Correlation bounds for discrete-time systems with banded dynamics

Jaganath Chandrasekar, Dennis S. Bernstein*

Department of Aerospace Engineering, University of Michigan, 1320 Beal St., Ann Arbor, MI 48109-2140, United States

Received 2 August 2005; received in revised form 25 June 2006; accepted 8 July 2006

Available online 14 September 2006

Abstract

We consider the steady-state error covariance for a discrete-time system with banded dynamics. Such systems frequently arise from the spatial and temporal discretization of partial differential equations. In such systems, the magnitudes of the entries of the steady-state covariance matrix typically decrease as the distance from the diagonal increases. We obtain a bound on the entries of the covariance matrix beyond a given distance from the diagonal.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Kalman filtering; Banded dynamics; High-dimensional systems; Error covariance approximation; Lyapunov equation

1. Introduction

For discrete-time linear systems driven by possibly non-Gaussian stochastic disturbances, the state covariance is determined by a Lyapunov difference equation. When measurements are available for reducing the state uncertainty through estimation, the covariance of the state-estimate error evolves according to a Riccati difference equation. The structure of the state estimator is determined by the Kalman filter, which is widely used for data assimilation [1,11].

For high-dimensional systems, the computational burden of updating the covariance of the state estimate is \(O(n^3)\), where \(n\) is the order of the state. Consequently, it is common practice to work with a sparse approximation of the error covariance. For example, computationally efficient data assimilation techniques are used in [7] to estimate electron densities on a global (Earth-wide) grid. In particular, the dense correlation matrix is replaced by a sparse approximation obtained by neglecting (that is, zeroing) correlations between quantities in spatially discretized cells that lie along different magnetic flux lines.

Similarly, data assimilation techniques developed in [3] take advantage of the block-tridiagonal structure of the dynamics in computing the error covariance. To reduce the computational burden, only covariance matrix entries located within a specified distance of the diagonal are updated at each time step. Extensions of these methods to more general block-banded dynamics are applied to an ocean-circulation model in [2].

Motivated by these works, the goal of the present paper is to bound the error incurred when the covariance matrix for a system with banded dynamics is replaced by a sparse, banded approximation. Banded dynamics are a direct consequence of nearest-neighbor interactions in discretized partial differential equations [6, pp. 88–98, 9,8]. Since the steady-state covariance matrix is the solution to a linear matrix equation, the structure of the inverse of a banded matrix is of interest. Relevant literature includes the results of [10] on banded positive-semidefinite matrices whose off-diagonal entries are nonpositive. A bound on the entries of the inverse of an arbitrary banded matrix is given in [5], where the magnitudes of the entries of the inverse are shown to decay exponentially with distance from the diagonal. Since the steady-state error covariance matrix is the product of the inverse of a banded matrix and another matrix which may not be banded, the bound in [5] cannot be used to obtain similar bounds on the entries of the steady-state error covariance matrix.

Rather than bounding every entry of the steady-state covariance matrix, we obtain bounds on the off-diagonal entries...
of the steady-state covariance matrix for a linear system whose
dynamics are asymptotically stable and banded. These bounds
are given in terms of the norm of the neglected matrix entries.
We demonstrate these results on a compartmental model driven
by white noise.

The present paper is limited to the analysis of the covariance
of stochastically driven linear systems. Extensions to covari-
ance approximation in data assimilation algorithms for systems
with banded dynamics will be explored in future work.

2. Banded matrices

Let \( A \in \mathbb{R}^{n \times n} \) and assume that the nonzero entries of \( A \)
are restricted to a banded region around the main diagonal. We
define the semi-width \( \omega(A) \) of \( A \) to be
\[
\omega(A) = \min \{ l : A_{i,j} = 0 \text{ for all } i, j \text{ such that } |i - j| > l \}. 
\]

(2.1)

For example, if \( A \) is diagonal, then \( \omega(A) = 0 \); if \( A \) is tridiagonal,
then \( \omega(A) = 1 \); and if \( A \) is pentadiagonal, then \( \omega(A) = 2 \). Clearly,
\( \omega(A) \leq n - 1 \). It is easy to see that \( \omega(AB) \leq \omega(A) + \omega(B) \).
More generally, we have the following observation.

Proposition 2.1. Let \( A_1, \ldots, A_p \in \mathbb{R}^{n \times n} \). Then,
\[
\omega(A_1 \ldots A_p) \leq \min \left\{ n - 1, \sum_{i=1}^{p} \omega(A_i) \right\}. 
\]

(2.2)

3. Correlation bounds

Consider the linear time-invariant discrete-time system
\[
x_{k+1} = Ax_k + w_k, 
\]

(3.1)

where \( x_k, w_k \in \mathbb{R}^n \) and \( w_k \) is zero-mean white noise with
covariance \( Q \). Furthermore, we assume that \( A \) is asymptotically
stable, that is,
\[
\text{sprad}(A) < 1, 
\]

(3.2)

for all \( A \in \mathbb{R}^{n \times n} \), the spectral radius of \( A \) is defined by
\[
\text{sprad}(A) = \max \{ |\lambda| : \lambda \in \text{spec}(A) \}. 
\]

(3.3)

The positive-semidefinite state covariance \( P_k \triangleq \mathbb{E}[x_k x_k^T] \), where \( \mathbb{E}[\cdot] \) denotes the expected value, is updated using
\[
P_{k+1} = AP_k A^T + Q. 
\]

(3.4)

Since \( A \) is asymptotically stable and \( Q \) is positive semidefinite,
\( P \triangleq \lim_{k \to \infty} P_k \) exists and satisfies the discrete-time Lyapunov
equation
\[
P = APA^T + Q. 
\]

(3.5)

Furthermore,
\[
P = \sum_{i=0}^{\infty} A^i Q A^T. 
\]

(3.6)

Let \( \epsilon > 0 \) satisfy
\[
\text{sprad}(A) < \epsilon < 1, 
\]

(3.7)

so that
\[
\text{sprad} \left( \frac{1}{\epsilon} A \right) = \frac{1}{\epsilon} \text{sprad}(A) < 1. 
\]

(3.8)

It thus follows from (3.6) that
\[
P = \sum_{i=0}^{\infty} \epsilon^{2i} Q_i, 
\]

(3.9)

where \( Q_0 = Q \) and, for all \( i = 1, 2, \ldots, Q_i \) is defined by
\[
Q_i \triangleq \left( \frac{A}{\epsilon} \right)^i Q \left( \frac{A^T}{\epsilon} \right)^i. 
\]

(3.10)

Since \( \omega(A) = \omega(A) = \omega(A^T) \), it follows from (2.2) that, for
all \( i = 0, 1, \ldots \),
\[
\omega(Q_i) \leq \min \{ n - 1, 2i \omega(A) + \omega(Q) \}. 
\]

(3.11)

Next, for \( i = 0, \ldots, n - 1 \), define \( H_i \in \mathbb{R}^{n \times n} \) by
\[
H_i \triangleq \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\
\vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & 1 \\
0 & \cdots & 0 & 1 & \cdots & \cdots \\
\end{bmatrix}, 
\]

(3.12)

where the semi-width of the band of ones is chosen such that
\( \omega(H_i) = i \).

(3.13)

Now, for \( i = 0, \ldots, n - 1 \), define \( P_i \) by
\[
P_i \triangleq H_i \circ P, 
\]

(3.14)

where \( \circ \) denotes the Schur product. Then the \( (k, l) \) entry of \( P_i \)
is given by
\[
(P_i)_{k,l} = \begin{cases} P_{k,l} & \text{if } |k - l| \leq i, \\
0 & \text{else}. \end{cases} 
\]

(3.15)

For all \( j = 0, 1, \ldots, i = 0, \ldots, n - 1 \), if \( \omega(Q_j) \leq \omega(H_i) \), then
\( (1_n - H_i) \circ Q_j = 0 \), where \( 1_n \) is the \( n \times n \) matrix whose entries are
all equal to 1. Therefore, for \( i = 0, \ldots, n - 1 \), taking the Schur
product of (3.9) with \( 1_n - H_i \) and using \( (1_n - H_i) \circ P = P - P_i \)
yields
\[
P - P_i = \sum_{j=L(i)}^{\infty} \epsilon^{2j} (1_n - H_i) \circ Q_j, 
\]

(3.16)

where \( L : \mathbb{N} \to \mathbb{N} \) is defined by
\[
L(i) \triangleq \max \left\{ 0, \left\lfloor \frac{i - \omega(Q)}{2\omega(A)} + 1 \right\rfloor \right\}. 
\]

(3.17)
Proposition 3.1. Assume that \( A \in \mathbb{R}^{n \times n} \) satisfies (3.2) and let \( \varepsilon > 0 \) satisfy \( \text{sprad}(A) < \varepsilon < 1 \). Let \( \| \cdot \| \) be a norm on \( \mathbb{R}^{n \times n} \). Then,
\[
\sigma_A \triangleq \max_{i \in \mathbb{N}} \frac{1}{\varepsilon} \| A^i \| (3.18)
\]
exists.

Proof. It follows from (3.8) that \( \lim_{i \to \infty} (1/\varepsilon) A^i = 0 \). Hence, \( \sigma_A \) exists. \( \square \)

Proposition 3.2. Assume that \( A \in \mathbb{R}^{n \times n} \) satisfies (3.2) and let \( \varepsilon > 0 \) satisfy \( \text{sprad}(A) < \varepsilon < 1 \). Let \( \| \cdot \| \) be a monotonic submultiplicative norm on \( \mathbb{R}^{n \times n} \). Then, for all \( i = 0, \ldots, n - 1, \)
\[
\| P - P_i \| \leq \frac{\varepsilon^{2L(i)}}{1 - \varepsilon} \| Q \|. (3.19)
\]

Proof. Since \( \| \cdot \| \) is monotonic, it follows that, for all \( i = 0, \ldots, n - 1 \) and \( j = 0, 1, \ldots, \)
\[
\| (1_n - H_i) \circ Q_j \| \leq \| Q_j \|. (3.20)
\]
Furthermore, since \( \| \cdot \| \) is submultiplicative, it follows that, for all \( j = 0, 1, \ldots, \)
\[
\| Q_j \| \leq \| Q \| \| A_j \|^2. (3.21)
\]
Hence, it follows from Proposition 3.1 that, for all \( j = 0, 1, \ldots, \)
\[
\| Q_j \| \leq \| Q \| \sigma_A^2. (3.22)
\]
Taking the norm of \( P - P_i \) in (3.16) and using (3.20) yields
\[
\| P - P_i \| \leq \varepsilon^{2L(i)} \| Q \| L(i) + \varepsilon^{2L(i) + 2} \| Q \| L(i) + 1 \| + \cdots. (3.23)
\]
It then follows from (3.22) that
\[
\| P - P_i \| \leq \sigma_A^2 \| Q \| (\varepsilon^{2L(i)} + \varepsilon^{2L(i) + 2} + \cdots). (3.24)
\]
Since \( 0 < \varepsilon < 1, \)
\[
\sum_{j=L(i)}^{\infty} \varepsilon^{2j} = \frac{\varepsilon^{2L(i)}}{1 - \varepsilon}. (3.25)
\]
Therefore, (3.24) and (3.25) imply (3.19). \( \square \)

4. Compartmental model example

We consider a system comprised of \( n \) compartments or subsystems that exchange energy through mutual interaction [4]. Applying conservation of energy yields, for \( i = 1, \ldots, n, \)
\[
x_i(k + 1) = x_i(k) - \beta x_i(k) + \varepsilon (x_{i+1}(k) - x_i(k)) - \varepsilon (x_i(k) - x_{i-1}(k)), (4.1)
\]

Table 1

<table>
<thead>
<tr>
<th>( x )</th>
<th>( \beta )</th>
<th>( \text{sprad}(A) )</th>
<th>( \varepsilon )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.8</td>
<td>0.2</td>
<td>0.4, 0.3, 0.21</td>
</tr>
</tbody>
</table>

Table 1 Parameters used in the compartmental model example

where \( 0 < \beta < 1 \) is the loss coefficient and \( 0 < \alpha < 1 \) is the flow coefficient. It follows from (4.1) that
\[
x(k + 1) = Ax(k), (4.2)
\]
where
\[
x = [x_1 \cdots x_n]^T (4.3)
\]
and \( A \in \mathbb{R}^{n \times n} \) is defined by
\[
A = \begin{bmatrix}
1 - \beta - \alpha & \alpha & 0 & \cdots & 0 \\
\alpha & 1 - \beta - 2\alpha & \alpha & \cdots & 0 \\
0 & \alpha & 1 - \beta - 2\alpha & \alpha & \cdots \\
& \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & 0 & \alpha & 1 - \beta - \alpha
\end{bmatrix}. (4.4)
\]

Since \( A \) is tridiagonal, \( \omega(A) = 1 \). We choose \( n = 20 \) and evaluate \( P \) using (3.5) with \( Q = I_n \) for \( (\alpha, \beta) = (0.1, 0.8) \). The spectral
radius of $A$, and the chosen value of $\varepsilon$ are shown in Table 1. We choose $\| \cdot \|$ to be the Frobenius norm $\| \cdot \|_F$.

Note that for $(\alpha, \beta) = (0.1, 0.8)$, $\text{sprad}(A) < 1$ and hence, $\sigma_A$ defined in (3.18) exists and is determined numerically. Next, for $i = 0, \ldots, 9$, we plot $(\varepsilon^{2L(i)}/(1 - \varepsilon^2))\sigma_A^2 \| Q \|_F$ and $\| P - P_i \|_F$ with $(\alpha, \beta) = (0.1, 0.8)$ in Fig. 1. Note that $\| Q \|_F = \sqrt{20}$. The magnitudes of the entries of the steady-state covariance $P$ for $(\alpha, \beta) = (0.1, 0.8)$ are plotted in Fig. 2. It can be seen that the magnitude of the entries of the covariance decrease as the distance from the diagonal increases.

References