

Approximating the eigenvalues of real symmetric Toeplitz matrices using the nearest circulants

Chris Salahub

University of Waterloo, csalahub@uwaterloo.ca

June 3, 2022

A number of methods exist to approximate the eigensystem of the symmetric Toeplitz matrix

$$\mathbf{\Sigma} = \begin{bmatrix} \rho_0 & \rho_1 & \rho_2 & \dots & \rho_{M-1} \\ \rho_1 & \rho_0 & \rho_1 & \dots & \rho_{M-2} \\ \rho_2 & \rho_1 & \rho_0 & \dots & \rho_{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{M-1} & \rho_{M-2} & \rho_{M-3} & \dots & \rho_0 \end{bmatrix}, \quad (1)$$

where $\rho_0, \dots, \rho_{M-1} \in \mathbb{R}$. $\mathbf{\Sigma}$ appears regularly in statistical applications such as the correlation matrix of the discrete autoregressive process of order one, in information theory in certain filtering tasks [Gray \(2006\)](#), and in genetics in the correlation between disjoint measured sequences [Salahub \(2022\)](#). In the case of the equidistant genetic survey of [Lander and Botstein \(1989\)](#) and the autoregressive time series, the matrix has the more specific form

$$\mathbf{\Sigma}_e = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{M-1} \\ \rho & 1 & \rho & \dots & \rho^{M-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{M-1} & \rho^{M-2} & \rho^{M-3} & \dots & 1 \end{bmatrix} \quad (2)$$

where $\rho \in [0, 1]$ is a real constant.

Of interest are the eigenvalues and eigenvectors of $\mathbf{\Sigma}$, call them $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_M$ and $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$ respectively. For brevity, the combination of these vectors and values

is here referred to as the *eigensystem* of Σ . [Gray \(2006\)](#) demonstrates the application of this particular eigensystem in signal processing, while [Cheverud \(2001\)](#), [Li and Ji \(2005\)](#), and [Galwey \(2009\)](#) use the eigenvalues to adjust for dependent multiple tests in the genomic context.

A number of approximations for this eigensystem exist. [Cheverud \(2001\)](#), [Li and Ji \(2005\)](#), and [Galwey \(2009\)](#) present different approximations based on the matrix

$$\mathbf{A}(\rho) = \rho \mathbf{1}\mathbf{1}^\top + (1 - \rho)\mathbf{I}_M. \quad (3)$$

$\mathbf{A}(\rho)$ has a known eigensystem for any ρ and matches the eigensystem of Σ for the edge cases where $\rho_1 = \rho_2 = \dots = 0$ or $\rho_1 = \rho_2 = \dots = 0$ and $\rho_0 = 1$.

[Gray \(2006\)](#) and [Grenander and Szegő \(1958\)](#) instead utilize circulant matrices to approximate the eigensystem of Σ . Asymptotically, Σ has the same eigensystem as certain circulant matrices and the eigensystem of any circulant matrix is known.

Absent from the symptotic approach of [Gray \(2006\)](#) or the myriad approaches of [Cheverud \(2001\)](#), [Li and Ji \(2005\)](#), and [Galwey \(2009\)](#) is a consideration of the finite, non-edge cases of Σ and how it might be best approximated. This work derives \mathbf{C}_Σ , the nearest circulant matrix to Σ in the weak norm. \mathbf{C}_Σ is symmetric circulant, and so has only real eigenvalues. Indeed, it is proven that a real circulant matrix has real eigenvalues if and only if it is a symmetric circulant. \mathbf{C}_Σ is additionally shown to be asymptotically equivalent to Σ by considering the limit of the remainder $\Sigma - \mathbf{C}_\Sigma$. Finally, the eigensystem of \mathbf{C}_Σ is stated explicitly.

1 Circulant Matrices

For clarity working with different matrices, let the function $\lambda_k(\mathbf{C})$ return the k^{th} eigenvalue of \mathbf{C} , which may not be ordered by magnitude.

A complex matrix $\mathbf{C} \in \mathbb{C}^{M \times M}$ is called circulant if the i^{th} row is given by the cyclic shift i elements rightward of a vector of M elements, typically denoted $(c_0, c_1, c_2, \dots, c_{M-1})$. Explicitly

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & c_2 & \dots & c_{M-2} & c_{M-1} \\ c_{M-1} & c_0 & c_1 & \dots & c_{M-3} & c_{M-2} \\ c_{M-2} & c_{M-1} & c_0 & \dots & c_{M-4} & c_{M-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ c_2 & c_3 & c_4 & \dots & c_0 & c_1 \\ c_1 & c_2 & c_3 & \dots & c_{M-1} & c_0 \end{bmatrix} \quad (4)$$

So every circulant matrix \mathbf{C} can be specified by its first row alone. Moreover, this first row corresponds to the coefficients in a convenient expression of \mathbf{C} as a matrix polynomial. Let \mathbf{P} be the circulant matrix with $c_0 = c_2 = c_3 = \dots = c_{M-1} = 0$ and $c_1 = 1$. That is,

$$\mathbf{P} = [\mathbf{e}_M | \mathbf{e}_1 | \dots | \mathbf{e}_{M-1}] \quad (5)$$

where \mathbf{e}_i is the i^{th} basis vector. Then \mathbf{P} is the permutation matrix corresponding to a cyclic shift of all elements of a vector $\mathbf{x} \in \mathbb{C}^M$ one to the right. Due to this cyclic shift property, it is also straightforward to note that

$$\mathbf{P}^m = \mathbf{P}\mathbf{P} \dots \mathbf{P} = [\mathbf{e}_{M-m+1} | \mathbf{e}_{M-m+2} | \dots | \mathbf{e}_M | \mathbf{e}_1 | \dots | \mathbf{e}_{M-m}]. \quad (6)$$

Using these \mathbf{P}^m , \mathbf{C} can be written

$$\mathbf{C} = c_0 \mathbf{I}_M + c_1 \mathbf{P} + c_2 \mathbf{P}^2 + \dots + c_{M-1} \mathbf{P}^{M-1}, \quad (7)$$

from which the eigensystem of \mathbf{C} can be derived from the eigensystem of \mathbf{P} . Using a cofactor expansion of $\det(\mathbf{P} - \lambda \mathbf{I})$ it can be shown that the eigenvalues of \mathbf{P} are the M^{th} roots of unity, that is

$$\lambda_k(\mathbf{P}) = \left(e^{\frac{2\pi i}{M}} \right)^k = \omega^k$$

where $k = 0, \dots, M-1$. The corresponding eigenvectors \mathbf{x}_k are then

$$\mathbf{x}_k = \begin{bmatrix} 1 \\ \omega^k \\ \omega^{2k} \\ \vdots \\ \omega^{(M-1)k} \end{bmatrix}, \quad (8)$$

which can be seen by considering $\mathbf{P}\mathbf{x}_k = \lambda_k(\mathbf{P})\mathbf{x}_k$. Note that any eigenvector of \mathbf{P} with eigenvalue λ is also an eigenvector of \mathbf{P}^m with eigenvalue λ^m , and so the eigenvalues of \mathbf{C} are given by

$$\lambda_k(\mathbf{C}) = c_0 + \sum_{m=1}^{M-1} c_m \omega^{mk} \quad (9)$$

with corresponding eigenvectors \mathbf{x}_k as above for $k = 0, \dots, M-1$. A particular circulant matrix structure is of interest.

Definition 1 (Symmetric circulant). *A circulant matrix $\mathbf{C} \in \mathbb{C}^{M \times M}$ is symmetric if elements in its first row $c_0, c_1, c_2, \dots, c_{M-1}$ satisfy $c_m = c_{M-m}$ for all $m \geq 1$.*

To provide a visual example of such a matrix, consider the circulant with $c_m = \min\{m, M - m\}$. When $M = 6$ we have

$$\begin{bmatrix} 0 & 1 & 2 & 3 & 2 & 1 \\ 1 & 0 & 1 & 2 & 3 & 2 \\ 2 & 1 & 0 & 1 & 2 & 3 \\ 3 & 2 & 1 & 0 & 1 & 2 \\ 2 & 3 & 2 & 1 & 0 & 1 \\ 1 & 2 & 3 & 2 & 1 & 0 \end{bmatrix}$$

This circulant is symmetric, and any symmetric circulant matrix will have an analogous structure. By basic results of linear algebra, it follows that any symmetric circulant matrix will have only real eigenvalues. Indeed, a circulant matrix will have only real eigenvalues for any M if and only if it is symmetric.

Theorem 1 (Real eigenvalues of symmetric circulants). *A circulant matrix $\mathbf{C} \in \mathbb{R}^{M \times M}$ has real eigenvalues if and only if it is a symmetric circulant matrix.*

Proof. Consider the eigensystem of \mathbf{C} . As it is circulant, it has eigenvalues

$$\lambda_k(\mathbf{C}) = c_0 + \sum_{m=1}^{M-1} c_m \omega^{mk},$$

or rather

$$\lambda_k(\mathbf{C}) = c_0 + \sum_{m=1}^{M-1} c_m \left(\cos \frac{2\pi mk}{M} + i \sin \frac{2\pi mk}{M} \right)$$

We can rewrite this to emphasize the real and imaginary components as

$$\lambda_k(\mathbf{C}) = \left(\sum_{m=0}^{M-1} c_m \cos \frac{2\pi mk}{M} \right) + i \left(\sum_{m=1}^{M-1} c_m \sin \frac{2\pi mk}{M} \right).$$

If $\lambda_k(\mathbf{C}) \in \mathbb{R}$ for all $k \in \{0, 1, \dots, M-1\}$, we must have

$$\sum_{m=1}^{M-1} c_m i \sin \frac{2\pi mk}{M} = 0 \quad \forall k \in \{0, 1, \dots, M-1\}.$$

But note that

$$i \sin \frac{2\pi mk}{M} = \frac{1}{2} \left(e^{\frac{2\pi mk}{M} i} - e^{-\frac{2\pi mk}{M} i} \right)$$

and

$$e^{-\frac{2\pi mk}{M}i} = e^{-\frac{2\pi mk}{M}i + 2\pi ki} = e^{\frac{2\pi(M-m)k}{M}i},$$

and so we require

$$\sum_{m=1}^{M-1} c_m \frac{1}{2} \left(e^{\frac{2\pi mk}{M}i} - e^{\frac{2\pi(M-m)k}{M}i} \right) = 0 \quad \forall k \in \{0, 1, 2, \dots, M-1\}.$$

However,

$$\sum_{m=1}^{M-1} c_m \frac{1}{2} \left(e^{\frac{2\pi mk}{M}i} - e^{\frac{2\pi(M-m)k}{M}i} \right) = 0$$

$$\iff \sum_{m=1}^{M-1} c_m e^{\frac{2\pi mk}{M}i} - \sum_{m=1}^{M-1} c_m e^{\frac{2\pi(M-m)k}{M}i} = 0$$

$$\iff \sum_{m=1}^{M-1} c_m e^{\frac{2\pi mk}{M}i} - \sum_{m=1}^{M-1} c_{M-m} e^{\frac{2\pi mk}{M}i} = 0$$

$$\iff \sum_{m=1}^{M-1} (c_m - c_{M-m}) e^{\frac{2\pi mk}{M}i} = \sum_{m=1}^{M-1} (c_m - c_{M-m}) \omega^{mk} = 0$$

for all $k \in \{0, 1, \dots, M-1\}$. In other words, the eigenvalues of \mathbf{C} are all real if and only if the vector of differences $(c_m - c_{M-m})_{m=1, \dots, M-1}$ is a vector in the null space of

$$\begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \omega & \omega^2 & \omega^3 & \dots & \omega^{M-1} \\ \omega^2 & \omega^4 & \omega^6 & \dots & \omega^{2(M-1)} \\ \omega^3 & \omega^6 & \omega^9 & \dots & \omega^{3(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \omega^{M-1} & \omega^{2(M-1)} & \omega^{3(M-1)} & \dots & \omega^{(M-1)^2} \end{bmatrix}.$$

Noting that these columns are eigenvectors of \mathbf{C} , it can quickly be recognized that the null space of this matrix is the line $x_1 = x_2 = \dots = x_M$, as this is the final orthogonal eigenvector of \mathbf{C} . Therefore, our differences must satisfy

$$c_m - c_{M-m} = a$$

for some $a \in \mathbb{C}$ for all $m \in \{1, 2, \dots, M-1\}$. If M is even, then $M/2 \in \{1, 2, \dots, M-1\}$, and so when $m = M/2$ we get the difference $c_{M/2} - c_{M-M/2} = c_{M/2} - c_{M/2} = 0$. Hence, $a = 0$ is the only solution. If M is odd, then we have in particular $c_{\lfloor M/2 \rfloor} - c_{M-\lfloor M/2 \rfloor} = c_{\lfloor M/2 \rfloor} - c_{\lfloor M/2 \rfloor+1} = a = c_{\lfloor M/2 \rfloor+1} - c_{\lfloor M/2 \rfloor} = c_{\lfloor M/2 \rfloor+1} - c_{M-\lfloor M/2 \rfloor-1}$, which is true only if $c_{\lfloor M/2 \rfloor+1} = c_{\lfloor M/2 \rfloor}$ and $a = 0$. Therefore, real eigenvalues are ensured if and only if

$$c_m - c_{M-m} = 0 \iff c_m = c_{M-m},$$

the definition of a symmetric circulant. □

2 The Nearest Circulant to Σ

Now consider the decomposition

$$\Sigma = \mathbf{C} + \mathbf{R} \tag{10}$$

where \mathbf{C} is a circulant matrix and $\mathbf{R} = \Sigma - \mathbf{C}$ is a matrix of the element-wise residuals between \mathbf{C} and Σ . This reframing moves the discussion from the space of asymptotic results to the features of \mathbf{C} and \mathbf{R} , a familiar framework for anyone accustomed to considering the residuals of a given approximation.

Of immediate and obvious interest is the closest circulant matrix to Σ , call it \mathbf{C}_Σ . Consider using the weak, or Frobenius, norm on matrices. First, introduce the *vectorization* operator on a matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$, denoted $\text{vec } \mathbf{A}$. This operator takes $\mathbf{A} \in \mathbb{C}^{M \times N}$ and converts it to a vector in \mathbb{C}^{MN} by appending columns in order. So, for example,

$$\text{vec} \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 2 \\ 5 \end{bmatrix}.$$

The Frobenius norm is then given by

$$\|\mathbf{A}\|_F = \sqrt{(\text{vec } \mathbf{A})^* (\text{vec } \mathbf{A})}$$

or equivalently

$$\|\mathbf{A}\|_F = \sqrt{\text{trace}(\mathbf{A}^* \mathbf{A})}$$

for a matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$ with complex conjugate \mathbf{A}^* . In the real case, this is

$$\|\mathbf{A}\|_F = \sqrt{\text{trace}(\mathbf{A}^\top \mathbf{A})}$$

Let \mathbf{C} be an $M \times M$ circulant matrix with first row $(c_0, c_1, c_2, \dots, c_{M-1}) \in \mathbb{R}^M$. Then by definition \mathbf{C}_Σ is given by $\arg \min_{\mathbf{C}} \|\Sigma - \mathbf{C}\|_F$, or equivalently $\arg \min_{\mathbf{C}} \|\Sigma - \mathbf{C}\|_F^2$. Taking the second of these, we have

$$\mathbf{C}_\Sigma = \arg \min_{\mathbf{C}} \|\Sigma - \mathbf{C}\|_F^2. \quad (11)$$

Considering that

$$\begin{aligned} \|\Sigma - \mathbf{C}\|_F^2 &= \text{trace} \left([\Sigma - \mathbf{C}]^\top [\Sigma - \mathbf{C}] \right) \\ &= \left(\text{trace} \Sigma^\top \Sigma - \text{trace} \Sigma^\top \mathbf{C} - \text{trace} \mathbf{C}^\top \Sigma + \text{trace} \mathbf{C}^\top \mathbf{C} \right) \end{aligned}$$

and $\text{trace} \Sigma^\top \Sigma$ is constant in \mathbf{C} , we need only consider minimizing

$$F(\mathbf{C}) = \text{trace} \mathbf{C}^\top \mathbf{C} - \text{trace} \Sigma^\top \mathbf{C} - \text{trace} \mathbf{C}^\top \Sigma. \quad (12)$$

The first term of Equation (12) is straightforward to express in terms of the c_m . As \mathbf{C} is circulant,

$$\text{trace} \mathbf{C}^\top \mathbf{C} = M \sum_{m=0}^{M-1} c_m^2.$$

The other terms can be evaluated by expressing \mathbf{C} as a matrix polynomial.

Equation (7) and the symmetry of Σ allow us to write

$$\Sigma^\top \mathbf{C} = \Sigma \left(c_0 \mathbf{I} + \sum_{m=1}^{M-1} c_m \mathbf{P}^m \right) = c_0 \Sigma + \sum_{m=1}^{M-1} c_m \Sigma \mathbf{P}^m$$

and similarly

$$\mathbf{C}^\top \Sigma = c_0 \Sigma + \sum_{m=1}^{M-1} c_m (\mathbf{P}^m)^\top \Sigma = c_0 \Sigma + \sum_{m=1}^{M-1} c_m (\Sigma \mathbf{P}^m)^\top.$$

Next, consider

$$\text{trace} \Sigma \mathbf{P}^m = \text{trace} (\Sigma [\mathbf{e}_{M-m+1} | \mathbf{e}_{M-m+2} | \dots | \mathbf{e}_{M-m}]),$$

which can be evaluated by considering the k^{th} row of Σ , σ_k . The first k elements of this row are the descending sequence $\rho_{k-1}, \rho_{k-2}, \dots, \rho_0$, and the remaining elements are the

ascending sequence $\rho_1, \rho_2, \dots, \rho_{M-k}$. Noting that the trace of a product of two matrices is simply a sum of the inner products of the rows of the first with the columns of the second, we obtain

$$\begin{aligned}
\text{trace } \Sigma \mathbf{P}^m &= \sum_{k=1}^m \boldsymbol{\sigma}_k^T \mathbf{e}_{M-m+k} + \sum_{k=1}^{M-m} \boldsymbol{\sigma}_{m+k}^T \mathbf{e}_k \\
&= \sum_{k=1}^m \rho_{M-m} + \sum_{k=1}^{M-m} \rho_m \\
&= m\rho_{M-m} + (M-m)\rho_m.
\end{aligned} \tag{13}$$

Equation (13) can then be substituted into Equation (12) using the decomposition of Equation (7) to give

$$F(\mathbf{C}) = M \sum_{m=0}^{M-1} c_m^2 - 2 \left(M c_0 \rho_0 + \sum_{m=1}^{M-1} c_m (m\rho_{M-m} + (M-m)\rho_m) \right). \tag{14}$$

As $\arg \min_{\mathbf{C}} F(\mathbf{C})$ is the same as $\arg \min_{\mathbf{C}} \|\Sigma - \mathbf{C}\|_F$, we can now consider the values which minimize Equation (14) in order to find the nearest circulant matrix to Σ . Taking

$$\frac{\partial}{\partial c_m} F(\mathbf{C}) = \begin{cases} 2Mc_0 - 2M\rho_0 & \text{for } m = 0, \\ 2Mc_m - 2(m\rho_{M-m} + (M-m)\rho_m) & \text{otherwise,} \end{cases}$$

and noting that the Hessian matrix is $2M\mathbf{I}$ and hence is positive definite so any solutions to $\arg \min F(\mathbf{C})$ must be minima, we obtain

$$c_m = \begin{cases} \rho_0 & \text{for } m = 0, \\ \frac{m}{M}\rho_{M-m} + \frac{M-m}{M}\rho_m & \text{otherwise,} \end{cases}$$

which can be re-expressed as

$$c_m = \begin{cases} \rho_0 & \text{for } m = 0, \\ \rho_m + \frac{m}{M}(\rho_{M-m} - \rho_m) & \text{otherwise,} \end{cases} \tag{15}$$

to make the relationship between c_m and ρ_m clearer. An important consequence of this system of equations is that $c_m = \frac{m}{M}\rho_{M-m} + \frac{M-m}{M}\rho_m = \frac{M-(M-m)}{M}\rho_{M-m} + \frac{M-m}{M}\rho_{M-(M-m)} = c_{M-m}$, and so \mathbf{C}_Σ is a symmetric circulant matrix with c_m defined as in Equation (15). Therefore, this nearest circulant will have only real eigenvalues.

So the optimal decomposition in the Frobenius norm is

$$\mathbf{\Sigma} = \mathbf{C}_\Sigma + \mathbf{R}_\Sigma \quad (16)$$

where \mathbf{C}_Σ is the circulant matrix defined by Equation (15), that is

$$c_m = \begin{cases} \rho_0 & \text{for } m = 0, \\ \rho_m + \frac{m}{M}(\rho_{M-m} - \rho_m) & \text{otherwise,} \end{cases}$$

and $\mathbf{R}_\Sigma = \mathbf{\Sigma} - \mathbf{C}_\Sigma$, and so the value in the m^{th} off-diagonal of \mathbf{R}_Σ is $\rho_m - c_m = \rho_m - (\rho_m + \frac{m}{M}(\rho_{M-m} - \rho_m)) = \frac{m}{M}(\rho_m - \rho_{M-m})$.

The eigenvalues of \mathbf{C}_Σ are given by a substitution of Equation (15) into Equation (9), giving

$$\lambda_k(\mathbf{C}_\Sigma) = \rho_0 + 2 \sum_{m=1}^{M-1} \frac{M-m}{M} \rho_m \cos \frac{2\pi mk}{M}. \quad (17)$$

The corresponding eigenvectors are given by Equation (8).

Consider \mathbf{R}_Σ briefly. For large M and small m , $\frac{m}{M}(\rho_{M-m} - \rho_m) \approx 0$ while when m is close to M , $\frac{m}{M}(\rho_m - \rho_{M-m}) \approx \rho_m - \rho_{M-m}$. This implies that in the case of large M , \mathbf{R}_Σ will have vanishingly small values for the central off-diagonals, and values in the corners of approximately $\rho_m - \rho_{M-m}$.

In the particular case of $\mathbf{\Sigma}_e$ from Equation (2), $\rho_m = \rho^m$. In this case, \mathbf{C}_{Σ_e} has entries

$$c_m = \rho^m + \frac{m}{M}(\rho^{M-m} - \rho^m)$$

and \mathbf{R}_{Σ_e} is $\frac{m}{M}(\rho^m - \rho^{M-m})$ for the m^{th} off-diagonal where $m \in \{0, 1, 2, \dots, M-1\}$. The eigenvalues of \mathbf{C}_{Σ_e} are therefore

$$\lambda_k(\mathbf{C}_{\Sigma_e}) = 1 + 2 \sum_{m=1}^{M-1} \frac{M-m}{M} \rho^m \cos \frac{2\pi mk}{M}.$$

3 Asymptotic Equivalence

The approximation matrix \mathbf{C}_Σ has been derived here based purely on minimizing $\|\Sigma - \mathbf{C}\|_F$ without any of the asymptotic guarantees of [Grenander and Szegö \(1958\)](#). [Gray \(2006\)](#) derives similar, but less general, results by considering the asymptotic equivalence of matrices in the weak norm. Matrices \mathbf{A} and \mathbf{B} in $\mathbb{C}^{M \times M}$ are said to be asymptotically equivalent in the weak norm if

$$\lim_{M \rightarrow \infty} \frac{1}{\sqrt{M}} \|\mathbf{A} - \mathbf{B}\|_F = 0.$$

Therefore, a natural consideration is the difference

$$\lim_{M \rightarrow \infty} \frac{1}{\sqrt{M}} \|\Sigma - \mathbf{C}_\Sigma\|_F = \lim_{M \rightarrow \infty} \frac{1}{\sqrt{M}} \|\mathbf{R}_\Sigma\|_F. \quad (18)$$

Before taking the limit, consider

$$\frac{1}{\sqrt{M}} \|\mathbf{R}_\Sigma\|_F = \sqrt{\frac{1}{M} \text{trace } \mathbf{R}_\Sigma^\top \mathbf{R}_\Sigma},$$

which has the square

$$\begin{aligned} \frac{1}{M} \|\mathbf{R}_\Sigma\|_F^2 &= \frac{1}{M} \text{trace } \mathbf{R}_\Sigma^\top \mathbf{R}_\Sigma \\ &= \frac{1}{M} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{|i-j|^2}{M^2} (\rho^{2|i-j|} - 2\rho^{|i-j|+M-|i-j|} + \rho^{2(M-|i-j|)}) \\ &= \frac{2}{M^3} \sum_{m=1}^{M-1} (M-m)m^2 (\rho^{2m} - 2\rho^M + \rho^{2(M-m)}) \end{aligned} \quad (19)$$

Evaluating this expression for $\rho < 1$ is made easier by considering the general sum

$$\sum_{m=1}^M m^k \rho^m.$$

Theorem 2 (Truncated geometric power series). *The finite sum*

$$\sum_{m=1}^M m^k \rho^m$$

can be expressed as

$$\frac{\rho}{1-\rho} \left[(1-\rho^M) \frac{d^k}{dt^k} G(t) \Big|_{t=0} - \rho^M \sum_{l=1}^k \binom{k}{l} M^l \frac{d^{k-l}}{dt^{k-l}} G(t) \Big|_{t=0} \right]$$

where

$$G(t) = E[e^{itX}]$$

is the moment generating function for $X \sim \text{Geo}(1-\rho)$.

Proof. Multiplying by $\frac{1-\rho}{1-\rho}$, we obtain

$$\sum_{m=1}^M m^k \rho^m = \frac{\rho}{1-\rho} \sum_{m=1}^M m^k \rho^{m-1} (1-\rho).$$

Now

$$\begin{aligned} \sum_{m=1}^M m^k \rho^{m-1} (1-\rho) &= \sum_{m=1}^{\infty} m^k \rho^{m-1} (1-\rho) - \rho^M \sum_{m=M+1}^{\infty} m^k \rho^{m-M-1} (1-\rho) \\ &= \sum_{m=1}^{\infty} m^k \rho^{m-1} (1-\rho) - \rho^M \sum_{m=1}^{\infty} (m+M)^k \rho^{m-1} (1-\rho) \\ &= \sum_{m=1}^{\infty} m^k \rho^{m-1} (1-\rho) - \rho^M \sum_{m=1}^{\infty} \sum_{l=0}^k \binom{k}{l} m^{k-l} M^l \rho^{m-1} (1-\rho) \\ &= (1-\rho^M) \sum_{m=1}^{\infty} m^k \rho^{m-1} (1-\rho) \\ &\quad - \rho^M \sum_{l=1}^k \binom{k}{l} M^l \sum_{m=1}^{\infty} m^{k-l} \rho^{m-1} (1-\rho), \end{aligned} \tag{20}$$

but

$$\sum_{m=1}^{\infty} m^k \rho^{m-1} (1-\rho)$$

is just the k^{th} moment of a geometric distribution with a probability of success of $1-\rho$. This distribution has a moment-generating function

$$G(t) = \frac{1-\rho}{e^{-t}-\rho} \tag{21}$$

which can be used to evaluate the k^{th} moment by taking the k^{th} derivative and evaluating it at $t = 0$, denoted

$$\left. \frac{d^k}{dt^k} G(t) \right|_{t=0}$$

and substituted into Equation 20 to give the more succinct

$$\sum_{m=1}^M m^k \rho^{m-1} (1 - \rho) = (1 - \rho^M) \left. \frac{d^k}{dt^k} G(t) \right|_{t=0} - \rho^M \sum_{l=1}^k \binom{k}{l} M^l \left. \frac{d^{k-l}}{dt^{k-l}} G(t) \right|_{t=0}. \quad (22)$$

This gives the result for the original sum

$$\sum_{m=1}^M m^k \rho^m = \frac{\rho}{1 - \rho} \left[(1 - \rho^M) \left. \frac{d^k}{dt^k} G(t) \right|_{t=0} - \rho^M \sum_{l=1}^k \binom{k}{l} M^l \left. \frac{d^{k-l}}{dt^{k-l}} G(t) \right|_{t=0} \right]. \quad (23)$$

□

As expected, taking the limit $M \rightarrow \infty$ reduces this expression to the k^{th} geometric moment multiplied by $\rho/(1 - \rho)$. and expresses it as a linear combination of the first k moments of the geometric distribution. In particular, for $k = 1$ and 2 we have

$$\sum_{m=1}^M m \rho^m = \frac{\rho}{(1 - \rho)^2} \left[1 - M \rho^{M-1} + (M - 1) \rho^M \right], \quad (24)$$

$$\begin{aligned} \sum_{m=1}^M m^2 \rho^m &= \frac{\rho}{(1 - \rho)^3} \left[1 + \rho - M^2 \rho^{M-1} + (3M^2 - (M + 1)^2) \rho^M \right. \\ &\quad \left. - 2(M - 1)^2 \rho^{M+1} \right]. \end{aligned} \quad (25)$$

Critically, Equations 24 and 25 demonstrate that as $M \rightarrow \infty$ Equation 23 approaches a value proportional to the k^{th} moment of the geometric distribution. These moments are finite for $k = 1$ and $k = 2$.

We can expand Equation 19 to give

$$\begin{aligned} \frac{1}{M} \|\mathbf{R}_\Sigma\|_F^2 &= \frac{2}{M^3} \sum_{m=1}^{M-1} (M - m) m^2 (\rho^{2m} - 2\rho^M + \rho^{2(M-m)}) \\ &= \frac{2}{M^2} \left[\sum_{m=1}^{M-1} (M^2 - 2Mm + 2m^2) \rho^{2m} - 2\rho^M \sum_{m=1}^{M-1} m^2 \right] \\ &\quad - \frac{2}{M^3} \left[\sum_{m=1}^{M-1} (M^3 - 3M^2m + 3Mm^2) \rho^{2m} - 2\rho^M \sum_{m=1}^{M-1} m^3 \right] \end{aligned}$$

Each term in this sum can then be considered separately. The final term is the sum of the first $M - 1$ squared integers multiplied by a constant and so

$$2\rho^M \sum_{m=1}^{M-1} m^2 = 2\rho^M \frac{(M-1)M(2M-1)}{6} = \frac{1}{3}\rho^M (M-1)M(2M-1). \quad (26)$$

Next,

$$\sum_{m=1}^{M-1} [2m^2 + M^2 - 2Mm] \rho^{2m}$$

can be split into three separate sums. The simplest of these to reduce is

$$M^2 \sum_{m=1}^{M-1} \rho^{2m} = M^2 \frac{\rho^2 - \rho^{2M}}{1 - \rho^2} = M^2 r (1 - \rho^{2(M-1)}) \quad (27)$$

when $\rho < 1$, where $r = \frac{\rho^2}{1-\rho^2}$ is useful shorthand. The restriction $\rho < 1$ corresponds to finite r . Considering next

$$\begin{aligned} 2M \sum_{m=1}^{M-1} m \rho^{2m} &= 2M \rho^2 \frac{d}{d\rho^2} \sum_{m=0}^{M-1} \rho^{2m} \\ &= 2M \rho^2 \frac{d}{d\rho^2} \frac{\rho^2 - \rho^{2M}}{1 - \rho^2} \\ &= 2M \rho^2 \left(\frac{1 - M\rho^{2(M-1)}}{1 - \rho^2} + \frac{\rho^2 - \rho^{2M}}{(1 - \rho^2)^2} \right) \\ &= 2Mr \left(1 - M\rho^{2(M-1)} + r - \frac{\rho^{2M}}{1 - \rho^2} \right) \\ &= 2Mr (1 + r - \rho^{2(M-1)}[M + r]). \end{aligned} \quad (28)$$

The same derivative trick can be applied to obtain

$$\begin{aligned}
\sum_{m=1}^{M-1} m^2 \rho^{2m} &= \sum_{m=1}^{M-1} m(m-1) \rho^{2m} + m \rho^{2m} \\
&= \rho^4 \frac{d^2}{(d\rho^2)^2} \sum_{m=1}^{M-1} \rho^{2m} + \rho^2 \frac{d}{d\rho^2} \sum_{m=1}^{M-1} \rho^{2m},
\end{aligned} \tag{29}$$

the second term of which can be determined easily using Equation (27) above. With a second differentiation, the first term is

$$\begin{aligned}
\rho^4 \frac{d^2}{(d\rho^2)^2} \sum_{m=1}^{M-1} \rho^{2m} &= \rho^4 \frac{d}{d\rho^2} \left(\frac{1 - M\rho^{2(M-1)}}{1 - \rho^2} + \frac{\rho^2 - \rho^{2M}}{(1 - \rho^2)^2} \right) \\
&= \rho^4 \left(\frac{-M(M-1)\rho^{2(M-2)}}{1 - \rho^2} + \frac{2(1 - M\rho^{2(M-1)})}{(1 - \rho^2)^2} + \frac{2(\rho^2 - \rho^{2M})}{(1 - \rho^2)^3} \right).
\end{aligned}$$

Substituting this and Equation (27) into Equation (29) gives

$$\begin{aligned}
\sum_{m=1}^{M-1} m^2 \rho^{2m} &= \rho^2 \left(\frac{1 - M\rho^{2(M-1)}}{1 - \rho^2} + \frac{\rho^2 - \rho^{2M}}{(1 - \rho^2)^2} \right) \\
&\quad + \rho^4 \left(\frac{-M(M-1)\rho^{2(M-2)}}{1 - \rho^2} + \frac{2(1 - M\rho^{2(M-1)})}{(1 - \rho^2)^2} + \frac{2(\rho^2 - \rho^{2M})}{(1 - \rho^2)^3} \right) \\
&= \rho^2 \left(\frac{1 - M^2\rho^{2(M-1)}}{1 - \rho^2} + \frac{3\rho^2 - (2M+1)\rho^{2M}}{(1 - \rho^2)^2} + \frac{2\rho^2(\rho^2 - \rho^{2M})}{(1 - \rho^2)^3} \right) \\
&= r \left(1 - M^2\rho^{2(M-1)} + 3r - (2M+1)r\rho^{2(M-1)} + 2r^2 - 2r^2\rho^{2(M-1)} \right) \\
&= r \left(1 + 3r + 2r^2 - [M^2 + (2M+1)r + 2r^2]\rho^{2(M-1)} \right) \\
&= r \left[(1+r)(1+2r) - \{(M+r)^2 + r(1+r)\}\rho^{M-1} \right]
\end{aligned} \tag{30}$$

So that Equation (??) becomes

$$\begin{aligned}
\frac{1}{M} \|\mathbf{R}_\Sigma\|_F^2 &= \frac{2}{M^3} \left(2r \left[(1+r)(1+2r) - \{(M+r)^2 + r(1+r)\} \rho^{M-1} \right] \right. \\
&\quad + M^2 r \left[1 - \rho^{2(M-1)} \right] - 2Mr \left[1 + r - \rho^{2(M-1)}(M+r) \right] \\
&\quad \left. + \frac{1}{3} \rho^M (M-1)M(2M-1) \right) \\
&= \frac{2}{M^3} \left(r \left[2(1+r)(1+2r-2M) + M^2 \right. \right. \\
&\quad \left. \left. - \{(M+r)^2 + r(2r+3)\} \rho^{2(M-1)} \right] \right. \\
&\quad \left. - \frac{1}{3} \rho^M M(M-1)(2M-1) \right) \tag{31} \\
&= \frac{2r}{M^3} \left[2(1+r)(1+2r-2M) + M^2 - \{(M+r)^2 + r(2r+3)\} \rho^{2(M-1)} \right] \\
&\quad - \frac{2\rho^M}{3M^2} (M-1)(2M-1) \\
&= 2 \left(r[1 - \rho^{2(M-1)}] + \rho^M \right) \frac{1}{M} - \left(4r[2 + 2r - r\rho^{2(M-1)}] - \frac{2}{3} \rho^M \right) \frac{1}{M^2} \\
&\quad + 2r(1+r) \left(1 + 2r - 3r\rho^{2(M-1)} \right) \frac{1}{M^3} - \frac{4}{3} \rho^M
\end{aligned}$$

when $\rho < 1 \iff r < \infty$. Under these constraints, the coefficients of Equation (31) are finite for all $M > 0$, and so

$$\lim_{M \rightarrow \infty} \frac{1}{M} \|\mathbf{R}_\Sigma\|_F^2 = 0 \implies \lim_{M \rightarrow \infty} \frac{1}{\sqrt{M}} \|\mathbf{R}_\Sigma\|_F = 0.$$

Therefore, \mathbf{C}_Σ and Σ are asymptotically equivalent in the weak norm.

4 Rate of Convergence

Gray (2006) suggests a slightly different approximation. Following Grenander and Szegö

(1958), the sum

$$f(x) = \sum_{k=-\infty}^{\infty} \rho_k e^{ikx} = \sum_{k=-\infty}^{\infty} \rho^{|k|} e^{ikx}$$

is known to be critical to the approximation of Σ . Gray (2006) considers circulant entries generated using the expression

$$c_m = \begin{cases} \rho_0 & \text{for } m = 0, \\ \frac{1}{M} \sum_{j=0}^{M-1} f\left(\frac{2\pi j}{M}\right) e^{\frac{2\pi i j m}{M}} & \text{otherwise,} \end{cases} \quad (32)$$

the second case can be expressed

$$\begin{aligned} \frac{1}{M} \sum_{j=0}^{M-1} f\left(\frac{2\pi j}{M}\right) e^{\frac{2\pi i j m}{M}} &= \frac{1}{M} \sum_{j=0}^{M-1} \sum_{k=-\infty}^{\infty} \rho^{|k|} e^{ik \frac{2\pi j}{M}} e^{\frac{2\pi i j m}{M}} \\ &= \frac{1}{M} \sum_{j=0}^{M-1} \sum_{k=-\infty}^{\infty} \rho^{|k|} e^{\frac{2\pi i j}{M}(m+k)} \\ &= \sum_{k=-\infty}^{\infty} \rho^{|k|} \frac{1}{M} \sum_{j=0}^{M-1} e^{\frac{2\pi i j}{M}(m+k)} \\ &= \sum_{k=-\infty}^{\infty} \rho^{|k|} I_{(m+k) \bmod M}(0) \end{aligned}$$

as the second sum is the sum of the squared M^{th} roots of unity, which are orthonormal.

Now

$$\begin{aligned}
\sum_{k=-\infty}^{\infty} \rho^{|k|} I_{(m+k) \bmod M}(0) &= \sum_{k=-\infty}^{\infty} \rho^{|-m+kM|} \\
&= \sum_{k=-\infty}^0 \rho^{m-kM} + \sum_{k=1}^{\infty} \rho^{-m+kM} \\
&= \rho^m \frac{1}{1-\rho^M} + \rho^{-m} \frac{\rho^M}{1-\rho^M} \\
&= \frac{1}{1-\rho^M} (\rho^m + \rho^{M-m})
\end{aligned}$$

when $\rho < 1$. Therefore, Equation (32) becomes

$$c_m = \begin{cases} \rho_0 & \text{for } m = 0, \\ \frac{1}{1-\rho^M} (\rho^m + \rho^{M-m}) & \text{otherwise,} \end{cases} \quad (33)$$

in the particular case of $\rho_m = \rho^m$. So, while \mathbf{C}_{Σ} has entries which are a weighted average of ρ^m and ρ^{M-m} , this approximation instead takes a simple sum scaled by $1 - \rho^M$. Define \mathbf{C}_{GS} as the circulant matrix with these entries, and let \mathbf{R}_{GS} be $\Sigma - \mathbf{C}_{GS}$. Then

$$\frac{1}{M} \|\mathbf{R}_{GS}\|_F^2 = \frac{2}{(1-\rho^M)^2} \left[\frac{2}{M(1-\rho^2)} + 2 \left(1 - \frac{1}{M} \right) \rho^{2M} - \frac{1}{M(1-\rho^2)} \rho^{4M} \right] \quad (34)$$

while an evaluation of Equation (31) gives

$$\begin{aligned}
\frac{1}{M} \|\mathbf{R}_{\Sigma}\|_F^2 &= \frac{2\rho^2}{M^3} \left[\frac{M^2}{1-\rho^2} + \frac{1-2M-\rho^{2M}}{(1-\rho^2)^2} + \frac{2\rho^2(1-\rho^M)}{(1-\rho^2)^3} \right. \\
&\quad \left. - 2\rho^M M(M-1)(2M-1) \right] \quad (35)
\end{aligned}$$

Both of these have a leading term approximately equal to

$$\frac{2}{M(1-\rho^2)}$$

for large M , and so their asymptotic behaviour is identical for all ρ . For small M , however, it is helpful to plot these values and compare these squared distances.

References

- James M Cheverud. A simple correction for multiple comparisons in interval mapping genome scans. *Heredity*, 87(1):52–58, 2001.
- Nicholas W Galwey. A new measure of the effective number of tests, a practical tool for comparing families of non-independent significance tests. *Genetic Epidemiology*, 33(7): 559–568, 2009.
- Robert M Gray. *Toeplitz and circulant matrices: A review*. now Publishers Inc., 2006.
- Ulf Grenander and Gabor Szegő. *Toeplitz forms and their applications*. University of California Press, 1958.
- Eric S Lander and David Botstein. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 121(1):185–199, 1989.
- J Li and L Ji. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity*, 95(3):221–227, 2005.
- Christopher Salahub. A structural model of genome-wide association studies, 2022. URL <https://arxiv.org/abs/2205.10391>.