Explicit diagonalization of a Cesaró matrix

Suvrit Sra.

Version of: 9th Mar, 2014

Abstract

We study a specific “anti-triangular” Cesaró matrix corresponding to a Markov chain. We derive closed forms for all the eigenvalues and eigenvectors of this matrix. As an application we obtain bounds on the operator norm of the positive definite matrices corresponding to the kernels \( \min(i, j) \) and \( 1/\max(i, j) \).

1 Introduction

Eigenvalues of Markov chains lend insight into the speed of convergence to an invariant measure (or stationary distribution). The corresponding eigenvector provides the distribution of the stationary state. In this paper, we study a particular Markov transition matrix, in particular by determining its eigenvalues and eigenvectors in closed form.

The matrix that we study is a Cesáro style matrix (we borrow this terminology from [2]—see also the remark below). The specific matrix that we study is anti-lower triangular; such matrices have also been studied by [6]—see also related work therein. It is worth mentioning here the connection to inverse eigenvalue problems for anti-bidiagonal matrices [4], and accurate numerical methods for anti-bidiagonal matrices [5].

Specifically, we investigate the problem of diagonalizing the following structured Markov transition matrix on \( n \) states:

\[
P = \begin{pmatrix}
1 & \frac{1}{2} & 1 & \frac{1}{2} & \cdots & \cdots & \frac{1}{2} & 1 \\
\frac{1}{n} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\frac{1}{n} & \cdots & \frac{1}{n}
\end{pmatrix}. \tag{1.1}
\]

We derive below explicit formulae for the eigenvectors \( S \) of \( P \) so that \( P = SAS^{-1} \), where \( \Lambda \) is the diagonal matrix of eigenvalues. One may ask, how do we even know that the matrix \( P \) is diagonalizable? Prop. 2.1 shows that \( P \) has \( n \) unique eigenvalues which ensures diagonalizability.

Remark: Unsurprisingly, diagonalizing the “Cesaró matrix” [2]

\[
C = \begin{pmatrix}
1 & \frac{1}{2} & 1 & \frac{1}{2} & \cdots & \cdots & \frac{1}{2} & 1 \\
\frac{1}{n} & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\frac{1}{n} & \cdots & \frac{1}{n}
\end{pmatrix},
\]

turns out to be much easier. In fact, it can be easily verified that \( VCV^{-1} = \text{Diag}([1/i]_{i=1}^n) \) for \( V \) given by (2.2); the anti-triangular form of (1.1) makes it harder to diagonalize explicitly.

*This work was done while the author was on leave from Max Planck Institute for Intelligent Systems, Tübingen, Germany at Carnegie Mellon University, Pittsburgh, USA
2 Explicit Diagonalization

Let $J$ denote the reverse identity matrix, i.e., the matrix with ones on its anti-diagonal. Then $JP$ is upper-triangular, so that eigenvalues of $JP$ can be read off of the diagonal. These eigenvalues are $\lambda_i(JP) = 1/i$. It turns out that the eigenvalues of $P$ are also $1/i$ multiplied with alternating signs. Let us prove this observation.

**Proposition 2.1.** Let $P$ be given by (1.1). Then, $\lambda_i(P) = (-1)^{i+1}/i$ for $i = 1, \ldots, n$.

**Proof.** It proves more convenient to analyze $P^{-1}$. An easy verification shows that $P^{-1}$ is the anti-bidiagonal matrix

$$P^{-1} = \begin{pmatrix}
1 & 2 - n & 1 - n & n \\
-1 & 2 & 1 & \\
& & \ddots & \\
& & & 1 - n
\end{pmatrix}.$$  \hspace{1cm} (2.1)

To obtain eigenvalues of $P^{-1}$ it suffices to find a triangular matrix similar to it. To that end, the following matrix proves useful (curiously, $L$ is the strict lower-triangular part of $JP^{-1}$):

$$L = \begin{pmatrix}
0 & 0 & \cdots & 0 \\
-1 & 0 & \cdots & 0 \\
0 & -2 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots \\
& & & 1 - n
\end{pmatrix}.$$ 

Setting $V = \exp(L)$, we see that $V$ is the following lower-triangular matrix:

$$V = [V_{ij}] = \begin{pmatrix}
(-1)^{i-j} \left( \frac{i-1}{j-1} \right)
\end{pmatrix} \text{ for } i \geq j.$$ \hspace{1cm} (2.2)

Explicitly carrying out the multiplication, we obtain the triangular matrix (using $V^{-1} = \exp(-L)$)

$$VP^{-1}V^{-1} = \begin{pmatrix}
1 & * & \cdots & * \\
-2 & * & \cdots & * \\
& & \ddots & \cdots \\
& & & (-1)^{n+1}n
\end{pmatrix},$$ \hspace{1cm} (2.3)

where $*$ represents unspecified entries. From (2.3) it is clear that $\lambda_i(P^{-1}) = (-1)^{i+1}i$. \hfill \Box

Obtaining eigenvectors that diagonalize $P^{-1}$ is harder and requires a new idea. A quick computation shows that $P^{-2}$ is tridiagonal but asymmetric, which rules out an easy solution. It turns out, however, that $VP^{-2}V^{-1}$ is a highly structured bidiagonal matrix. Indeed,

$$B := VP^{-2}V^{-1} = \begin{pmatrix}
1 & -2(n-1) & -3(n-2) & -4(n-3) & \cdots \\
4 & 9 & 16 & \cdots & \\
& 25 & 49 & \cdots & \\
& & \ddots & \cdots & \\
& & & (n-1)^2 & -n(n-1)
\end{pmatrix}.$$ 

If we succeed in diagonalizing $B$, then we are done. Suppose therefore that there is a matrix $S$ that diagonalizes $B$, that is $SBS^{-1} = \Lambda$ with $\Lambda = [i^2]_{i=1}^n$ diagonal. Then,

$$S^{-1}\Lambda S = VP^{-2}V^{-1} \implies SVP^{-2}V^{-1}S^{-1} = \Lambda = \text{Diag}([i^2]_{i=1}^n),$$ \hspace{1cm} (2.4)

which shows that $SV$ diagonalizes $P^{-2}$, completing the answer.
3 APPLICATION TO KERNELS

Theorem 2.2. Let $S = M^T$, where $M$ is lower-triangular, whose nonzero entries in column $j$ are

$$m_{kj} := \frac{(j+1)_{k-j}(n-k+1)_{k-j}}{(2j+1)_{k-j}(k-j)!}, \quad k \geq j \quad (1 \leq j \leq n). \quad (2.5)$$

This choice of $S$ diagonalizes $B$, that is, it satisfies (2.4).

Proof. To find $S$ we need to solve the system of equations:

$$SB = \Lambda S \iff B^T M = M \Lambda.$$

Consider the $j$th eigenvalue $\lambda_j = j^2$; denote the corresponding column of $M$ by $m$ and its $k$th entry by $m_k$. To obtain $m$ we must solve the linear system

$$B^T m = j^2 m, \quad \implies m_1 = j^2 m_1$$

$$-(k+1)(n-k)m_k + (k+1)^2m_{k+1} = j^2m_{k+1}, \quad 1 \leq k \leq n.$$

Since $B$ is bidiagonal, a brief reflection shows that $M$ is lower-triangular with 1s on its diagonal—the 1s come from the equation corresponding to the index $k+1 = j$. The subsequent entries of $m$ are nonzero. Symbolic computation with a few different values of $j$ suggests the general solution (which can be formally proved using an easy but tedious induction):

$$m_k = \frac{(j+1)_{k-j}(n-k+1)_{k-j}}{(2j+1)_{k-j}(k-j)!}, \quad k \geq j,$$

where $(x)_k := x(x+1) \cdots (x+k-1)$ denotes the Pochhammer symbol (rising factorial). \hfill \Box

Remark 2.3 (Added Oct 31, 2014). One can similarly diagonalize the closely related transition matrix

$$Z = \begin{pmatrix}
\frac{1}{n-1} & \frac{1}{n} & \cdots & \frac{1}{n-1} & \frac{1}{n} \\
\frac{1}{n} & \cdots & \frac{1}{n-1} & 0 \\
\frac{1}{n-1} & \cdots & \frac{1}{n-1} & \frac{1}{n} \\
\frac{1}{n} & \cdots & \frac{1}{n} & 0 \\
\frac{1}{n} & \cdots & \frac{1}{n} & \frac{1}{n-1}
\end{pmatrix}. \quad (2.6)$$

Clearly, from a diagonalization of $P^{-1}$ we can recover a diagonalization of $Z$ above. To see why, observe that $P^{-1} = (J^T Z^{-1} J)^T$ where $J$ is the “reverse identity” (anti-diagonal) matrix.

3 Application to kernels

In this section we provide some additional observations about the matrix $P$ from (1.1) that may be of independent interest.

3.1 Brownian bridge kernel

It turns out that the matrix $P$ is closely related to the Brownian bridge kernel $K = [k_{ij}] = \min(i,j)$ [3]. Proposition 3.1 makes this connection precise and highlights a well-known property of this kernel.

Proposition 3.1. The kernel matrix $K = PP^T$ is infinitely divisible (i.e., $[k_{ij}^r] \geq 0$ for $r \geq 0$).
Proof. We show that the Schur power $K^r = [k_{ij}^r]$, for $r > 0$ is positive definite. This claim follows after we realize that

$$K = [k_{ij}] = \left[ \frac{1}{\max(i, j)} \right], \quad 1 \leq i, j \leq n,$$

which is well-known to be infinitely divisible [1, Ch. 5]. We include a short proof below. Let $D = \text{Diag}([i^{-1}]_{i=1}^n)$; also let $M := [\min(i, j)]$. As $K = DMD$, it suffices to establish infinite divisibility of $M$. We prove a more general statement. Let $f$ be a positive monotonic function, and for any set $C \subset \mathbb{R}$, define $1_C(x) = 1$ if $x \in C$ and 0 otherwise. Then,

$$f(m_{ij}) = \min(f(i), f(j)) = \int_0^\infty 1_{[0, f(i)]}(x)1_{[0, f(j)]}(x)dx,$$

which is nothing but an inner-product; thus $[f(m_{ij})]$ is a Gram matrix and hence positive definite. Setting $f(t) = t^r$, $r > 0$, infinite divisibility of $M$ (and hence of $K$) follows. \hfill \Box

### 3.2 Operator Norm bounds

When using the kernels $\min(i, j)$ or $1/\max(i, j)$ in an application, one may need to bound their operator norms (e.g., for approximation or optimization). To that end Proposition 2.1 proves of great value. But before we see how, let us illustrate a bound on $\|K\|$ that we might have obtained without noticing the connection exposed in Proposition 3.1.

Since $K = DMD$, an easy bound is $\|K\| \leq \|D\|^2\|M\| = \|M\|$ (as $\|D\| = 1$). Proposition 3.2 bounds $\|M\|$.

**Proposition 3.2.** Let $M = [m_{ij}] = [\min(i, j)]$; define $\theta_k := \frac{2k\pi}{2n+1}$. Then, $M = VAV^{-1}$ with

$$\lambda_k = (2 + 2\cos \theta_k)^{-1}, \quad k = 1, 2, \ldots, n$$

$$v_{jk} = \frac{2}{\sqrt{2n+1}} \sin(k - \frac{1}{2})\theta_j, \quad \theta_j \neq \pi, \quad j, k = 1, 2, \ldots, n. \quad (3.1)$$

Consequently, $\|M\| \approx 2n/\pi$.

**Proof.** The key is to observe that $M$ has the Cholesky factorization $M = LL^T$, where $L$ is the all ones lower-triangular matrix

$$L = \begin{pmatrix} 1 & 1 & \cdots & \cdots & 1 \\ 1 & 1 & \cdots & \cdots & 1 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & \cdots & \cdots & 1 \\ \vdots & \vdots & \cdots & \cdots & \vdots \end{pmatrix}.$$

We wish to explicitly diagonalize $LL^T$, a task that becomes much easier if we consider the inverse $M^{-1} = L^{-T}L^{-1}$, as this is a perturbed Toeplitz-tridiagonal matrix

$$L^{-T}L^{-1} = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \ddots \\ -1 & & & -1 & 2 \end{pmatrix}. \quad (3.2)$$

Applying the derivation of [7, Thm. 3.2-(ix)]\(^1\), we find that the eigenvalues of (3.2) are $\lambda_k^{-1} = 2 + 2\cos \theta_k$ where $\theta_k = \frac{2k\pi}{2n+1}$. In fact $M^{-1}$ can also be diagonalized explicitly: its eigenvectors are given by (again by resorting to arguments of [7, Thm. 3.3]):

$$\frac{\sqrt{2n+1}}{2} v_{jk} := \sin(k - \frac{1}{2})\theta_j, \quad j, k = 1, 2, \ldots, n. \quad (3.3)$$

\(^1\)There seems to be a typo in the cited theorem; the cases (viii) and (ix) stated in that paper seem to be switched.
3.2 Operator Norm bounds

Figure 1: Numerical values of $\|P\|$ and $\|P^{-1}\|$. The bounds from Prop. 3.5 are reasonably tight.

With $V = [v_{jk}]$ and noting $V^{-1} = V^T$, we obtain $V^T(LL^T)^{-1}V = \text{Diag}(\lambda_k^{-1})$. Hence, it immediately follows that $\|M\| = \lambda_{\max}(LL^T) = (2 + 2\cos\theta_n)^{-1} \approx 2n/\pi$.

The eigendecomposition of $M$ derived above is no surprise; $M = [\min(i,j)]$ is essentially the Brownian bridge covariance function whose spectrum is well-studied [3]. But it is worth noting that our derivation uses elementary linear algebra, compared with a more advanced Fourier-analytic derivation typically employed when studying eigenfunctions of kernels.

Remark 3.3. Using $\|K\| = \|DMD\| \leq \|D\|^2 \|M\| \leq \|M\| \sim 2n/\pi$, we get a very pessimistic bound on $\|K\|$. Remark 3.4 analyzes $\|K\|$ directly, yielding a much better dependence on $n$. Finally, Proposition 3.5 actually provides a dimension independent bound on $\|K\|$.

Remark 3.4. The bound from Remark 3.3 can be greatly improved rather easily. Indeed,

$$\|K\| = \|PP^T\| = \lambda_{\max}(PP^T) \leq \text{tr}(PP^T) = \sum \frac{1}{i} = H_n,$$

(3.4)

where $H_n$ denotes the $n$-th Harmonic number. However, even bound (3.4) is suboptimal as it depends on the dimension of the matrix $P$.

**Proposition 3.5.** Let $K = [1/\max(i,j)]$ for $1 \leq i, j \leq n$. Then, $\|K\| \leq 4$, while $\|K^{-1}\| \leq 4n^2$.

**Proof.** The key ingredient is provided by Proposition 2.1. Since $K = PP^T$, we analyze $\|K\| = \lambda_{\max}(PP^T)$, where $P$ is defined by (1.1).

From Proposition 2.1 we know that $|\lambda_j(P)| = 1/j$. Thus, the spectral radius $\rho(P) = 1$. Since the operator norm is upper-bounded by twice the spectral radius, i.e., $\|X\| \leq 2\rho(X)$, we see that $\|P\| \leq 2$. Similarly, from Proposition 2.1 it follows that $\rho(P^{-1}) = n$, which combined with $\|P^{-1}\| \leq 2\rho(P^{-1})$ yields $\|P^{-1}\| \leq 2n$. Subsequently, since $\|K\| = \|PP^T\| = \|P\|^2$ we obtain $\|K\| \leq 4$; and similarly, that $\|K^{-1}\| \leq 4n^2$.

The bounds proved by Proposition 3.5 are asymptotically accurate. The upper bound is somewhat loose for small $n$, but for larger $n$ it is fairly accurate. The lower-bound is quite accurate for almost all values of $n$ (Fig. 1 illustrates these bounds numerically).
Acknowledgments

I would like to thank David Speyer, whose comment to an initial answer of mine on MathOverflow prodded me to think more carefully about the problem, which ultimately led to this paper.

References

[1] R. Bhatia. Positive Definite Matrices. Princeton University Press, 2007.

[2] M.-D. Choi. Tricks or treats with the Hilbert matrix. American Mathematical Monthly, pages 301–312, 1983.

[3] T. Hofmann, B. Schölkopf, and A. J. Smola. A review of kernel methods in machine learning. Technical Report 156, Max Planck Institute for Biological Cybernetics, 2006.

[4] O. Holtz. The inverse eigenvalue problem for symmetric anti-bidiagonal matrices. Linear algebra and its applications, 408:268–274, 2005.

[5] P. Koev and F. Dopico. Accurate eigenvalues of certain sign regular matrices. Linear Algebra and its Applications, 424:435–447, 2007.

[6] C. Mehl. Anti-triangular and anti-m-Hessenberg forms for Hermitian matrices and pencils. Linear Algebra and its Applications, 317(1):143–176, 2000.

[7] W.-C. Yueh and S. S. Cheng. Explicit Eigenvalues and Inverses of Tridiagonal Toeplitz Matrices With Four Perturbed Corners. The ANZIAM Journal, 49(03):361, Aug. 2008.