Distributed Control of Positive Systems

Rantzer, Anders

Published in: Proceedings of 50th IEEE Conference on Decision and Control

2011

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Distributed Control of Positive Systems

Anders Rantzer

Abstract—Stabilization and optimal control is studied for state space systems with nonnegative coefficients (positive systems). In particular, we show that a stabilizing distributed feedback controller, when it exists, can be computed using linear programming. The same methods are also used to minimize the closed loop input-output gain. An example devoted to distributed control of a vehicle platoon is examined.

I. INTRODUCTION

The study of matrices with nonnegative coefficients has a long history dating back to the Perron-Frobenius Theorem in 1912. A classic book on the topic is [2]. The theory has been used in Leontief economics, where the states denote nonnegative quantities of commodities, in the study of Markov chains, where the states denote nonnegative probabilities and compartment models, where the states could denote quantities of chemical species in an organism.

Positive systems have received increasing attention in the control literature during the last decade. In particular, stabilization of positive linear systems was studied in [4], [7]. Basic control theory for monotone systems, the nonlinear counterpart of positive systems, was developed in [1]. The importance of nonnegative matrices for consensus algorithms has been widely recognized (see e.g. [6]) and can be traced back to the fact that stability of a nonnegative matrix can be verified using a linear Lyapunov function [5].

A recent remarkable result by [8] shows that decentralized controllers can be optimized for positive systems using semi-definite programming. The criterion is the closed loop $H_\infty$ norm and the authors show that diagonal quadratic storage functions can be used without conservatism.

This paper can be viewed as an extension and alternative to [8]. First we demonstrate that stability analysis and synthesis of stabilizing controllers can be performed by distributed linear programming. Secondly, we show that several notions of system gain are equivalent for positive systems and can be optimized using the same methods that solved the stabilization problem.

II. NOTATION AND PRELIMINARIES

Let $\mathbb{R}_+$ denote the set of nonnegative real numbers. The inequality $X > 0$ ($X \geq 0$) means that all elements of the matrix (or vector) $X$ are positive (nonnegative). For a symmetric matrix $X$, the inequality $X > 0$ means that the matrix is positive definite. The matrix $A \in \mathbb{R}^{n \times n}$ is said to be Schur if all eigenvalues are strictly inside the unit circle. It is Metzler if all eigenvalues have positive real part. Finally, the matrix is said to be Metzer if all off-diagonal elements are nonnegative.

III. DISTRIBUTED STABILITY VERIFICATION

Proposition 1: Let $A \in \mathbb{R}^{n \times n}_+$. Then the following statements are equivalent:

(i) The matrix $A$ is Schur
(ii) There is a $x \in \mathbb{R}_+^n$ such that $Ax < x$
(iii) There is a diagonal matrix $P > 0$ such that $A^T P A < P$

Remark 1. If $A$ is Schur, then the same is true for $A^T$, so the proposition guarantees existence of $p \in \mathbb{R}_+^n$ with $A^T p < p$. Hence the linear function $V(x) = p^T x$ is a Lyapunov function for the dynamics $x^t = Ax$.

Proposition 1 is well known [2, Theorem 6.2.3], but we give a proof for completeness.

Proof. If $A$ is Schur, take any vector $\xi > 0$ and define $x = (I - A)^{-1} \xi$. Then $x = \xi + A \xi + A^2 \xi + A^3 \xi + \cdots > 0$ and $x - Ax = \xi > 0$, so (ii) holds. On the other hand, if (ii) holds, choose $\epsilon > 0$ such that $Ax < (1 - \epsilon)x$. Then for every $z \in \mathbb{R}^n$ with $0 < x < x$ we have

$$0 \leq A'z \leq A'x < (1 - \epsilon) x \quad \text{for } t = 1, 2, 3, \ldots$$

In particular, $\lim_{t \to \infty} A'tz$, so $A$ must be Schur.

The implication from (iii) to (i) is standard, so it remains to prove the opposite direction. Hence, assuming that $A$ (and therefore $A^T$) is Schur, we use (ii) to pick $x$ and $y$ in $\mathbb{R}_+^{n \times 1}$ with

$$Ax < x \quad A^T y < y$$

Define the matrix $P = \text{diag}(p_1, \ldots, p_n)$ by $p_k = y_k/x_k$ and the vector $z \in \mathbb{R}^n$ by $z_k = \sqrt{x_ky_k}$ for $k = 1, \ldots, n$. Then

$$P^{-1/2} A^T P A P^{-1/2} z = P^{-1/2} A^T P A x < P^{-1/2} A^T P x = P^{-1/2} A^T y < P^{-1/2} y = z$$

so the symmetric matrix $P^{-1/2} A^T P A P^{-1/2}$ is Schur. Hence

$$P^{-1/2} A^T P A P^{-1/2} < I$$

A. Rantzer is with Automatic Control LTH, Lund University, Box 118, SE-221 00 Lund, Sweden, rantzer at control.1th.se.
and (iii) follows.

The corresponding continuous time result can be stated as follows:

**Proposition 2:** Let $A \in \mathbb{R}^{n \times n}$ be Metzler. Then the following statements are equivalent:

(i) The matrix $A$ is Hurwitz

(ii) There exists a $x \in \mathbb{R}^{n \times 1}$ such that $Ax < 0$

(iii) There is a diagonal matrix $P > 0$ such that

$$A^T P + PA < 0$$

**Example 1** Consider a dynamical system inter-connected according to the graph illustrated in Figure 1:

$$\begin{align*}
\dot{x}_1 &= \alpha_1 x_1 + \ell_{12} (x_2 - x_1) \\
\dot{x}_2 &= \alpha_2 x_2 + \ell_{21} (x_1 - x_2) + \ell_{23} (x_3 - x_2) \\
\dot{x}_3 &= \alpha_3 x_3 + \ell_{32} (x_2 - x_3) + \ell_{34} (x_4 - x_3) \\
\dot{x}_4 &= \alpha_4 x_4 + \ell_{43} (x_3 - x_4)
\end{align*}$$

(1)

The model could for example be used to describe a platoon of four vehicles using distance measurements for position adjustment. The parameter $\ell_{ij} \geq 0$ can be used to stabilize the dynamics even in situations where some of the $\alpha_i$ parameters are positive. Notice that the dynamics can be written as $\dot{x} = M x$ where $M$ is a Metzler matrix. Hence, by Proposition 2, stability is equivalent to existence of numbers $x_1, \ldots, x_4 > 0$ such that

$$\begin{align*}
(\alpha_1 - \ell_{12}) x_1 + \ell_{12} x_2 &< 0 \\
(\ell_{21} - \ell_{23}) x_1 + (\alpha_2 - \ell_{21} - \ell_{23}) x_2 + \ell_{23} x_3 &< 0 \\
(\ell_{32} - \ell_{34}) x_2 + (\alpha_3 - \ell_{32} - \ell_{34}) x_3 + \ell_{34} x_4 &< 0 \\
(\ell_{43} x_3 + (\alpha_4 - \ell_{43}) x_4 &< 0
\end{align*}$$

This can be implemented as a distributed test where the first node verifies the first inequality, the second node verifies the second inequality and so on.

**IV. DISTRIBUTED STABILIZATION BY LINEAR PROGRAMMING**

Given the distributed stability test, the next step is to search for stabilizing feedback laws using distributed optimization. This can be done using the following theorem:

**Theorem 1:** Let $A \in \mathbb{R}^{n \times n}$, $E \in \mathbb{R}^{p \times m}$, $F \in \mathbb{R}^{m \times n}$ and let $D$ be the set of $m \times m$ diagonal matrices with entries in $[0, 1]$. Suppose that $A + ELF$ is non-negative for all $L \in D$. Then, the following are equivalent:

(i) There is $L \in D$ such that $A + ELF$ is Schur.

(ii) There exists $p \in \mathbb{R}^m_+$, $q \in \mathbb{R}^m_+$ with $q \leq E^T p$ and $A^T p + F^T q < p$.

**Proof.** Suppose (i) holds. Let $A + ELF$ be Schur and define $p \in \mathbb{R}^m_+$ with $(A + ELF)^T p < p$. Then $q \leq E^T p$ and $A^T p + F^T q = (A + ELF)^T p < p$.

Conversely, suppose that (ii) holds. Choose $L \in D$ to get $q = LE^T p$. Then

$$(A + ELF)^T p = A^T p + F^T q < p$$

so $A + ELF$ is Schur.

**Remark 2.** It is natural to compare the expression $A + ELF$ with the “state feedback” expression $A + BK$ of standard linear quadratic optimal control. A major difference is the matrix $F$, which makes the optimization of $A + ELF$ into a problem of “static output feedback” rather than state feedback. Another difference is the diagonally structured $L$ instead of the full matrix $K$. The diagonal structure gives a much higher degree of flexibility, particularly in the specification of distributed controllers.

**Remark 3.** If the diagonal elements of $D$ are restricted to $\mathbb{R}_+$ instead of $[0, 1]$, then the condition $q \leq E^T p$ is replaced by $0 \leq E^T p$

**Remark 4.** Each row of the inequalities $q \leq E^T p$ and $A^T p + F^T q < 0$ can be verified separately to get a distributed test.

**Example 2** Given $\alpha_1, \ldots, \alpha_4$ consider the system (1) and the problem to find feedback gains $\ell_{ij} \in [0, 1]$ that stabilize the system. The problem can be solved by applying distributed linear programming to condition (ii) in Theorem 2 with

$$A = \text{diag}\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$$

$$E = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$L = \text{diag}\{\ell_{12}, \ell_{21}, \ell_{23}, \ell_{32}, \ell_{34}, \ell_{43}\}$$

$$F = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

$$M = A + ELF$$

is the state space matrix of (1).
V. DISTRIBUTED PERFORMANCE OPTIMIZATION

Given the formulas for stabilizing feedback found in the previous section, it is natural move beyond stability and also optimize input-output performance. This can be done using the following theorem.

**Theorem 3:** Let $G(z) = C (z I - A)^{-1} B + D$ where $A \in \mathbb{R}_+^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}_+^{1 \times n}$ and $D \in \mathbb{R}_+$. Then the following two conditions are equivalent:

(i) The matrix $A$ is Schur and $\|G\|_\infty < \gamma$.

(ii) The matrix \[
\begin{bmatrix}
A & B \\
\gamma^{-1} C & \gamma^{-1} D
\end{bmatrix}
\] is Schur.

Moreover, $\|G\|_\infty = C (I - A)^{-1} B + D$ .

**Proof.** First note that the maximum $\max_\omega |G(e^{i \omega})|$ must be attained at $\omega = 0$ since

\[
|G(e^{i \omega})| = \left| D + \sum_{i=1}^{\infty} C (e^{i \omega} A)^{i-1} B \right|
\]

\[
\leq |D| + \sum_{i=1}^{\infty} |C A^{i-1} B| = C (I - A)^{-1} B + D = G(1)
\]

Hence $\|G\|_\infty < \gamma$ may equivalently be written

\[
C (I - A)^{-1} B + D < \gamma
\]

Suppose that (i) holds. By Proposition 1 there exists $\xi \in \mathbb{R}_+^{n \times 1}$ such that $A \xi < \xi$. Define $x = \xi + (I - A)^{-1} B$. Then $x > 0$ and

\[
Ax + B = A x - A \xi - \xi < 0
\]

If $\xi$ is sufficiently small, we also get $C x + D < \gamma$ so

\[
\begin{bmatrix}
A & B \\
\gamma^{-1} C & \gamma^{-1} D
\end{bmatrix}
\begin{bmatrix}
x \\
1
\end{bmatrix} < \begin{bmatrix}
x \\
1
\end{bmatrix}
\]

and (ii) holds. Conversely, (ii) implies that $Ax < x$ and

\[
(I - A)^{-1} B < x
\]

\[
0 \leq C (I - A)^{-1} B + D < C x + D < \gamma
\]

so (i) follows. \hfill \square

A continuous time version is given without proof:

**Theorem 4:** Suppose that $G(s) = C (s I - A)^{-1} B + D$ where $A \in \mathbb{R}_+^{n \times n}$ is Metzler, while $B \in \mathbb{R}_+^{n \times 1}$, $C \in \mathbb{R}_+^{1 \times n}$ and $D \in \mathbb{R}_+$. Then the following two conditions are equivalent:

(i) The matrix $A$ is Hurwitz and $\|G\|_\infty < \gamma$.

(ii) The matrix \[
\begin{bmatrix}
A & B \\
C & D - \gamma
\end{bmatrix}
\] is Hurwitz.

Moreover, $\|G\|_\infty = D - CA^{-1} B$ .

Combining this theorem with Theorem 2 gives a linear programming formulation of the problem to minimize input-output gain:

\[
\text{minimize input-output gain:}
\]

\[
\begin{align*}
\text{subject to:} & \\
& A x + B < x \\
& 0 \leq C (I - A)^{-1} B + D < C x + D < \gamma
\end{align*}
\]

is stabilized, while the gain from $u$ to $x_1$ is minimized. The solution is obtained by distributed linear programming solving the inequalities of condition (ii) in Theorem 5 with matrices $A$, $E$, $F$ and $L$ specified in Example 2 and

\[
B^T = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \\
C = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \\
D = 0
\]

Once $p$ and $q$ are determined, $L$ can be obtained from the equation $q = (EL)^T p$. \hfill \square
VI. ALL INDUCED NORMS ARE THE SAME

In this section we will verify that for positive systems, several notions of system gain are equivalent.

For a sequence \( w = \{w(t)\}_{t=0}^{\infty} \) define the norms

\[
\|w\|_p = \left( \sum_{t=0}^{\infty} |w(t)|^p \right)^{1/p} \quad \|w\|_{\infty} = \sup_t |w(t)|
\]

for \( 1 \leq p \). Given matrices \( A, B, C, D \), define the transfer function \( G(z) = C(zI - A)^{-1}B + D \) with corresponding impulse response \( g = \{g(t)\}_{t=0}^{\infty} = (D, CB, CAB, CA^2B, \ldots) \). Let \( g \ast w \) denote the convolution of \( g \) and \( w \) and define

\[
\|g\|_{p-ind} = \sup_w \frac{\|g \ast w\|_p}{\|w\|_p}
\]

for \( p = 1, 2 \) and \( \infty \). It is well known that if \( A \) is Schur and \( g \) scalar, then \( \|g\|_{p-ind} = \max_\omega |G(e^{i\omega})| \). Another notation for this norm is \( \|G\|_\infty \). When \( g(t) \geq 0 \), the maximum must be attained at \( \omega = 0 \) since

\[
|G(e^{i\omega})| = \left| D + \sum_{t=1}^{\infty} C(e^{-i\omega}A)^{-1}B \right| \\
\leq |D| + \sum_{t=1}^{\infty} |CA^{-1}B| = \sum_{t} g(t) = G(1)
\]

In fact, all \( p \)-induced norms are equal for \( 1 \leq p \leq \infty \):

**Theorem 6:** Suppose that \( g = \{g(t)\}_{t=0}^{\infty} \) is nonnegative with \( \sum_{t=0}^{\infty} g(t) < \infty \). Then

\[
\|g\|_{p-ind} = \sum_{t} g(t) \quad \text{for } 1 \leq p \leq \infty
\]

**Proof:** Let \( y = g \ast w \). Then

\[
\|y(t)\|_1 = \sum_{\tau=0}^{\infty} \left| \sum_{\tau=0}^{t} g(t-\tau)w(\tau) \right| \\
\leq \sum_{\tau=0}^{\infty} \left( \sum_{\tau=0}^{\infty} g(t-\tau) \right) |w(\tau)| \\
\leq \left( \sum_{t=0}^{\infty} g(t) \right) \|w\|_1
\]

with equality when \( w(t) \geq 0 \) for all \( t \). Moreover,

\[
|y(t)| = \left| \sum_{\tau=0}^{\infty} g(\tau)w(t-\tau) \right| \leq \sum_{\tau=0}^{\infty} g(\tau) |w(t-\tau)| \\
\leq \sum_{\tau=0}^{\infty} g(\tau) \|w\|_\infty
\]

with equality if \( w \) is constant. Hence the desired equality

\[
\|g\|_{p-ind} = \sum_{t} g(t)
\]

has been proved for \( p = 1 \) and \( p = \infty \) and it follows for intermediate values of \( p \) from the Riesz-Thorin convexity theorem [3, Theorem 7.1.12].

VII. ACKNOWLEDGMENT

The author is grateful to Cedric Langbort for useful suggestions and comments. The work has been supported by the Swedish Research Council through the Linnaeus Center LCCC.

REFERENCES