynamical system including an output vector:

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{u} \\ \boldsymbol{y} &= C\boldsymbol{x} \end{aligned} \tag{5.1}$$

Let us investigate the *storage function* $V(t)$ associated with the system. The storage function can be assumed to be quadratic: $V(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T P\boldsymbol{x} > 0$. From (5.1) we obtain:

$$\begin{aligned} 0 &= \dot{\boldsymbol{x}} - A\boldsymbol{x} - B\boldsymbol{u} \\ 0 &= \boldsymbol{x}^T P\dot{\boldsymbol{x}} - \boldsymbol{x}^T PA\boldsymbol{x} - \boldsymbol{x}^T PB\boldsymbol{u} \\ \boldsymbol{u}^T\boldsymbol{y} &= \boldsymbol{x}^T P\dot{\boldsymbol{x}} - \boldsymbol{x}^T PA\boldsymbol{x} - (\boldsymbol{x}^T PB\boldsymbol{u})^T + \boldsymbol{u}^T\boldsymbol{y} \\ \boldsymbol{u}^T\boldsymbol{y} &= \boldsymbol{x}^T P\dot{\boldsymbol{x}} - \boldsymbol{x}^T PA\boldsymbol{x} - \boldsymbol{u}^T B^T P\boldsymbol{x} + \boldsymbol{u}^T C\boldsymbol{x} \end{aligned} \tag{5.2}$$

The derivative of $V(\boldsymbol{x})$ is

$$\dot{V}(\boldsymbol{x}) = \frac{\mathrm{d}V(\boldsymbol{x})}{\mathrm{d}t} = \frac{1}{2}\dot{\boldsymbol{x}}^T P\boldsymbol{x} + \frac{1}{2}\boldsymbol{x}^T P\dot{\boldsymbol{x}} = \left(\frac{1}{2}\dot{\boldsymbol{x}}^T P\boldsymbol{x}\right)^T + \frac{1}{2}\boldsymbol{x}^T P\dot{\boldsymbol{x}} = \boldsymbol{x}^T P\dot{\boldsymbol{x}}$$

Integrating the last equation of (5.2) between 0 and a positive real number t with respect to time gives us the relation

$$\begin{aligned} \int_0^t \boldsymbol{u}(s)^T\boldsymbol{y}(s)\ \mathrm{d}s =& V(t) - V(0) - \frac{1}{2}\int_0^t \boldsymbol{x}(s)^T(PA + A^T P)\boldsymbol{x}(s)\ \mathrm{d}s - \\ &- \int_0^t \boldsymbol{u}(s)^T(B^T P - C)\boldsymbol{x}(s)\ \mathrm{d}s \end{aligned} \tag{5.3}$$

Let us claim that the last term simply disappears. Then we get the so-called *dissipation equality*:

$$\overbrace{V(t)}^{(a)} = \overbrace{V(0)}^{(b)} + \overbrace{\frac{1}{2}\int_0^t \boldsymbol{x}^T(s)(PA + A^T P)\boldsymbol{x}(s)\ \mathrm{d}s}^{(c)} + \overbrace{\int_0^t \boldsymbol{u}(s)^T\boldsymbol{y}(s)\ \mathrm{d}s}^{(d)}$$

A system that satisfies the above equation along its trajectories is called dissipative. There is a very intuitive physical interpretation of the above equation. The term (a) on the left-hand side tells us how much energy there is at time t. This energy depends on (b) the initial energy, (c) the absorbed energy and (d) the energy that is produced externally. The term (c) is never positive, and hence if the last term in (5.3) disappears we get the *dissipation inequality*:

$$V(t) - V(0) \leq \int_0^t \boldsymbol{u}(s)^T \boldsymbol{y}(s) \, \mathrm{d}s$$

If the above holds then the system is called *passive*, meaning that it does not produce energy via the input-output process. Returning to the infinitesimal form gives us:

$$\boldsymbol{x}^T P \dot{\boldsymbol{x}} \leq \boldsymbol{u}^T C \boldsymbol{x}$$

and hence

$$\begin{aligned} 2\boldsymbol{x}^T P(A\boldsymbol{x} + B\boldsymbol{u}) - 2\boldsymbol{u}^T C\boldsymbol{x} = \boldsymbol{x}^T(A^T P + PA)\boldsymbol{x} + 2\boldsymbol{x}^T(PB - C^T)\boldsymbol{u} = \\ = \begin{pmatrix} A^T P + PA & PB - C^T \\ B^T P - C & 0 \end{pmatrix} \leq 0 \end{aligned} \tag{5.4}$$

It can be shown that the above has a solution $P \in \mathcal{S}^n_{++}$ if and only if there exists an $H \in \mathbb{R}^{(n+m)\times n}$ such that

$$\begin{pmatrix} A^T P + PA & PB - C^T \\ B^T P - C & 0 \end{pmatrix} = -H^T H \tag{5.5}$$

for some $P \in \mathcal{S}^n_{++}$. So if the system (5.1) is passive then the above equation and the LMI (5.4) are feasible. To show that the opposite implication holds, set $H = (L^T \quad W)$. Then the above becomes

$$\begin{pmatrix} A^T P + PA & PB - C^T \\ B^T P - C & 0 \end{pmatrix} = - \begin{pmatrix} LL^T & LW \\ W^T L^T & W^T W \end{pmatrix}$$

From this we get $B^T P - C = 0$ and hence the last term in (5.3) disappears which leads to the dissipation inequality.

We have now found two different ways to determine whether a system is passive. To find a third one, let us consider the *transfer function* $H : \mathbb{C} \to \mathcal{H}^m$ of the system. The transfer function is defined as the ratio between the output $\boldsymbol{y}$ and input $\boldsymbol{u}$. It can be shown to be equal to

$$H(s) = C(sI_n - A)^{-1} B$$

Notice that if we have the initial condition $x(\mathbf{0}) = \mathbf{0}$ then the dissipation inequality takes the form $\int_0^t \boldsymbol{u}(s)^T \boldsymbol{y}(s) \, \mathrm{d}s \geq 0$. Let us define

$$\boldsymbol{u}_{\mathcal{T}}(s) = \begin{cases} \boldsymbol{u}(s) & 0 \leq s \leq t \\ 0 & \text{otherwise} \end{cases}$$

It follows from Parseval's theorem and some other manipulations that we can write

$$\int_0^t \boldsymbol{u}(s)^T \boldsymbol{y}(s) \, \mathrm{d}s = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{Re}(H(i\omega)) |\boldsymbol{u}_{\mathcal{T}}(\omega)|^2 \, \mathrm{d}\omega$$

If $\mathrm{Re}(H(i\omega)) \geq 0$ for every $\omega \in \overline{\mathbb{R}}$, then the right-hand side of the above equation is nonnegative and the system is passive. The opposite implication also holds. Hence the system (5.1) is passive if and only if the transfer function satisfies the following:

$$\mathrm{Re}\, H(i\omega) \geq 0 \quad \text{for all} \quad \omega \in \overline{\mathbb{R}} \tag{5.6}$$

For a more thorough discussion, see the proof of Theorem 2.6 in [22]. Summing up, one can show passivity by using any of the conditions (5.4), (5.5) or (5.6). ◊

Remark. We have not taken into account all necessary details in the above example, such as assumptions on controllability and the poles of the transfer function. ◊

The equivalence of the three conditions that guarantee passivity of the system in Example 5.1 is essentially what the Kalman-Yakubovich-Popov Lemma looks like. But since similar results arise in other applications it is not desirable to restrict our attention to this special case.

Set

$$\Lambda'(P) = \begin{pmatrix} A^T P + PA & PB \\ B^T P & 0 \end{pmatrix} \tag{5.7}$$

and pick any matrix $G \in \mathcal{S}^{n+m}$. Consider the system

$$\Lambda'(P) - G = -H^T H \qquad P \in \mathcal{S}^n \qquad H \in \mathbb{R}^{m\times(n+m)} \tag{5.8}$$

This is known as the Lur'e equation. It is easily seen that (5.5) is a special case of the above. The only difference is that P is not assumed to be positive definite. This follows from other assumptions that we do not wish to include in the general form of the KYP-lemma. The Lur'e equation is feasible if and only if the following linear matrix inequality is feasible:

$$\Lambda'(P) - G \leq 0 \qquad P \in \mathcal{S}^n \tag{5.9}$$

This expression is the "associated LMI" that was mentioned in the introduction. There are results involving strict inequalities, but we shall only discuss the semidefinite case.

The third equivalent system usually included in the KYP-lemma is the so-called frequency condition. In Example 5.1 it was (5.6). Let Γ be the set of purely imaginary numbers. It then follows from (5.6) that the system (5.1) is passive if and only if:

$$H(\lambda) + H(\lambda)^* \geq 0 \quad \text{for all} \quad \lambda \in \Gamma$$

However, this expression is not good since it is defined by means of the matrix C. Instead we want to have a condition expressed by means of G in order to connect the frequency condition to the general case (5.8). We get

$$\begin{aligned} H(\lambda) + H(\lambda)^* &= C(\lambda I_n - A)^{-1}B + ((\lambda I_n - A)^{-1}B)^* C^T = \\ &= \begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix}^* \begin{pmatrix} 0 & C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix} = \\ &= \begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix}^* G \begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix} \geq 0 \end{aligned}$$

Notice that λ must be chosen so that $\lambda I_n - A$ is invertible. This inconsistency rises from the fact that we were not very exact in example 5.1. The above expression is well-defined as long as no $\lambda \in \Gamma$ is not an eigenvalue of A, that is $\operatorname{Sp} A \cap \Gamma = \varnothing$.

Returning to the general case the frequency condition with arbitrary $G \in \mathcal{S}^{n+m}$ takes the form

$$\begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix}^* G \begin{pmatrix} (\lambda I_n - A)^{-1}B \\ I_m \end{pmatrix} \geq 0 \qquad \text{for all } \lambda \in \Gamma \tag{5.10}$$

We are now ready to present the KYP-lemma.

Lemma 5.2 (KYP-Lemma)**.** [24] *Suppose that the pair* $\{A, B\}$ *is controllable and that the condition* $\mathrm{Sp}(A \cap \Gamma) = \varnothing$ *holds. Pick any matrix* $G \in \mathcal{S}^{n+m}$*. Then the systems* (5.8)*,* (5.9) *and* (5.10) *are equivalent.*

Example 5.3 (Positive real lemma and ARE's)**.** Sometimes the Lur'e equation is given as a set of equalities. The system

$$\begin{aligned}\dot{\boldsymbol{x}} &= A\boldsymbol{x} + B\boldsymbol{u}\\ \boldsymbol{y} &= C\boldsymbol{x} + D\boldsymbol{u}\end{aligned}$$

can be shown to be passive in and only if there exists $P = P^T > 0$ and $H \in \mathbb{R}^{(n+m)\times n}$ satisfying the following Lur'e equation:

$$\begin{pmatrix} A^TP + PA & PB - C^T \\ B^TP - C & -D - D^T \end{pmatrix} = -H^TH \tag{5.11}$$

Setting $H = \begin{pmatrix} L^T & W \end{pmatrix}$ gives us the following equations:

$$\begin{aligned} A^TP + PA &= -LL^T\\ PB - C^T &= -LW\\ D + D^T &= W^TW\end{aligned}$$

The feasibility of the above system of equations is equivalent to the transfer function $H(s) = C(sI_n - A)^{-1}B + D$ fulfilling (5.6). This result is a special case of the KYP-lemma and it often goes by the name *positive real lemma*.

Another interesting point is that the Lur'e equation (5.11) has the same solutions as the following algebraic Riccatti equation:

$$-PA - A^TP - (C - B^TP)^T(D + D^T)^{-1}(C - B^TP) \geq 0$$

assuming that $D + D^T > 0$. $\Diamond$

5.2 Generalizations

We shall now start generalizing the definitions from the previous section. Remember that so far we have used $\Gamma = i\mathbb{R}$. In his article [25] Churilov shows that the KYP-lemma holds in the case in which Γ is a circle on the complex plane or any line vertical with the imaginary axis. We let Γ to be defined by means of a real 2×2 matrix Θ such that $\det\Theta < 0$ in the following manner: Consider the function $\varphi(\lambda) = \begin{pmatrix}\lambda & 1\end{pmatrix}\Theta\begin{pmatrix}\lambda & 1\end{pmatrix}^*$ defined on $\mathbb{C}$. Then

$$\Gamma = \{\lambda \in \mathbb{C} \mid \varphi(\lambda) = 0\} \tag{5.12}$$

By setting $\lambda = a + bi$ we obtain

$$\begin{aligned}\begin{pmatrix}\lambda & 1\end{pmatrix}\Theta\begin{pmatrix}\lambda^*\\ 1\end{pmatrix} &= \begin{pmatrix}\lambda & 1\end{pmatrix}\begin{pmatrix}\theta_{11} & \theta_{12}\\ \theta_{12} & \theta_{22}\end{pmatrix}\begin{pmatrix}\lambda^*\\ 1\end{pmatrix} =\\ &= \theta_{11}a^2 + \theta_{11}b^2 + 2a\theta_{12} + \theta_{22} = 0\end{aligned}$$

If $\theta_{11} = 0$ then Γ is a line parallel with the imaginary axis. If $\theta \neq 0$ then it is a circle centered at $a = -\theta_{12}/\theta_{11}$, $b = 0$ and with radius $\theta_{12}^2/\theta_{11}^2 - \theta_{22}/\theta_{11}$. The assumption set on the determinant of Θ guarantees that the radius is positive. Γ divides the complex plane into two disjoint sets:

$$\begin{aligned}\Omega_\Theta^+ &= \{\lambda \in \mathbb{C} \mid \varphi(\lambda) > 0\}\\ \Omega_\Theta^- &= \{\lambda \in \mathbb{C} \mid \varphi(\lambda) < 0\}\end{aligned} \tag{5.13}$$

Now, assume that two matrices M and N in $\mathbb{R}^{n\times(n+m)}$ are given. Consider the function $\Lambda : \mathcal{H}^{n+m} \to \mathcal{S}^n$ given by

$$\Lambda(S) = \begin{pmatrix} M & N \end{pmatrix} (\Theta \otimes S) \begin{pmatrix} M^T \\ N^T \end{pmatrix} \tag{5.14}$$

To see that Λ really is real-valued, see the proof of Lemma 5.4 in the appendix. In the case $m = 0$, $M = A$ and $\Theta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ we simply get the Lyapunov operator: $\Lambda(S) = AS + SA^T$. Therefore (5.14) is called the generalized Lyapunov operator. The adjoint operator of (5.14) is given by

$$\Lambda'(P) = \begin{pmatrix} M^T & N^T \end{pmatrix} (\Theta^T \otimes P) \begin{pmatrix} M \\ N \end{pmatrix} \tag{5.15}$$

and it is a function from $\mathcal{S}^n$ to $\mathcal{S}^{n+m}$.

Suppose now that $m > 0$ and consider the following system:

$$\exists F^{\pm} \in \mathbb{R}^{m\times(n+m)} \qquad \det \begin{pmatrix} N \\ F^{\pm} \end{pmatrix} \neq 0 \qquad \operatorname{Sp}\left(M \begin{pmatrix} N \\ F^{\pm} \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix} \right) \subseteq \Omega_{\Theta}^{\pm} \tag{5.16}$$

We are interested in the case in which (5.16) is feasible.

To understand how the above definitions are connected to those made in the previous section, consider the following special case:

$$\Theta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad N = \begin{pmatrix} I_n & 0 \end{pmatrix} \tag{5.17}$$

This special case has certain advantages toward the general form. We shall see that there always exists a transformation that allows us to express our function in the form (5.17). If we denote

$$M = \begin{pmatrix} A & B \end{pmatrix} \tag{5.18}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ then we obtain

$$\begin{aligned} \Lambda'(P) &= \begin{pmatrix} A^T & I_n \\ B^T & 0 \end{pmatrix} \begin{pmatrix} 0 & P \\ P & 0 \end{pmatrix} \begin{pmatrix} A & B \\ I_n & 0 \end{pmatrix} = \\ &= \begin{pmatrix} P & A^T P \\ 0 & B^T P \end{pmatrix} \begin{pmatrix} A & B \\ I_n & 0 \end{pmatrix} = \\ &= \begin{pmatrix} A^T P + PA & PB \\ B^T P & 0 \end{pmatrix} \end{aligned} \tag{5.19}$$

This expression is consistent with our previous definition of the adjoint operator, see equation (5.7). Also, we get

$$\Lambda(S) = \begin{pmatrix} M & N \end{pmatrix} \begin{pmatrix} 0 & S \\ S & 0 \end{pmatrix} \begin{pmatrix} M^T \\ N^T \end{pmatrix} = NSM^T + MSN^T = 0 \tag{5.20}$$

If Θ is given in this standard way we get

$$\varphi(\lambda) = \begin{pmatrix} \lambda & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda^* \\ 1 \end{pmatrix} = \lambda + \lambda^* = 2 \operatorname{Re} \lambda$$

Hence Γ is simply the imaginary axis. Ω_Θ^+ therefore contains all complex numbers with positive real part and Ω_Θ^- those with negative real part. If we use (5.18) it follows from condition (5.16) that

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} I_n & 0 \\ F^\pm & \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix} = \begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} I_n & * \\ K^\pm & * \end{pmatrix} \begin{pmatrix} I_n \\ 0 \end{pmatrix} =$$
$$= \begin{pmatrix} A + BK^\pm & * \end{pmatrix} \begin{pmatrix} I_n \\ 0 \end{pmatrix} = A + BK^\pm$$

Hence the feasibility of (5.16) translates to the existence of matrices $K^\pm \in \mathbb{R}^{m\times n}$ such that $A + BK^-$ is Hurwitz and $A + BK^+$ is antihurwitz. This is equivalent to assuming that the pair $\{A, B\}$ is controllable.

We shall now present some auxiliary result needed in the next section.

Lemma 5.4. *If* (5.16) *is feasible, then* $\Lambda(\mathcal{S}_+^{n+m}) = \mathcal{S}^n$.

Lemma 5.5. *Suppose* (5.16) *is feasible. Then*

(a) If the matrix $S \in \mathcal{S}_+^{n+m}$ *solves the equation* $\Lambda(S) = 0$ *then there exists vectors* $\boldsymbol{z}_i \in \mathbb{C}^{n+m}$, $i = 1, 2, \ldots, n+m$ *solving* $\Lambda(\boldsymbol{z}_i \boldsymbol{z}_i^*) = 0$ *such that*

$$S = \sum_{i=1}^{n+m} \boldsymbol{z}_i \boldsymbol{z}_i^*$$

(b) Let

$$A_1 = \{\boldsymbol{z} \in \mathbb{C}^{n+m} \mid \Lambda(\boldsymbol{z}\boldsymbol{z}^*) = 0\}$$
$$A_2 = \{\boldsymbol{z} \in \mathbb{C}^{n+m} \mid (\lambda N - M)\boldsymbol{z} = 0 \text{ for some } \lambda \in \Gamma\}$$

Then $A_1 = \operatorname{cl} A_2$

The proofs can be found in the appendix.

We shall now prove a duality result for matrices. As was mentioned in the remark following Theorem 4.10, the result holds even for matrices. Since the scalar product for matrices is defined as trace, the support function in the theorem takes the form:

$$\Psi_D(P) = \inf_{S \in D} \operatorname{tr}(SP)$$

Theorem 5.6. *Let* D *be a nonempty closed convex set in* $\mathcal{S}^n$. *Suppose that* (5.16) *is feasible. Then we have the following duality relation:*

$$\inf_{\substack{S \in \mathcal{S}_+^{n+m} \\ \Lambda(S) \in D}} \operatorname{tr}(GS) = \sup_{\substack{P \in \operatorname{dom} \Psi_D \\ \Lambda'(P) - G \leq 0}} \Psi_D(P)$$

where Λ *is given by* (5.14). *If there exists* $P \in \operatorname{dom} \Psi_D$ *satisfying*

$$\Lambda'(P) - G \leq 0$$

then the supremum attained.

Proof. Pick $f(S) = \operatorname{tr}(GS)$, $\boldsymbol{g}(S) = \Lambda(S)$ and $C = \mathcal{S}_+^{n+m}$ in Theorem 4.10. The condition (4.7) holds by Lemma 6.3 and $\Omega(\mathcal{S}_+^{n+m})$ is a convex cone because the mappings f and $\boldsymbol{g}$ are linear. If there exists $P \in \operatorname{dom} \Psi_D$ satisfying $\Lambda'(P) - G \leq 0$ then (4.9) is fulfilled and the supremum is attained. ∎

5.3 Extended Version

We are now ready to prove a more general version of the celebrated Kalman-Yakubovich-Popov Lemma.

Theorem 5.7. [[1], Theorem 3] *Let $M, N \in \mathbb{R}^{n\times(n+m)}$ and $\Theta \in \mathcal{S}^2$ such that $\det\Theta < 0$. Suppose that that (5.16) is feasible. Then for any $G \in \mathcal{S}^{n+m}$ the following statements are equivalent:*

(1) (Lur'e equation) There exist $P \in \mathcal{S}^n$ and $H \in \mathbb{R}^{m\times(n+m)}$ such that

$$\Lambda'(P) - G = -H^T H \tag{5.21}$$

(2) (Corresponding LMI) There exists $P \in \mathcal{S}^n$ such that

$$\Lambda'(P) - G \leq 0$$

(3) (Frequency condition) For every $\boldsymbol{z} \in \mathbb{C}^{n+m}$ for which there exists a $\lambda \in \Gamma$ such that

$$(\lambda N - M)\boldsymbol{z} = 0$$

the following inequality holds:

$$\boldsymbol{z}^* G \boldsymbol{z} \geq 0$$

(4) For every $\boldsymbol{z} \in \mathbb{C}^{n+m}$ such that $\Lambda(\boldsymbol{z}\boldsymbol{z}^*) = 0$ we have

$$\boldsymbol{z}^* G \boldsymbol{z} \geq 0$$

(5) For all $S \in \mathcal{S}_+^{n+m}$ that satisfy $\Lambda(S) = 0$ we have

$$\operatorname{tr}(GS) \geq 0$$

(6) There exist $Q \in \mathcal{S}^n$ such that

$$\inf_{\substack{S\in\mathcal{S}_+^{n+m}\\ \Lambda(S)=Q}} \operatorname{tr}(GS) > -\infty$$

(7) If D is a bounded closed convex nonempty set in $\mathcal{S}^n$ then the following duality result holds:

$$\inf_{\substack{S\in\mathcal{S}_+^{n+m}\\ \Lambda(S)\in D}} \operatorname{tr}(GS) = \sup_{\substack{P\in\mathcal{S}^n\\ \Lambda'(P)-G\leq 0}} \Psi_D(P) \tag{5.22}$$

where the supremum is attained.

Proof. We shall show the implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (4) $\Rightarrow$ (5) $\Rightarrow$ (6) $\Rightarrow$ (7) $\Rightarrow$ (1).

The implication (1) $\Rightarrow$ (2) is self-evident.

To show (2) $\Rightarrow$ (3), suppose that (2) holds and pick $\boldsymbol{z} \in \mathbb{C}^{n+m}$ for which the equality $\lambda N\boldsymbol{z} = M\boldsymbol{z}$ holds for some $\lambda \in \Gamma$. Then we get

$$\begin{aligned}\Lambda(\boldsymbol{z}\boldsymbol{z}^*) &= \begin{pmatrix} M & N\end{pmatrix}(\Theta \otimes \boldsymbol{z}\boldsymbol{z}^*)\begin{pmatrix} M & N\end{pmatrix}^T = \begin{pmatrix} M\boldsymbol{z} & N\boldsymbol{z}\end{pmatrix}\Theta\begin{pmatrix} M\boldsymbol{z} & N\boldsymbol{z}\end{pmatrix}^* \\ &= N\boldsymbol{z}\begin{pmatrix}\lambda & 1\end{pmatrix}\Theta\begin{pmatrix}\lambda & 1\end{pmatrix}^*(N\boldsymbol{z})^* = 0\end{aligned}$$

The last equality follows from the definition of Γ, see equation (5.12). Using the above, we get

$$\boldsymbol{z}^*G\boldsymbol{z} = \operatorname{tr}(G\boldsymbol{z}\boldsymbol{z}^*) = \operatorname{tr}(G\boldsymbol{z}\boldsymbol{z}^*) - \overbrace{\operatorname{tr} P\Lambda(\boldsymbol{z}\boldsymbol{z}^*)}^{=0} = \operatorname{tr}(G\boldsymbol{z}\boldsymbol{z}^* - P\Lambda(\boldsymbol{z}\boldsymbol{z}^*))$$

where $P \in \mathcal{S}^n$ is a matrix satisfying (2), that is $G - \Lambda'(P) \geq 0$. We then obtain

$$\boldsymbol{z}^*G\boldsymbol{z} = \operatorname{tr}(G\boldsymbol{z}\boldsymbol{z}^* - \Lambda'(P)\boldsymbol{z}\boldsymbol{z}^*) = \operatorname{tr}((G - \Lambda'(P))\boldsymbol{z}\boldsymbol{z}^*) = \boldsymbol{z}^*(G - \Lambda'(P))\boldsymbol{z} \geq 0$$

as desired.

(3) $\Rightarrow$ (4) follows directly from Lemma 5.5 (b).

We shall now prove the implication (4) $\Rightarrow$ (5). Suppose that $P \in \mathcal{H}_+^{n+m}$ fulfills $\Lambda(P) = 0$. By Lemma 5.5 (a) we can express P as a sum

$$P = \sum_{i=1}^{n+m} \boldsymbol{z}_i\boldsymbol{z}_i^*$$

where $\boldsymbol{z}_i \in \mathbb{C}^{n+m}$, $i = 1, 2, \ldots, n+m$ solve the equation $\Lambda(\boldsymbol{z}\boldsymbol{z}^*) = 0$. Using general rules for trace and the condition (4) we get

$$\operatorname{tr}(GS) = \operatorname{tr}\left(G\sum_{i=1}^{n+m} \boldsymbol{z}_i\boldsymbol{z}_i^*\right) = \sum_{i=1}^{n+m} \operatorname{tr}(G\boldsymbol{z}_i\boldsymbol{z}_i^*) = \sum_{i=1}^{n+m} \boldsymbol{z}_i^*G\boldsymbol{z}_i \geq 0$$

This completes the proof of the implication.

(5) $\Rightarrow$ (6) is trivially true: it suffices to choose $Q = 0$.

To show (6) $\Rightarrow$ (7), suppose that (6) holds for some $Q \in \mathcal{S}^n$. The duality relation (5.22) follows directly from Theorem 5.6 and the fact that $\operatorname{dom}\Psi_D = \mathcal{S}^n$ since D is bounded. It remains to show that the supremum is attained.

Consider the set $\{Q\}$. We have

$$\Psi_{\{Q\}}(P) = \inf_{S=Q} \operatorname{tr}(SP) = \operatorname{tr}(QP)$$

Notice also that $\operatorname{dom}\Psi_{\{Q\}} = \mathcal{S}^n$. It now follows from Theorem 5.6 and the assertion (6) that

$$-\infty < \inf_{\substack{S \in \mathcal{S}_+^{n+m} \\ \Lambda(S)=Q}} \operatorname{tr}(GS) = \sup_{\substack{P \in \mathcal{S}^n \\ \Lambda'(P)-G \leq 0}} \operatorname{tr}(QP)$$

In the above we see that there must exist a symmetric P that satisfies $\Lambda'(P) - G \leq 0$. Again, by Theorem 5.6 this implies that the supremum is attained.

It remains to show that (7) $\Rightarrow$ (1) holds. As was seen above, the condition (7) implies that there exists a P satisfying $\Lambda'(P) - G \leq 0$. This in turn implies (1). ■

Remark. There is also a similar result involving only strict inequalities. ◊

Remark. We have taken away some components from the original formulation of condition (7). The reason is that proving the complete version requires using results from the strict version of the lemma. As the components are consequences of the given duality relation, it is really not necessary to include them.

The missing items state that there exist $P^{\pm} \in \mathcal{S}^n$ such that:

$$\arg\max\{\operatorname{tr}(P) \mid P \in \mathcal{S}^n,\ \Lambda'(P) - G \leq 0\} = \{P^+\}$$
$$\arg\min\{\operatorname{tr}(P) \mid P \in \mathcal{S}^n,\ \Lambda'(P) - G \leq 0\} = \{P^-\}$$

Also, any matrix $P \in \mathcal{S}^n$ that satisfies the condition $\Lambda'(P) - G \leq 0$ also fulfills

$$P^- \leq P \leq P^+$$

Furthermore, there exists $H^{\pm} \in \mathbb{R}^{m \times (n+m)}$ such that the pairs (P^+, H^+) and (P^-, H^-) satisfy the Lur'e equation (5.21). $\Diamond$

6 Appendix: Proofs

Transformation

In this section we state and prove a proposition that shall simplify the proofs of Lemmas 5.4 and 5.5 significantly. We show that there always exists a transformation that allows us to express Θ and N in the special form (5.17).

Consider a transformation $\mathcal{T}$ determined by two nonsingular matrices $\Pi \in \mathbb{R}^{2\times 2}$ and $T \in \mathbb{R}^{(n+m)\times(n+m)}$ such that

$$\begin{aligned} \Theta_{\mathcal{T}} &= \Pi^{-1}\Theta(\Pi^{-1})^T \\ \begin{pmatrix} M_{\mathcal{T}} & N_{\mathcal{T}} \end{pmatrix} &= \begin{pmatrix} M & N \end{pmatrix} (\Pi \otimes I_{n+m})(I_2 \otimes T^{-1}) \end{aligned} \tag{6.1}$$

Furthermore, a vector $\boldsymbol{z} \in \mathbb{C}^{n+m}$ and a matrix $F \in \mathbb{R}^{(n+m)\times n}$ become

$$\begin{aligned} \boldsymbol{z}_{\mathcal{T}} &= T\boldsymbol{z} \\ F_{\mathcal{T}} &= TF \end{aligned} \tag{6.2}$$

The matrices G and S in $\mathcal{S}^{n+m}$ that define the function Λ are be given by

$$\begin{aligned} G_{\mathcal{T}} &= (T^{-1})^T G T^{-1} \\ S_{\mathcal{T}} &= TST^T \end{aligned} \tag{6.3}$$

and the function Λ itself by

$$\Lambda_{\mathcal{T}}(P) = \begin{pmatrix} M_{\mathcal{T}} & N_{\mathcal{T}} \end{pmatrix} (\Theta_{\mathcal{T}} \otimes P) \begin{pmatrix} M_{\mathcal{T}}^T \\ N_{\mathcal{T}}^T \end{pmatrix} \tag{6.4}$$

Set $\rho(\nu,\mu) = \begin{pmatrix} \nu & \mu \end{pmatrix} \Theta \begin{pmatrix} \nu & \mu \end{pmatrix}^*$. This function becomes:

$$\begin{aligned} \rho_{\mathcal{T}}(\nu,\mu) &= \begin{pmatrix} \nu & \mu \end{pmatrix} \Theta_{\mathcal{T}} \begin{pmatrix} \nu & \mu \end{pmatrix}^* \\ \begin{pmatrix} \nu_{\mathcal{T}} & \mu_{\mathcal{T}} \end{pmatrix} &= \begin{pmatrix} \nu & \mu \end{pmatrix} \Pi \end{aligned} \tag{6.5}$$

The above conditions lead to the following relations:

$$\operatorname{tr}(G_{\mathcal{T}} S_{\mathcal{T}}) = \operatorname{tr}(GS) \tag{6.6a}$$

$$\Lambda_{\mathcal{T}}(S_{\mathcal{T}}) = \Lambda(S) \tag{6.6b}$$

$$\rho_{\mathcal{T}}(\nu_{\mathcal{T}}, \mu_{\mathcal{T}}) = \rho(\nu,\mu) \tag{6.6c}$$

We shall need the following two lemmas:

Lemma 6.1. *For any $F \in \mathbb{C}^{(n+m)\times n}$ the following two conditions are equivalent:*

(a) $\det(NF) \neq 0$ *and* $\operatorname{Sp}[(NF)^{-1}MF] \subset \Omega_{\Theta}^{\pm}$

(b) $\det(NF) \neq 0$ *and* $\det(\nu NF - \mu MF) \neq 0$ *for all* $\pm\rho(\nu, \mu) \leq 0$.

Lemma 6.2. *Let* $F \in \mathbb{R}^{(n+m)\times n}$. *If* $F_{\mathcal{T}} = TF$ *then*

$$\nu_{\mathcal{T}} N_{\mathcal{T}} F_{\mathcal{T}} - \mu_{\mathcal{T}} M_{\mathcal{T}} F_{\mathcal{T}} = \det(\Pi)(\nu NF - \mu MF)$$

Similar result holds for any vector $\boldsymbol{z} \in \mathbb{C}^{n+m}$.

Proof. Remember that it is generally true that

$$(A \otimes B)(C \otimes D) = AC \otimes BD$$

if the matrices involved have suitable dimensions. Notice also that the following equality holds:

$$\begin{pmatrix} -\mu_{\mathcal{T}} \\ \nu_{\mathcal{T}} \end{pmatrix} = \det(\Pi)\Pi^{-1} \begin{pmatrix} -\mu \\ \nu \end{pmatrix}$$

We now get

$$\begin{aligned}
\nu_{\mathcal{T}} N_{\mathcal{T}} F_{\mathcal{T}} - \mu_{\mathcal{T}} M_{\mathcal{T}} F_{\mathcal{T}} &= \begin{pmatrix} M_{\mathcal{T}} & N_{\mathcal{T}} \end{pmatrix} \left(\begin{pmatrix} -\mu_{\mathcal{T}} \\ \nu_{\mathcal{T}} \end{pmatrix} \otimes I_{n+m} \right) F_{\mathcal{T}} = \\
&= \det(\Pi) \begin{pmatrix} M & N \end{pmatrix} (\Pi \otimes I_{n+m})(I_2 \otimes T^{-1}) \left(\Pi^{-1} \begin{pmatrix} -\mu \\ \nu \end{pmatrix} \otimes I_{n+m} \right) TF = \\
&= \det(\Pi) \begin{pmatrix} M & N \end{pmatrix} (\Pi \otimes T^{-1}) \left(\Pi^{-1} \begin{pmatrix} -\mu \\ \nu \end{pmatrix} \otimes I_{n+m} \right) TF = \\
&= \det(\Pi) \begin{pmatrix} M & N \end{pmatrix} \left(\Pi\Pi^{-1} \begin{pmatrix} -\mu \\ \nu \end{pmatrix} \otimes T^{-1} \right) TF = \\
&= \det(\Pi)(\nu NF - \mu MF)
\end{aligned}$$

The proof for $\boldsymbol{z}$ is almost identical. ■

Proposition 6.3. *If* (5.16) *is feasible, there exists a transformation* $\mathcal{T}$ *determined by two non-singular matrices* $\Pi \in \mathbb{R}^{2\times 2}$ *and* $T \in \mathbb{R}^{(n+m)\times(n+m)}$ *that satisfy the conditions* (6.1) – (6.5) *such that we can express* Θ *and* N *in the form*

$$\begin{aligned}
\Theta_{\mathcal{T}} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
N_{\mathcal{T}} &= \begin{pmatrix} I_n & 0 \end{pmatrix}
\end{aligned} \tag{6.7}$$

Furthermore, the matrices $\Theta_{\mathcal{T}}$, $M_{\mathcal{T}}$ *and* $N_{\mathcal{T}}$ *satisfy the condition* (5.16).

Proof. We can always write

$$\Theta = \Pi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Pi^T \tag{6.8}$$

for some invertible $\Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$. This gives us the desired transformation for Θ:

$$\Theta_{\mathcal{T}} = \Pi^{-1}\Theta(\Pi^{-1})^T = \Pi^{-1}\Pi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \Pi^T(\Pi^{-1})^T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

We shall now construct two transformations $\mathcal{T}_1$ and $\mathcal{T}_2$ and determine $\mathcal{T}$ as a combination of these. The variables associated with transformation $\mathcal{T}_1$ shall have subscript 1, and those associated with $\mathcal{T}_2$ will be marked by 2.

Now consider $\mathcal{T}_1$ given by $\Pi_1 = \Pi$ and $T_1 = I_{n+m}$. As we saw above, we get $\Theta_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and by (6.1) we gain

$$\begin{aligned} M_1 &= (\pi_{21} N + \pi_{11} M) \\ N_1 &= (\pi_{22} N + \pi_{12} M) \end{aligned}$$

Furthermore, from (6.5) we get the relations

$$\begin{aligned} \begin{pmatrix} \nu_1 & \mu_1 \end{pmatrix} &= \begin{pmatrix} \nu & \mu \end{pmatrix} \Pi \\ \rho_1(\nu, \mu) &= \begin{pmatrix} \nu & \mu \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \nu^* \\ \mu^* \end{pmatrix} = 2\,\mathrm{Re}(\nu \mu^*) \end{aligned}$$

Let $F^\pm$ be solutions to (5.16). Define $H^\pm \in \mathbb{R}^{(n+m)\times n}$ and $K^\pm \in \mathbb{R}^{(n+m)\times m}$ by

$$\begin{pmatrix} N \\ F^\pm \end{pmatrix}^{-1} = \begin{pmatrix} H^\pm & K^\pm \end{pmatrix}$$

It follows from

$$\begin{aligned} \begin{pmatrix} N \\ F^\pm \end{pmatrix} \begin{pmatrix} N \\ F^\pm \end{pmatrix}^{-1} &= \begin{pmatrix} N \\ F^\pm \end{pmatrix} \begin{pmatrix} H^\pm & K^\pm \end{pmatrix} = \\ &= \begin{pmatrix} NH^\pm & NK^\pm \\ F^\pm H^\pm & F^\pm K^\pm \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ 0 & I_m \end{pmatrix} \end{aligned} \tag{6.9}$$

that $NH^\pm = I_n$. This, together with the fact that

$$\mathrm{Sp}\left(M \begin{pmatrix} N \\ F^\pm \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix} \right) = \mathrm{Sp}(MH^\pm) \subset \Omega_\Theta^\pm$$

allows us to use Lemma 6.1. Hence

$$\det(\nu N H^\pm - \mu M H^\pm) \neq 0 \qquad \text{for all } \pm\rho(\nu, \mu) \leq 0 \tag{6.10}$$

Using (6.8) we get that $\rho(\pi_{22}, -\pi_{12}) = 0$ and hence

$$\det((\pi_{22} N + \pi_{12} M) H^\pm) = \det(N_1 H^\pm) \neq 0 \tag{6.11}$$

We also get from the equation (6.9) that

$$\begin{pmatrix} N_1 \\ F^\pm \end{pmatrix} \begin{pmatrix} H^\pm & K^\pm \end{pmatrix} = \begin{pmatrix} N_1 H^\pm & N_1 K^\pm \\ 0 & I_m \end{pmatrix} \tag{6.12}$$

Equation (6.11) implies that the right-hand side of (6.12) is nonsingular and hence we obtain $\det \begin{pmatrix} N_1 \\ F^\pm \end{pmatrix} \neq 0$. We can therefore define matrices $H_1^\pm \in \mathbb{R}^{(n+m)\times n}$ and $K_1^\pm \in \mathbb{R}^{(n+m)\times m}$ by the relations

$$\begin{pmatrix} N_1 \\ F^\pm \end{pmatrix}^{-1} = \begin{pmatrix} H_1^\pm & K_1^\pm \end{pmatrix}$$

Combining the above with (6.12) yields $H_1^\pm = H^\pm (N_1 H^\pm)^{-1}$. Keeping in mind that $T = I_{n+m}$ we now get from Lemma 6.2 that

$$\det(\nu_1 N_1 H_1^\pm - \mu_1 M_1 H_1^\pm) = \det(\nu N H^\pm - \mu M H^\pm) \det(\Pi) \det(N_1 H^\pm)^{-1}$$

We already know that the determinants of Π and $(N_1H^{\pm})^{-1}$ are nonzero. From(6.10) and (6.6c) we get that $\det(\nu_1 N_1 H_1^{\pm} - \mu_1 M_1 H_1^{\pm}) \neq 0$ whenever $\pm\rho_1(\nu_1, \mu_1) \leq 0$. Remember that Γ is simply the imaginary axis whenever Θ is of the standard form (5.17). Since $N_1 H_1^{\pm} = I_n$ in the same way as in (6.9) it follows from Lemma 6.1 that

$$\mathrm{Sp}(M_1 F_1^{\pm}) \subset \mathbb{C}^{\pm} \tag{6.13}$$

Let now $\mathcal{T}_2$ be given by $\Pi_2 = I_2$ and $T_2 = \begin{pmatrix} N_1 \\ F^- \end{pmatrix}$ and consider the transformation $\mathcal{T}$ determined by $\Theta = \Theta_1\Theta_2$ and $T = T_1T_2$. Then $\Theta_{\mathcal{T}}$ and $N_{\mathcal{T}}$ are of the desired form (6.7). Also, we get $M_{\mathcal{T}} = M_1T^{-1}$. It remains to show that (5.16) holds for these matrices.

Consider $F_{\mathcal{T}}^{\pm} = F^{\pm}T^{-1}$. Then

$$\det \begin{pmatrix} N_{\mathcal{T}} \\ F_{\mathcal{T}}^{\pm} \end{pmatrix} = \det \begin{pmatrix} N_1 \\ F^{\pm} \end{pmatrix} \det(T^{-1}) \neq 0$$

Let us choose

$$H_{\mathcal{T}}^{\pm} = \begin{pmatrix} N_{\mathcal{T}} \\ F_{\mathcal{T}}^{\pm} \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix}$$

We the get $M_{\mathcal{T}}H_{\mathcal{T}}^{\pm} = M_1H_1^{\pm}$ and the result follows from (6.13). ■

Lemma 5.4

Let us first note that it follows from the relation (6.6b) that

$$\Lambda(S) = \Lambda_{\mathcal{T}}(S_{\mathcal{T}}) = N_{\mathcal{T}}S_{\mathcal{T}}M_{\mathcal{T}}^T + M_{\mathcal{T}}S_{\mathcal{T}}N_{\mathcal{T}}^T$$

A matrix plus its Hermitian conjugate is a real symmetric matrix, from which it follows that the right-hand side is real symmetric. Hence the choice of codomain to Λ is justified.

Let $Q \in \mathcal{S}^n$ be given. By Lemma 6.3 and the relation (6.6b) we may restrict our attention to the special case (5.17). We shall show that there exists $S \in \mathcal{S}_+^{n+m}$ such that $\Lambda(S) = Q$. This shall prove the lemma.

Let $F^{\pm}$ be matrices that fulfill the condition (5.16). If we set

$$H^{\pm} = \begin{pmatrix} N \\ F^{\pm} \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix}$$

then the condition (5.16) implies that MH^- is Hurwitz and MH^+ antihurwitz. Notice that

$$NH^{\pm} = \begin{pmatrix} I_n & 0 \end{pmatrix} \begin{pmatrix} I_n & 0 \\ F^{\pm} \end{pmatrix}^{-1} \begin{pmatrix} I_n \\ 0 \end{pmatrix} = \begin{pmatrix} I_n & 0 \end{pmatrix} \begin{pmatrix} I_n & * \\ * & * \end{pmatrix} \begin{pmatrix} I_n \\ 0 \end{pmatrix} = I_n \tag{6.14}$$

Since Q is symmetric we may express it as a sum $Q = Q^+ + Q^-$ where $Q^+ \geq 0$ and $Q^- \leq 0$. As seen in Example 4.7 and the discussion following it, the Lyapunov equations

$$(MH^{\pm})P + P(MH^{\pm})^* = Q^{\pm} \tag{6.15}$$

have solutions $P^{\pm}$ in $\mathcal{S}_+^n$. Let $S = H^+P^+(H^+)^* + H^-P^-(H^-)^* \in \mathcal{S}_+^{n+m}$. The result follows by inserting S in (5.20) and using relations (6.14) and (6.15). The proof is now complete.

Lemma 5.5

In order to show the result we need the following two lemmas from [24]:

Lemma 6.4. *Let N and M be complex matrices of the same size. Then $FG^* + GF^* = 0$ if and only if there exists a matrix U such that $UU^* = I$ and $F(I + U) = G(I - U)$.*

Lemma 6.5. *Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{C}^n$ and $\boldsymbol{y} \neq \boldsymbol{0}$. Then $\boldsymbol{x}\boldsymbol{y}^* + \boldsymbol{y}\boldsymbol{x}^* = 0$ if and only if there exists $a \in \mathbb{R}$ such that $\boldsymbol{x} = ia\boldsymbol{y}$.*

To show that (a) holds, suppose that S solves the equation $\Lambda(S) = 0$. By Lemma 6.3 and the relation (6.6b) we can again restrict our attention to the special case (5.17). We get from (5.20) that

$$\Lambda(P) = NPM^T + MPN^T = NP^{1/2}(MP^{1/2})^T + MP^{1/2}(NP^{1/2})^T = 0$$

By Lemma 6.4 there exists a matrix U such that

$$NP^{1/2}(I + U) = MP^{1/2}(I - U) \tag{6.16}$$

The condition $UU^* = I$ means that U is unitary, and hence there exists $\theta_j \in \mathbb{R}$, $\boldsymbol{u}_j \in \mathbb{C}^{n+m}$, $j = 1, 2, \ldots, n + m$ such that $\sum_{j=1}^{n+m} \boldsymbol{u}_j \boldsymbol{u}_j^* = I$ and $U = \sum_{j=1}^{n+m} e^{i\theta_j} \boldsymbol{u}_j \boldsymbol{u}_j^*$. If we set $\boldsymbol{z}_j = P^{1/2}\boldsymbol{u}_j$ then we obtain $\sum_{j=1}^{n+m} \boldsymbol{z}_j \boldsymbol{z}_j^* = P$ as desired. Let us now show that

$$\Lambda(\boldsymbol{z}_j \boldsymbol{z}_j^*) = N\boldsymbol{z}_j (M\boldsymbol{z}_j)^* + M\boldsymbol{z}_j (N\boldsymbol{z}_j)^* = 0$$

We have

$$\begin{aligned} N\boldsymbol{z}_j(1 + e^{i\theta_j}) = NP^{1/2}\boldsymbol{u}_j(1 + e^{i\theta_j}) = NP^{1/2}(I + U)\boldsymbol{u}_j = \\ = MP^{1/2}(I - U)\boldsymbol{u}_j = M\boldsymbol{z}_j(1 - e^{i\theta_j}) \end{aligned}$$

Applying Lemma 6.4 once again gives the desired result. We have now proved the first part of our lemma.

To show (b) it is not sufficient to consider the case (5.17). The reason is that whether or not A_2 is compact depends on the nature of Γ. Let $\mathcal{T}$ be a transformation as in Lemma 2.9 and suppose that $\boldsymbol{z} \in \mathbb{C}^{n+m} \in A_1$.

Consider the point $\boldsymbol{z}_\mathcal{T} = T\boldsymbol{z}$. We shall show that the equation

$$(\nu_\mathcal{T} N_\mathcal{T} - \mu_\mathcal{T} M_\mathcal{T})\, \boldsymbol{z}_\mathcal{T} = 0 \tag{6.17}$$

is satisfied for some $\mu_\mathcal{T}, \nu_\mathcal{T} \in \mathbb{C}$ such that

$$\begin{aligned} &|\mu_\mathcal{T}| + |\nu_\mathcal{T}| \neq 0 \\ &\rho_\mathcal{T}(\nu_\mathcal{T}, \mu_\mathcal{T}) = 2\,\mathrm{Re}(\mu_\mathcal{T}\nu_\mathcal{T}^*) = 0 \end{aligned} \tag{6.18}$$

If $N_\mathcal{T}\boldsymbol{z}_\mathcal{T} = \boldsymbol{0}$, then we may choose $\mu_\mathcal{T} = 0$ and $\nu_\mathcal{T} = i$. Suppose that $N_\mathcal{T}\boldsymbol{z}_\mathcal{T} \neq \boldsymbol{0}$. We have

$$0 = \Lambda(\boldsymbol{z}\boldsymbol{z}^*) = \Lambda_\mathcal{T}(\boldsymbol{z}_\mathcal{T}\boldsymbol{z}_\mathcal{T}^*) = N_\mathcal{T}\boldsymbol{z}_\mathcal{T}(M\boldsymbol{z}_\mathcal{T})^* + M_\mathcal{T}\boldsymbol{z}_\mathcal{T}(N_\mathcal{T}\boldsymbol{z}_\mathcal{T})^* = 0$$

Set $\boldsymbol{x} = N_\mathcal{T}\boldsymbol{z}_\mathcal{T}$ and $\boldsymbol{y} = M_\mathcal{T}\boldsymbol{z}_\mathcal{T}$. By Lemma 6.5 there exist an $a \in \mathbb{R}$ such that $\mathrm{Re}(a) \geq 0$ and $M\boldsymbol{z} = iaN\boldsymbol{z}$. Hence the points $\mu_\mathcal{T} = 1$ and $\nu_\mathcal{T} = ai$ satisfy (6.17) and the conditions (6.18).

By the Lemma 6.2 we get from (6.17) that

$$(\nu N - \mu M)\, \boldsymbol{z} = 0 \tag{6.19}$$

Here μ and ν are determined by the relation (6.5). The relations (6.18) translate directly to

$$\begin{aligned} |\mu| + |\nu| &\neq 0 \\ \rho\,(\nu, \mu) &= 0 \end{aligned} \tag{6.20}$$

Pick any ν, μ satisfying the equations (6.19) and (6.20). We shall now show that $\boldsymbol{z} \in \operatorname{cl} A_2$. There are two cases to consider: (1) $\mu \neq 0$ and (2) $\mu = 0$.

(1) If $\mu \neq 0$ then the result follows by inserting $\lambda = \nu/\mu$ in (6.19). The condition $\rho\,(\nu, \mu) = 0$ implies that $\rho\,(\lambda, 1) = \varphi(\lambda) = 0$ and hence $\lambda \in \Gamma$. This means that $\boldsymbol{z} \in A_2$.

(2) Suppose that $\mu = 0$. We get from (6.19) that

$$\rho\,(\nu, 0) = \begin{pmatrix} \nu & 0 \end{pmatrix} \Theta \begin{pmatrix} \nu^* \\ 0 \end{pmatrix} = \begin{pmatrix} \nu & 0 \end{pmatrix} \begin{pmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{pmatrix} \begin{pmatrix} \nu^* \\ 0 \end{pmatrix} = \nu\nu^*\pi_{11} = 0$$

From this it follows that $\theta_{11} = 0$, otherwise we would have $|\mu| + |\nu| = 0$, a contradiction. As mentioned earlier, this implies that Γ is a line in $\mathbb{C}$ and not a circle. Hence we can find a sequence $\{\lambda_i\}_{i=1}^{\infty} \subset \Gamma$ such that $\lim_{i\to\infty} \lambda_i$ is $+\infty$ or $-\infty$.

We shall now construct a sequence $\{\boldsymbol{z}_i\}_{i=1}^{\infty} \subset \mathbb{C}^{n+m}$ in A_2 that converges to $\boldsymbol{z}$. Let F^- satisfy (5.16) and consider

$$H^- = \begin{pmatrix} N \\ F^- \end{pmatrix} \begin{pmatrix} I_n \\ 0 \end{pmatrix}$$

In the same way as in (6.14) we get $NH^- = I_n$. From (5.16) we also obtain

$$\det(\lambda_i I_n - MH^-) \neq 0, \qquad i = 1, 2, \ldots$$

Let us now construct the sequence $\{\boldsymbol{z}_i\}_{i=1}^{\infty}$ by setting:

$$\boldsymbol{z}_i = \boldsymbol{z} + H^-(\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z}$$

Then we get

$$\begin{aligned} \lambda_i N\boldsymbol{z}_i - M\boldsymbol{z}_i &= \lambda_i I_n(\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z} - M\boldsymbol{z} - MH^-(\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z} = \\ &= (\lambda_i I_n - MH^-)(\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z} - M\boldsymbol{z} = M\boldsymbol{z} - M\boldsymbol{z} = 0 \end{aligned}$$

Hence each $\boldsymbol{z}_i$ belongs to A_2. The result follows by noting that

$$\lim_{i\to\infty} \boldsymbol{z}_i = \boldsymbol{z} + H^- \lim_{i\to\infty} (\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z} = \boldsymbol{z}$$

since $\lim_{i\to\infty}(\lambda_i I_n - MH^-)^{-1} M\boldsymbol{z} = 0$. The proof is now complete.

References

[1] S. V. Gusev, "The Fenchel Duality, S-Procedure, and the Yakubovich-Kalman Lemma," *Automation and Remote Control*, vol. 67, no. 2, pp. 293–310, 2006.

[2] R. T. Rockafellar, *Convex Analysis*. Princeton University Press, 1970.

[3] M. S. Bazaraa, H. D. Sherali, and C. Shetty, *Nonlinear Programming*. Wiley-Interscience, 2006.

[4] I. Pólik and T. Terlaky, "A Survey of the S-Lemma," *Society for Industrial and Applied Mathematics*, vol. 49, no. 3, pp. 371–418, 2007.

[5] A. L. Fradkov and V. A. Yakubovich, "The S-Procedure and Duality Relations in certain Quadratic Programming Problems," *Vestn. Leningr. Gos. Univ.*, no. 1, pp. 71–76, 1973.

[6] D. P. Bertsekas, *Convex Analysis and Optimization*. Athena Scientific, 2003.

[7] R. T. Rockafellar, *Conjugate Duality and Optimization*. SIAM, 1974.

[8] T. Magnanti, "Fenchel and Lagrange Duality Are Equivalent," *Mathematical Programming*, vol. 7, pp. 253–258, 1974.

[9] R. T. Rockafellar, "Duality and Optimality in Multistage Stochastic Programming," *Annals of Operations Research*, vol. 85, pp. 1–19, 1999.

[10] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*. Springer, 2000.

[11] K. Murota, *Discrete Convex Analysis*. SIAM, 2003.

[12] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[13] E. D. Sontag, *Mathematical Control Theory*. Springer, 1990.

[14] A. J. Laub, *Matrix Analysis for Scientists and Engineers*. SIAM, 2005.

[15] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*. Prentice Hall, 1996.

[16] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, 1996.

[17] U. T. Jönsson, "A Lecture on the S-Procedure," 2006.

[18] P. Finsler, "Über das Vorkommen definiter und semidefiniter Formen in Scharen quadratischer Formen," *Commentarii Mathematici Helvetici*, vol. 9, no. 1, pp. 188–192, 2006.

[19] L. L. Dines, "Om the Mappings of Quadratic Forms," *Bull. Amer. Math. Soc.*, vol. 49, pp. 494–498, 1941.

[20] A. L. Fradkov, "Duality Theorems for Certain Nonconvex Extremal Problems," *Sib. Math. Zh*, vol. 14, no. 2, pp. 355–383, 1973.

[21] S. V. Gusev and A. L. Likhtarnikov, "Kalman-Popov-Yakubovich Lemma and the *S*-Procedure: A Historical Essay," *Automation and Remote Control*, vol. 67, no. 11, pp. 1768–1810, 2006.

[22] B. Brogliato, *Dissipative Systems Analysis and Control*. Springer, 2007.

[23] J. C. Willems, "Least Squares Stationary Optimal Cotrol and the Algebraic Riccatti Equation," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 621–634, 1971.

[24] A. Rantzer, "On the Kalman-Yakubovich-Popov Lemma," *Systems & Control Letters*, vol. 28, pp. 7–10, 1996.

[25] A. N. Churilov, "Solubility of Matrix Inequalities," *Mathematical notes of the Academy of Sciences of the USSR*, vol. 36, no. 5, pp. 862–866, 1984.

[26] R. E. Kalman, "Lyapunov Functions for the Problem of Lur'e in Automatic Control," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 143, no. 6, pp. 201–205, 1963.