Decay Rates for Inverses of Band Matrices

By Stephen Demko*, William F. Moss and Philip W. Smith**

Abstract. Spectral theory and classical approximation theory are used to give a new proof of the exponential decay of the entries of the inverse of band matrices. The rate of decay of $A^{-1}$ can be bounded in terms of the (essential) spectrum of $AA^*$ for general $A$ and in terms of the (essential) spectrum of $A$ for positive definite $A$. In the positive definite case the bound can be attained. These results are used to establish the exponential decay for a class of generalized eigenvalue problems and to establish exponential decay for certain sparse but nonbanded matrices. We also establish decay rates for certain generalized inverses.

1. Introduction. The exponential decay of the entries of inverse of band matrices has been of some use in establishing local rates of convergence of spline approximations [12], [11], [6] and in bounding the $L_\infty$-norm of the orthogonal projection onto spline spaces [4] and [15]. Kershaw proved a result of this nature for tridiagonal matrices and Descloux's paper [7] contains such a result for Grammian matrices arising in finite element approximations although exponential decay is not explicitly mentioned. For general banded invertible matrices the first proof appeared in [6]. A later proof in [3], [5] gave explicit estimates for the rate of decay. In this paper we use spectral theory and a result of Chebyshev on the best approximation of $(x - a)^{-1}$ by polynomials to give a new proof. The bounds on the rate of decay obtained from this proof appear to be sharper than those previously known and are actually attained in some cases. In addition, the method of proof easily extends to certain generalized inverses and certain nonbanded matrices. We show that the rate of decay for $A^{-1}$ given by our method depends on only the essential spectrum of $AA^*$ and is, thus, stable under banded compact perturbations. This fact is used to establish the exponential decay of the eigenvectors of certain generalized eigenvalue problems.

There are two key ideas. First use spectral theory to write

$$\|A^{-1} - p(A)\| = \max_{x \in \sigma(A)} |1/x - p(x)|$$

for any positive definite operator $A$ and any real polynomial $p$. Secondly, use approximation theory to estimate the best error.
2. Exponential Decay. We begin with a few preliminary remarks and definitions. Let $H$ be a complex, separable, Hilbert space and let $B(H)$ denote the Banach algebra of all bounded linear operators on $H$. If $A \in B(H)$ then we can represent $A$ as a matrix with respect to any complete orthonormal set. Once a particular representation is chosen, we may regard $A$ as an element of $B(l^2(S))$ where $S = \{1, \ldots, N\}$, $Z^+$, or $Z$. In this setting the usual matrix product defines the action of $A$. Throughout this paper we will assume that the above identifications have been made and will reserve the letter $A$ for a matrix representing a bounded operator in $B(l^2(S))$. For such matrices $A$ we will say that $A$ is $m$-banded if there is an index $l$ so that

$$A(i, j) = 0 \quad \text{if} \quad j \not\in [i - l, i - l + m].$$

We will say that $A$ is centered and $m$-banded if $m$ is even and the $l$ above may be chosen to be $m/2$. Thus, for a centered $m$-banded matrix one has

$$A(i, j) = 0 \quad \text{if} \quad |i - j| > m/2.$$

Notice that selfadjoint matrices are naturally centered and, for example, a tridiagonal selfadjoint matrix is centered and 2-banded.

As mentioned in the introduction we combine certain approximation theoretic results with the spectral theorem in order to obtain estimates for the exponential decay rate of $A^{-1}$. Let $\pi_n$ denote the polynomials of degree less than or equal to $n$. If $K$ is a subset of the complex plane $\mathbb{C}$ and $f$ is a fixed complex-valued function on $K$ we define

$$\|f\|_K := \sup\{|f(z)| : z \in K\}, \quad e_n(K) := \inf\{\|f - p\|_K : p \in \pi_n\}.$$

We are now ready to state a proposition which is just a corollary of a result of Chebyshev [14, p. 33], but which is of fundamental importance to all that follows.

**Proposition 2.1.** Let $f(x) = 1/x$ and let $0 < a < b$. Set $r = b/a$ and

$$q := q(r) := (\sqrt{r} - 1)/(\sqrt{r} + 1).$$

Then

$$e_n([a, b]) = \frac{(1 + r^{1/2})^2}{2ar} q^{n+1}.$$

This exponential rate of approximation readily yields exponential decay of the inverse of a banded positive definite matrix as the next proposition shows. We will let $\sigma(A)$ denote the spectrum of the matrix $A$.

**Proposition 2.2.** Let $A$ be a positive definite, $m$-banded, bounded and boundedly invertible matrix in $l^2(S)$. Let $[a, b]$ be the smallest interval containing $\sigma(A)$. Set $r = b/a$, $q = q(r)$ as in (2.1), and set $C_0 = (1 + r^{1/2})^2/(2ar)$ and $\lambda = q^2/m$. Then we have

$$|A^{-1}(i, j)| \leq C\lambda^{|i-j|}$$

where

$$C := C(a, r) := \max\{a^{-1}, C_0\}.$$

**Proof.** Since $A$ is positive definite and invertible we have $0 < a < b$ and we know that $A$ is centered. Thus $A^k$ is centered and $km$-banded for $k = 0, 1, 2, \ldots$. Thus if
$p \in \pi_k$ then $p(A)$ is $km$-banded and centered. From Proposition 2.1 we know there exists a sequence of polynomials $p_n \in \pi_n$ satisfying

$$\|1/x - p_n\|_{a,b} = C_0 q^{n+1}.$$ 

An application of the spectral theory [16] yields

$$\|A^{-1} - p_n(A)\| = \|1/x - p_n\|_{\sigma(A)} \leq C_0 q^{n+1}.$$ 

Now writing $|i - j| = nm/2 + k$ for $k = 1, \ldots, m/2$ and $i \neq j$, we see that $|i - j|2/m \leq (n + 1)$ and hence

$$|A^{-1}(i, j)| = |A^{-1}(i, j) - p_n(A)(i, j)| \leq \|A^{-1} - p_n(A)\| \leq C_0 \lambda_{i-j}.$$ 

In case $i = j$ note that $1/a = \|A^{-1}\|$, and (2.3) follows. This completes the proof of Proposition 2.2.

The phenomenon of exponential decay is certainly not restricted to positive definite band matrices. In fact, we can use the preceding proposition to prove a more universal result. For any matrix $A$, we will say that $A$ is quasi-centered if the central diagonal (i.e., $(i, i)$) is contained within the nontrivial bands. For example if $S = \{1, \ldots, N\}$ or $Z^+$ then $A \in B(l^2(S))$ is invertible only if $A$ is quasi-centered. This is of course not true for $A \in l^2(Z)$.

**Proposition 2.3.** Let $A$ be $m$-banded, bounded and boundedly invertible on $l^2(S)$. Let $[a, b]$ be the smallest interval containing $\sigma(\lambda_1 AA^*)$. Then setting $r = b/a$, $q = q(r)$ as in (2.1), and $\lambda_1 = q^{1/m}$, there is a constant $C_1$ depending on $A$ so that

$$|A^{-1}(i, j)| \leq C_1 \lambda_{i-j}.$$ 

If $A$ is quasi-centered then we may choose $C_1 = (m + 1)\|A\|^{-m}C(a, r)$.

This result follows immediately from Proposition 2.2, the observation that $A^{-1} = A^*(\lambda_1 AA^*)^{-1}$ and the fact that $\|A\| = \|A^*\|$.

We collect these results in a theorem below which will allow us to compare them with earlier estimates. First let us note that if $A$ is positive definite and invertible then $\|A\| = \max\{s : s \in \sigma(A)\}$. We will set $\text{cond}(A) := \|A\| \|A^{-1}\|$. Note that since $B(l^2(S))$ is a $B^*$-algebra, then

$$\text{cond}(\lambda_1 AA^*) = [\text{cond}(A)]^2.$$ 

**Theorem 2.4.** Let $A$ and $A^{-1}$ be in $B(l^2(S))$. Then if $A$ is positive definite and $m$-banded we have

$$|A^{-1}(i, j)| \leq C_1 \lambda_{i-j},$$ 

where

$$\lambda = \left(\sqrt{\text{cond}(A) - 1} - \sqrt{\text{cond}(A)} + 1\right)^{2/m}$$

and

$$C = \|A^{-1}\| \max\left\{1, \left(1 + \sqrt{\text{cond}(A)}\right)^2/(2 \text{cond}(A))\right\}.$$
If \( A \) fails to be positive definite but is still \( m \)-banded, quasi-centered, bounded, and boundedly invertible then

\[
A^{-1}(i, j) \leq C_1 \lambda_1^{i-j},
\]

where

\[
(2.9) \quad \lambda_1 = \left( \frac{\text{cond}(A) - 1}{\text{cond}(A) + 1} \right)^{1/m}
\]

and

\[
(2.10) \quad C_1 = (m + 1) \lambda_1^{-m} \|A^{-1}\| \text{cond}(A) \max \left\{ 1, \left[ \frac{1 + \text{cond}(A)}{\text{cond}(A)} \right]^2 \right\}.
\]

We remind the reader that if \( S = \mathbb{Z}^+ \) or \( \{1, \ldots, N\} \) then the quasi-centered hypothesis is redundant, and if \( S = \mathbb{Z} \), we may reindex \( A \) by a shift so that it is quasi-centered.

3. Examples and Comparisons. In the previous section we derived several results concerning exponential decay of the inverse of band matrices, culminating in Theorem 2.4. In this section we will present some examples indicating the precision of the estimates in Theorem 2.4 as well as comparing this result with the earlier estimates of [5].

We first show that the exponential rate in (2.5) can indeed be attained. Let \( 0 < s < 1 \) and consider the tridiagonal Toeplitz matrix \( A = \text{trid}(s, 1 + s^2, s) \) in \( l^2(\mathbb{Z}) \) with symbol

\[
A(z) = (z + s)(z^{-1} + s).
\]

Now it is well known [9] that

\[
\sigma(A) = \{ A(z) : |z| = 1 \} = \left[ (1 - s)^2, (1 + s)^2 \right].
\]

Thus from Proposition 2.2 or Theorem 2.4 (2.5) we have for \( i \neq j \)

\[
|A^{-1}(i, j)| \leq \frac{2}{(1 - s^2)^2} s^{\|i-j\|}.
\]

On the other hand we can compute \( A^{-1} \) directly using the Laurent expansion for \( 1/A(z) \) which converges in a neighborhood of the unit circle yielding

\[
A^{-1}(i, j) = (1 - s^2)^{-1} (-s)^{\|i-j\|}.
\]

Thus we predict the precise exponential decay rate in this case but our constant is pessimistic. This should come as no surprise since there is clearly something lost in the inequality \( |A^{-1}(i, j)| \leq \|A^{-1} - p_n(A)\| \) which is used in the proof of Proposition 2.2.

Next, consider the positive definite matrix \( A \) with symbol

\[
A(z) = (z + \frac{1}{2})(z^{-1} + \frac{1}{2})(z + \frac{1}{4})(z^{-1} + \frac{1}{4}).
\]

In this case one may verify that

\[
\sigma(A) = [9/64, 225/64].
\]
Thus Proposition 2.2 or Theorem 2.4 would predict an exponential decay rate of \( \sqrt[3]{2/3} \approx 0.82 \) whereas it follows from the Laurent series for \( 1/A(z) \) that
\[
|A^{-1}(i, j)| \leq C2^{-|i-j|}.
\]
In this case we see that a conservative estimate has been obtained.

In the \( l^2(S) \) setting [5] obtained the estimate
\[
\left( \frac{(\text{cond}(A))^2 - 1}{(\text{cond}(A))^2 + 1} \right)^{1/2m}
\]
for the decay rate of the inverse of an arbitrary \( m \)-banded matrix. We note that the \( \lambda_1 \) in (2.9) is smaller. However, de Boor also obtained results for all \( p \). In particular, as Demko has pointed out [6], once it is known that a banded matrix \( A \) is bounded and boundedly invertible on some \( l^p(S) (1 \leq p \leq \infty) \) then in fact it is bounded and boundedly invertible on all \( l^p(S) \). Thus setting
\[
\text{cond}_p(A) = \|A\|_p\|A^{-1}\|_p
\]
de Boor showed [5] that the exponential decay rate for the inverse of an \( m \)-banded matrix is bounded by
\[
\left( \frac{(\text{cond}_p(A))^p - 1}{(\text{cond}_p(A))^p + 1} \right)^{1/pm}
\]
Thus, for instance, if \( A \) is totally positive Toeplitz then one can see [10, Theorem 6] by a result of de Boor that \( \text{cond}_p(A) \) is constant in \( p \) and hence the best choice for \( p \) in the above setting would be \( p = 1 \). Interestingly, this reduces precisely to our \( \lambda_1 \) in (2.9). If we had assumed in addition that \( A \) was symmetric then we would have had the better estimate \( \lambda \) in (2.6). It is not easy to compare these results but we feel confident that Theorem 2.4 is superior if for no other reason than finding the best \( p \) in (3.1) seems to be a difficult task in general.

4. Extensions and Eigenvectors. This section is devoted to certain refinements of Theorem 2.4. The results in this section were motivated by some eigenvalue computations of the second and third authors which will be reported elsewhere. For the rest of this section we will assume that \( S \) is either \( Z \) or \( Z^* \).

As mentioned earlier \( B := B(l^2(S)) \) is a \( B^* \)-algebra and the set \( \mathcal{C} := \mathcal{C}(l^2(S)) \) of compact operators is a two-sided ideal in \( B(l^2(S)) \). Then the following facts are known [1] and [8]. The quotient space \( B/\mathcal{C} \) is a \( B^* \)-algebra and the quotient map
\[
\nu: B \rightarrow B/\mathcal{C}
\]
satisfies
\[
\sigma_e(A) := \sigma(\nu(A)) \subseteq \sigma(A)
\]
where we use the symbol \( \sigma_e \) to denote the essential spectrum of \( A \), that is, the spectrum of the element \( \nu(A) \) in \( B/\mathcal{C} \). The object \( B/\mathcal{C} \) is known as the Calkin algebra. In addition the norm in the Calkin algebra is called the essential norm and will be denoted by
\[
\|A\|_e := \|\nu(A)\| := \inf\{\|A - K\| : K \in \mathcal{C}\}.
\]
Of course, the essential condition number will be
\[ \text{cond}_e(A) := \|A\|_e \|A^{-1}\|_e. \]
Finally we remark that if \( A \) is selfadjoint and \( U \) is an open neighborhood of \( \sigma_e(A) \), then the complement of \( U \) contains at most finitely many points of \( \sigma(A) \).

Recall that in the proof of Proposition 2.2 we used the inequality
\[ \|A^{-1} - P_n(A)\| = \|1/x - P_n\|_{\sigma(A)} \leq \|1/x - P_n\|_{[a,b]} \]
where \( \sigma(A) \subset [a, b] \). Now it is quite possible that the inequality above does not always yield sharp estimates. This is indeed the case as will be illustrated by the following theorem whose proof is quite similar to the proof of Theorem 2.4.

**Theorem 4.1.** Let \( A \) and \( A^{-1} \) be in \( B(l^2(S)) \). Then if \( A \) is positive definite and \( m \)-banded set
\[
\lambda = \left( \frac{\sqrt{\text{cond}_e(A)} - 1}{\sqrt{\text{cond}_e(A)} + 1} \right)^{2/m}.
\]
For any \( \gamma > \lambda \) there is a constant \( C_2 = C_2(\gamma, A) \) so that
\[
|A^{-1}(i, j)| \leq C_2 \gamma^{i-j}.
\]
If \( A \) fails to be positive definite but is quasi-centered, \( m \)-banded, bounded and boundedly invertible set
\[
\lambda_1 = \left( \frac{\text{cond}_e(A) - 1}{\text{cond}_e(A) + 1} \right)^{1/m}.
\]
For any \( \gamma > \lambda_1 \) there is a constant \( C_3 = C_3(\gamma, A) \) so that
\[
|A^{-1}(i, j)| \leq C_3 \gamma^{i-j}.
\]

**Proof.** Of course, (4.4) follows from (4.2) by considering \( A^{-1} = A^*(AA^*)^{-1} \) just as in Proposition (2.3). In order to prove (4.2) we use an approximation lemma.

**Lemma 4.2.** Let \( D = [a, b] \cup \{x_1, \ldots, x_k\} \) be a subset of the positive real axis and let \( f(x) = 1/x \). Then there is a constant \( R = R(x_1, \ldots, x_k, a, b) \) so that
\[
e_n(D) \leq R q^n.
\]
where \( q = q(b/a) \) as in (2.1).

**Proof.** Let \( r \in \pi_{k-1} \) interpolate \( f \) at \( x_1, \ldots, x_k \) and let \( \psi(x) = (x - x_1) \cdots (x - x_k) \). Note that
\[
\frac{1 - xr}{\psi} = -f[x_1, \ldots, x_k] = F_k.
\]
Now compute
\[
\inf_{p \in \pi_n} \|f - p\|_D \leq \inf_{p \in \pi_{n-k}} \|f - (r + \psi p)\|_D \leq \inf_{p \in \pi_{n-k}} \|\psi\|_{[a,b]} \inf_{p \in \pi_{n-k}} \|\psi\|_{[a,b]} \|\frac{f - r}{\psi} - p\|_{[a,b]} = \|\psi\|_{[a,b]} F_k \inf_{p \in \pi_{n-k}} \|f - p\|_{[a,b]} = \|\psi\|_{[a,b]} F_k e_{n-k}([a,b]) \leq R q^{n-k+1},
\]
where the last inequality follows from Proposition 2.1. This completes the proof of the lemma.
Returning to the proof of (4.2), let $\gamma > \lambda$ be given, let $[a, b]$ be the smallest interval containing $\sigma_e(A)$, and choose $\epsilon > 0$ so that $0 < a - \epsilon$ and
\[
\left( q\left( \frac{b + \epsilon}{a - \epsilon} \right) \right)^{2/m} < \gamma.
\]
Then there exists a number $k$ and $0 < x_1 < \cdots < x_k$ so that
\[
\sigma(A) \subset (a - \epsilon, b + \epsilon) \cup x_1 \cup \cdots \cup x_k = D.
\]
Thus by Lemma 4.2 we can find polynomials $p_n \in \pi_n$ satisfying
\[
\|f - p_n\|_D \leq \left( Rq^{-k} \right) q^{n+1}.
\]
The proof of Theorem 4.1 is now completed in the same manner as Proposition 2.2.

Theorem 4.1 can yield significantly better estimates than Theorem 2.4. For example

**Corollary 4.3.** Let $A = I + K$ be banded where $K$ is compact and $I$ is the identity operator. Then if $I + K$ is invertible we have for any $\gamma > 0$: there is a constant $C = C(\gamma, A)$ so that
\[
|A^{-1}(i, j)| \leq C\gamma^{l|i-j|}.
\]

We refer to such behavior as decay faster than any exponential. This corollary follows from Theorem 4.1 by noting that $\text{cond}_e(A) = 1$ and hence $\lambda_1 = 0$ in (4.3).

Exponential decay can also be observed in certain generalized eigenvalue problems.

**Corollary 4.4.** Let $A$ be $m$-banded, bounded and boundedly invertible. Let $K$ be compact so that $A + K$ is $m$-banded. Suppose $\mu \neq 0$ and $x \in l^2(S)$, $x \neq \theta$, satisfy $(A - \mu K)x = \theta$. Then for any $\gamma > \lambda_1$ ($\lambda_1$ as in (4.3)), there is a constant $C = C(\gamma, A, K)$ so that
\[
|x(i)| \leq C\gamma^{|i|}.
\]

**Proof.** Define the projections $P_N$ by
\[
P_N e_j = \begin{cases} 
e_j & \text{if } |j| \leq N, \\ \theta & \text{otherwise.} \end{cases}
\]

Here $e_j$ are the standard unit vectors, $e_j(i) = \delta_{ij}$. Set $K_N = P_N K P_N$ and note that $K_N$ is $m$-banded and $\|K - K_N\| \to 0$ as $N \to \infty$. It follows that for some $N$, $A - \mu(K - K_N)$ is boundedly invertible and hence
\[
x = \mu \left( A - \mu(K - K_N) \right)^{-1}(K_N x).
\]

This yields the result since
\[
\text{cond}_e(A - \mu(K - K_N)) = \text{cond}_e(A)
\]
and $K_N x$ is finitely supported so that $x$ can be seen to be a finite linear combination of the columns of $(A - \mu(K - K_N))^{-1}$.

We close this section with a result on generalized inverses of band matrices which is valid in both finite and infinite dimensions.

**Corollary 4.5.** Let $A$ be an $m$-banded, bounded linear operator from $l^2(S_1)$ onto $l^2(S_2)$ where $S_1$ and $S_2$ are chosen from $\{1, \ldots, N\}$, $Z^+$ or $Z$ and let $A^\dagger$ denote the
Moore-Penrose inverse. Let $\lambda_1$ be as in (4.3) if $S_2$ is infinite dimensional or as in (2.9) otherwise. Then for any $\gamma > \lambda_1$ there is a constant $C = C(\gamma, A)$ so that

$$|A^t(i, j)| \leq C\gamma^{k-n}.$$

Proof. By the open mapping theorem $AA^*$ is boundedly invertible on $l^2(S)$. Now $A^* = A^*(AA^*)^{-1}$ and the result follows just as in Proposition 2.3. For more information on generalized inverses the reader may consult [2].

5. Remarks. It has been recognized for some time that exponential decay is manifested in the inverses of sparse (but not necessarily banded) matrices, see, for example, [3]. The technique of Section 2 (and its extensions in Section 4) yields such a result in a simple way. For any matrix $A \in l^2(S)$ let us define the support sets

$$S_n(A) := \bigcup_{k=0}^{n} \{(i, j) : A^k(i, j) \neq 0\}$$

and the decay sets

$$D_n(A) := (S \times S) \setminus S_n(A).$$

Proposition 5.1. Let $A$ be positive definite bounded and boundedly invertible on $l^2(S)$. Then

$$\sup_{(i, j) \in D_n(A)} \{|A^{-1}(i, j)| : (i, j) \in D_n(A)\} \leq C_0q^{n+1},$$

where $C_0$ and $q$ are as in Proposition 2.2.

The proof of this proposition follows the lines of Proposition 2.2 and we omit the details. The following examples illustrate the generality of Proposition 5.1.

Example 1. Suppose $A$ is positive definite and block banded where each block is itself banded. Then the sets $S_n(A)$ and hence $D_n(A)$ would give a much better idea of the decay of the entries of $A^{-1}$ than would the results of Section 2.

Example 2. $A$ is positive definite and tridiagonal with the exception that $(1, n)$ and $(n, 1)$ entries are nonzero. Then $A^{-1}$ decays away from the main diagonal and away from the corners (at least until $D_n \neq \emptyset$). Such matrices arise in spline interpolation of periodic data.

Example 3. If $\{f_i\}_{i \in S}$ is a local well-conditioned basis in $L_2$, for a space of finite element approximations, then taking $A = ((f_i, f_j))$ where $\langle \cdot, \cdot \rangle$ denotes the inner product in $L_2$, we see that the biorthogonal functions

$$\phi_i(x) = \sum_j A^{-1}(i, j)f_j$$

satisfy

$$\langle \phi_i, f_j \rangle = \delta_{ij}$$

and decay exponentially to 0 as $x$ moves away from the support of $f_i$, provided that the supports of the $f_j$ satisfy a global mesh ratio restriction. A precise formulation can be found in a paper of Descloux [7], in which the idea of exponential decay is already implicit.

We conclude this paper with some remarks concerning the generality of Propositions 2.1 and 2.2. There are two key ideas here. First use spectral theory to write $(f(x) = 1/x)$

$$\|A^{-1} - p(A)\| = \|f - p\|_{\sigma(A)}$$
and secondly use approximation theory to estimate the best error $\|f - p\|_{\sigma(A)}$. In all our propositions we arranged it so that $\sigma(A)$ was a subset of the positive real axis (i.e., $A$ (or $AA^*$) positive definite). The spectral theory applies more generally to normal matrices and Bernstein's theorem (see [13, p. 114]) can be used in place of Proposition 2.1, but we were unable to see how to use this extra freedom quantitatively.

Georgia Institute of Technology
Atlanta, Georgia 30332

Department of Mathematical Sciences
Old Dominion University
Norfolk, Virginia 23508