Positive state controllability of positive linear systems

Chris Guiver\textsuperscript{1,2}, Dave Hodgson\textsuperscript{3} and Stuart Townley\textsuperscript{4}

Systems & Control Letters, Vol. 65, pp. 23–29, March 2014,
\url{http://dx.doi.org/10.1016/j.sysconle.2013.12.002}

Abstract

Controllability of positive systems by positive inputs arises naturally in applications where both external and internal variables must remain positive for all time. In many applications, particularly in population biology, the need for positive inputs is often overly restrictive. Relaxing this requirement, the notion of positive state controllability of positive systems is introduced. A connection between positive state controllability and positive input controllability of a related system is established and used to obtain Kalman–like controllability criteria. In doing so we aim to encourage further study in this underdeveloped area.

Keywords: controllability, linear system, discrete time, positive system, constrained system, population ecology.

Mathematics Subject Classification (2010): 93C05 , 93C55 , 93B05 , 93B03 , 15B48 , 92D25

1 Introduction

Controllability is one of the most fundamental concepts in control theory. For finite–dimensional, linear, time–invariant, continuous time systems the notions of reachability, controllability and null controllability are all equivalent. The formulation of these concepts dates back to Kalman \cite{kalman1960} and, as is well–known, their appeal lies in the interplay between analytic and algebraic concepts. For instance, the existence of a control steering the system to a desired state is equivalent to the reachability matrix having full rank.

Controllability does not \textit{a priori} respect any (componentwise) nonnegativity of a system. This is problematic for many physically motivated applications, where state and input variables correspond to quantities that cannot take negative values. The need for nonnegative variables motivated the development of positive systems theory and there now exist several textbooks on the subject (for example, \cite{bryce1991}–\cite{farid2013}). Naturally, controllability, that is \textit{positive input controllability}, in such a framework is more limited than the general case, but the situation is well understood (\cite{guiver2004,guiver2010} and the references therein). A key feature of positive systems theory is the notion that both the state and the input variables must be nonnegative.

It is of interest, however, to consider

\[
x(t + 1) = Ax(t) + Bu(t), \quad t \in \mathbb{N}_0,
\]

where $A$ and $B$ are componentwise nonnegative under the constraints that just the state must be nonnegative – what might be termed \textit{positive state controllability}. There are conceivably many applications

\textsuperscript{1}Environment and Sustainability Institute, College of Engineering Mathematics and Physical Sciences, University of Exeter, Penryn Campus, Cornwall, TR10 9FE, UK, \texttt{c.guiver@ex.ac.uk}, \texttt{s.b.townley@ex.ac.uk}.

\textsuperscript{2}corresponding author

\textsuperscript{3}Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn Campus, Cornwall, TR10 9FE, UK, \texttt{d.j.hodgson@ex.ac.uk}.

1
of such a framework, for example in economic or logistic type models (see, for example, Miller & Blair [7]). Our primary example of where such a framework is necessary, however, is population ecology. Here matrix models are often used (see, for example, Caswell [8] or Cushing [9]) with the nonnegative state $x$ denoting a stage– or age–structured population, and the control $u$ denoting a conservation strategy or a form of pest control or harvesting. There are many papers (including, for example, [10]–[13]) where the model (1.1) is suitable for describing the addition or removal of individuals from a population and for a full description of these actions we require that $u$ can take negative values.

The framework of positive state controllability places a nonnegativity constraint on the codomain, and not on the domain, of the input–to–state map and it is not immediately clear that the positive input controllability theory is applicable. Here we demonstrate that under certain assumptions (reasonable for applications to population ecology) the problem of positive state controllability is equivalent to positive input controllability of a related positive system. Using this approach we characterise both the set of reachable states and the set of null controllable states of the pair $(A, B)$ under the constraint that the state must remain nonnegative. We demonstrate that, for example, the class of Leslie matrices [14] (with suitable control) that is frequently used in ecological modelling are positive state controllable, but often with negative control signals. We believe that there is seemingly a non–trivial ‘middle ground’ between controllability of linear systems and positive input controllability of positive systems that is worthy of in depth study.

2 Positive state control

For $n \in \mathbb{N}$, $\mathbb{R}^n_+$ denotes the nonnegative orthant in $\mathbb{R}^n$ and $e_i \in \mathbb{R}^n$ is the $i$th standard basis vector. For vectors $x$ and matrices $X$, $x \geq 0$ (also $0 \leq x$) and $X \geq 0$ (also $0 \leq X$) denotes componentwise nonnegativity. The superscript $^T$ denotes matrix transposition. We are interested in the pair $(A, B)$ generating the controlled system (1.1) where $A, B \geq 0$ and the state $x$ is nonnegative.

Our main result is Theorem 2.6 which relates nonnegative state trajectories with possibly nonpositive inputs to nonnegative state trajectories with nonnegative inputs of a related system. Such a connection allows us to appeal to existing positive input control results for this related system. Our key assumption is the following

(A) Given the pair $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ with $A, B \geq 0$ there exists $F \in \mathbb{R}^{m \times n}$ such that with $\tilde{A} := A - BF$ both $\tilde{A} \geq 0$ and if $v \in \mathbb{R}^n_+$, $w \in \mathbb{R}^m$ satisfy $Av + Bw \geq 0$ then $w \geq 0$.

The idea of assumption (A) is that by decomposing $A$ into $\tilde{A} + BF$, then negative controls $u$ in $Ax + Bu$ can be can be absorbed as $\tilde{A}x + B(Fx + u)$. Lemma 2.1 below gives a constructive characterisation of assumption (A) and demonstrates that if (A) holds then if holds for precisely one $F$ which can be calculated explicitly.

Lemma 2.1. Assumption (A) holds for $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$ with $A, B \geq 0$ if, and only if, there exist $m$ rows of $B$ such that the $m \times m$ submatrix, denoted $\underline{B}$, formed by taking these $m$ rows from $B$ is a positive monomial matrix and

$$A - \underline{B}^{-1} \underline{A} \geq 0. \quad (2.1)$$

Here $\underline{A}$ is formed of the $m$ rows of $A$ that appear in $\underline{B}$. Consequently, (A) holds if, and only if, it holds with $F = \underline{B}^{-1} \underline{A}$ so that $\tilde{A} := A - BF \geq 0$.

To prove Lemma 2.1 we first need an intermediate result.

Lemma 2.2. Given a pair $(A, B)$ with $A, B \geq 0$ then assumption (A) holds for $(A, B)$ if, and only if, there exists $F, H \in \mathbb{R}^{m \times n}$ such that the following four conditions hold

$$\begin{align*}
(a) & \quad F, H \geq 0, \\
(b) & \quad A - BF \geq 0, \\
(c) & \quad HA = F, \\
(d) & \quad HB = I_m.
\end{align*} \quad (2.2)$$

Here $I_p$ with $p \in \mathbb{N}$ denotes the $p \times p$ identity matrix.
Proof. Assumption (A) can be written as, there exists $F \in \mathbb{R}^{m \times n}$, $F \geq 0$ such that $\tilde{A} = A - BF \geq 0$ and for all $v \in \mathbb{R}^n$, $w \in \mathbb{R}^m$

$$
\begin{bmatrix}
\tilde{A} & B \\
I_n & 0
\end{bmatrix}
\begin{bmatrix}
v \\
w
\end{bmatrix}
\geq 0 \quad \Rightarrow \quad w = [0 \quad I_m] \begin{bmatrix}
v \\
w
\end{bmatrix}
\geq 0.
$$

By \[15\], (2.3) is equivalent to the existence of $\tilde{H} \in \mathbb{R}^{m \times (n+m)}$, $\tilde{H} \geq 0$ such that

$$
\begin{bmatrix}
\tilde{H}_1 & \tilde{H}_2 \\
I_n & 0
\end{bmatrix}
\begin{bmatrix}
\tilde{A} & B \\
I_n & 0
\end{bmatrix}
= [0 \quad I_m] \quad \Leftrightarrow \quad \tilde{H}_1 \tilde{A} + \tilde{H}_2 = 0, \quad \text{and} \quad \tilde{H}_1 B = I_m.
$$

Since we require that $\tilde{H} \geq 0$ we see immediately from (2.4) that we can always take $\tilde{H}_2 = 0$ and hence by writing $H = \tilde{H}_1$ it follows that (A) is equivalent to the existence of $F, H \in \mathbb{R}^{m \times n}$ such that

$$
F, H \geq 0, \quad A - BF \geq 0, \quad H \tilde{A} = 0, \quad HB = I_m.
$$

Using the fact that $\tilde{A} = A - BF$ so that $H \tilde{A} = HA - HBF$ we have that (2.5) is equivalent to (2.2), as required.

Proof of Lemma 2.1: First assume that $B$ contains an $m \times m$ positive monomial submatrix, which we denote by $B$. Let $A$ denote the $m \times n$ submatrix of $A$ formed by taking the $m$ rows $\{i_1, \ldots, i_m\}$ of $A$ that appear in $B$. Since $B$ is positive monomial, it follows (from, for example, p. [16, p. 68]) that $B$ has a positive inverse. Define

$$
F := B^{-1}A \geq 0,
$$

the nonnegativity following as $B^{-1}, A \geq 0$. Furthermore, by assumption $A - BF \geq 0$ and hence

$$
\tilde{A} := A - BF = A - BB^{-1}A \geq 0.
$$

Now assume that $v \in \mathbb{R}^n, w \in \mathbb{R}^m$ are such that $\tilde{A}v + Bw \geq 0$. By restricting attention to rows $\{i_1, \ldots, i_m\}$ (where $\tilde{A} - BF = 0$) we have that

$$
(\tilde{A} - BF)v + Bw \geq 0 \quad \Rightarrow \quad Bw \geq 0 \quad \Rightarrow \quad w \geq 0,
$$

as $B^{-1} \geq 0$. We conclude that (A) holds.

To prove the converse we use the characterisation of (A) from Lemma 2.2. Suppose that (A) holds so that there exists $H, F \in \mathbb{R}^{m \times n}$ such that (a)–(d) hold. We also need a corollary to [16, Lemma 4.3, p. 68], which we repeat here. A real nonnegative $n \times m$ matrix $X$ of rank $m$ has a nonnegative left inverse if, and only if, $X$ contains an $m \times m$ monomial submatrix.

Assumptions (a) and (d) imply that $B$ has a nonnegative left inverse $H$ and thus $B$ must contain at least one (although possibly many) $m \times m$ positive monomial submatrix (submatrices), as must $H$. From the above arguments there must be a set of $m$ rows

$$
\{i_1, \ldots, i_m\} \subseteq \{1, 2, \ldots, n\},
$$

of $B$ that give rise to a monomial submatrix $B$ where the corresponding columns $\{i_1, \ldots, i_m\}$ of $H$ must each have precisely one nonzero entry. The columns must each have at least one nonzero entry so that the product $HB$ does not have a zero column. They cannot have more than one else $HB = I_m$ cannot hold.

The equalities $HA = F$ and $HB = I_m$ together imply that $H(A - BF) = 0$ and as $H \geq 0$, $H \neq 0$, $A - BF \geq 0$ it follows that the rows $\{i_1, \ldots, i_m\}$ of $A - BF$ must be zero. Therefore, restricting to the rows $\{i_1, \ldots, i_m\}$ we have that

$$
\tilde{A} - BF = 0,
$$

(2.7)

where $\tilde{A}$ is $m \times n$ submatrix formed from rows $\{i_1, \ldots, i_m\}$ of $A$. From (2.7) it follows that $F = B^{-1}A$ which by construction yields $A - BB^{-1}A = A - BF \geq 0$, as required. \[ \Box \]
Matlab code for verifying whether (A) holds for a given \((A,B)\) is available as online supplementary material. We comment here that (A) holds for any \(A \geq 0\) in the single input case \(B = b = c_0 e_1\), \(c_1 > 0\) and the corresponding multiple input version case when \(B\) is a combination of \(e_i\), that is, \(B = [c_{i1} e_{i1}, \ldots, c_{im} e_{im}]\) for positive \(c_{ik}\). These two cases are arguably the most important for applications. The following corollary interprets Lemma 2.1 in the single input case.

**Corollary 2.3.** Let \(A \geq 0\) with \(i\)th row denoted by \(r_i\) and \(B = b\) be given by

\[
b = \sum_{k=1}^{n} c_{ik} e_{ik} \quad \text{with} \quad c_{ik} > 0.
\]

Assumption (A) holds for \((A,b)\) if, and only if, there exists \(i_k \in \{i_1, \ldots, i_n\}\) such that

\[
r_{ij} - \frac{c_{ik} r_{ik}}{c_{ik}} \geq 0, \quad \forall i_j \in \{i_1, \ldots, i_n\},
\]

and in this case \(F = f^T = \frac{r_{ik}}{c_{ik}}\), where \(i_k\) is as in (2.8).

**Remark 2.4.** (i) Lemma 2.1 provides an algorithm for checking assumption (A). First, we see that \(B\) containing an \(m \times m\) monomial submatrix is necessary for (A). Second, there are then only finitely many \(B\) (formed from the monomial rows of \(B\)) to check whether (A) holds by verifying whether \(A - B B^T A \geq 0\).

(ii) The condition (2.8) is the single input version of (2.1). In words, it requires that \(A\) has a smallest row (in the sense of (2.8)) over all the non–zero rows of \(b\).

**Example 2.5.** Consider the systems

\[
(a) \quad A_1 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \\ 1 & 2 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},
\]

\[
(b) \quad A_2 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \\ 3 & 2 & 2 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix},
\]

\[
(c) \quad A_3 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \\ 3 & 2 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]

Assumption (A) holds in (a), fails in (b) and holds in (c). We proceed to prove these claims. For (a), if we take \(f_1^T := \begin{bmatrix} 1 & 2 \end{bmatrix}\) so that

\[
\tilde{A}_1 := A_1 - b f_1^T = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix},
\]

then whenever \(v \in \mathbb{R}_n^*, w \in \mathbb{R}\) are such that \(\tilde{A}_1 v + bw \geq 0\), by inspection of the third component we see that \(w \geq 0\), which by definition is (A). Alternatively, with \(i_1 = 1\) and \(i_2 = 3\) we compute

\[
r_1 - r_3 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 \end{bmatrix} \geq 0.
\]

Thus Corollary 2.3 applies with \(i_k = i_2 = 3\), so that \(f_1 = \begin{bmatrix} 1 & 2 \end{bmatrix}\), as in (2.9). However, repeating this process in (b) gives

\[
r_1 - r_3 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 2 & 2 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 2 \end{bmatrix} \not\geq 0,
\]

and also that \(r_3 - r_1 \not\geq 0\). We conclude from Corollary 2.3 that (A) does not hold for (b). For (c) we note that \(B\) contains two \(2 \times 2\) positive monomial submatrices

\[
B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad B_2 = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & 0 \end{bmatrix}.
\]
formed from rows one and two, and rows two and three of $B$ respectively. Taking the corresponding submatrices from $A$ gives
\[
\hat{A}_1 = \begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{and} \quad \hat{A}_2 = \begin{bmatrix} 1 & 2 & 1 \\ 3 & 2 & 2 \end{bmatrix}.
\]
from which we compute
\[
A - BB^{-1}\hat{A}_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & \frac{1}{2} & 0 \end{bmatrix} \geq 0
\]
and
\[
A - BB^{-1}\hat{A}_2 = \begin{bmatrix} -4 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq 0.
\]
We conclude that (A) holds for (c) with $F = [2 \ 3 \ 4]$. 

**Reachability**

**Theorem 2.6.** Let the pair $(A, B)$ satisfy (A) and denote $\hat{A} := A - BF$. Then the state trajectories of $(A, B)$ from initial state $x_0 \in \mathbb{R}^n_+$ with nonnegative state are precisely the state trajectories of $(\hat{A}, B)$ from initial state $x_0 \in \mathbb{R}^n_+$ with nonnegative control.

**Proof.** Fix $x_0 \in \mathbb{R}^n_+$ and suppose that $x = x(\cdot; x_0, u)$ is a nonnegative state trajectory of $(A, B)$ with initial state $x_0$ and control $u$, that is, $x(t) \geq 0$ for every $t \geq 0$. Note that the control $u$ need not always take nonnegative values. Defining
\[
\hat{u} := u + Fx,
\]
we see that
\[
0 \leq x(1) = Ax_0 + Bu(0) = \hat{A}x_0 + B(u(0) + Fx_0)
\]
\[
= \hat{A}x_0 + B\hat{u}(0),
\]
and so (A) implies that $\hat{u}(0) \geq 0$. Arguing now by induction using (2.10) we see that for $k \in \mathbb{N}_0$
\[
0 \leq x(k+1) = Ax(k) + Bu(k) = \hat{A}x(k) + B\hat{u}(k),
\]
and so as before, (A) implies that $\hat{u}(k) \geq 0$. We conclude that $x = x(\cdot; x_0, \hat{u})$, a nonnegative state trajectory of the pair $(\hat{A}, B)$ with initial state $x_0$ and nonnegative control $\hat{u}$. The converse argument reverses these steps. Specifically, for fixed $x_0 \in \mathbb{R}^n_+$, let $x = x(\cdot; x_0, \hat{u})$ denote a (necessarily nonnegative) state trajectory of $(\hat{A}, B)$ with initial state $x_0$ and nonnegative control $\hat{u}$. We define the control sequence $u$ by $u := \hat{u} - Fx$ which, by construction, satisfies for $k \in \mathbb{N}_0$
\[
x(k+1) = \hat{A}x(k) + B\hat{u}(k) = (A - BF)x(k) + B\hat{u}(k)
\]
\[
= Ax(k) + Bu(k).
\]
Therefore $x = x(\cdot; x_0, u)$, a nonnegative state trajectory of the pair $(A, B)$ with initial state $x_0$ and control $u$ which although not necessarily always nonnegative valued, nevertheless preserves nonnegativity of the state.

The consequence of Theorem 2.6 is that (under assumption (A)) results for positive state controllability of the pair $(A, B)$ follow from existing results for positive input controllability results for the pair $(A, B)$. We discuss reachability, null controllability and then controllability.

In order to formalise reachability with nonnegative state – positive state reachability – we introduce the following definition.

**Definition 2.7.** Given the pair $(A, B)$ with $A, B \geq 0$, we say that $x_T \in \mathbb{R}^n_+$ is positive state reachable in finite time if there exists a control sequence that steers the state $x$ of $(A, B)$ from 0 to $x_T$ in $N$ steps and additionally maintains nonnegativity of $x$. The collection of all such $x_T \in \mathbb{R}^n_+$ is called the positive state reachable set (in finite time). We say that $(A, B)$ is positive state reachable in finite time if this set is $\mathbb{R}^n_+$. 

5
We note that the positive state reachable set is a convex cone of the linear space $\mathbb{R}^n$ over $\mathbb{R}$, and not a linear subspace. For nonnegative vectors $x_1, x_2, \ldots, x_k \in \mathbb{R}^n_+$ we let $\langle x_1, x_2, \ldots, x_k \rangle_+ \subseteq \mathbb{R}^n_+$ denote their nonnegative linear span. For nonnegative matrices $X_1, X_2, \ldots, X_k \in \mathbb{R}^{n \times n}_+$ we let $\langle X_1, X_2, \ldots, X_k \rangle_+ \subseteq \mathbb{R}^n_+$ denote the nonnegative linear span of their columns.

**Corollary 2.8.** Let the pair $(A, B)$ satisfy (A). Then the positive state reachable set of the pair $(A, B)$ in finite time is precisely

$$
\bigcup_{N \in \mathbb{N}} \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+ ,
$$

(2.11)

where $\tilde{A}$ is as in (A). The positive state reachable set of the pair $(A, B)$ in infinite time is precisely

$$
\bigcup_{N \in \mathbb{N}} \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+ .
$$

(2.12)

**Proof.** This follows immediately from Theorem 2.6 and the definition of positive state reachable. We note that the sets in (2.11) and (2.12) are denoted in [5, p. 41] as $R_\infty(\tilde{A}, B)$ and $R(\tilde{A}, B)$ respectively.

**Remark 2.9.** (i) For standard reachability of single-input systems $(A, B)$ the Cayley–Hamilton Theorem implies that every reachable state is reachable in at most $n$ steps, where $n$ is the dimension of $A$. It is known ([5, p. 42]) that this is not the case for positive systems, and by Corollary 2.8 we see that the same is true for positive state reachability. Example 2.10 contains a pair $(A, B)$ that is positive state reachable, but only in infinite time.

(ii) Clearly, classical positive input reachability of the nonnegative pair $(A, B)$ pair implies positive state reachability of $(A, B)$. This is apparent from Corollary 2.8 as for each $N \in \mathbb{N}$

$$
\langle B, AB, \ldots, A^{N-1}B \rangle_+ \subseteq \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+ .
$$

Example 3.1 contains a pair $(A, B)$ where the above inclusion is strict.

**Example 2.10.** Consider the $3 \times 3$ nonnegative matrix and control vector

$$
A = \begin{bmatrix} a_1 & 1 & 0 \\ 0 & a_1 & 1 \\ 0 & 0 & a_2 \end{bmatrix}, \quad b = e_3 ,
$$

with $1 > a_1 > 0$ and $a_2 \geq 0$. By Corollary 2.3 it follows that (A) applies to the pair $(A, b)$ with $F = f^T$ the third row of $A$. A calculation shows that $\tilde{A}b = e_2$ and for $k \geq 2$

$$
\tilde{A}^k b = [(k-1)a_1^{k-2} a_1^{k-1} 0]^T .
$$

(2.13)

The positive state reachable set in $k + 1$ steps is all nonnegative linear combinations of these vectors, which here is strictly increasing with increasing $k$ and note does not include $e_1$ for finite $k$. However, by noting that for $k \geq 2$

$$
\frac{1}{(k-1)a_1^{k-2}} \cdot \tilde{A}^k b = [1 \quad a_1 \quad 0]^T \rightarrow e_1 ,
$$

as $k \rightarrow \infty$, it follows that the pair $(A, b)$ is positive state reachable in infinite time.

**Corollary 2.11.** A pair $(A, B)$ satisfying (A) is positive state reachable in finite time if, and only if, for some $N \in \mathbb{N}$ the matrix $[B \quad AB \quad \ldots \quad \tilde{A}^{N-1}B]$ contains an $n \times n$ monomial submatrix.

**Proof.** This follows immediately from Theorem 2.6, Corollary 2.8 and [5, Proposition 3].

The next lemma provides sufficient conditions for when, given a pair $(A, B)$ satisfying (A), the positive state reachable set is achieved in finite time.

**Lemma 2.12.** Let the pair $(A, B)$ satisfy (A). Then the statements
(i) $\tilde{A}$ is nilpotent,
(ii) the characteristic polynomial of $\tilde{A}$ is given by $t^n - \sum_{i=0}^{n-1} \beta_i t^i$, with $\beta_i \geq 0$ for each $i$,
(iii) the minimal polynomial of $\tilde{A}$ is given by $t^m - \sum_{i=0}^{m-1} \gamma_i t^i$, with $\gamma_i \geq 0$ for each $i$, are each sufficient for

$$\bigcup_{N \in \mathbb{N}} \langle B, \tilde{A}^N B, \ldots, \tilde{A}^{N-1}B \rangle_+ = \langle B, \tilde{A} B, \ldots, \tilde{A}^M B \rangle_+, \quad (2.14)$$

for some $M \in \mathbb{N}$. Equality (2.14) implies that the positive state reachable set of the pair $(A, B)$ is attained in finite time.

**Proof.** Each of the statements (i)--(iii) imply that the sequence of cones $(\langle B, \tilde{A}^N B, \ldots, \tilde{A}^{N-1}B \rangle_+ )_{n \in \mathbb{N}}$ which is nested and increasing is in fact constant for some $M \in \mathbb{N}$ and hence all larger $n \in \mathbb{N}$, which is condition (2.14).

**Remark 2.13.** It is important to note that given $A, B \geq 0$, when $A = \tilde{A} + BF$ with $\tilde{A}, F \geq 0$ then, even without assumption (A), every state trajectory of $(\tilde{A}, B)$ with nonnegative control is a state trajectory of $(A, B)$ with nonnegative state (this is one of the implications in Theorem 2.6). Consequently, the positive input reachable set of the pair $(\tilde{A}, B)$ is a subset of the positive state reachable set of the pair $(A, B)$. Although not giving the complete picture of positive state control, this connection can be used as an intermediate stage between positive state control and positive input control. For example, consider the pair $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ where we have

$$\langle B, AB \rangle_+ = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix} +.$$  

(2.15)

More generally an induction argument shows that the positive input reachable space is $\langle e_1, e_1 + e_2 \rangle_+$; the hatched area in Figure 1. The pair $(A, B)$ in (2.15) do not satisfy assumption (A) (as $B$ contains no $2 \times 2$ monomial submatrices). However, taking $F = B$ gives

$$\tilde{A} := A - BF = A - B^2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \geq 0,$$

and so the positive state reachable set contains

$$\langle B, \tilde{A} B \rangle_+ = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \mathbb{R}_+^2.$$  

We conclude that, although assumption (A) does not hold, the pair $(A, B)$ is positive state reachable in finite time. Note that in this instance the choice $F = B$ is in no sense unique; the same conclusions are reached for the pair $(A, B)$ with $F = \begin{bmatrix} 1 & 0 \end{bmatrix}$ as here $\tilde{A} = A - BF = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \geq 0$.

Notwithstanding the above, it is true that if $F_1 \geq F_2$ then $\tilde{A}_1 := A - BF_1 \leq A - BF_2 =: \tilde{A}_2$ and thus for each $k \in \mathbb{N}$

$$\langle B, \tilde{A}_2 B, \ldots, \tilde{A}_2^k B \rangle_+ \subseteq \langle B, \tilde{A}_1 B, \ldots, \tilde{A}_1^k B \rangle_+.$$
Consequently, to describe positive state control for a pair \((A, B)\) when assumption \((A)\) fails, the above suggests considering \(\tilde{A} := A - BF\), where \(F\) is chosen as (componentwise) large as possible so that \(\tilde{A} \geq 0\). Such a process can simplify calculations considerably. For example, consider the pair \(A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix}\) and \(B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}\)

\[
\langle B, AB, A^2B \rangle_+ = \left( \begin{bmatrix} 1 & 1 & 4 & 2 & 19 & 6 \\ 1 & 0 & 5 & 1 & 25 & 6 \\ 1 & 0 & 6 & 1 & 33 & 7 \end{bmatrix} \right)_+ .
\]  

Choosing \(F = B^T\) gives \(\tilde{A} := A - BF = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \geq 0\) so that

\[
\langle B, \tilde{A}B, \tilde{A}^2B \rangle_+ = \left( \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 4 & 0 \\ 1 & 0 & 3 & 0 & 9 & 0 \end{bmatrix} \right)_+ .
\]

We see by inspection of columns two and three in the matrices of the right hand sides of (2.16) and (2.17) that in this example the difference between positive state control and positive input control is that in the former the directions \(e_1\) and \(2e_2 + 3e_3\) have been ‘decoupled’.

**Null controllability**

**Definition 2.14.** Given the pair \((A, B)\) with \(A, B \geq 0\), we say that \(x_0 \in \mathbb{R}_n^+\) is positive state null controllable in finite time if there exists a control sequence that steers the state \(x\) of \((A, B)\) from \(x_0\) to 0 in \(N\) steps and additionally maintains nonnegativity of \(x\). The collection of all such \(x_0 \in \mathbb{R}_n^+\) is called the positive state null controllable set (in finite time). We say that \((A, B)\) is positive state null controllable in finite time if this set is \(\mathbb{R}_n^+\).

Appealing to Theorem 2.6, we now characterise the positive state null controllable set of the pair \((A, B)\).

**Corollary 2.15.** Let the pair \((A, B)\) satisfy \((A)\). Then the positive state null controllable set of the pair \((A, B)\) in finite time is precisely

\[
\mathbb{R}_n^+ \cap \ker \tilde{A}^n.
\]

The set of positive null controllable states in infinite time is precisely

\[
\mathbb{R}_n^+ \cap (\ker \tilde{A}^n + E(\tilde{A})),
\]

where \(E(\tilde{A})\) is the sum of (generalised) eigenspaces corresponding to the stable eigenvalues of \(\tilde{A}\) (that is, the eigenvalues \(\lambda\) of \(\tilde{A}\) with \(|\lambda| < 1\)).

**Proof.** The claims follow from Theorem 2.6. More specifically, \(\ker \tilde{A}^k\) is exactly the set of states that are positive state null controllable in \(k\) steps. The increasing sequence \((\ker \tilde{A}^k)_{k \in \mathbb{N}}\) is constant after some \(m\) with \(0 \leq m \leq n\) and so a state is null controllable in finite time if and only if it is null controllable in at most \(n\) steps. 

**Corollary 2.16.** Let the pair \((A, B)\) satisfy \((A)\). Then

(i) \((A, B)\) is positive state null controllable in finite time if, and only if, \(\tilde{A}\) is nilpotent.

(ii) \((A, B)\) is positive state null controllable in infinite time if, and only if, \(\tilde{A}\) is Schur (that is, every eigenvalue \(\lambda\) of \(\tilde{A}\) has \(|\lambda| < 1\)).

**Proof.** The claims are straightforward to establish from (2.18) and (2.19) in Corollary 2.15, or can be proven using arguments as in [5, Proposition 1] and [5, Proposition 2] respectively. 

8
Example 2.17. Using the above results it follows that for a pair \((A, B)\) satisfying \((A)\), \(e_i\) is positive state null controllable in \(k\) steps if, and only if, \(A^k\) has \(i^{th}\) column zero.

Remark 2.18. Analogously to Remark 2.13, when assumption \((A)\) fails for a pair \((A, B)\) then it follows from Theorem 2.6 and Corollary 2.15 that the positive state null controllable set in finite time contains \(\mathbb{R}^n_+ \cap \ker A\), where \(A = A - BF \geq 0\) for some \(F \geq 0\).

### Controllability

We now comment on positive state controllability, which as the underlying system (1.1) is linear, is the combination of positive state reachability and positive state null controllability.

**Definition 2.19.** Given the pair \((A, B)\) with \(A, B \geq 0\), we say that \(x_0 \in \mathbb{R}^n_+\) is positive state controllable to \(x_T \in \mathbb{R}^n_+\) in finite time if there exists a control sequence that steers the state \(x = (A, B)\) from \(x_0\) to \(x_T\) in \(N\) steps and additionally maintains nonnegativity of \(x\). We say that \((A, B)\) is positive state controllable in finite time if every \(x_0 \in \mathbb{R}^n_+\) can be controlled to every \(x_1 \in \mathbb{R}^n_+\) in finite time with nonnegative state.

The proof of the next result is elementary and is therefore not provided. Note that assumption \((A)\) is not required here—the result follows from the definitions and the linearity of (1.1).

**Corollary 2.20.** The pair \((A, B)\) with \(A, B \geq 0\) is positive state controllable in (in)finite time if, and only if, \((A, B)\) is both reachable and null controllable with positive state in (in)finite time.

If \(x_0 \in \mathbb{R}^n_+\) is positive state null controllable and \(x_T \in \mathbb{R}^n_+\) is positive state reachable, then \(x_0\) can be controlled to \(x_T\) by ‘travelling via the zero state’. The next lemma shows that under assumption \((A)\), this is in fact more or less everything.

**Lemma 2.21.** Let \((A, B)\) denote a pair satisfying \((A)\). If \(x_0 \in \mathbb{R}^n_+\) is positive state controllable to \(x_T \in \mathbb{R}^n_+\) in \(N\) \(\in \mathbb{N}\) steps then either

\[
\begin{align*}
(i) \quad & x_0 \in \ker \tilde{A}^N \text{ and } x_T \in \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+, \text{ or,} \\
(ii) \quad & x_T - \tilde{A}^Nx_0 \geq 0 \text{ and } x_T - \tilde{A}^Nx_0 \in \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+.
\end{align*}
\]

Proof. By hypothesis there exist \(N \in \mathbb{N}\) and control \(u = (u(k))_{k=0}^{N-1}\) such that the state \(x = (\cdot; x_0, u)\) of the pair \((A, B)\) satisfies \(x(0) = x_0\), \(x(k) \geq 0\) for \(k \in \{0, 1, \ldots, N\}\) and \(x(N) = x_T\). By Theorem 2.6, we can choose a control sequence \(\tilde{u} = (\tilde{u}(k))_{k=0}^{N-1}\) (given by (2.10) in fact) such that \(x = (\cdot; x_0, \tilde{u})\) of the pair \((\tilde{A}, B)\) satisfies

\[
0 \leq x_T = \tilde{A}^N x_0 + \sum_{k=0}^{N-1} \tilde{A}^{N-1-k} B \tilde{u}(k). \tag{2.20}
\]

Now either \(\tilde{A}^N x_0 = 0\) and so from (2.20) condition (i) holds or,

\[
0 \leq x_T - \tilde{A}^N x_0 = \sum_{k=0}^{N-1} \tilde{A}^{N-1-k} B \tilde{u}(k) \in \langle B, \tilde{A}B, \ldots, \tilde{A}^{N-1}B \rangle_+,
\]

which is condition (ii). 

\[\square\]
3 Examples

One motivation for studying positive state controllability of the linear system (1.1) is the class of systems that arise in population ecology. Here we present some examples.

Example 3.1. We recall that an $n \times n$ Leslie [14] matrix has the following structure

$$\begin{bmatrix} f_1 & f_2 & \cdots & f_n \\ s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & s_{n-1} & 0 \end{bmatrix}, \quad (3.1)$$

which models a population partitioned into discrete, increasing age-stages. Correspondingly, $f_i \geq 0$ denote reproductive rates and $s_i \geq 0$ denote survival rates, the latter as proportions are each no greater than one. For ecologically meaningful models [17], we will always assume that the Leslie matrix (3.1) has $s_1, \ldots, s_{n-1} > 0$, $f_1, \ldots, f_n \geq 0$ and there exists at least one $i \in \{1, 2, \ldots, n\}$ such that $f_i > 0$. Noting that the $n \times n$ positive diagonal matrix

$$T = \text{diag} \left( 1, \frac{1}{s_1}, \frac{1}{s_1 s_2}, \ldots, \frac{1}{s_1 \cdots s_{n-1}} \right), \quad \text{has } T^{-1} \geq 0,$$

for positive state controllability it is sufficient to consider the similarity transformed pair $(T^{-1}AT, T^{-1}B)$. Since $T^{-1}AT$ has the same structure as $A$ with ones on the subdiagonal and top row with entries $\tilde{f}_i > 0$ (which we abuse notation and write as $f_i$), when considering controllability with positive state there is no loss of generality in assuming that a Leslie matrix has $s_j = 1$ for each $j$.

We consider the (transformed) pair $(A, b)$ and single input $b = e_j$, $j \in \{1, 2, \ldots, n\}$. Assumption (A) is always satisfied for such a pair, with $F = f^T$ the $j^{th}$ row of $A$.

When $j = 1$ so that $b = e_1$ it follows that $\tilde{A}$ is nilpotent with $\tilde{A}^n = 0$. The positive state reachable set is therefore

$$\langle b, \tilde{A}b, \ldots, \tilde{A}^{n-1}b \rangle_+ = \langle e_1, e_2, \ldots, e_n \rangle_+ = \mathbb{R}^n_+,$$

and by Corollary 2.8 the pair $(A, b)$ is positive state reachable. Furthermore, $\tilde{A}^n = 0$, $\implies \mathbb{R}^n_+ \cap \ker \tilde{A}^n = \mathbb{R}^n_+$, and so the pair $(A, b)$ is positive state null controllable. In this very special case it follows that positive state controllability and standard controllability coincide: the unique control $u$ that steers $x$ between any two nonnegative states in $n$ steps is such that the state remains in $\mathbb{R}^n_+$. Such a control can take negative values and thus is not permitted in a positive input framework.

For example, consider the pair $A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}$, $b = e_1$. Trivially $e_1$ is reachable from zero with positive state in one step and the control $u(0) = 1$, $u(1) = -2$ steers the state from zero to $e_2$ in two steps. The resulting state trajectories are plotted in Figure 2(a). By taking suitable linear combinations of these inputs all of $\mathbb{R}^2_+$ can be reached with nonnegative state. If we restrict attention to $(A, b)$ with only positive inputs, then the positive input reachable space is spanned by $b = e_1$ and $Ab = 2e_1 + e_2$, and is depicted in Figure 2(b): note that not all of $\mathbb{R}^2_+$ is reachable. Even in this very simple example there is a difference between positive state controllability and classical positive input controllability.

For $b = e_j$, $j > 1$ the situation is somewhat different. Assumption (A) holds with $F = f^T$ the $j^{th}$ row of $A$ and we note that the characteristic polynomial of $\tilde{A}$ is given by

$$t^n - \sum_{k=1}^{j-1} f_k t^{n-k},$$

which follows easily from, for example, the expression on [4, p. 121]. Consequently, by Lemma 2.12 (ii) the positive state reachable set of the pair $(A, b)$ is achieved in finite time; indeed, in $n$ steps. However,
the pair \((A, b)\) is not, in general, positive state reachable. In the case \(n = 3\) and \(b = e_2\) the positive state reachable set in finite time is contained in

\[
\langle b, \tilde{A}b, \tilde{A}^2b \rangle_+ = \left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} f_2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} f_3 + f_1f_2 \\ 0 \\ 0 \end{bmatrix} \right\rangle_+.
\]

If \(f_2 > 0\) then \(e_3\) cannot be steered to whilst maintaining nonnegative state (in finite time), but the vectors \(e_1\) and \(e_2\) can (in the former case provided that \(f_1f_2 + f_3 > 0\)). For positive state null controllability we see that

\[
\tilde{A}^2 = \begin{bmatrix} f_1^2 & f_1f_2 + f_3 & f_1f_3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \tilde{A}^3 = f_1\tilde{A}^2.
\]

The important term here is \(f_1\), reproduction of individuals in the first stage class. If \(f_1 = 0\) then \(\tilde{A}^3 = 0\) and so every state can be steered to zero with nonnegative state. However, if \(f_1, f_2, f_3 > 0\) then the top row of \(\tilde{A}^3\) is positive and thus there are no nontrivial positive state null controllable states in finite time!

**Remark 3.2.** The conclusions of Example 3.1 are biologically sensible. When state \(i \geq 1\) is controlled then the structure of \(A\) means that it is not possible to remove individuals from the earlier stage classes. Consequently we can steer to the later stage classes \(e_j, j > i\) which appear in the terms \(\tilde{A}^i b\), but also with a contribution from stages one to \(i - 1\). When \(i = 1\) then it follows that all stages are positive state reachable.

Leslie matrices of course have a very simple structure, but already the results presented here demonstrate the differences that arise between positive state control and positive input control. Our next example considers an application of positive state control.

**Example 3.3.** Discrete time matrix models for the invasive weed *Cirsium vulgare* (spear thistle) in Nebraska, USA, are considered in [18], and also [19, Section 3.1]. Here time-steps correspond to years and a four stage model is used with states one to four corresponding to the seed bank, small plants, medium plants and large plants respectively (see [18]). The nominal uncontrolled system has \(A\) given by

\[
A = \begin{bmatrix}
0 & 0 & f_1 \\
0 & f_3 & f_4 \\
0 & s_2 & s_3 \\
0 & s_4 & s_5 & s_6
\end{bmatrix},
\]

with

\[
s_1 = 0.0077, \quad s_2 = 0.12, \quad s_3 = 0.11, \quad s_4 = 0.02,
\]

\[
s_5 = 0.27, \quad s_6 = 0.17, \quad \tilde{f}_1 = 93.1, \quad \tilde{f}_2 = 423,
\]

\[
\tilde{f}_3 = 6.74, \quad \tilde{f}_4 = 30.6.
\]
As with Leslie matrices, the $s_i$ denote survival and growth parameters and the $\tilde{f}_i$ are reproductive values. We note that as $s_4 > 0$, small plants can grow into large plants in one year. Furthermore, $\tilde{f}_3, \tilde{f}_4 > 0$ means that in a given year both medium and large plants can produce seeds that germinate and grow into small plants (in addition to seeds that germinate the following year).

The uncontrolled population is unstable as the spectral radius of $A$ is $r(A) = 1.57 > 1$. We seek to reduce the weed population by using an additive management strategy, so that the system is of the form (1.1). As a management strategy we add or remove large plants so that $B = b = e_4$. When this action is performed (shortly) before the census or measurement then the resulting model is well described by (1.1). Here assumption (A) holds with $F = f^T$ the fourth row of $A$ so that $\tilde{A} := A - bf^T$ is given by

$$\tilde{A} = \begin{bmatrix} 0 & 0 & \tilde{f}_1 & \tilde{f}_2 \\ s_1 & 0 & \tilde{f}_3 & \tilde{f}_4 \\ 0 & s_2 & s_3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$ 

As $\tilde{A}^k$ has no zero columns for any $k \in \mathbb{N}$ we see that no state is null controllable in finite time. Here $r(\tilde{A}) = 1.0024 > 1$ and although $\tilde{A}$ is not primitive (or even irreducible), $r(\tilde{A})$ is a simple eigenvalue and the following limit holds

$$\lim_{k \to \infty} (r(\tilde{A}))^k x_0 = \frac{v^T x_0}{v^T w} w,$$

where $v^T$ and $w$ are left and right eigenvectors of $\tilde{A}$ corresponding to $r(\tilde{A})$ respectively (which are both positive once positively scaled and satisfy $v^T w \neq 0$). When $x_0 \geq 0$ and $x_0 \neq 0$ the right hand side of (3.3) is positive and hence there are no non–trivial states that are positive state null controllable in infinite time. Equivalently, the negative control $u(t) = -f^T x(t)$ does not stabilise any nonzero initial population. Furthermore, the characterisation from Theorem 2.6 shows that this system cannot be stabilised by positive state control.

If instead the control action is in fact performed (shortly) after the census or measurement, then a more accurate model is

$$x(t + 1) = A(x(t) + bu(t)) = Ax(t) + Abu(t), \quad t \in \mathbb{N}_0,$$

and so we replace $b = e_4$ by $Ab = [\tilde{f}_2 \ \tilde{f}_4 \ 0 \ s_6]^T$. Corollary 2.3 can be applied to check whether assumption (A) holds for the pair $(A, Ab)$. Of rows one, two and four (the nonzero rows of $Ab$) the only possible candidate ‘smallest’ row of $A$ (in the sense of (2.8)) is the first (as rows two and four have nonzero entries that are zero in the first row). A straightforward calculation shows that (A) holds with $F = f^T = [0 \ 0 \ f_1/f_2 \ 1]$ if, and only if,

$$\tilde{f}_1 f_4 \leq \tilde{f}_2 f_3 \quad \text{and} \quad \tilde{f}_1 s_6 \leq \tilde{f}_2 s_5.$$ 

Both of these conditions are satisfied for the parameters in (3.2). Therefore, if $F = f^T = [0 \ 0 \ f_1/f_2 \ 1]$ then

$$\tilde{A} := A - Abf^T = \begin{bmatrix} 0 & 0 & \frac{-\tilde{f}_1 f_4}{f_2} & 0 \\ s_1 & 0 & \frac{-\tilde{f}_1 f_4}{f_3} & \frac{-\tilde{f}_1 f_4}{f_3} \\ 0 & s_2 & s_3 & 0 \\ 0 & s_4 & s_5 & \frac{-\tilde{f}_1 f_4}{f_6} \end{bmatrix} \geq 0,$$

and hence positive state control for the pair $(A, Ab)$ is precisely positive input control for the pair $(\tilde{A}, Ab)$. As the fourth column of $\tilde{A}$ is zero, we have that $x = e_4$ is null controllable (in finite time) and as $r(\tilde{A}) = 0.1153 < 1$, every state is positive state null controllable in infinite time. The above observations suggest that when control actions act on large weeds, organising these actions to take place post census is preferable to pre census. This is not biologically surprising because, loosely speaking, the fourth stage class is the most reproductive and the post census control strategy limits to a greater extent reproduction in this stage class.

Acknowledgements

Chris Guiver is fully supported and Dave Hodgson and Stuart Townley are partially supported by EPSRC grant EP/I019456/1.
References


