THE WONDERFUL WORLD OF EIGENVALUES

ARUP BOSE*
Stat-Math Unit
Indian Statistical Institute
Kolkata 700108

ALOKE DEY†
Stat-Math Unit
Indian Statistical Institute
New Delhi 110016

May 28, 2010

1 Introduction

Linear transformations can be represented by matrices which act on vectors. Eigenvalues, eigenvectors and eigenspaces are properties of a matrix.[1] They capture all the essential properties of the matrix or the corresponding transformation. Historically, the importance of eigenvalues and the corresponding eigenvectors arose from studies in physics and in the study of quadratic forms and differential equations. The concepts of eigenvalues and vectors extend to linear transformations in more general spaces. These have applications in many different areas of science; in particular, in economics, engineering, finance, quantum mechanics, mathematics and statistics.

In this article, we first describe some basic facts and mathematical results on eigenvalues and also provide the eigenvalues of k-circulants. We then provide a glimpse of a few applications–the role of eigenvalues and eigenvectors in principal component analysis, some simulations of eigenvalue distributions of a few random matrices and, application of the extended idea of eigenvalues for general linear functionals to an important result in asymptotic theory of statistics.

2 Basic concepts and results

2.1 Notation and preliminaries

For simplicity, we initially restrict attention to matrices with real entries. The symbols $\mathbb{R}^n$ and $\mathbb{R}$ will stand respectively for the n-dimensional Euclidean space and the set of reals. All vectors will be written as column vectors and denoted by bold-face lower case letters or numerals. A prime over a matrix or a vector will denote its transpose. For instance, A' is the transpose of A. For positive integers s, t, I_s is the identity matrix of order s and $\mathbf{0}_{st}$ is the $s \times t$ null matrix. The $s \times 1$ vector $\mathbf{0}_{s1}$ will be denoted simply by $\mathbf{0}_s$. The subscripts indicating the order of matrices and vectors may not always be written out when it is obvious what they are in a given context. The length (or Euclidean norm) $||\boldsymbol{x}||$, of a vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)' \in \mathbb{R}^n$,

*Research supported by J.C.Bose National Fellowship, Dept. of Science and Technology, Govt. of India. This work was done while visiting the Department of Economics, University of Cincinnati.

†Research supported by the Indian National Science Academy under its Senior Scientist scheme.

[1]The prefix *eigen* is the German word for innate, distinct, self or proper.

is the positive square root of $(x_1^2 + x_2^2 + \cdots + x_n^2)$. The diagonal matrix

$$\begin{bmatrix} a_1 & 0 & 0 & \cdots & 0 \\ 0 & a_2 & 0 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & a_n \end{bmatrix}$$

will be abbreviated as $\mathrm{diag}(a_1, a_2, \ldots, a_n)$. A square matrix A is said to be *symmetric* if $A' = A$ and *idempotent*, if $A^2 = A$. For a square matrix A, $\det(A)$ and $\mathrm{tr}(A)$ will respectively, denote its determinant and trace (the sum of the entries on the principal diagonal of A). A pair of vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ are said to be orthogonal to each other if $\boldsymbol{x}'\boldsymbol{y} = 0$. A square matrix P of order n is said to be *orthogonal* if $PP' = I_n$. As a consequence, for an orthogonal matrix P, we also have $P'P = I_n$. Two $n \times n$ matrices $A,\ B$ are called *similar* if there exist a non-singular matrix S such that $A = S^{-1}BS$.

2.2 Linear transformation and matrices

A transformation (or map) $T : \mathbb{R}^n \to \mathbb{R}^m$ associates to every vector $\boldsymbol{x} \in \mathbb{R}^n$, a unique vector $\boldsymbol{y} \in \mathbb{R}^m$. T is said to be a *linear map* if $T(\alpha\boldsymbol{x} + \boldsymbol{y}) = \alpha T(\boldsymbol{x}) + T(\boldsymbol{y})$ for each pair $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and every $\alpha \in \mathbb{R}$.

Linear maps are intimately connected to matrices. Let $B = (b_{ij})$ be an $m \times n$ matrix and let $\boldsymbol{y} = B\boldsymbol{x}$ where $\boldsymbol{x} \in \mathbb{R}^n$, and consider the map

$$T_B(\boldsymbol{x}) = B\boldsymbol{x}. \tag{1}$$

It is obvious that T_B is a linear transformation. In particular, any $n \times n$ matrix A may be viewed as a linear map from $\mathbb{R}^n$ into itself. Conversely, one can show the following.

Lemma 1. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Then there exists an $m \times n$ matrix B such that $T(\boldsymbol{x}) = B\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$.*

2.3 Eigenvector, eigenvalue and algebraic multiplicity

Suppose A is an $n \times n$ (real) matrix. Then a non-null $n \times 1$ vector $\boldsymbol{x}$ is called a (right) *eigenvector* (of A) if there is a scalar λ (possibly complex) satisfying

$$A\boldsymbol{x} = \lambda\boldsymbol{x}. \tag{2}$$

The scalar λ is called an *eigenvalue* of A associated with the eigenvector $\boldsymbol{x}$. From equation (2) it is clear that λ is an eigenvalue of A if and only if there is a non-null solution to the equation

$$(A - \lambda I_n)\boldsymbol{x} = \mathbf{0}_n. \tag{3}$$

Recall that equation (3) has a non-null solution (i.e., $\boldsymbol{x} \neq \mathbf{0}_n$) if and only if the matrix $A - \lambda I_n$ is *not* invertible, i.e., if and only if

$$\det(A - \lambda I_n) = 0. \tag{4}$$

Note that $\det(A - \lambda I_n)$ is a polynomial of degree n. It is called the *characteristic polynomial* of A and shall be denoted by $\mathrm{ch_A}(\lambda)$. Equation (4) is called the *characteristic equation* of A. Hence the eigenvalues of A are simply the roots of the characteristic polynomial and are also known as the characteristic roots of A. The *algebraic multiplicity* of an eigenvalue is its multiplicity as a root of equation (4).

If λ is an eigenvalue of A, then it turns out that there exists a (left) eigenvector $\boldsymbol{y}$ such that $\boldsymbol{y}'A = \lambda\boldsymbol{y}'$. This implies that A and A' have identical left and right eigenvectors corresponding to an eigenvalue. In this article all eigenvectors are taken to be right eigenvectors.

Equation (2) implies that the vector $\boldsymbol{x}$ has the property that its direction is unchanged by the transformation A and it is only scaled by a factor of λ. Geometrically, this means that eigenvectors experience changes in magnitude and sign and $|\lambda|$ is the amount of "stretch" (if $|\lambda| > 1$) or "shrink" (if $|\lambda| < 1$) to which the eigenvector is subjected to when transformed by A.

Since we are restricting attention to matrices with real entries, all coefficients of $\mathrm{ch_A}(\lambda)$ are real. Its roots of course need not necessarily be real.

Example 1. Let

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \end{bmatrix}.$$

The characteristic polynomial of A is

$$\mathrm{ch_A}(\lambda) = (1-\lambda)(\lambda^2 + \lambda + 1).$$

Hence the eigenvalues of A are 1, $(-1 - i\sqrt{3})/2$ and $(-1 + i\sqrt{3})/2$, where $i^2 = -1$.

2.4 Eigenspace and geometric multiplicity

There may be two (or more) different eigenvectors corresponding to the same eigenvalue. If $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are eigenvectors corresponding to the same eigenvalue λ of A, then for $\alpha, \beta \in \mathbb{R}$, $\alpha\boldsymbol{x}_1 + \beta\boldsymbol{x}_2$ is also an eigenvector with the same eigenvalue λ. In particular if $\boldsymbol{x}$ is an eigenvector then for any α, $\alpha\boldsymbol{x}$ is also an eigenvector with the same eigenvalue.

Consider the set of eigenvectors of A corresponding to the same eigenvalue λ of A, along with the null vector. It follows from the above discussion that, these vectors form a vector subspace, called the *eigenspace* of λ. The dimension of the eigenspace of an eigenvalue is called its *geometric multiplicity*.

Recall that for any $m \times n$ matrix B, the collection of $n \times 1$ vectors $\{\boldsymbol{x}\}$ satisfying $B\boldsymbol{x} = \mathbf{0}_m$ forms a subspace of $\mathbb{R}^n$, called the null space of B, denoted by $\mathcal{N}(B)$. The dimension of $\mathcal{N}(B)$ is called the nullity of B, denoted by $\nu(B)$. Furthermore, the *rank-nullity theorem* states that

$$\mathrm{Rank}(B) + \mathrm{Nullity}(B) = n, \text{ the number of columns in } B.$$

One can verify easily that the geometric multiplicity of an eigenvalue λ of A equals the nullity of $A - \lambda I_n$.

Example 2. Let

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

The characteristic equation of A is easily seen to be

$$(1-\lambda)^3 = 0,$$

which shows that 1 is an eigenvalue with algebraic multiplicity 3.

The geometric multiplicity of 1 equals

$$\begin{aligned} \nu(A - I) &= 3 - \mathrm{Rank}(A - I), \text{ by the rank-nullity theorem,} \\ &= 3 - 2 = 1. \end{aligned}$$

Thus, here the algebraic multiplicity of 1 is strictly larger than its geometric multiplicity.

In general, the following holds.

Theorem 1. *For any eigenvalue λ, the geometric multiplicity cannot exceed the algebraic multiplicity.*

2.5 Elementary results on eigenvalues

Lemma 2. *The sum of the eigenvalues of A is $\mathrm{tr}(A)$ and their product is $\det(A)$.*

Lemma 3. *If A is upper or lower triangular, then its entries on the principal diagonal are all its eigenvalues.*

Lemma 4. *Similar matrices have the same characteristic polynomial and hence, the same set of eigenvalues.*

Lemma 5. *Let $f(\cdot)$ be a polynomial and λ be an eigenvalue of A. Then $f(\lambda)$ is an eigenvalue of $f(A)$. Hence, if A is idempotent then each of its eigenvalue is 0 or 1.*

Lemma 6. *Let $A_{m\times n}$ and $B_{n\times m}$ be a pair of matrices, where $m \leq n$. Then, $\mathrm{ch}_{BA}(\lambda) = \lambda^{n-m}\mathrm{ch}_{AB}(\lambda)$.*

It immediately follows from Lemma 6 that the non-zero eigenvalues of AB and BA are same and the algebraic multiplicity of 0 as an eigenvalue of BA is at least $n - m$. In particular, if $m = n$ then the characteristic polynomials of AB and BA are identical.

2.6 Eigenvalues of a symmetric matrix and the spectral decomposition

As we have observed in Example 1, the eigenvalues of a (real) matrix in general can be complex numbers. However, for symmetric matrices the following holds.

Theorem 2. *If A is symmetric then all its eigenvalues are real and the eigenvectors can be chosen to have real entries.*

Theorem 3. *For a symmetric matrix, the eigenvectors are mutually orthogonal.*

We emphasize that in Theorem 3 the orthogonality holds not only for eigenvectors corresponding to distinct eigenvalues but also for eigenvectors corresponding to the eigenvalues with algebraic multiplicities greater than 1. The following example illustrates this.

Example 3. Let

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}.$$

The characteristic equation is $(\lambda + 1)^2(\lambda - 5) = 0$. Hence, the eigenvalues are $\lambda_1 = 5$ with (algebraic) multiplicity 1 and $\lambda_2 = -1$ with multiplicity 2. It is easy to see that $\boldsymbol{x}_1 = (1, 1, 1)'$ is an eigenvector corresponding to λ_1. Also, it can be seen that $\boldsymbol{x}_{21} = (-2, 1, 1)'$ and $\boldsymbol{x}_{22} = (0, -1, 1)'$ are each an eigenvector corresponding to λ_2. The vectors $\boldsymbol{x}_1, \boldsymbol{x}_{21}$ and $\boldsymbol{x}_{22}$ are clearly, pairwise orthogonal.

Theorem 4. *Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of a symmetric matrix A, including multiplicities, and $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$ be the corresponding mutually orthogonal eigenvectors each chosen to be of unit length. Then*

$$A = \lambda_1\boldsymbol{\xi}_1\boldsymbol{\xi}_1' + \cdots + \lambda_n\boldsymbol{\xi}_n\boldsymbol{\xi}_n'. \tag{5}$$

Such a representation of a symmetric matrix A is called its *spectral representation*. Equivalently, the above can be written as

$$A = P\Delta P' \tag{6}$$

where the matrix P, with columns $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n$, is orthogonal, and $\Delta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

From equation (6), remembering that for an $n \times n$ orthogonal matrix P, $PP' = I_n = P'P$, it is easily seen that $P'AP = \Delta$ and thus, A can be diagonalized by an orthogonal transformation. In other words, A is orthogonally similar to a diagonal matrix, the diagonal entries being the eigenvalues of A.

2.7 Quadratic forms

Quadratic forms arise in the study of conics in geometry and energy in physics. They also have wide applications in several other areas, including statistics. A quadratic form in n variables $x_1, \ldots, x_n$, is an expression of the form

$$Q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{1 \le i < j \le n} \beta_{ij} x_i x_j, \tag{7}$$

where the α_i's and β_{ij}'s belong to $\mathbb{R}$. Writing $\boldsymbol{x} = (x_1, \ldots, x_n)'$, the right side of (7) can be written as $\boldsymbol{x}'A\boldsymbol{x}$, where the $n \times n$ symmetric matrix $A = (a_{ij})$ is given by

$$a_{ij} = \begin{cases} \alpha_i & \text{if } i = j \\ \beta_{ij}/2 & \text{if } i < j \\ \beta_{ji}/2 & \text{if } i > j. \end{cases}$$

Conversely, if $A = (a_{ij})$ is any symmetric matrix of order n, then $\boldsymbol{x}'A\boldsymbol{x}$ is the right side of (7) with $\alpha_i = a_{ii}$ and for $i < j$, $\beta_{ij} = 2a_{ij}$. Thus, any quadratic form can be uniquely written as $\boldsymbol{x}'A\boldsymbol{x}$, where A is a symmetric matrix. The symmetric matrix A associated with the quadratic form $\boldsymbol{x}'A\boldsymbol{x}$ is called the matrix of the quadratic form. Quadratic forms are usually classified as follows:

(i) nonnegative definite (n.n.d.) if $\boldsymbol{x}'A\boldsymbol{x} \ge 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$;

(ii) positive definite (p.d.) if $\boldsymbol{x}'A\boldsymbol{x} > 0$ for all $\boldsymbol{x} \ne \boldsymbol{0}_n$;

(iii) nonpositive definite (n.p.d.) if $\boldsymbol{x}'A\boldsymbol{x} \le 0$ for all $\boldsymbol{x} \in \mathbb{R}^n$;

(iv) negative definite (n.d.) if $\boldsymbol{x}'A\boldsymbol{x} < 0$ for all $\boldsymbol{x} \ne \boldsymbol{0}_n$.

A quadratic form is called indefinite if it is neither n.n.d nor n.p.d. It is customary to call the matrix of a quadratic form as n.n.d., p.d., etc., if the associated quadratic form is n.n.d., p.d., etc. The following are some important results that characterize a p.d. (or, n.n.d.) matrix.

Theorem 5. *A symmetric matrix A is p.d. (respectively, n.n.d.) if and only if each eigenvalue of A is positive (respectively, nonnegative).*

Theorem 6. *A (real) matrix A is n.n.d. if and only if there exists a real matrix B such that $A = B'B$. A is p.d. if and only if $A = B'B$ for some real non-singular matrix B.*

We now state a result on the bounds of $\boldsymbol{x}'A\boldsymbol{x}/\boldsymbol{x}'\boldsymbol{x}$, called the *Rayleigh quotient*.

Theorem 7. *Let $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ be the eigenvalues of a real symmetric matrix A. Then*

$$\lambda_1 \le \frac{\boldsymbol{x}'A\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}} \le \lambda_n. \tag{8}$$

Since the extrema of $\boldsymbol{x}'A\boldsymbol{x}/\boldsymbol{x}'\boldsymbol{x}$ given by (8) are attained by choosing $\boldsymbol{x}$ to be an eigenvector corresponding to λ_1 or λ_n, one can write

$$\begin{aligned} \lambda_1 &= \min_{\boldsymbol{x}} \frac{\boldsymbol{x}'A\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}, \\ \lambda_n &= \max_{\boldsymbol{x}} \frac{\boldsymbol{x}'A\boldsymbol{x}}{\boldsymbol{x}'\boldsymbol{x}}. \end{aligned}$$

For many other results on the extrema of quadratic forms in terms of the eigenvalues of the matrix of the form, one may refer to Magnus and Neudecker [23].

2.8 QR algorithm for computing eigenvalues and eigenvectors

The characteristic polynomial is of degree n and the complexity of calculating the eigenvalues increases with increasing n. For large n, one has to resort to appropriate numerical methods to find approximate roots. The roots of a polynomial equation are extremely sensitive functions of the coefficients and thus computational procedures can often be very inaccurate in the presence of round-off errors.

An efficient and accurate method to compute eigenvalues and eigenvectors is based on the QR algorithm discovered in 1961, independently by Francis [16, 17] and Kublanovskaya [22]. We briefly describe the essentials of this algorithm as applied to $n \times n$ non-singular matrices. We first state the following important and well-known result.

Theorem 8. *Any non-singular matrix A of order n can be factorized as $A = QR$, where Q is an orthogonal matrix and R is an upper triangular matrix. The factorization is unique if all the diagonal entries of R are assumed to be positive.*

Example 4. Let

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -2 \\ -1 & 2 & 3 \end{bmatrix}.$$

Then, with Q and R as given below, one can verify that $A = QR$:

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{4}{\sqrt{42}} & \frac{2}{\sqrt{14}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{42}} & -\frac{3}{\sqrt{14}} \\ -\frac{1}{\sqrt{3}} & \frac{5}{\sqrt{42}} & -\frac{1}{\sqrt{14}} \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{3} & -\frac{1}{\sqrt{3}} & -\sqrt{3} \\ 0 & \frac{\sqrt{14}}{\sqrt{3}} & \frac{\sqrt{21}}{\sqrt{2}} \\ 0 & 0 & \frac{\sqrt{7}}{\sqrt{2}} \end{bmatrix}.$$

The first step in the QR algorithm is to factorize A as

$$A = A_1 = Q_1 R_1$$

into a product of an orthogonal matrix Q_1 and an upper triangular matrix R_1 with all diagonal entries positive, by using a numerically stable *Gram-Schmidt orthogonalization* procedure. Next, multiply the two factors in the *reverse order* to obtain

$$A_2 = R_1 Q_1.$$

These two steps are then repeated. Thus, we factor $A_2 = Q_2 R_2$, again using the Gram-Schmidt process and then multiply the factors in the reverse order to get

$$A_3 = R_2 Q_2.$$

The complete algorithm can be written as

$$A = Q_1 R_1, \;\; R_k Q_k = A_{k+1} = Q_{k+1} R_{k+1}, \; k = 1, 2, 3, \ldots,$$

where Q_k, R_k come from the previous step, and the subsequent orthogonal and upper triangular matrices Q_{k+1} and R_{k+1}, respectively, are computed by using a numerically stable form of the Gram-Schmidt algorithm. The surprising fact is that for many matrices, the iterates A_k converge to an upper triangular matrix U whose diagonal entries are the eigenvalues of A. This means that after a sufficient number of iterations, say k', the matrix $A_{k'}$ will have very small entries below the diagonal and the entries on the diagonal are the approximate eigenvalues of A. For each eigenvalue, the computation of the corresponding eigenvectors can be accomplished by solving the appropriate linear homogeneous system of equations.

If the original matrix is (real) symmetric, then the limiting matrix $A_k \to \Lambda$, where Λ is a diagonal matrix containing the eigenvalues of A on its diagonal. Furthermore, if we recursively define

$$S_k = S_{k-1}Q_k = Q_1Q_2 \dots Q_{k-1}Q_k, \;\; S_1 = Q_1,$$

then the limit of S_k is S, an orthogonal matrix whose columns are the orthonormal eigenvector basis of A.

Software packages like `MAPLE` or `MATLAB` can be used to obtain the eigenvalues and corresponding eigenvectors.

2.9 Circulant matrices

There is a class of useful matrices, for which a formula for the eigenvalues is available. This is the class of 1-circulants and its variations.

2.9.1 The 1-Circulant matrix

This is the usual circulant matrix. An $n \times n$ 1-circulant matrix is given by

$$C_n = \begin{bmatrix} a_0 & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \dots & a_{n-3} & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \dots & a_{n-4} & a_{n-3} \\ & & & \vdots & & \\ a_1 & a_2 & a_3 & \dots & a_{n-1} & a_0 \end{bmatrix}.$$

Note that the entries in a row are obtained by shifting the elements of the previous row by one position. Its eigenvalues $\{\lambda_i\}$ are (see e.g., Davis[12] or Brockwell and Davis [10]),

$$\lambda_k = \sum_{l=0}^{n-1} a_l e^{i\omega_k l} = b_k + ic_k \;\; \forall\, k = 1, 2, \cdots, n,$$

where

$$\omega_k = \frac{2\pi k}{n}, \;\; b_k = \sum_{l=0}^{n-1} a_l \cos(\omega_k l), \;\; c_k = \sum_{l=0}^{n-1} a_l \sin(\omega_k l). \tag{9}$$

For $k = 1, 2, \dots, n$, the corresponding eigenvectors are given by $\boldsymbol{\xi}_k = (x_k^0, x_k, x_k^2, \dots, x_k^{n-1})'$, where $x_1, \dots, x_n$ are the n roots of unity.

2.9.2 The k-circulant matrix

These are circulant matrices where, instead of shifting the elements in subsequent rows by one position, they are shifted by k positions. Suppose $\boldsymbol{a} = \{a_l\}_{l \geq 0}$ is a sequence of real numbers. For positive integers k and n, a k-circulant matrix of order $n \times n$ with *input* sequence $\{a_l\}$ is defined as

$$A_{k,n}(\boldsymbol{a}) = \begin{bmatrix} a_0 & a_1 & \dots & a_{n-1} \\ a_{n-k} & a_{n-k+1} & \dots & a_{n-k-1} \\ a_{n-2k} & a_{n-2k+1} & \dots & a_{n-2k-1} \\ & & \vdots & \end{bmatrix}.$$

We write $A_{k,n}(\boldsymbol{a}) = A_{k,n}$. The subscripts appearing in the matrix entries above are calculated modulo n and the convention is to start the row and column indices from zero. Thus, the 0-th row of $A_{k,n}(\boldsymbol{a})$ is

$(a_0,\ a_1,\ a_2,\ \ldots,\ a_{n-1})$. For $0 \leq j < n-1$, the $(j+1)$-th row of $A_{k,n}$ is a right-circular shift of the j-th row by k positions (equivalently, k mod n positions). Without loss of generality, k may always be reduced modulo n.

Note that $A_{1,n}$ is the 1-circulant matrix C_n. It can be checked that the k-circulant matrix is symmetric for any $\{a_i\}$ if and only if $k = n-1$. This $(n-1)$-circulant matrix is also known as the *reverse circulant matrix* has identical elements on each of the *anti-diagonals* and is given by

$$RC_n = \begin{bmatrix} a_0 & a_1 & a_2 & \ldots & a_{n-2} & a_{n-1} \\ a_1 & a_2 & a_3 & \ldots & a_{n-1} & a_0 \\ a_2 & a_3 & a_4 & \ldots & a_0 & a_1 \\ & & & \vdots & & \\ a_{n-1} & a_0 & a_1 & \ldots & a_{n-3} & a_{n-2} \end{bmatrix}_{n \times n} .$$

2.9.3 Why study k-circulants?

The k-circulant matrix and its block versions arise in many different areas of mathematics and statistics– in multi-level supersaturated design of experiment (Georgiou and Koukouvinos [18]), spectra of de Bruijn graphs (Strok [27]) and in $(0,1)$-matrix solutions to $A^m = J_n$ (Wu, Jia and Li [33]). See also the book by Davis [12] and the article by Pollock [25].

A matrix is said to be a *Toeplitz* matrix (see Section 3.2 for a definition) if each diagonal has identical elements. Circulant matrices have deep connection with the Toeplitz matrix. As we have seen, the 1-circulant matrix has an explicit spectral decomposition. However, the spectral analysis of the latter is much harder and challenging in general. If the input sequence $\{a_l\}_{l \geq 0}$ is *square summable*, then the circulant approximates the corresponding Toeplitz in various senses when the dimension $n \to \infty$. See Gray [19] for a recent and relatively easy account. For an introduction to many interesting results, specially on the eigenvalue structure of Toeplitz matrices, see Grenander and Szegö [20].

The eigenvalues of the circulant matrices crop up crucially in time series analysis. For example, the *periodogram* of a sequence $\{a_l\}_{l \geq 0}$ is defined as $n^{-1} I_n(\omega_k)$ where

$$I_n(\omega_k) = \Big| \sum_{t=0}^{n-1} a_t e^{-it\omega_k} \Big|^2, \ \ k = 0, 1, \ldots, n-1, \tag{10}$$

and $\omega_k = 2\pi k/n$ are the Fourier frequencies. This is a simple function of the eigenvalues of the corresponding circulant matrix. The study of the properties of periodogram is fundamental in the spectral analysis of time series. See for instance Fan and Yao [15]. The maximum of the periodogram, in particular, has been studied in Davis and Mikosch [13].

The k-circulant matrices with random input sequence also serves as an important class of "patterned" matrices. Deriving the asymptotic properties of the spectrum of general patterned matrices has drawn significant attention in the recent literature. See Section 3.2 for some simulations on the spectrum of a few patterned random matrices, including the (random) 1-circulant.

2.9.4 Description of eigenvalues of k circulant

One can give a formula for the eigenvalues of the k-circulant for general k. This formula solution was derived by Zhou [34] and is given below in Theorem 9. A proof may also be found in Bose and Sen [9]. Recall the definition of Fourier frequencies $\{\omega_k\}$ defined in equation (10) and the eigenvalues $\lambda_t = \sum_{l=0}^{n-1} a_l \omega^{tl}, \ \ 0 \leq t < n$ of the 1-circulant $A_{1,n}$. Let $p_1 < p_2 < \cdots < p_c$ be all the common

prime factors of n and k. Then we may write,

$$n = n' \prod_{q=1}^{c} p_q^{\beta_q} \text{ and } k = k' \prod_{q=1}^{c} p_q^{\alpha_q}. \tag{11}$$

Here $\alpha_q,\ \beta_q \geq 1$ and n', k', p_q are pairwise relatively prime. For any positive integer m, let

$$\mathbb{Z}_m = \{0, 1, 2, \ldots, m-1\}.$$

We introduce the following family of sets

$$S(x) := \left\{xk^b \bmod n' : b \geq 0\right\}, \quad x \in \mathbb{Z}_{n'}. \tag{12}$$

Let $g_x = \#S(x)$. Note that $g_0 = 1$. It is easy to see that

$$S(x) = \{xk^b \mod n' : 0 \leq b < g_x\}.$$

Clearly $x \in S(x)$ for every x. Suppose $S(x) \cap S(y) \neq \emptyset$. Then, $xk^{b_1} = yk^{b_2} \bmod n'$, for some integers $b_1, b_2 \geq 1$. Multiplying both sides by $k^{g_x - b_1}$ we see that, $x \in S(y)$ so that, $S(x) \subseteq S(y)$. Hence, reversing the roles, $S(x) = S(y)$. Thus, the distinct sets in $\{S(x)\}_{x \in \mathbb{Z}_{n'}}$ form a partition, called the *eigenvalue partition*, of $\mathbb{Z}_{n'}$. Denote the partitioning sets and their sizes by

$$\mathcal{P}_0 = \{0\}, \mathcal{P}_1, \ldots, \mathcal{P}_{\ell-1} \text{ and } k_j = \#\mathcal{P}_j,\ 0 \leq j < \ell. \tag{13}$$

Define

$$\Pi_j := \prod_{t \in \mathcal{P}_j} \lambda_{tn/n'},\ \ j = 0, 1, \ldots, \ell - 1. \tag{14}$$

Theorem 9. *(Zhou [34]). The characteristic polynomial of $A_{k,n}$ is given by*

$$\mathrm{ch}_{A_{k,n}}(\lambda) = \lambda^{n-n'} \prod_{j=0}^{\ell-1} \left(\lambda^{k_j} - \Pi_j\right). \tag{15}$$

In particular, the above formula can be specialised to RC_n. The eigenvalues of RC_n may also be obtained directly (see Bose and Mitra [7]) and are given by

$$\begin{cases} \lambda_0 & = \sum_{t=0}^{n-1} a_t \\ \lambda_{n/2} & = \sum_{t=0}^{n-1} (-1)^t a_t, \text{ if } n \text{ is even} \\ \lambda_k = -\lambda_{n-k} & = \sqrt{I_n(\omega_k)},\ 1 \leq k \leq [\frac{n-1}{2}]. \end{cases}$$

where $\{\omega_k\}$ and $\{I_n(\omega_k)\}$ are as defined in equation (10).

3 Some selected applications

3.1 Principal Component Analysis

Scientific phenomena are often quite complex in nature and because of the complexities involved, investigators often collect observations on many different variables. Multivariate statistical analysis deals with a set of methods to elicit information from such data sets. In this section, we briefly discuss one aspect of

multivariate analysis, called *principal component analysis* because of its connection with the eigenvalues and eigenvectors of a matrix.

In multivariate situations, an investigator selects $p > 1$ variables or characters to record, the values of these variables being recorded on each individual, item or experimental unit. Then the multivariate data obtained from n individuals on each of the p variables can be represented by an $n \times p$ matrix, say $X = (x_{ij})$, where x_{ij} is the value of the variable j on the individual i, $1 \leq i \leq n, 1 \leq j \leq p$. The variability among the p variables is conveniently described by a $p \times p$ *variance-covariance* (or *dispersion*) matrix, say Σ, whose ith diagonal element is the variance of the ith variable, $\mathrm{Var}(X_i)$ and for $i \neq j$, the (i, j)th element represents the covariance between the ith and jth variables, written as $\mathrm{Cov}(X_i, X_j)$. Since $\mathrm{Cov}(U, V) = \mathrm{Cov}(V, U)$ for any $U,\ V$, the dispersion matrix Σ is real symmetric matrix. It is also n.n.d. We shall henceforth assume Σ to be positive definite.

Now consider a situation where p is large. Although p components are required to reproduce the total variability in the system, often much of the variability can be accounted for by a small number k of the *principal* components, which are linear combinations of the original variables. The k principal components can then replace the original p variables and, in such a case the original $n \times p$ data matrix can be reduced to an $n \times k$ data matrix. Thus, in this sense, the method of principal components can be visualized as a procedure to reduce the dimensionality of the problem.

Algebraically, principal components are specific linear combinations of the p original random variables $X_1, \ldots, X_p$. Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system (with $X_1, \ldots, X_p$ as the coordinate axes). The new axes represent the directions with maximum variability and provide a simpler and more transparent description of the variance-covariance structure.

Suppose the random vector $X = (X_1, X_2, \ldots, X_p)'$ has the dispersion matrix Σ and suppose $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ are the eigenvalues of Σ. Let $Y_i = \boldsymbol{a}_i' X,\ 1 \leq i \leq p$, be a set of linear combinations of the X_i's. Then, it can be easily seen that

$$\begin{aligned} \mathrm{Var}(Y_i) &= \boldsymbol{a}_i' \Sigma \boldsymbol{a}_i, 1 \leq i \leq p, \\ \mathrm{Cov}(Y_i, Y_j) &= \boldsymbol{a}_i' \Sigma \boldsymbol{a}_j, 1 \leq i < j \leq p. \end{aligned}$$

The first principal component is the linear combination with maximum variance, that is, it maximizes $\mathrm{Var}(Y_i) = \boldsymbol{a}_i' \Sigma \boldsymbol{a}_i$ over i. But, the variance $\boldsymbol{a}_i' \Sigma \boldsymbol{a}_i$ can be made arbitrarily large by multiplying $\boldsymbol{a}_i$ by some constant. In order to eliminate this indeterminacy, it is convenient to restrict attention to vectors of unit length. Thus, the first principal component is defined to be that linear combination $\boldsymbol{a}_1' X$ say, that maximizes $\boldsymbol{a}_i' \Sigma \boldsymbol{a}_i$ subject to $\boldsymbol{a}_1' \boldsymbol{a}_1 = 1$. The second principal component is that linear combination $\boldsymbol{a}_2' X$ that maximizes $\boldsymbol{a}_i' \Sigma \boldsymbol{a}_i$ subject to the conditions $\boldsymbol{a}_2' \boldsymbol{a}_2 = 1$ and $\mathrm{Cov}(\boldsymbol{a}_1' X, \boldsymbol{a}_2' X) = 0$. Proceeding in this way, the mth principal component is $\boldsymbol{a}_m' X$ that maximizes $\mathrm{Var}(\boldsymbol{a}_i' X)$ subject to $\boldsymbol{a}_m' \boldsymbol{a}_m = 1$ and $\mathrm{Cov}(\boldsymbol{a}_m' X, \boldsymbol{a}_k' X) = 0$ for all $k < m$.

We then have the following result.

Theorem 10. *Let Σ denote the dispersion matrix of a $p \times 1$ random vector $X = (X_1, \ldots, X_p)'$. Suppose $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$ are the eigenvalues of Σ and let $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_p$ be the corresponding normalized eigenvectors. Then, the ith principal component is given by $Y_i = \boldsymbol{\xi}_i' X,\ 1 \leq i \leq p$. If multiplicities exist among the eigenvalues, then the principal components Y_i are not unique.*

From Theorem 10, it is seen that the principal components are *uncorrelated* random variables and have variances equal to the eigenvalues of Σ. We next have the following result.

Theorem 11. *In the set up of Theorem 10, let $Y_i = \boldsymbol{\xi}_i' X$ be the ith principal component, $1 \leq i \leq p$. Then*

$$\sigma_{11} + \sigma_{22} + \cdots + \sigma_{pp} = \sum_{i=1}^{p} \mathrm{Var}(X_i) = \lambda_1 + \lambda_2 + \cdots + \lambda_p = \sum_{i=1}^{p} \mathrm{Var}(Y_i),$$

where for $1 \leq i \leq p$, σ_{ii} *is the* i*th diagonal element of* Σ.

Theorem 11 shows that the total variability, $\sum_{i=1}^{p} \sigma_{ii} = \sum_{i=1}^{p} \lambda_i$. In view of this, *the proportion of total variance explained by the* k*th principal component* is

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_p}, \ 1 \leq k \leq p.$$

The components of the vectors $\boldsymbol{\xi}_i$ appearing in the computation of principal components also have an interesting implication as given in the next result.

Theorem 12. *If for* $1 \leq i \leq p$, $Y_i = \boldsymbol{\xi}_i' X$ *are the principal components obtained from a dispersion matrix* Σ *then the correlation coefficient between* Y_i *and* X_k *is given by*

$$\frac{\xi_{ik}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}},$$

where $\boldsymbol{\xi}_i = (\xi_{i1}, \xi_{i2} \ldots, \xi_{ip})'$.

Example 5. Consider three random variables $X = (X_1, X_2, X_3)'$ whose dispersion matrix is given by

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

For this Σ, we have the eigenvalues and (normalized) eigenvectors as given below:

$$\begin{array}{rclcl} \lambda_1 & = & 5.83, & \boldsymbol{\xi}_1' & = & (0.383, -0.924, 0); \\ \lambda_2 & = & 2.00, & \boldsymbol{\xi}_2' & = & (0, 0, 1); \\ \lambda_3 & = & 0.17, & \boldsymbol{\xi}_3' & = & (0.924, 0.383, 0). \end{array}$$

The principal components are thus

$$\begin{array}{rcl} Y_1 & = & \boldsymbol{\xi}_1' X = 0.383X_1 - 0.924X_2, \\ Y_2 & = & \boldsymbol{\xi}_2' X = X_3, \\ Y_3 & = & \boldsymbol{\xi}_3' X = 0.924X_1 + 0.383X_2. \end{array}$$

The proportion of the total variability accounted for by the first principal component is $\lambda_1/(\lambda_1 + \lambda_2 + \lambda_3) = 5.83/8 = 0.73$. Furthermore, the first two principal components explain $(5.83 + 2)/8 = 0.98 = 98\%$ of the variability. Thus, in this case, the principal components Y_1, Y_2 could replace the original three variables with nearly no loss of information.

3.2 Random matrices

In random matrices the entries are random variables. Such matrices arise in nuclear physics, mathematics, signal processing, statistics and wireless communication. In *random matrix theory*, different properties of these matrices are studied. In particular, properties of eigenvalues of random matrices when the dimension increases, are attempted to be captured through the *empirical distribution*, the *spectral radius*, the *extreme eigenvalues* and the spacings between successive eigenvalues.

In this section we define a few important random matrices and show a few simulations for their empirical spectral distribution. The *empirical spectral distribution* (ESD) of the matrix $A_{n\times n}$ is defined as

$$F_n(x) = n^{-1} \sum_{i=1}^{n} I\{\lambda_i \leq x\}.$$

Thus it is a random probability distribution with mass $1/n$ at each λ_i.

Some of the more popular and/or important random matrices are the Wigner matrix, the sample variance covariance matrix, the sample autocovariance matrix, the Toeplitz and Hankel matrices, matrices of the form XX' used in signal processing and wireless communication models, the so called IID matrix which is probabilistically interesting and, the circulant and k circulant matrices. There are variants of these matrices such as the tridiagonal matrices, band matrices, balanced matrices etc. which are also objects of study for various reasons.

For more information on eigenvalue properties of random matrices and their applications, the reader may consult the books Anderson et al. [1], Bai and Silverstein [3], Mehta [24] and Tulino and Verdu [29] and the articles Bai [2], Bose, Gangopadhyay and Sen [6], Bose and Sen [9] and Bose, Hazra and Saha [4].

1-circulant matrix. For some detailed information on the eigenvalue distribution of k-circulant random matrices, see for example Bose, Hazra and Saha [5] and Bose, Mitra and Sen [8]. Here we provide a simulation result for the 1-circulant matrix.

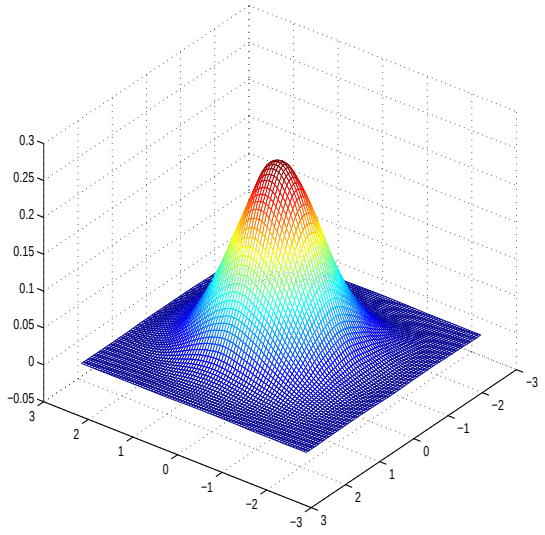

Figure 1: Distribution of eigenvalues for the scaled 1-circulant matrix of order 400 with Bernoulli entries.

Wigner matrix. This is a symmetric matrix where the entries on and above the diagonal are i.i.d. with zero mean and variance 1. We scale by dividing by $n^{1/2}$.

$$W_n = n^{-1/2} \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1(n-1)} & x_{1n} \\ x_{12} & x_{22} & x_{23} & \dots & x_{2(n-1)} & x_{2n} \\ & & & \vdots & & \\ x_{1n} & x_{2n} & x_{3n} & \dots & x_{(n-1)n} & x_{nn} \end{bmatrix}. \tag{16}$$

This matrix was introduced by Wigner [30, 31] in nuclear physics. His works eventually led to the entire subject of random matrices. See Anderson et al [1] for a thorough mathematical treatment of these matrices.

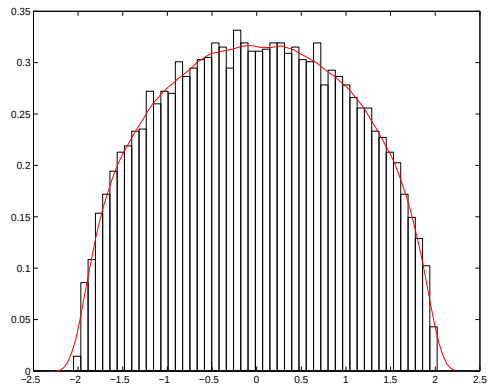

Figure 2: Histogram and kernel density estimate for the ESD for 15 scaled Wigner matrices of order 400 with iid Bernoulli entries.

IID matrix. This could be thought of the asymmetric version of the Wigner matrix. Here the entries are *independent and identically distributed* (i.i.d.) with mean zero and variance 1. For the latest results on the spectral distribution of these matrices, see Tao, Vu and Krishnapur [28].

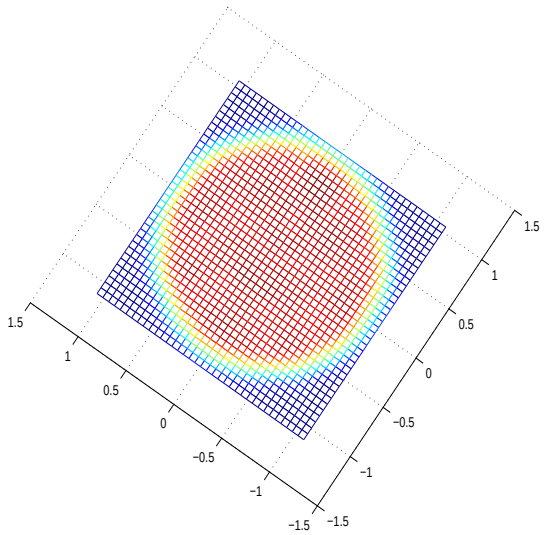

Figure 3: Distribution of eigenvalues for the scaled iid matrix of order 400 with Bernoulli entries.

Sample variance covariance matrix (S matrix). This was originally studied by Wishart [32]. It is an estimate of the population dispersion matrix and as seen earlier, is fundamental in statistical analysis of multivariate data. Suppose $\{x_{jk}, 1 \leq j \leq p, 1 \leq k \leq n\}$ is a double array of i.i.d. random variables with mean zero and variance 1. Write $\boldsymbol{x}_j = (x_{j1}, \ldots, x_{jn})$ and define the $p \times n$ matrix $X = [\boldsymbol{x}_1' \ \boldsymbol{x}_2' \ \cdots \ \boldsymbol{x}_p']'$. The *sample variance covariance matrix*, in short the S matrix, is

$$S_n = n^{-1} X X'.$$

If p is held fixed, then standard asymptotic theory applies. The situation becomes interesting if $p \to \infty$ such that p/n converges or, $p/n \to \infty$. For detailed results on this matrix, see Bai [2] and Bose, Gangopadhyay and Sen [6].

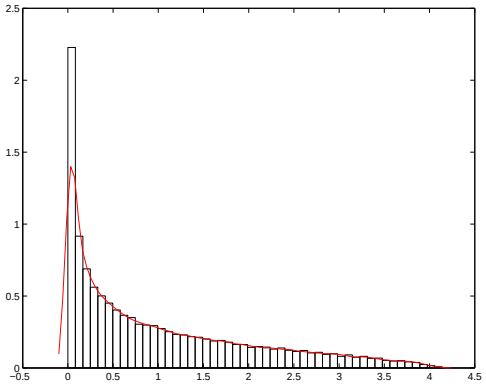

Figure 4: Histogram and kernel density estimate for the ESD for 15 scaled sample var-covar matrices with $n = p = 400$ with iid $exp(1)$ entries.

Toeplitz and Hankel matrix. Let $\{x_0, x_1, \ldots x_n, \ldots\}$ be a sequence of i.i.d. real random variables with mean zero and variance 1. The Toeplitz and Hankel matrices are defined respectively, as

$$T_n = \begin{bmatrix} x_0 & x_1 & x_2 & \ldots & x_{n-2} & x_{n-1} \\ x_1 & x_0 & x_1 & \ldots & x_{n-3} & x_{n-2} \\ x_2 & x_1 & x_0 & \ldots & x_{n-4} & x_{n-3} \\ & & & \vdots & & \\ x_{n-1} & x_{n-2} & x_{n-3} & \ldots & x_1 & x_0 \end{bmatrix},$$

$$H_n = \begin{bmatrix} x_0 & x_1 & x_2 & \ldots & x_{n-2} & x_{n-1} \\ x_1 & x_2 & x_3 & \ldots & x_{n-1} & x_n \\ x_2 & x_3 & x_4 & \ldots & x_n & x_{n+1} \\ & & & \vdots & & \\ x_{n-1} & x_n & x_{n+1} & \ldots & x_{2n-3} & x_{2n-2} \end{bmatrix}.$$

These matrices and the corresponding operators are important objects in operator theory. For probabilistic properties of these matrices, see Bryc, Dembo and Jiang [11], Hammond and Miller [21] and Bose, Gangopadhyay and Sen [6].

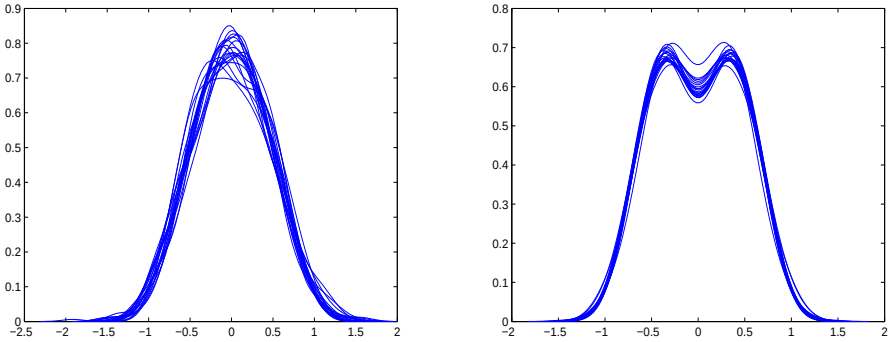

Figure 5: Individual kernel density estimate for the ESD for 15 simulations from the scaled Toeplitz (left) and Hankel (right) matrices of order 400 with Bernoulli entries.

Sample autocovariance matrix. Let $x_0, x_1, x_2, \ldots x_n, \ldots$ be a sequence of random variables. Let the autocovariances be defined as

$$\hat{\gamma}(k) = n^{-1} \sum_{i=1}^{n-|k|} x_i x_{i+|k|}.$$

The *sample autocovariance matrix* is

$$\hat{\Gamma}_n = ((\hat{\gamma}(i-j)))_{1 \le i,\, j \le n}$$

Note that this matrix has the Toeplitz structure. The autocovariances and the autocovariance matrix are fundamental in time series analysis.

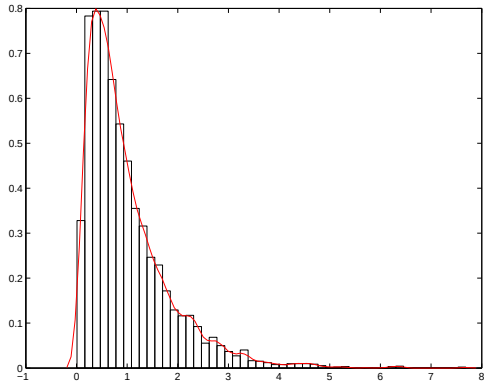

Figure 6: Histogram with kernel density estimate of the ESD of 15 realizations of the scaled sample autocovariance of order 400 with Bernoulli entries.

4 Eigenvalues and eigenfunctions in function spaces

The concepts of eigenvalue and eigenvector that we have discussed for matrices (i.e., finite dimensional linear transformations) have deep and significant extensions to *linear operators* in abstract spaces. See Dunford and Schwartz [14] for this development. Here we present an illustrative situation.

Consider the interval $C = [0,\ 1]$ and the space L_2 of all square integrable (w.r.t. the Lebesgue measure) functions on C. Let $K(x, y)$ be a square integrable (w.r.t. the two dimensional Lebesgue measure) function on $C \times C$. Given K, define T_K

$$T_K(g(x)) = \int_0^1 K(x, y) g(y) dy.$$

It can be shown that $T : L_2 \to L_2$ is a linear map. Extending the notion of eigenvalue and eigenvector, we may now say that λ is an eigenvalue of T with *eigenfunction* $f \in L_2$ if

$$T_K(f(x)) = \lambda f(x) \ \text{ for almost all } \ x \in [0,\ 1].$$

It can be shown that corresponding to any K, there is a (possibly infinite) sequence $\{\lambda_i, f_i(\cdot)\}$ which are all the eigenvalues and eigenvectors of T_K with analogous properties that we have described earlier for the finite dimensional case. In particular, the eigenfunctions are *orthogonal*, that is, $\int_0^1 f_k(x) f_{k'}(x) dx = 0$ whenever $k \neq k'$. In addition, they can also be chosen to be *orthonormal*, that is, they further satisfy $\int_0^1 f_k^2(x) dx = 1$ for all k. The *spectral representation* of the operator takes shape as

$$K(x, y) = \sum_{k=1}^{\infty} \lambda_k f_k(x) f_k(y) \ \text{ (in the } \ L_2 \ \text{ sense)}.$$

Example 6. As an example, let

$$K(x,y) = \int_0^1 [I(x \le t) - t][I(y \le t) - t]dt, \;\; x, y \in [0,\ 1]. \tag{17}$$

Since K is a bounded function it is square integrable. The eigenvalues and eigenfunctions of T_K are given by

$$\lambda_k = \frac{1}{k^2\pi^2}, \;\; f_k(x) = 2\sin(k\pi x), \; k = 1, 2, \ldots.$$

To see this, first note that $T_K(c) = 0$ where c is the constant function. This implies that if f is an eigenfunction, then so is $g(\cdot) = f(\cdot) - \int_0^1 f(x)dx$ (since $f \in L_2$, $\int_0^1 f(x)dx$ is well defined). Note that $\int_0^1 g(x)dx = 0$. Thus, without loss of generality, we can assume that the eigenfunctions f which we are seeking, satisfy $\int_0^1 f(x)dx = 0$. Now start with the eigenvalue equation to obtain

$$\begin{aligned}
T_K(f(x)) &= \int_0^1 \int_0^1 [I(x \le t) - t][I(y \le t) - t]f(y)dydt, \;\; x \in [0,\ 1] && (18)\\
&= \int_0^1 \int_0^1 [I(\max(x,y) \le t) - tI(x \le t) - tI(y \le t) + t^2]f(y)dydt, \;\; x \in [0,\ 1] && (19)\\
&= \int_0^1 \int_0^1 [I(\max(x,y) \le t) - tI(x \le t) - tI(y \le t)]f(y)dydt, \;\; x \in [0,\ 1] && (20)\\
&= \int_0^1 \int_x^1 I(y \le t)f(y)dydt - \int_0^1 \int_0^x I(x \le t)f(y)dydt - 0 - \int_0^1 \int_0^t tf(y)dydt && (21)\\
&= \int_x^1 (1-y)f(y)dy + (1-x)\int_0^x f(y)dy + C_f \;\text{ where }\; C_f \;\text{ is a constant} && (22)\\
&= \lambda f(x) && (23)
\end{aligned}$$

Now for the moment, assume that the eigenfunction we seek are smooth enough so that we can take derivatives. Thus taking derivatives in the above relation,

$$\begin{aligned}
\lambda f'(x) &= -(1-x)f(x) + (1-x)f(x) - \int_0^x f(y)dy && (24)\\
&= -\int_0^x f(y)dy && (25)
\end{aligned}$$

Taking another derivative, we see that

$$-f(x) = -\lambda f''(x). \tag{26}$$

It is easy to see that $f_k(x) = a_k \sin(\pi k x), \;\; k = 1, 2, \ldots$ where $\{a_k\}$ are constants, satisfy the above differential equation. Clearly these eigenfunctions are mutually orthogonal, that is $\int_0^1 f_k(x)f_k'(x)dx = 0$ whenver $k \neq k'$. They are further, orthonormal if we choose $a_k \equiv 2$ for all k. The eigenvalues are easy to obtain from equation (26) and come out as

$$\lambda_k = \frac{1}{k^2\pi^2}, \; k = 1, 2, \ldots. \tag{27}$$

It may be checked that this yields the spectral representation (that is there are no other eigenfunctions).

4.1 Cramér–von Mises statistic

Suppose $X_1, \ldots, X_n \ldots$ are independent uniform and identically distributed over the interval $(0,\ 1)$. The Cramer–von Mises statistic is defined as (we are considering only the "null" distribution):

$$V_n = n \int_0^1 [F_n(x) - x]^2 dx$$

where F_n is defined to be the *empirical* distribution function

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x).$$

The distribution of V_n is needed for statistical testing purposes but unfortunately cannot be found out in a closed form. After some simple algebra,

$$V_n = \frac{1}{n} \sum_{i=1}^{n} K(X_i, X_i) + \frac{2}{n} \sum_{1 \leq i < j \leq n} K(X_i, X_j)$$

where K is as defined in equation (17). As a consequence, the *asymptotic* distribution of V_n is closely connected to the linear transformation T_K and the eigenvalues described above. It can be shown that the limiting distribution of V_n is the same as $\sum_{k=1}^{\infty} \lambda_k Y_k$ where $\{\lambda_k\}$ is as in equation (27) and $\{Y_k\}$ are i.i.d. chi-squared variables each with one degree of freedom. The reader may consult Serfling [26] for further details.

Acknowledgement. We are thankful to Sujatha Ramdorai for inviting us to write this article. We also thank Subhajit Dutta, Debashis Pal and Koushik Saha for the help extended during the preparation of this manuscript.

References

[1] Anderson, G. W., Guionnet, A. and Zeitouni, O. (2009). *An Introduction to Random Matrices*. Cambridge University Press.

[2] Bai, Z. D. (1999). Methodologies in spectral analysis of large dimensional random matrices, a review (with discussion). *Statist. Sinica* **9**, 611–677.

[3] Bai, Z. D. and Silverstein, J. (2006). *Spectral Analysis of Large Dimensional Random Matrices.* Science Press, Beijing.

[4] Bose, A., Hazra, R. S. and Saha, K. (2010). Patterned random matrices and method of moments. *Proc. Internat. Congr. Mathematicians.* (to appear).

[5] Bose, A., Hazra, R. S. and Saha, K. (2010). Product of exponentials and spectral radius of random k circulants. Technical Report R2/2010, www.isical.ac.in/ ∼statmath. **Submitted** (21 pages).

[6] Bose, A., Gangopadhyay, S. and Sen, A. (2009). Limiting spectral distribution of XX' matrices. *Ann. Inst. Henri Poincare Probab. Stat.* (to appear). Available at http://imstat.org/aihp/accepted.html.

[7] Bose, A. and Mitra, J. (2002). Limiting spectral distribution of a special circulant. *Stat. Probab. Lett.* **60**, 111–120.

[8] Bose, A., Mitra, J. and Sen, A. (2008/2009). Large dimensional random k circulants. Technical Report R10/2008, www.isical.ac.in/∼statmath. **Submitted** (21 pages).

[9] Bose, A. and Sen, A. (2008). Another look at the moment method for large dimensional random matrices. *Elec. J. Probab.* **13**, 588–628.

[10] Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*, 2nd. ed. Springer-Verlag, New York.

[11] Bryc, W.; Dembo, A. and Jiang, T. (2006). Spectral measure of large random Hankel, Markov and Toeplitz matrices. *Ann. Probab.* **34**, 1–38.

[12] Davis, P. J. (1979). *Circulant Matrices*. Wiley, New York.

[13] Davis, R. A. and Mikosch, T. (1999). The maximum of the periodogram of a non-Gaussian sequence. *Ann. Prob.* **27**, 522–536.

[14] Dunford, N. and Schwartz, J. T. (1988). *Linear Operators, Part I, General Theory*. Wiley-Interscience.

[15] Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric and Parametric Methods.* Springer-Verlag, New York.

[16] Francis, J. G. F. (1961). The QR Transformation, I. *The Computer Journal* **4**, 265–271.

[17] Francis, J. G. F. (1962). The QR Transformation, II. *The Computer Journal* **4**, 332–345.

[18] Georgiou, S. and Koukouvinos, C. (2006). Multi-level k-circulant supersaturated designs. *Metrika* **64**, 209–220.

[19] Gray, Robert M. (2009). *Toeplitz and Circulant Matrices: A Review*. Now Publishers, Norwell, Massachusetts.

[20] Grenander, U. and Szegö, G. (1984). *Toeplitz Forms and Their Applications*, 2nd ed. Chelsea, New York.

[21] Hammond, C. and Miller, S. J. (2005). Distribution of eigenvalues for the ensemble of real symmetric Toeplitz matrices. *J. Theoret. Probab.* **18**, 537–566.

[22] Kublanovskaya, Vera N. (1963). On some algorithms for the solution of the complete eigenvalue problem. *USSR Computational Mathematics and Mathematical Physics* **1**, 637–657. (received Feb 1961).

[23] Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, New York.

[24] Mehta, M. L. (1967). *Random Matrices and The Statistical Theory of Energy Levels*. Academic Press, New York.

[25] Pollock, D. S. G. (2002). Circulant matrices and time-series analysis. *Internat. J. Math. Ed. Sci. Tech.* **33**, 213–230.

[26] Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley.

[27] Strok, V. V. (1992). Circulant matrices and the spectra of de Bruijn graphs. *Ukrainian Math. J.* **44**, 1446–1454.

[28] Tao, T., Vu, V. and Krishnapur M. (2010). Random Matrices: Universality of the ESDs and the circular law. *Ann. Probab.* (to appear).

[29] Tulino, A. M. and Verdu, S. (2004). *Random Matrix Theory and Wireless Communications*, Now Publishers, Norwell, Massachusetts.

[30] Wigner, E. P. (1955). Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.* (2) **62**, 548–564.

[31] Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Ann. Math.* (2) **67**, 325–327.

[32] Wishart, J. (1928). The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A**, 32–52.

[33] Wu, Y. K. and Jia, R. Z. and Li, Q. (2002). g-Circulant solutions to the $(0, 1)$ matrix equation $A^m = J_n$. *Lin. Alg. Appln.* **345**, 195–224.

[34] Zhou, J. T. (1996). A formula solution for the eigenvalues of g circulant matrices. *Math. Appl. (Wuhan)* **9**, 53–57.