On the Kalman-Yakubovich-Popov Lemma for Positive Systems

Rantzer, Anders

Published in:
Proceedings of 51st IEEE Conference on Decision and Control

2012

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
On the Kalman-Yakubovich-Popov Lemma for Positive Systems

Anders Rantzer

Abstract—The classical Kalman-Yakubovich-Popov lemma gives conditions for solvability of a certain inequality in terms of a symmetric matrix. The lemma has numerous applications in systems theory and control. Recently, it has been shown that for positive systems, important versions of the lemma can equivalently be stated in terms of a diagonal matrix rather than a general symmetric one. This paper generalizes these results and a new proof is given.

It was shown in [7] that the Kalman-Yakubovich-Popov lemma [8], [5], [3] (also known as the bounded real lemma) can be considerably simplified for “internally positive” systems. In particular, the matrix inequality can be restricted to diagonal matrices. This result enabled the authors of [7] to design decentralized control laws by convex optimization.

Earlier this year, a discrete time version was proved by [4] with weaker assumptions. The purpose of this note is to generalize and unify these results and to give a new proof. After some notation, we state the result both for continuous time and discrete time:

The matrix $A \in \mathbb{R}^{n \times n}$ is said to be Hurwitz if all eigenvalues have negative real part. It is Schur if all eigenvalues are strictly inside the unit circle. Finally, a matrix is said to be Metzler if all off-diagonal elements are non-negative. The expression $\|x\|^2_P$ is short for $x^TPx$.

**Theorem 1:** Let $A \in \mathbb{R}^{n \times n}$ be Metzler and Hurwitz, while $B \in \mathbb{R}^{n \times m}$ and the pair $(A,B)$ is controllable. Suppose that all entries of $Q \in \mathbb{R}^{(n+m) \times (n+m)}$ are non-negative, except for the last $m$ diagonal elements. Then the following statements are equivalent:

1. For $\omega \in [0, \infty)$ is is true that
   
   $$
   \begin{bmatrix}
   (\omega I - A)^{-1}B \\
   I
   \end{bmatrix}^* \begin{bmatrix}
   (\omega I - A)^{-1}B \\
   I
   \end{bmatrix} \preceq 0
   $$

2. There exist a diagonal $P \succeq 0$ such that
   
   $$
   Q + \begin{bmatrix}
   A^TP + PA & PB \\
   B^TP & 0
   \end{bmatrix} \succeq 0.
   $$

Moreover, if all inequalities are taken to be strict, then the equivalences hold even without the controllability assumption.

**Remark 1.** For $A = -1$, $B = 0$, $Q = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, condition (1.1) holds, but not (1.3). This demonstrates that the controllability assumption is essential when the inequalities are non-strict.

**Theorem 2:** Let $A \in \mathbb{R}^{n \times n}$ be Schur, while $B \in \mathbb{R}^{n \times m}$ and the pair $(A,B)$ is controllable. Suppose that all entries of $Q \in \mathbb{R}^{(n+m) \times (n+m)}$ are non-negative, except for the last $m$ diagonal elements. Then the following statements are equivalent:

1. For $\omega \in [0, \infty)$ is is true that
   
   $$
   \begin{bmatrix}
   (\omega I - A)^{-1}B \\
   I
   \end{bmatrix}^* \begin{bmatrix}
   (\omega I - A)^{-1}B \\
   I
   \end{bmatrix} \preceq 0
   $$

2. There exist a diagonal $P \succeq 0$ such that
   
   $$
   Q + \begin{bmatrix}
   A^TPA - P & A^TBP \\
   B^TPA & B^TBP
   \end{bmatrix} \succeq 0
   $$

Moreover, if all inequalities are taken to be strict, then the equivalences hold even without the controllability assumption.

**Remark 2.** The results of [7] and [4] are recovered by the strict inequality versions of Theorem 1 and Theorem 2 with

$$
Q = \begin{bmatrix}
I & 0 \\
0 & -I
\end{bmatrix}
$$

The proofs of Theorem 1 and Theorem 2 will be based on the following extension of [1, Theorem 3.1]:

**Proposition 1 (Positive Quadratic Programming):**

Suppose $M_0, \ldots, M_K$ are Metzler and $b_1, \ldots, b_K \in \mathbb{R}$. Then

$$
\max_{x \in \mathbb{R}^n_+} x^T M_0 x = \max_{x \geq 0} \text{trace}(M_0 X)
$$

$$
x^T M_k x \geq b_k \quad X \geq 0 \quad \text{trace}(M_k X) \geq b_k 
$$

$$
k = 1, \ldots, K
$$

(1)

The value on the right hand side remains the same if the condition $X \geq 0$ is relaxed to $X \in \mathbb{X}$, where $\mathbb{X}$ is the set of symmetric matrices $(x_{ij}) \in \mathbb{R}^{n \times n}$ satisfying $x_{ii} \geq 0$ and $x_{ij}^2 \leq x_{ii} x_{jj}$ for all $i, j$. Moreover, if there exists a matrix $X$ in the interior of $\mathbb{X}$ with $\text{trace}(M_k X) \geq b_k$ for every $k$, then the maximum of (1) equals the minimum of $-\sum_k \tau_k b_k$ over $\tau_1, \ldots, \tau_K \geq 0$ such that $M_0 + \sum_k \tau_k M_k \preceq 0$. 

A. Rantzer is with Automatic Control LTH, Lund University, Box 118, SE-221 00 Lund, Sweden, rantzer at control.lth.se.
Remark 3. The problem on the right of (1) is always convex and readily solvable by semidefinite programming. The problem on the left is generally not a convex program, since the matrices $M_k$ may be indefinite. However, the maximization on the left is concave in $(x_1^r, \ldots, x_n^r)$ [2]. This is because every product $x_i x_j$ is the geometric mean of two such variables, hence concave.

A proof is included here for completeness:

Proof of Proposition 1. Every $x$ satisfying the constraints on the left hand side of (1) corresponds to a matrix $X = xx^T$ satisfying the constraints on the right hand side. This shows that the right hand side of (1) is at least as big as the left.

On the other hand, let $X = (x_{ij})$ be a positive definite matrix. In particular, the diagonal elements $x_{11}, \ldots, x_{nn}$ are non-negative and $x_{ij} \leq \sqrt{x_{ii} x_{jj}}$. Let $x = (\sqrt{x_{11}}, \ldots, \sqrt{x_{nn}})$. Then the matrix $xx^T$ has the same diagonal elements as $X$, but has off-diagonal elements $\sqrt{x_{ii} x_{jj}}$ instead of $x_{ij}$. The fact that $xx^T$ has off-diagonal elements at least as big as those of $X$, together with the assumption that the matrices $M_k$ are Metzler, gives $x^T M_k x \geq \text{trace}(M_k X)$ for $k = 1, \ldots, K$. This shows that the left hand side of (1) is at least as big as the right. Nothing changes if $X$ is not positive definite but $X \in \mathbb{X}$, so the second statement is also proved.

For the last statement, note that the conditions $\text{trace}(M_k X) \geq b_k$ are linear in $X$, so strong duality holds [6, Theorem 28.2] and the right hand side of (1) has a finite maximum if and only if $M_0 + \sum_{k=1}^K \tau_k M_k \succeq 0$ for some $\tau_1, \ldots, \tau_K \geq 0$.

Proof of Theorem 1. One at a time, we will prove the implications $(1.1) \Rightarrow (1.2) \Rightarrow (1.3) \Rightarrow (1.1)$. Putting $\omega = 0$ immediately gives (1.2) from (1.1).

Assume that (1.2) holds. The matrix $-A^{-1}$ is non-negative (because $A$ is Hurwitz and Metzler), so $x^T Q x \leq 0$ for all $x \in \mathbb{R}_+^n$, $w \in \mathbb{R}_+^m$ with

$$x \leq -A^{-1} B w$$

(2)

The inequality (4) follows (by multiplication with $-A^{-1}$ from the left) from the constraint $0 \leq Ax + Bw$, which can also be written $0 \leq A_i x + B_i w$ for $i = 1, \ldots, n$, where $A_i$ and $B_i$ denote the $i$th rows of $A$ and $B$ respectively. For non-negative $x$ and $w$, this is equivalent to

$$0 \leq x_i (A_i x + B_i w) \quad i = 1, \ldots, n$$

(3)

Hence (1.2) implies that $x^T Q x \leq 0$ for $x \in \mathbb{R}_+^n$, $w \in \mathbb{R}_+^m$ satisfying (5). Proposition 1 with $b_1 = \cdots = b_n = 0$ will next be used to verify existence of $\tau_1, \ldots, \tau_n \geq 0$ such that the quadratic form

$$\sigma(x, w) = x^T Q x + \sum_i \tau_i x_i (A_i x + B_i w)$$

is negative semi-definite. However, it remains to verify the “Slater condition”; existence of a positive definite $X$ such that all diagonal elements of

$$[A \ B] X \begin{bmatrix} I \\ 0 \end{bmatrix}$$

are non-negative. The pair $(A, B)$ is controllable, so there exists $K$ that make all eigenvalues of $A + BK$ unstable and therefore $(A + BK) Z + Z(A + BK)^T = I$ has a symmetric positive definite solution $Z$. Hence the desired $X$ can be constructed as

$$X = \begin{bmatrix} Z & ZK^T \\ KZ & * \end{bmatrix}$$

where the lower right corner is chosen big enough to make $X \succ 0$.

Define $P = \frac{1}{2} \text{diag}(\tau_1, \ldots, \tau_n) \geq 0$. Then $\sigma$ being negative definite means that

$$Q + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \preceq 0$$

so (1.3) follows.

Finally, assume that (1.3) holds. Integrating $\sigma(x(t), w(t))$ over time gives

$$0 \geq \int_0^\infty \left( x^T Q x + x^T P (Ax + Bw) \right) dt$$

For square integrable solutions to $\dot{x} = Ax + Bw$, $x(0) = 0$ we get

$$0 \geq \int_0^\infty \left( x^T Q x + \frac{d}{dt} (x^T P x/2) \right) dt$$

$$= \int_0^\infty \left( x(t)^T Q x(t) + \frac{d}{dt} (x(t)^T P x(t)/2) \right) dt$$

which in frequency domain implies (1.1). Hence (1.1) $\Rightarrow$ (1.2) $\Rightarrow$ (1.3) $\Rightarrow$ (1.1).

For strict inequalities, the proofs that (1.3) $\Rightarrow$ (1.1) $\Rightarrow$ (1.2) remain the same. Assuming that (1.2) holds with strict inequality, we get

$$\begin{bmatrix} -A^{-1} B^T \\ I \end{bmatrix}^* (Q + \epsilon I) \begin{bmatrix} -A^{-1} B^T \\ I \end{bmatrix} \leq 0$$

for some scalar $\epsilon > 0$. Hence, there exists a diagonal $P \succ 0$ such that

$$Q + \epsilon I + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \leq 0$$

Adding a small multiple of the identity to $P$ gives $P > 0$ such that

$$Q + \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} < 0$$

so also (1.3) holds with strict inequality. Hence the proof of Theorem 1 is complete. \qed
Proof of Theorem 2. In analogy with the previous proof, we will prove the implications (2.1)⇒(2.2)⇒(2.3)⇒(2.1). Putting \( \omega = 0 \) immediately gives (2.2) from (2.1).

Assume that (2.2) holds. The matrix \((I - A)^{-1}\) is non-negative (because \(A\) is non-negative and Schur), so
\[
\begin{bmatrix} x \n w \end{bmatrix}^T Q \begin{bmatrix} x \n w \end{bmatrix} \leq 0 \quad \text{for all} \quad x \in \mathbb{R}_+^n, \quad w \in \mathbb{R}_+^m
\]
with
\[
x \leq (I - A)^{-1} B w
\]  
(4)
The inequality (4) follows from \(x \leq Ax + Bw\), which can also be written \(x_i \leq A_i x + B_i w\) for \(i = 1, \ldots, n\), where \(A_i\) and \(B_i\) denote the \(i\)-th rows of \(A\) and \(B\) respectively. For non-negative \(x\) and \(w\), this is equivalent to
\[
x_i^2 \leq (A_i x + B_i w)^2 \quad i = 1, \ldots, n
\]  
(5)
Hence (2.2) implies that \(\begin{bmatrix} x \n w \end{bmatrix}^T Q \begin{bmatrix} x \n w \end{bmatrix} \leq 0\) for \(x \in \mathbb{R}_+^n, \quad w \in \mathbb{R}_+^m\) satisfying (5). Proposition 1 with \(b_1 = \cdots = b_n = 0\) will next be used to verify existence of \(\tau_1, \ldots, \tau_n \geq 0\) such that the quadratic form
\[
\sigma(x, w) = \begin{bmatrix} x \\
 w \end{bmatrix}^T Q \begin{bmatrix} x \\
 w \end{bmatrix} + \sum \tau_i[(A_i x + B_i w)^2 - x_i^2]
\]
is negative semi-definite. However, the application of Proposition 1 requires existence of a positive definite \(X\) such that all diagonal elements of
\[
\begin{bmatrix} A \n B \end{bmatrix} X \begin{bmatrix} A \n B \end{bmatrix}^T - \begin{bmatrix} I \n 0 \end{bmatrix} X \begin{bmatrix} I \n 0 \end{bmatrix}^T
\]
are non-negative. The pair \((A, B)\) is controllable, so there exists \(K\) that puts all eigenvalues of \(A + BK\) outside the unit disc and therefore \((A + BK)Z(A + BK)^T = Z + I\) has a symmetric positive definite solution \(Z\). Hence the desired \(X\) can be constructed as
\[
X = \begin{bmatrix} Z & ZK^T \\
KZ & * \end{bmatrix}
\]
where the lower right corner is chosen big enough to make \(X > 0\).

Define \(P = \text{diag}(\tau_1, \ldots, \tau_n) \geq 0\). Then \(\sigma\) being negative definite means that
\[
Q + \begin{bmatrix} A^T P A - P & A^T P B \\
B^T P A & B^T P B
\end{bmatrix} \preceq 0
\]
so (2.3) follows.

Finally, assume that (2.3) holds. Summing \(\sigma(x(t), w(t))\) over \(t\) gives
\[
0 \geq \sum_{0}^{\infty} \begin{bmatrix} x(t) \\
 w(t) \end{bmatrix}^T Q \begin{bmatrix} x(t) \\
 w(t) \end{bmatrix} + |A x(t) + Bw(t)|^2_p - |x(t)|^2_p
\]
For square summable solutions to \(x^+ = Ax + Bw\), \(x(0) = 0\) the telescope sum gives
\[
0 \geq \sum_{0}^{\infty} \begin{bmatrix} x(t) \\
 w(t) \end{bmatrix}^T Q \begin{bmatrix} x(t) \\
 w(t) \end{bmatrix}
\]
which in frequency domain implies (2.1). Hence (2.1)⇒(2.2)⇒(2.3)⇒(2.1).

For strict inequalities, the proofs that (2.3)⇒(2.1)⇒(2.2) remain the same. Assuming that (2.2) holds with strict inequality, we get
\[
\left[ (I - A)^{-1} B \right]^T (Q + \epsilon I) \left[ (I - A)^{-1} B \right] < 0
\]
for some scalar \(\epsilon > 0\). Hence, there exists a diagonal \(P > 0\) such that
\[
Q + \epsilon I + \begin{bmatrix} A^T P A - P & A^T P B \\
B^T P A & B^T P B
\end{bmatrix} < 0
\]
Adding a small multiple of the identity to \(P\) gives \(P > 0\) such that
\[
Q + \begin{bmatrix} A^T P A - P & A^T P B \\
B^T P A & B^T P B
\end{bmatrix} < 0
\]
so also (2.3) holds with strict inequality. Hence the proof of Theorem 2 is complete. \(\square\)

I. ACKNOWLEDGMENT

The author is grateful for suggestions and comments by numerous colleagues. The work has been supported by the Swedish Research Council through the LCCC Linnaeus Center.

REFERENCES