CONDITIOING OF THE STABLE, DISCRETE-TIME LYAPUNOV OPERATOR*  
MICHAEL K. TIPPETT†, STEPHEN E. COHN‡, RICARDO TODLING§, AND DAN MARCHESI††

Abstract. The Schatten \( p \)-norm condition of the discrete-time Lyapunov operator \( L_A \) defined on matrices \( P \in \mathbb{R}^{n \times n} \) by \( L_A P \equiv P - A P A^T \) is studied for stable matrices \( A \in \mathbb{R}^{n \times n} \). Bounds are obtained for the norm of \( L_A \) and its inverse that depend on the spectrum, singular values and radius of stability of \( A \). Since the solution \( P \) of the discrete-time algebraic Lyapunov equation (DALE) \( L_A P = Q \) can be ill-conditioned only when either \( L_A \) or \( Q \) is ill-conditioned, these bounds are useful in determining whether \( P \) admits a low-rank approximation, which is important in the numerical solution of the DALE for large \( n \).

Key words. Lyapunov matrix equation, condition estimates, large-scale systems, radius of stability.

AMS subject classifications. 15A12, 93C55, 93A15, 47B65

1. Introduction. Properties of the solution \( P \) of the discrete algebraic Lyapunov equation (DALE), \( P = A P A^T + Q \), are closely related to the stability properties of \( A \). For instance, the DALE has a unique solution \( P = P^T > 0 \) for any \( Q = Q^T > 0 \) if \( A \) is stable [11], a fact also true in infinite-dimensional Hilbert spaces [18]. In the setting treated here with \( A, Q, P \in \mathbb{R}^{n \times n} \), \( A \) is stable if its eigenvalues \( \lambda_i(A), i = 1, \ldots, n \), lie inside the unit circle; the eigenvalues are ordered so that \( |\lambda_1(A)| \geq |\lambda_2(A)| \geq \cdots \geq |\lambda_n(A)| \). Here \( A \) is always assumed to be stable.

In applications where the dimension \( n \) is very large, direct solution of the DALE or even storage of \( P \) is impractical or impossible. For instance, in numerical weather prediction applications \( A \) is the matrix that evolves atmospheric state perturbations. The DALE and its continuous-time analogs can be solved directly for simplified atmospheric models [6, 23], but in realistic models \( n \) is about \( 10^6 - 10^7 \) and even the storage of \( P \) is impossible. Krylov subspace [5] and Monte Carlo [9] methods have been used to find low-rank approximations of the right-hand side of the DALE and of the solution of the DALE [10].

The solution \( P \) of the DALE can be well approximated by a rank-deficient matrix if \( P \) has some small singular values. Therefore, it is useful to identify properties of \( A \) or \( Q \) that lead to \( P \) being ill-conditioned. If \( A \) is normal then

\[
\frac{\lambda_1(P)}{\lambda_n(P)} \leq \frac{\lambda_1(Q)}{\lambda_n(Q)} \frac{1 - |\lambda_n(A)|^2}{1 - |\lambda_1(A)|^2}
\]

(1.1)

the conditioning of \( P \) is controlled by that of \( Q \) and by the spectrum of \( A \). In the general case, the conditioning of \( Q \) and of the discrete-time Lyapunov operator \( L_A \) defined by \( L_A P \equiv P - A P A^T \) determine when \( P \) may be ill-conditioned.

**THEOREM 1.1.** Let \( A \) be a stable matrix and suppose that \( L_A P = Q \) for \( Q = Q^T > 0 \). Then

\[
\|P\|_p \|P^{-1}\|_p \leq \|L_A\|_p \|L_A^{-1}\|_p \|Q\|_p \|Q^{-1}\|_p, \quad p = \infty,
\]

(1.2)

* This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Grants 91.0029/95-4, 381737/97-7 and 30.0204/83-3, Financiadora de Estudos e Projetos (FINEP) Grant 77.97.0315.00, and the NASA EOS Interdisciplinary Project on Data Assimilation.
†IRI, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964-8000, USA (tippet-t@iri.ldeo.columbia.edu). This work was done while the author was with the Centro de Previsão de Tempo e Estudos Climáticos, Cachoeira Paulista, SP, Brazil.
‡Data Assimilation Office, Code 910.3, NASA/GSFC, Greenbelt, MD 20771, USA (cohn@dao.gsfc.nasa.gov).
§General Sciences Corp./SAIC, Code 910.3, NASA/GSFC/DAO, Greenbelt, MD 20771, USA (todling@dao.gsfc.nasa.gov).
††Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, Brazil (marchesi@impa.br).
where \( \| \cdot \|_p \) is the Schatten p-norm (see Eq. 2.2).

Theorem 1.1 (see proof in Appendix) follows from \( L^{-1}_A \) and its adjoint being positive operators. Therefore the same connection between rank-deficient approximate solutions and operator conditioning exists for matrix equations such as the continuous algebraic Lyapunov equation. We note that Theorem 1.1 also holds for \( 1 < p < \infty \) if either \( A \) is singular or \( \sigma_1^2(A) > 2 \); \( \sigma_1(A) \) is the largest singular value of \( A \).

Here we characterize the Schatten p-norm condition of \( L_A \). The main results are the following. Theorem 3.1 bounds \( \| L_A \|_p \) in terms of the singular values of \( A \). A lower bound for \( \| L_A^{-1} \|_p \) depending on \( \lambda_1(A) \) is presented in Theorem 4.1, generalizing results of [7]. Theorem 4.2 gives lower bounds for \( \| L_A^{-1} \|_1 \) and \( \| L_A^{-1} \|_{\infty} \) in terms of the singular values of \( A \). Theorem 4.6 gives an upper bound for \( \| L_A^{-1} \|_p \) depending on the radius of stability of \( A \) and generalizes results in [20]. Three examples illustrating the results are included. The issue of whether \( L_A \) and \( L_A^{-1} \) achieve their norms on symmetric, positive definite matrices is addressed in the concluding remarks.

2. Preliminaries. We investigate the condition number \( \kappa(L_A) = \| L_A \| \| L_A^{-1} \| \), where \( \| \cdot \| \) is a norm on \( \mathbb{R}^{n \times n} \) induced by a matrix norm on \( \mathbb{R}^{n \times n} \). Specifically, for \( M \in \mathbb{R}^{n \times n} \) we consider norms defined by

\[
\| M \|_p = \max_{S \neq 0} \frac{\| MS \|_p}{\| S \|_p}, \quad 1 \leq p < \infty,
\]

where the Schatten matrix p-norm for \( S \in \mathbb{R}^{n \times n} \) is defined by

\[
\| S \|_p = \left( \sum_{i=1}^{n} (\sigma_i(S))^p \right)^{1/p}.
\]

\( \sigma_i(S) \) are the singular values of \( S \) with ordering \( \sigma_1(S) \geq \sigma_2(S) \geq \cdots \geq \sigma_n(S) \geq 0 \). On \( \mathbb{R}^{n \times n} \), \( \| \cdot \|_2 \) is the Frobenius norm and \( \| \cdot \|_\infty = \sigma_1(\cdot) \). If \( S = S^T \geq 0 \) then \( \| S \|_1 = \text{tr} S \).

The following lemma about the Schatten p-norms follows from their being unitarily invariant [1, p. 94].

**Lemma 2.1.** For any three matrices \( X, Y \) and \( Z \in \mathbb{R}^{n \times n} \),

\[
\| XYZ \|_p \leq \| X \|_\infty \| Y \|_p \| Z \|_\infty, \quad 1 \leq p \leq \infty.
\]

The \( p = 2 \) Schatten norm on \( \mathbb{R}^{n \times n} \) is equivalently defined as \( \| S \|_2^2 = (S, S) \), where \( (\cdot, \cdot) \) is the inner product on \( \mathbb{R}^{n \times n} \) defined by \( (S_1, S_2) = \text{tr} S_1^T S_2 \). This norm corresponds to the usual Euclidean norm on \( \mathbb{R}^n \) since \( \| S \|_2^2 \) is equal to the sum of the squares of the entries of \( S \). As a consequence \( \kappa_2(L_A) = \sigma_1(L_A)/\sigma_n(L_A) \), where \( \sigma_1(L_A) \) and \( \sigma_n(L_A) \) are respectively the largest and smallest singular values of \( L_A \). The adjoint of \( L_A \) is given by \( L_A^* S = L_A^T S = S - A^T A \).

We now state some lemmas about mappings \( M \in \mathbb{R}^{n \times n} \) and about the spectra of \( L_A \) and \( A \).

**Lemma 2.2 ((15) of [2]).** \( \| M \|_p \leq \| M \|_1^{1/p} \| M \|_\infty^{1-1/p}, \quad 1 \leq p \leq \infty \).

**Lemma 2.3.** \( \| M \|_1 = \| M^* \|_\infty \).

**Lemma 2.4** (See proof of Theorem 1, [2]). If \( MS > 0 \) for all \( S \in \mathbb{R}^{n \times n} \) such that \( S > 0 \), then \( \| M \|_\infty = \| M^* \|_\infty \).

**Lemma 2.5** ([13, 14]). The \( n^2 \) eigenvalues of \( L_A \) are \( 1 - \lambda_i(A) \bar{\lambda}_j(A), \ 1 \leq i, j \leq n \).
3. The norm of the Lyapunov operator. If A is normal, then $L_A$ is normal, and its conditioning in the $p = 2$ Schatten norm depends only on its eigenvalues. Therefore when A is normal,

$$\|L_A^{-1}\|_2 = \frac{1}{\sigma_{n^2}(L_A)} = \frac{1}{|\lambda_{n^2}(L_A)|} = \frac{1}{1 - |\lambda_1(A)|^2},$$

and

$$\|L_A\|_2 = \sigma_1(L_A) = |\lambda_1(L_A)| = \max_{i,j} |1 - \lambda_i(A)\bar{\lambda}_j(A)|.$$  

For general A, the following theorem bounds $\|L_A\|_p$ in terms of the singular values of A.

**Theorem 3.1.**

$$|1 - \sigma_j^2(A)| \leq \max_j |1 - \sigma_j^2(A)| \leq \|L_A\|_p \leq 1 + \sigma_1^2(A), \quad 1 \leq p \leq \infty.$$  

**Proof.** Note that $L_A v_j v_j^T = v_j v_j^T - \sigma_j^2 u_j u_j^T$, where $u_j$ and $v_j$ are respectively the $j$-th left and right singular vectors of A such that $Av_j = \sigma_j u_j$. The lower bound follows from $\|u_j u_j^T\|_p = \|v_j v_j^T\|_p = 1$ and

$$\|L_A\|_p \geq \|v_j v_j^T - \sigma_j^2 u_j u_j^T\|_p \geq \|v_j v_j^T\|_p - \|\sigma_j^2 u_j u_j^T\|_p = |1 - \sigma_j^2|.$$  

The upper bound follows from

$$\|L_A P\|_p \leq \|P\|_p + \|APA^T\|_p \leq \|P\|_p + \|A\|_\infty^2 \|P\|_p.$$  

If A is normal, $\sigma_1(A)$ can be replaced by $|\lambda_1(A)|$ in Theorem 3.1, and $\|L_A\|_p \leq 1 + |\lambda_1(A)|^2$. If A is normal and $(-\lambda_1(A))$ is an eigenvalue of A, then $1 + |\lambda_1(A)|^2$ is an eigenvalue of $L_A$ and $\|L_A\|_p = 1 + |\lambda_1(A)|^2$.

Theorem 3.1 shows that $\|L_A\|_p$ is large and contributes to ill-conditioning if and only if $\sigma_1(A)$ is large, a situation that occurs in various applications [3, 22]. If $\sigma_1(A) \gg 1$ and $|\lambda_1(A)| < 1$, A is highly nonnormal [8, p. 314] and as Corollary 4.8 will show, close to an unstable matrix.

4. The norm of the inverse Lyapunov operator. We first show that a sufficient condition for $\|L_A^{-1}\|_p$ to be large is that $\lambda_1(A)$ be near the unit circle. The condition is necessary when A is normal.

**Theorem 4.1.** Let A be a stable matrix. Then

$$\|L_A^{-1}\|_p \geq \frac{1}{1 - |\lambda_1(A)|^2}, \quad 1 \leq p \leq \infty,$$

with equality holding if A is normal.

**Proof.** To obtain the lower bound, let $z_1$ be the leading eigenvector of A, $Az_1 = \lambda_1(A)z_1$, and note that $L_A z_1 z_1^H = (1 - |\lambda_1(A)|^2) z_1 z_1^H$ where $(\cdot)^H$ denotes conjugate transpose. Either $\text{Re} z_1 z_1^H \neq 0$ or $\text{Im} z_1 z_1^H \neq 0$ is an eigenvector of $L_A$, and it follows that $\|L_A^{-1}\|_p \geq (1 - |\lambda_1(A)|^2)^{-1}$. Finally, if A is normal, then

$$L_A^{-1} I = L_A^{-1} I = \sum_{i=1}^{n} \frac{1}{1 - |\lambda_i(A)|^2} z_i z_i^H.$$
and \( \|L_A^{-1}\|_\infty = \|L_A^{-1}\|_1 = (1 - \lambda_1(A)^2)^{-1} \). Using Lemma 2.2 gives \( \|L_A^{-1}\|_p \leq (1 - \lambda_1(A)^2)^{-1} \) when \( A \) is normal, and therefore \( \|L_A^{-1}\|_p = (1 - \lambda_1(A)^2)^{-1} \).

When \( A \) is nonnormal, \( \|L_A^{-1}\|_p \) can be large without \( \lambda_1(A) \) being near the unit circle. For instance, if \( \sigma_1(A) \) is large or more generally if \( \|A^k\|_\infty \) converges to zero slowly as a function of \( k \), then \( \|L_A^{-1}\|_p \) is large. We show this fact first for \( p = 1, \infty \).

**Theorem 4.2.** Let \( A \) be a stable matrix. For all \( m \geq 1 \),

\[
\begin{align*}
\|L_A^{-1}\|_1 &= \left\| \sum_{k=0}^{\infty} (A^k)^T A^k \right\|_1 \\
&\geq \left\| \frac{1}{m+1} \sum_{k=0}^{m} (A^k)^T A^k \right\|_1 + \frac{\sigma_n^{2(m+1)}(A)}{1 - \sigma_n^2(A)}, \\
\|L_A^{-1}\|_\infty &= \left\| \sum_{k=0}^{\infty} A^k (A^k)^T \right\|_\infty \\
&\geq \left\| \frac{1}{m+1} \sum_{k=0}^{m} A^k (A^k)^T \right\|_\infty + \frac{\sigma_n^{2(m+1)}(A)}{1 - \sigma_n^2(A)}. 
\end{align*}
\]

In particular,

\[
\|L_A^{-1}\|_p \geq 1 + \sigma_1^2(A) + \frac{\sigma_n^4(A)}{1 - \sigma_n^2(A)}, \quad p = 1, \infty. 
\]

**Proof.** The operator \( L_A^{-1} \) applied to \( S \in \mathbb{R}^{n \times n} \) can be expressed as [18]

\[
L_A^{-1} S = \sum_{k=0}^{\infty} A^k S (A^k)^T.
\]

Applying Lemma 2.4 gives \( \|L_A^{-1}\|_\infty = \|L_A^{-1}I\|_\infty \), with the inequality in (4.4) being a consequence of

\[
\left\| \sum_{k=0}^{\infty} A^k (A^k)^T \right\|_\infty \geq \left\| \sum_{k=0}^{m} A^k (A^k)^T \right\|_\infty + \lambda_n \left( \sum_{k=m+1}^{\infty} A^k (A^k)^T \right), \quad p = 1, \infty
\]

and

\[
\lambda_n \left( \sum_{k=m+1}^{\infty} A^k (A^k)^T \right) \geq \sum_{k=m+1}^{\infty} \lambda_n (A^k (A^k)^T) \geq \sum_{k=m+1}^{\infty} \sigma_n^{2k}(A) = \frac{\sigma_n^{2(m+1)}(A)}{1 - \sigma_n^2(A)},
\]

where we have used the facts that for matrices \( W, X, Y \in \mathbb{R}^{n \times n} \) with \( X, Y \) being symmetric positive semi-definite, \( \lambda_i(X + Y) \geq \lambda_i(X) + \lambda_n(Y) \) and \( \lambda_i(WXWT) \geq \sigma_n^2(W)\lambda_i(X) \) [17]. Likewise the \( p = 1 \) results follow from \( \|L_A^{-1}\|_1 = \|L_A^{-1}I\|_\infty \).

Lower bounds for \( 1 < p < \infty \) follow trivially, e.g.,

\[
\|L_A^{-1}\|_p = \left\| \frac{L_A^{-1}I}{I} \right\|_p \geq \frac{\|L_A^{-1}I\|_p}{\|I\|_p} \geq n^{-1/p} \|L_A^{-1}\|_\infty,
\]

but give little information when \( n \) is large. A lower bound for \( 1 \leq p \leq \infty \) depending on \( \sigma_1(A) \) and independent of \( n \) is given in Corollary 4.9.

We now relate \( \|L_A^{-1}\|_p \) to the distance from \( A \) to the set of unstable matrices as measured by its **radius of stability** [15].

**Definition 4.3.** For any stable matrix \( A \in \mathbb{R}^{n \times n} \) define the radius of stability \( r(A) \) by

\[
r(A) \equiv \min_{0 \leq \theta \leq 2\pi} \| (e^{i\theta} I - A)^{-1} \|_\infty = \min_{0 \leq \theta \leq 2\pi} \| R(e^{i\theta}, A) \|_\infty^{-1},
\]

\[ (4.10) \]
where the resolvent of \( A \) is \( R(\lambda, A) = (\lambda I - A)^{-1} \).

If \( A \) is normal and stable, then \( r(A) = 1 - |\lambda_1(A)| \). However, if \( A \) is nonnormal and if its eigenvalues are sensitive to perturbations, then \( r(A) \ll 1 - |\lambda_1(A)| \). The sensitivity of the eigenvalues of \( A \) is most completely described by its pseudospectrum [21]. The radius of stability \( r(A) \) is the largest value of \( \epsilon \) such that the \( \epsilon \)-pseudospectrum of \( A \) lies inside the unit circle; \( r(A) \) being small indicates that the \( \epsilon \)-pseudospectrum of \( A \) is close to the unit circle for small \( \epsilon \). The following theorem shows that when \( r(A) \) is small, \( \|L_A^{-1}\|_p \) must be large.

**THEOREM 4.4** (Proven for \( p = \infty \) in [7]). Let \( A \) be a stable matrix. Then

\[
\|L_A^{-1}\|_p \geq \frac{1}{2r(A) + r^2(A)}, \quad 1 \leq p \leq \infty .
\] (4.11)

**Proof.** There exists a matrix \( E \in \mathbb{R}^{n \times n} \) with \( |\lambda_1(A + E)| = 1 \) and \( \|E\|_{\infty} = r(A) \). Therefore there exists a vector \( x \) with \( x^H x = 1 \) such that \( (A + E)x = e^{i\theta}x \) for some \( 0 \leq \theta \leq 2\pi \). Using \( \|xx^H\|_p = 1 \) and Lemma 2.1 gives

\[
\|L_A xx^H\|_p = \|E x x^H E^T + e^{i\theta} x x^H E^T + e^{-i\theta} E x x^H\|_p
\leq \|E x x^H E^T\|_p + \|x x^H E^T\|_p + \|E x x^H\|_p
\leq \|E\|^2_\infty + 2\|E\|_{\infty} = r^2(A) + 2r(A),
\] (4.12)

and we have

\[
\|L_A^{-1}\|_p \geq \frac{\|L_A^{-1} L_A xx^H\|_p}{\|L_A xx^H\|_p} = \frac{1}{\|L_A xx^H\|_p} \geq \frac{1}{2r(A) + r^2(A)}.
\] (4.13)

A consequence of Theorem 4.4 is the following lower bound for \( r(A) \) in terms of \( \|L_A^{-1}\|_p \).

**COROLLARY 4.5.** Let \( A \) be a stable matrix. Then

\[
r(A) \geq \frac{\|L_A^{-1}\|_p^{-1}}{1 + \sqrt{1 + \|L_A^{-1}\|_p^{-1}}}, \quad 1 \leq p \leq \infty .
\] (4.14)

Bounds for \( r(A) \) are useful in robust stability [12] and in the study of perturbations of the discrete algebraic Riccati equation (DARE) [19]. In [19, Lemma 2.2] the bound

\[
r(A) \geq \frac{\|L_A^{-1}\|_\infty^{-1}}{\sigma_1(A) + \sqrt{\sigma_1^2(A) + \|L_A^{-1}\|_\infty^{-1}}},
\] (4.15)

was used to formulate conditions under which a perturbed DARE has a unique, symmetric, positive definite solution. Since the lower bound in (4.14) with \( p = \infty \) is sharper than that in (4.15) when \( \sigma_1(A) > 1 \), it can be used to show existence of a unique, symmetric, positive definite solution of the perturbed DARE for a larger class of perturbations [19, Theorem 4.1].

We generalize to Schatten \( p \)-norms the conjecture of [7] proven in [20] for the Frobenius norm.

**THEOREM 4.6.** Let \( A \) be a stable matrix. Then

\[
\|L_A^{-1}\|_p \leq \frac{1}{r^2(A)}, \quad 1 \leq p \leq \infty .
\] (4.16)
Proof. $L_A^{-1}I$ can be expressed as [20, 13],

$$L_A^{-1}I = \frac{1}{2\pi} \int_0^{2\pi} R(e^{i\theta}, A)R(e^{i\theta}, A)^H d\theta.$$  \hspace{1cm} (4.17)

Therefore, from Lemma 2.4,

$$\|L_A^{-1}\|_\infty = \|L_A^{-1}I\|_\infty \leq \frac{1}{2\pi} \int_0^{2\pi} \|R(e^{i\theta}, A)\|_\infty^2 d\theta \leq \frac{1}{r^2(A)}.$$ \hspace{1cm} (4.18)

The inequality (4.16) for $p = 1$ follows from $\|L_A^{-1}\|_1 = \|L_A^{-1}I\|_\infty$ and $r(A) = r(A^T)$. The theorem follows from Lemma 2.2.

As a consequence, any solution of the DALE can be used to obtain an upper bound for $r(A)$.  

**Corollary 4.7.** Let $A$ be a stable matrix and let $L_A P = Q$. Then

$$r^2(A) \leq \frac{\|Q\|_p}{\|P\|_p}, \hspace{1cm} 1 \leq p \leq \infty.$$ \hspace{1cm} (4.19)

Theorem 4.6 can be combined with any lower bound for $\|L_A^{-1}\|_p$ to obtain an upper bound for $r(A)$. For instance, from Theorem 4.2 we get the following upper bound.

**Corollary 4.8.** Let $A$ be a stable matrix. Then

$$r^2(A) \leq \frac{1}{1 + \sigma_1^2(A)}.$$ \hspace{1cm} (4.20)

Combining Corollary 4.8 and Theorem 4.4 gives a lower bound for $\|L_A^{-1}\|_p$.

**Corollary 4.9.** Let $A$ be a stable matrix. Then

$$\|L_A^{-1}\|_p \geq \frac{1 + \sigma_1^2(A)}{1 + 2\sqrt{1 + \sigma_1^2(A)}}, \hspace{1cm} 1 \leq p \leq \infty.$$ \hspace{1cm} (4.21)

**5. Examples.** We present three examples that illustrate how ill-conditioning of $L_A$ leads to low-rank approximate solutions of the DALE.

**Example 1. Almost unit eigenvalues.** Take $A = \lambda zz^T$ where $\lambda$ and $z$ are real, $0 < \lambda < 1$ and $z^Tz = 1$. The matrix $A$ is symmetric and $L_A$ is self-adjoint. The eigenvalues of $A$ are $(\lambda, 0, \ldots, 0)$. The operator $L_A$ has singular values (and eigenvalues) $(1, \ldots, 1, 1 - \lambda^2)$. Therefore $\|L_A\|_2 = 1$ and $1 \leq \|L_A\|_p \leq 1 + \lambda^2$ from Theorem 3.1. The norm of the inverse Lyapunov operator is

$$\|L_A^{-1}\|_p = \frac{1}{1 - \lambda^2}, \hspace{1cm} 1 \leq p \leq \infty,$$ \hspace{1cm} (5.1)

according to Theorem 4.1. As the eigenvalue $\lambda$ approaches the unit circle, $L_A$ is increasingly poorly conditioned. The solution of the DALE for this choice of $A$ is:

$$P = \frac{\lambda^2}{1 - \lambda^2} (z^T Q z) zz^T + Q.$$ \hspace{1cm} (5.2)

A "natural" rank-1 approximation $\bar{P}$ of $P$ is $\bar{P} = \lambda^2 (1 - \lambda^2)^{-1} (z^T Q z) zz^T$. As the eigenvalue $\lambda$ approaches the unit circle, if $(z^T Q z)$ is nonzero, $P$ is increasingly well-approximated by $\bar{P}$ in the sense that $\|P - \bar{P}\|_p/\|P\|_p$ approaches zero.
EXAMPLE 2. Large singular values. Take \( A = \sigma yz^T \) where \( \sigma > 0 \) and \( y \) and \( z \) are real unit \( n \)-vectors. The matrix \( A \) has at most one nonzero eigenvalue, namely \( \lambda = \sigma(y^Tz) \), taken to be less than one in absolute value. The sensitivity \( s \) of the eigenvalue \( \lambda \) is the cosine of the angle between \( y \) and \( z \), i.e., \( s = \lambda/\sigma \) for \( \lambda \neq 0 \), indicating that \( \lambda \) is sensitive to perturbations to \( A \) when \( \sigma \) is large [8].

Theorem 3.1 gives that \( 1 + \sigma^2 \geq \|L_A\|_p \geq |1 - \sigma^2| \), showing that \( \|L_A\|_p \) is large when \( \sigma \) is large. From Lemmas 2.3 and 2.4,

\[
\|L_A^{-1}\|_1 = \|L_A^{-1}\|_\infty = 1 + \frac{\sigma^2}{1 - \lambda^2},
\]

and it follows from Lemma 2.2 that \( \|L_A^{-1}\|_p \leq 1 + \sigma^2/(1 - \lambda^2) \). A lower bound for the \( p = 2 \) norm is

\[
\|L_A^{-1}\|_2 \geq \|L_A^{-1}zz^T\|_2 = \sqrt{1 + 2\frac{\lambda^2}{1 - \lambda^2} + \frac{\sigma^4}{(1 - \lambda^2)^2}}.
\]

The matrix \( A \) is near an unstable matrix when either \( |\lambda| \) is near unity or when \( \sigma \) is large since

\[
\left\| (e^{it}I - \sigma yz^T)^{-1} \right\|_\infty = \left\| e^{-it}I + \frac{\sigma e^{-2it}}{1 - \lambda^2} e^{it}yz^T \right\|_\infty \geq 1 + \frac{2|\lambda|}{1 - |\lambda|} + \frac{\sigma^2}{(1 - |\lambda|)^2}.
\]

Therefore \( r(A) \leq (1 - |\lambda|)/\sigma \) and a lower bound on \( \|L_A^{-1}\|_p \) follows from Theorem 4.4. When either \( |\lambda| \) is close to unity or when \( \sigma \) is large, \( r(A) \) is small and \( \kappa_p(L_A) \) is large.

The solution of the DALE is

\[
P = \frac{\sigma^2}{1 - \lambda^2} (z^TQz) yy^T + Q.
\]

When \( L_A \) is ill-conditioned and \( (z^TQz) \neq 0 \), the rank-1 matrix \( \tilde{P} = \sigma^2(1 - \lambda^2)^{-1}(z^TQz)yy^T \) is a good approximation of \( P \) in the sense that \( \|P - \tilde{P}\|_p/\|P\|_p \) is small.

EXAMPLE 3. Sensitive eigenvalues. Consider the dynamics arising from the one-dimensional advection equation, \( wt + w_x = 0 \) for \( 0 \leq x \leq n \), with boundary condition \( w(0, t) = 0 \). The matrix \( A \) that advances the \( n \)-vector \( w(x = 1, 2, \ldots, n, t) \) to \( w(x = 1, 2, \ldots, n, t = t_0 + 1) \) is the \( n \times n \) matrix with ones on the sub-diagonal and zero elsewhere, i.e., the transpose of an \( n \times n \) Jordan block with zero eigenvalue. Adding stochastic forcing with covariance \( Q \) at unit time intervals leads to the DALE, \( L_A P = Q \), where \( P \) is the steady-state covariance of \( w \).

Since \( \sigma_1(A) = 1 \), Theorem 3.1 yields \( 1 \leq \|L_A\|_p \leq 2 \). Further, since \( \|L_A\|_1 \geq \|L_Ae_1e_1^T\|_1 = \|e_1e_1^T - e_2e_2^T\|_1 = 2 \), where \( e_j \) is the \( j \)-th column of the identity matrix, \( \|L_A\|_1 = 2 \). A similar argument with \( L_A^T \) gives \( \|L_A\|_\infty = 2 \). Calculating \( L_A^{-1}I \) and \( L_A^{-1}I \) gives \( \|L_A^{-1}\|_\infty = \|L_A^{-1}\|_1 = n \). Therefore, using Lemma 2.2, \( \|L_A^{-1}\|_p \leq n \). Also,

\[
\|L_A^{-1}\|_2 \geq \frac{\|L_A^{-1}e_1e_1^T\|_2}{\|e_1e_1^T\|_2} = \sqrt{n}.
\]

A direct calculation shows that

\[
\left\| (e^{it}I - A)^{-1} \right\|_2 = \left\| \sum_{k=0}^{n-1} A^k e^{-i(k+1)t} \right\|_2^2 = \frac{n(n + 1)}{2},
\]
for any real \( \theta \). Since \( \sqrt{n}\| (e^{i\theta} I - A)^{-1} \|_\infty \geq \| (e^{i\theta} I - A)^{-1} \|_2 \), we have \( r^2(A) \leq 2/(n + 1) \). Theorem 4.4 then gives a lower bound for \( \| A^{-1} \|_p \), \( 1 \leq p \leq \infty \). Thus as \( n \) becomes large, that is, as the domain becomes large with respect to the advection length scale, \( A \) is increasingly ill-conditioned.

The elements \( P_{ij} \) of the solution \( P \) of the DALE are

\[
P_{ij} = e_i^T P e_j = \sum_{k=0}^{n-1} e_i^T A^k Q(A^T)^k e_j = \sum_{k=0}^{\min{(i-1,j-1)}} Q_{i-k,j-k}.
\] (5.9)

Therefore if \( Q = Q^T > 0 \), a “natural” rank-\( m \) approximation of \( P \) is the matrix \( \tilde{P} \) defined by

\[
\tilde{P}_{i,j} = \begin{cases} P_{i,j}, & n - m < i, j \leq n \\ 0, & \text{otherwise} \end{cases}.
\] (5.10)

When \( Q \) is diagonal, \( P \) is also diagonal and

\[
P_{ii} = \sum_{k=1}^{i} Q_{kk}.
\] (5.11)

In this case, each \( Q_{kk} > 0 \) and \( \tilde{P} \) is the best rank-\( m \) approximation of \( P \) in the sense of minimizing \( \| P - \tilde{P} \|_p \). We note that \( \tilde{P} \) is associated with the left-most part of the domain \( 0 \leq x \leq n \).

6. Concluding Remarks. Results about \( \| A^{-1} \|_p \) translate into bounds for solutions of the DALE. For instance, the solution \( P \) of the DALE for \( Q = Q^T \geq 0 \) satisfies

\[
\text{tr} \ P \leq \| A^{-1} \|_1 \text{tr} \ Q,
\] (6.1)

and the upper bound is achieved for \( Q = w_1 w_1^T \), where \( w_1 \) is the leading eigenvector of \( A^{-1} I \). In the \( p = \infty \) norm, \( A^{-1} \) achieves its norm on the identity. In the \( p = 2 \) norm, \( A^{-1} \) does not in general achieve its norm on the identity, and the question arises whether it achieves its norm on any symmetric, positive semi-definite matrix. The forward operator \( A \) does not in general assume its norm on a symmetric, positive semi-definite matrix. The following theorem states that \( A^{-1} \) does achieve its \( p = 2 \) norm on a symmetric, positive semi-definite matrix.

**Theorem 6.1.** There exists a matrix \( S = S^T \geq 0 \) such that \( \| A^{-1} S \|_2 / \| S \|_2 = \| A^{-1} \|_2 \).

**Proof.** Theorem 8 of [4] states that the inverse of the stable, continuous-time Lyapunov operator achieves its \( p = 2 \) norm on a symmetric matrix. The proof is easily adapted to give that \( A^{-1} \) achieves its \( p = 2 \) norm on a symmetric matrix. We now show that if \( A^{-1} \) achieves its \( p = 2 \) norm on a symmetric matrix, it does so on a symmetric, positive semi-definite matrix. Suppose that \( \| A^{-1} S \|_2 / \| S \|_2 = \| A^{-1} \|_2 \) and \( S \) is symmetric with Schur decomposition \( S = U D U^T \). Define the symmetric, positive semi-definite matrix \( S^+ = U |D| U^T \). Then \( \| S \|_2 = \| S^+ \|_2 \) and \(-S^+ \leq S \leq S^+\). The positiveness of the stable, discrete-time inverse Lyapunov operator mapping implies that \(-A^{-1} S^+ \leq A^{-1} S \leq A^{-1} S^+\), which implies that \( \| A^{-1} S \|_2 \leq \| A^{-1} S^+ \|_2 \). Therefore

\[
\frac{\| A^{-1} S \|_2}{\| S \|_2} = \frac{\| A^{-1} S \|_2}{\| S^+ \|_2} \leq \frac{\| A^{-1} S^+ \|_2}{\| S^+ \|_2}.
\] (6.2)
Additional information about the leading singular vectors of $L^{-1}_A$ could be useful for determining low-rank approximations of $P$. The power method can be applied to $L^{-1}_A L^{-1}_A$ to calculate the leading right singular vector and singular value of $L^{-1}_A$ [7]. However, this approach requires solving two DALEs at each iteration, which may be impractical for large $n$. If it is practical to store $P$ and to apply $L_A$ and $L_T$, a Lanczos method could be used to compute the trailing eigenvectors of $L_A L_T$ while avoiding the cost of solving any DALEs.

**Appendix.** Proof of Theorem 1.1. By definition, $||P||_p \leq ||L^{-1}_A||_p ||Q||_p$, and it remains to show that $||P^{-1}||_\infty \leq ||L^{-1}_A||_\infty ||Q^{-1}||_\infty$. Since $P = P^T > 0$, there is a nonzero $x \in \mathbb{R}^n$ such that

$$||P^{-1}||_\infty = \frac{1}{\lambda_n(P)} = \frac{x^T x}{x^T (L^{-1}_A Q) x} = \frac{\text{tr} xx^T}{\text{tr} \left( (L^{-1}_A)^{-1} xx^T \right) Q}.$$  (A.1)

Let $B = L^{-1}_A (xx^T)$ and note $B = B^T > 0$. Then using Lemma 2.3 and $\text{tr} Q Q \geq \lambda_n(Q) \text{tr} B$ gives

$$||P^{-1}||_\infty = \frac{\text{tr} L^{-1}_A B}{\text{tr} B} \leq \frac{\text{tr} L^{-1}_A B}{\lambda_n(Q)} \leq ||L^{-1}_A||_1 ||Q^{-1}||_\infty = ||L^{-1}_A||_\infty ||Q^{-1}||_\infty. \quad \Box$$  (A.2)

Theorem 1.1 holds for $1 \leq p \leq \infty$ given some restrictions on $A$. From [16], $\lambda_1(P) \geq \lambda_1(Q) + \sigma_1^2(A) \lambda_n(P)$, and it follows that $||P^{-1}||_p \leq ||Q^{-1}||_p$ for $1 \leq p \leq \infty$. From Theorem 3.1, $||L^{-1}_A||_p \geq 1$ if either $A$ is singular or $\sigma_1^2(A) \geq 2$. Therefore if either $A$ is singular or $\sigma_1^2(A) \geq 2$, $||P^{-1}||_p \leq ||L^{-1}_A||_p ||Q^{-1}||_p$, $1 \leq p \leq \infty$.  (A.3)

**Acknowledgments.** The authors thank Greg Gaspari for valuable observations and notation suggestions and the reviewer for useful comments.

**REFERENCES**


10

M. K. TIPPETT ET AL.