Approximating the inverse of a diagonally dominant matrix with positive elements

Ting Yan
Central China Normal University

Abstract

For an \( n \times n \) diagonally dominant matrix \( T = (t_{i,j})_{n \times n} \) with positive elements satisfying certain bounding conditions, we propose to use a diagonal matrix \( S = (s_{i,j})_{n \times n} \) to approximate the inverse of \( T \), where \( s_{i,j} = \delta_{i,j}/t_{i,i} \) and \( \delta_{i,j} \) is the Kronecker delta function. We derive an explicitly upper bound on the approximation error, which is in the magnitude of \( O(n^{-2}) \). It shows that \( S \) is a very good approximation to \( T^{-1} \).

Key words: Approximation error, Diagonally dominant, Inverse.

Mathematics Subject Classification: 15A09, 15B48.

1 Introduction

In this paper, we consider the approximate inverse of an \( n \times n \) diagonally dominant matrices \( T = (t_{i,j})_{n \times n} \) with positive elements satisfying certain bounding conditions, i.e.,

\[
t_{i,j} > 0, \quad t_{i,i} \geq \sum_{j=1,j\neq i}^{n} t_{i,j}, \quad i = 1, \ldots, n.
\]  

It is easy to show that \( T \) must be positive definite. We propose to use a diagonal matrix \( S = (s_{i,j})_{n \times n} \) to approximate the inverse of \( T \), where

\[
s_{i,j} = \frac{\delta_{i,j}}{t_{i,i}}.
\]  

*Department of Statistics, Central China Normal University, Wuhan, 430079, China. Email: tingyanty@mail.ccnu.edu.cn.
and $\delta_{i,j}$ is the Kronecker delta function. We obtain an explicitly upper bound on the approximation error in terms of $\max_{i,j}|(T^{-1} - S)_{ij}|$, which has the magnitude of $1/n^2$. This shows that $S$ is a very good approximation to $T^{-1}$.

The problems on inverses of nonnegative matrices have been extensively investigated; see Berman and Plemmons (1994); Loewy and London (1978); Egleston et al. (2004). It has applications to solving a large system of linear equations, in which a good approximate inverse of the coefficient matrix plays an important role in establishing fast convergence rates of iterative algorithms Axelsson (1985); Benzi (2002); Bruaset (1995); Zhang et al. (2009). Within statistics, Yan (2019) use the approximate inverse of $T$ to obtain a fast geometric rate of convergence of an iterative sequences for solving the estimate of parameters in the node-parameter network models with dependent structures. Further, it is used to derive the asymptotic representation of an estimator of the model parameter.

## 2 An explicit bound on the approximation error

For a general matrix $A = (a_{i,j})$, define the matrix maximum norm:

$$
\|A\| := \max_{i,j} |a_{i,j}|.
$$

We measure the approximation error of using $S$ to approximate $T^{-1}$ in terms of $\|T^{-1} - S\|$. Some notations are defined as follows:

$$
m := \min_{1 \leq i < j \leq n} t_{i,j}, \quad \Delta_i := t_{i,i} - \sum_{j=1, j \neq i}^n t_{i,j}, \quad M := \max\{\max_{1 \leq i < j \leq n} t_{i,j}, \max_{1 \leq i \leq n} \Delta_i\}.
$$

Note that $M \geq m > 0$. Let

$$
C(m, M) = \frac{2(n - 2)m}{nM + (n - 2)m} - \frac{(n - 2)Mm}{[(n - 2)m + M][(n - 2)m + 2M]} - \frac{M}{m(n - 1)}.
$$

The approximate error is formally stated below.

**Theorem 1.** If $C(m, M) > 0$, then for $n \geq 3$, we have

$$
\|T^{-1} - S\| \leq \frac{M}{m^2(n - 1)^2 C(m, M)}.
$$

**Proof.** Let $I_n$ be the $n \times n$ identity matrix. Define

$$
F = T^{-1} - S, \quad V = (v_{ij}) = I_n - TS, \quad W = (w_{ij}) = SV.
$$
Then, we have the recursion:

$$F = T^{-1} - S = (T^{-1} - S)(I_n - TS) + S(I_n - TS) = FV + W. \quad (3)$$

A direct calculation gives that

$$v_{i,j} = \delta_{i,j} - \sum_{k=1}^{n} t_{i,k} s_{k,j} = \delta_{i,j} - \sum_{k=1}^{n} t_{i,k} \frac{\delta_{k,j}}{t_{j,j}} = (\delta_{i,j} - 1) \frac{t_{i,j}}{t_{j,j}}, \quad (4)$$

and

$$w_{i,j} = \sum_{k=1}^{n} s_{i,k} v_{k,j} = \sum_{k=1}^{n} \frac{\delta_{i,k}}{t_{i,i}} [(\delta_{k,j} - 1) \frac{t_{k,j}}{t_{j,j}}] = \frac{(\delta_{i,j} - 1) t_{i,j}}{t_{i,i} t_{j,j}}. \quad (5)$$

Recall that $m \leq t_{i,j} \leq M$ and $(n-1)m \leq t_{i,i} \leq nM$. When $i \neq j$, we have

$$0 < \frac{t_{i,j}}{t_{i,i} t_{j,j}} \leq \frac{M}{m^2(n-1)^2},$$

such that for three different subscripts $i, j, k$,

$$|w_{i,i}| = 0, \quad |w_{i,j}| \leq \frac{M}{m^2(n-1)^2},$$

$$|w_{i,j} - w_{i,k}| \leq \frac{M}{m^2(n-1)^2}, \quad |w_{i,i} - w_{i,k}| \leq \frac{M}{m^2(n-1)^2}.$$

It follows that

$$\max(|w_{i,j}|, |w_{i,j} - w_{i,k}|) \leq \frac{M}{m^2(n-1)^2}, \quad \text{for all } i, j, k. \quad (6)$$

We use the recursion (3) to obtain a bound of the approximate error $\|F\|$. By (3) and (4), for any $i$, we have

$$f_{i,j} = \sum_{k=1}^{n} f_{i,k} [(\delta_{k,j} - 1) \frac{t_{k,j}}{t_{j,j}}] + w_{i,j}, \quad j = 1, \ldots, n. \quad (7)$$

Thus, to prove Theorem 1, it is sufficient to show that for any $i, j$,

$$|f_{i,j}| \leq \frac{M}{m^2C(M,m)(n-1)^2}.$$ 

Define $f_{i,\alpha} = \max_{1 \leq k \leq n} f_{i,k}$ and $f_{i,\beta} = \min_{1 \leq k \leq n} f_{i,k}$.
First, we will show that $f_{i,\beta} \leq 0$. Since for any fixed $i$,

$$
\sum_{k=1}^{n} f_{i,k} t_{k,i} = \sum_{k=1}^{n} \left( [T^{-1}]_{i,k} - \delta_{i,k} \frac{t_{k,i}}{t_{i,i}} \right) t_{k,i} = 1 - 1 = 0,
$$

we have

$$
f_{i,\beta} \sum_{k=1}^{n} t_{k,i} \leq \sum_{k=1}^{n} f_{i,k} t_{k,i} = 0.
$$

It follows that $f_{i,\beta} \leq 0$. With similar arguments, we have that $f_{i,\alpha} \geq 0$.

Recall that $\Delta_{\alpha} = t_{\alpha,\alpha} - \sum_{k=1, k \neq \alpha}^{n} t_{k,\alpha}$. Since $t_{\alpha,\alpha} = \left\{ \sum_{k=1}^{n} \left[ (\delta_{k,\alpha} - 1) t_{k,\alpha} - \delta_{k,\alpha} \Delta_{\alpha} \right] \right\}$, we have the identity

$$
f_{i,\beta} = -\sum_{k=1}^{n} f_{i,\beta} \frac{(\delta_{k,\alpha} - 1) t_{k,\alpha} - \delta_{k,\alpha} \Delta_{\alpha}}{t_{\alpha,\alpha}}.
$$

Similarly, we have

$$
f_{i,\beta} = -\sum_{k=1}^{n} f_{i,\beta} \frac{(\delta_{k,\beta} - 1) t_{k,\beta} - \delta_{k,\beta} \Delta_{\beta}}{t_{\beta,\beta}}.
$$

By combining (7) and (9), where we set $i = \alpha$ in (7), it yields that

$$
f_{i,\alpha} + f_{i,\beta} = \sum_{k=1}^{n} \left( f_{i,k} - f_{i,\beta} \right) \frac{[\delta_{k,\alpha} - 1] t_{k,\alpha} - \delta_{k,\alpha} \Delta_{\alpha}}{t_{\alpha,\alpha}} + w_{i,\alpha}.
$$

Again, by combining (7) and (10), we have

$$
2 f_{i,\beta} = \sum_{k=1}^{n} \left( f_{i,k} - f_{i,\beta} \right) \frac{[\delta_{k,\beta} - 1] t_{k,\beta} - \delta_{k,\beta} \Delta_{\beta}}{t_{\beta,\beta}} + w_{i,\beta}.
$$

By subtracting (12) from (11), we get

$$
\begin{align*}
\frac{f_{i,\alpha} - f_{i,\beta}}{f_{i,\beta}} &= \sum_{k=1}^{n} \left( f_{i,k} - f_{i,\beta} \right) \frac{[\delta_{k,\alpha} - 1] t_{k,\alpha} - \delta_{k,\alpha} \Delta_{\alpha}}{t_{\alpha,\alpha}} - \left\{ \frac{\Delta_{\alpha}}{t_{\alpha,\alpha}} - \frac{\Delta_{\beta}}{t_{\beta,\beta}} \right\} f_{i,\beta}.
\end{align*}
$$

Let $\Omega = \{ k : (1 - \delta_{k,\beta}) t_{k,\beta} / t_{\beta,\beta} \geq (1 - \delta_{k,\alpha}) t_{k,\alpha} / t_{\alpha,\alpha} \}$ and define $\lambda := |\Omega|$. Note that
\[1 \leq \lambda \leq n - 1. \text{ Then,} \]
\[
\sum_{k=1}^{\lambda} (f_{i,k} - f_{i,\beta})[(\delta_{k,a} - 1)\frac{t_{k,a}}{t_{a,a}} - (\delta_{k,\beta} - 1)\frac{t_{k,\beta}}{t_{\beta,\beta}}] \\
\leq \sum_{k \in \Omega} (f_{i,k} - f_{i,\beta})[(1 - \delta_{k,\beta})\frac{t_{k,\beta}}{t_{\beta,\beta}} - (1 - \delta_{k,a})\frac{t_{k,a}}{t_{a,a}}] \\
\leq (f_{i,a} - f_{i,\beta})\left[\frac{\lambda M}{\lambda M + (n - 1 - \lambda)m} - \frac{(\lambda - 1)m}{(\lambda - 1)m + (n - \lambda)M + M}\right].
\] (14)

We will obtain the maximum value of the expression in the above bracket through dividing it into two functions \(f(\lambda)\) and \(g(\lambda)\) of \(\lambda\), where

\[
f(\lambda) = \frac{\lambda M}{\lambda M + (n - 1 - \lambda)m} - \frac{(\lambda - 1)m}{(\lambda - 1)m + (n - \lambda)M + M}, \quad g(\lambda) = \frac{(\lambda - 1)m}{(\lambda - 1)m + (n - \lambda)M + M}.
\]

We first derive the maximum value of \(f(\lambda)\). There are two cases to consider the maximum value of \(f(\lambda)\) in the range of \(\lambda \in [1, n - 1]\).

Case I: When \(M = m\), it is easy to show \(f(\lambda) = 1/(n - 1)\).

Case II: \(M \neq m\). A direct calculation gives that

\[
f'(\lambda) = \frac{(n - 1)Mm}{[\lambda M + (n - 1 - \lambda)m]^2} - \frac{(n - 1)Mm}{[(\lambda - 1)m + (n - \lambda)M]^2} \\
= \frac{(n - 1)Mm[(n - 2\lambda)(M - m)][\lambda M + (n - 1 - \lambda)m + (\lambda - 1)m + (n - \lambda)M]}{[\lambda M + (n - 1 - \lambda)m]^2[(\lambda - 1)m + (n - \lambda)M]^2}
\]

\[
f''(\lambda) = -2(M - m)Mm(n - 1)\left(\frac{1}{[\lambda M + (n - 1 - \lambda)m]^3} + \frac{1}{[(\lambda - 1)m + (n - \lambda)M]^3}\right).
\]

Since \(f''(\lambda) \leq 0\) when \(\lambda \in [1, n - 1]\), \(f(\lambda)\) is a convex function of \(\lambda \in [1, n - 1]\) such that \(f(\lambda)\) takes its maximum value at \(\lambda = n/2\) when \(1 \leq \lambda \leq n - 1\). Note that

\[
f\left(\frac{n}{2}\right) = \frac{nM - (n - 2)m}{nM + (n - 2)m}.
\]
So we have

\[
\sup_{\lambda \in [0, n-1]} f(\lambda) \leq \frac{nM - (n-2)m}{nM + (n-2)m}.
\] (15)

Next, we obtain the maximum value of \(g(\lambda)\). Since

\[
g'(\lambda) = \frac{Mm[M^2((n-\lambda)^2 + 2(n-\lambda)(\lambda - 1) + n - 1) + (2Mm - m^2)(\lambda - 1)^2]}{[(\lambda - 1)m + (n-\lambda)M]^2[(\lambda - 1)m + (n-\lambda)M + M]^2},
\]

\(g'(\lambda) > 0\) when \(1 \leq \lambda \leq n - 1\). So \(g(\lambda)\) is an increasing function on \(\lambda\) such that

\[
0 \leq \sup_{\lambda \in [1, n-1]} g(\lambda) \leq g(n - 1) = \frac{(n-2)Mm}{[(n-2)m + M][(n-2)m + 2M]}.
\] (16)

By combining (15) and (16), we have

\[
\sup_{1 \leq \lambda \leq n-1} \left[ \frac{\lambda M}{\lambda M + (n-1 - \lambda)m} - \frac{(\lambda - 1)m}{(\lambda - 1)m + (n-\lambda)M + M} \right]
\leq \sup_{1 \leq \lambda \leq n-1} \left[ f(\lambda) + \sup_{1 \leq \lambda \leq n-1} g(\lambda) \right]
\leq \frac{1}{n-1} I(M = m) + \frac{nM - (n-2)m}{nM + (n-2)m} I(M \neq m) + \frac{(n-2)Mm}{[(n-2)m + M][(n-2)m + 2M]}
\leq \frac{nM - (n-2)m}{nM + (n-2)m} + \frac{(n-2)Mm}{[(n-2)m + M][(n-2)m + 2M]},
\] (17)

where \(I(\cdot)\) is an indicator function. By combining (13), (14) and (17), we have

\[
f_{i,\alpha} - f_{i,\beta} \leq \left\{ \frac{nM - (n-2)m}{nM + (n-2)m} + \frac{(n-2)Mm}{[(n-2)m + M][(n-2)m + 2M]} \right\} (f_{i,\alpha} - f_{i,\beta}) + |w_{i,\alpha} - w_{i,\beta}| + \left| \frac{\Delta_{i,\beta} - \Delta_{i,\alpha}}{t_{\beta,\alpha} - t_{\alpha,\alpha}} \right| |f_{i,\beta}|.
\] (18)

Since \(f_{i,\alpha} \geq |f_{i,\beta}|\) and \(f_{i,\beta} \leq 0\), we have

\[
\left| \frac{\Delta_{i,\beta} - \Delta_{i,\alpha}}{t_{\beta,\alpha} - t_{\alpha,\alpha}} \right| |f_{i,\beta}| \leq \frac{\Delta_{i,\beta} - \Delta_{i,\alpha}}{t_{\beta,\alpha} - t_{\alpha,\alpha}} (f_{i,\alpha} - f_{i,\beta}) \leq \frac{M}{m(n-1)} (f_{i,\alpha} - f_{i,\beta}).
\] (19)

Recall the definition of \(C(m, M)\) in (2). By combining (18) and (19), it yields

\[
(f_{i,\alpha} - f_{i,\beta}) C(m, M) \leq |w_{i,\alpha} - w_{i,\beta}| \leq \frac{M}{m^2(n-1)^2}.
\]

Consequently,

\[
\max_{j=1, \ldots, n} |f_{i,j}| \leq f_{i,\alpha} - f_{i,\beta} \leq \frac{M}{m^2(n-1)^2 C(M, m)}.
\]
This completes the proof.

We discuss the condition $C(m, M) > 0$. $C(m, M)$ can be represented as

$$C(m, M) = \frac{2(n - 2)m}{nM + (n - 2)m} - \frac{(n - 2)(M/m)}{[(n - 2) + M/m][(n - 2) + 2M/m]} - \frac{M/m}{n - 1}. $$

So if $M/m = o(n)$, then for large $n$

$$C(m, M) = \frac{2m}{M + m} + o(1). $$

Then we immediately have the corollary.

**Corollary 1.** If $M/m = o(n)$, then for large $n$,

$$\|T^{-1} - S\| = O\left(\frac{M^2}{m^3 n^2}\right).$$

### 3 Discussion

The bound on the approximation error in Theorem 1 depends on $m$, $M$ and $n$. When $m$ and $M$ are bounded by a constant, all the elements of $T^{-1} - S$ are of order $O(1/n^2)$ as $n \to \infty$, uniformly. Therefore we conjecture that $T$ may belong to inverse $M$-matrices.

The interested readers can refer to Berman and Plemmons (1994); Foregger (1990).

We illustrate by an example that the bound on the approximation error in Theorem 1 is optimal in the sense that any bound in the form of $K(m, M)/f(n)$ requires $f(n) = O(n^2)$ as $n \to \infty$. Assume that the matrix $T$ consists of the elements: $t_{i,i} = (n - 1)M, i = 1, \cdots, n - 1; t_{n,n} = (n - 1)m$ and $t_{i,j} = m, i, j = 1, \cdots, n; i \neq j$, which satisfies (1). By the Sherman-Morrison formula, we have

$$(T^{-1})_{i,j} = \frac{\delta_{i,j}}{(n - 1)M - m} - \frac{m}{[(n - 1)M - m]^2}, i, j = 1, \cdots, n - 1$$

$$(T^{-1})_{n,j} = \frac{\delta_{n,j}}{(n - 2)m} - \frac{1}{(n - 2)[(n - 1)M - m]}, j = 1, \cdots, n.$$ 

In this case, the elements of $S$ are

$$S_{i,j} = \frac{\delta_{i,j}}{(n - 1)M} - \frac{1}{n(n - 1)m}, i, j = 1, \cdots, n - 1; i \neq j,$$

$$S_{n,j} = \frac{\delta_{n,j}}{(n - 1)m} - \frac{1}{n(n - 1)m}, j = 1, \cdots, n.$$ 

It is easy to show that the bound of $\|T^{-1} - S\|$ is $O(\frac{1}{n(n - 1)m})$. This suggests that the rate
$1/(n - 1)^2$ is optimal. On the other hand, there is a gap between $1/m$ and $O(M^2/m^3)$ which implies that there might be space for improvement. It is interesting to see if the bounds in Theorem 1 can be further relaxed.

**References**

Axelsson, O. (1985). A survey of preconditioned iterative methods for linear systems of algebraic equations. *BIT Numerical Mathematics*, 25(1):165–187.

Benzi, M. (2002). Preconditioning techniques for large linear systems: A survey. *Journal of Computational Physics*, 182(2):418 – 477.

Berman, A. and Plemmons, R. (1994). *Nonnegative Matrices in the Mathematical Sciences*. Society for Industrial and Applied Mathematics.

Bruaset, A. M. (1995). *A Survey of Preconditioned Iterative Methods*. Longman Scientific & Technical.

Egleston, P. D., Lenker, T. D., and Narayan, S. K. (2004). The nonnegative inverse eigenvalue problem. *Linear Algebra and its Applications*, 379:475 – 490. Special Issue on the Tenth ILAS Conference (Auburn, 2002).

Foregger, T. H. (1990). Review of nonnegative matrices: By henryk minc. *Linear Algebra and its Applications*, 134:181 – 183.

Loewy, R. and London, D. (1978). A note on an inverse problem for nonnegative matrices. *Linear and Multilinear Algebra*, 6(1):83–90.

Yan, T. (2019). Moment estimation in the node-parameter network models with dependent edges. *Manuscript*.

Zhang, Y., Huang, T.-Z., Liu, X.-P., and Gu, T.-X. (2009). A class of approximate inverse preconditioners for solving linear systems. *International Journal of Computer Mathematics*, 86(7):1243–1252.