

**LARGEST EIGENVALUES AND SAMPLE COVARIANCE
MATRICES. TRACY-WIDOM AND PAINLEVÉ II:
COMPUTATIONAL ASPECTS AND REALIZATION IN S-PLUS
WITH APPLICATIONS.**

ANDREI IU. BEJAN

ABSTRACT. Distributions of the largest eigenvalues of Wishart covariance matrices corresponding to the large size data matrices are studied in this paper by reviewing the most recent results on this topic. Special attention is paid to the corresponding asymptotic laws – Tracy-Widom distributions, and their realizations in *S-Plus*. A further look at the practical aspects of the results studied in the paper is taken by considering some examples and discussing appearing difficulties and open problems.

1. INTRODUCTION

1.1. **Preliminary notes.** This paper is devoted to the study of the recent developments in the theory of random matrices, – a topic of importance in mathematical physics, engineering, multivariate statistical analysis, – developments, related to distributions of largest eigenvalues of sample covariance matrices.

From the statistical point of view this particular choice in the class of random matrices may be explained by the importance of the Principal Component Analysis (PCA), where covariance matrices act as principal objects, and behaviour of their eigenvalues accounts for the final result of the technique's use.

Nowadays statistical work deals with data sets of huge sizes. In these circumstances it becomes more difficult to use classical, often exact, results from the fluctuation theory of the extreme eigenvalues. Thus, in some cases, asymptotic results stated in more simple forms are preferable.

The recent asymptotic results on the extreme eigenvalues of the real Wishart matrices are studied here. Johnstone (2001) has established that it is the Tracy-Widom law of order one that appears as a limiting distribution of the largest eigenvalue of a Wishart matrix with identity covariance in the case when the ratio of the data matrix' dimensions tends to a constant. The Tracy-Widom distribution which appears here is stated in terms of the solution to the Painlevé II differential equation

Date: November 20, 2005.

1991 *Mathematics Subject Classification.* Primary 62-02, 65C60; Secondary 65L05, 65L07, 68W25, 62Q05.

Key words and phrases. Largest eigenvalue, sample covariance matrix, Tracy-Widom laws, Painlevé II transcedents.

The support provided by OSI / Chevening / The University of Warwick Scholarship 2004/2005 is acknowledged.

– one of the six exceptional nonlinear second-order ordinary differential equations discovered by Painlevé a century ago.

The known exact distribution of the largest eigenvalue in the null case follows from a more general result of Constantine (1963) and is expressed in terms of hypergeometric function of a matrix argument. This hypergeometric function represents a zonal polynomial series. Until recently such representation was believed to be inapplicable for the efficient numeric evaluation. However, the most recent results of Koev and Edelman (2005) deliver new algorithms that allow to approximate efficiently the exact distribution. The present work is mostly devoted to the study of the asymptotic results.

In this section a general introduction into the topic is given. It also contains a literature review, historical aspects of the problem and some arguments in favour of its importance.

Section 2 represents a summarized, and necessarily selective account on the topic from the theoretical point of view: the exact and limiting distributions of the largest eigenvalue in the white Wishart case are particularly discussed. The limiting results are viewed in connection with similar results for the Gaussian Orthogonal, Gaussian Unitary, and Symplectic Ensembles. The content of the section also serves as a preparative tool for the subsequent exposition and developments.

Application aspects of the discussed results in statistics are deferred to Section 4. We discuss on applications in functional data analysis, PCA, sphericity tests.

Large computational work has been done for the asymptotic case and some novel programming routines for the common statistical use in the *S-Plus* package are presented. This work required a whole bag of tricks, mostly related to peculiarities of the forms in which the discussed distributions are usually presented. Section 3 contains the corresponding treatment to assure the efficient computations. Special effort was put to provide statistical tools for using in *S-Plus*. The codes are publicly downloadable from www.vitrum.md/andrew/MScWrwck/codes.txt. Refer further to this document as *Listing*.

The computational work allowed to tabulate the Tracy-Widom distributions (of order 1, 2, and 4). Particularly, a table containing standard tail p -values for the Tracy-Widom distribution is given. It is ready for the usual statistical look-up use.

The final section concludes the paper with a discussion on some open problems and suggestions on the further work seeming to be worth of doing.

Acknowledgement. Professor D. Firth and my colleague Charalambos Charalambous have kindly shared a data set containing lizards' spectral reflectance measurements. This contribution is appreciated. The data set serves as an example of a studied technique.

I am deeply grateful to the Department of Statistics, The University of Warwick for providing comfortable study and research facilities.

The support provided by OSI/Chevening/Warwick Scholarship 2004-2005 is also acknowledged.

1.2. Sample covariance matrices and their eigenvalues in statistical applications. Multivariate statistical analysis has been one of the most productive

areas of large theoretical and practical developments during the past eighty years. Nowadays its consumers are from very diverse fields of scientific knowledge. Many techniques of the multivariate analysis make use of sample covariance matrices, which are, in some extent, to give an idea about the statistical interdependence structure of the data.

Theoretically, covariance matrices are the objects which represent the true statistical interdependence structure of the underlying population units. At the same time, sample (or empirical) covariance matrices based on experimental measurements only give some picture of that interdependence structure. One of the very important questions is the following: in what extent this picture, drawn from a sample covariance matrix, gives a right representation about the juncture?

As mathematical objects covariance matrices originate many other scions which can be well considered as functions of a matrix argument: *trace, determinant, eigenvalues, eigenvectors, etc.* Not surprisingly, the study of these is of great importance in multivariate statistical analysis.

There is a special interest in behaviour of eigenvalues of the sample covariance matrices.

Thus, in the *Principal Component Analysis* (PCA), which is, perhaps, one of the most employed multivariate statistical techniques, reduction of the data dimensionality is heavily based on the behaviour of the largest eigenvalues of the sample covariance matrix. The reduction can be achieved by finding the (orthogonal) variance maximizing directions in the space containing the data. The formalism of the technique is introduced here briefly.

In Principal Component Analysis, known also as the *Karhunen-Loève transform* [e.g. see Johnstone (2001)], one looks for the successively orthogonal directions that maximally cover the variation in the data. Explicitly, in the PCA the following problem is set to be solved

$$(1.1) \quad \begin{aligned} \text{var}(\gamma_i' \mathbf{X}) &\rightarrow \max, \quad i = 1, \dots, p, \\ \gamma_k &\perp \gamma_{k-r}, \quad r = 1, \dots, (k-1) \quad \forall k = 2, \dots, p, \end{aligned}$$

where \mathbf{X} is the p -variate population vector and $\{\gamma_i\}_1^p$ are the orthonormal vectors to be found. The solution of this problem is nearly related to the spectral decomposition of the population covariance matrix Σ . Vectors $\{\gamma_i\}_1^p$ satisfy the equation

$$\Sigma = \begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_p \end{pmatrix} \Lambda \begin{pmatrix} \gamma_1 & \gamma_2 & \dots & \gamma_p \end{pmatrix}',$$

where $\Lambda \equiv \text{diag}(\lambda_i)$ is the diagonal matrix containing eigenvalues of Σ . Moreover, these are equal to the variances of the linear combinations from (1.1).

Thus, the orthogonal directions of the maximal variation in the data are given by the eigenvectors of a population covariance matrix and are called to be the *principal component directions*, and the variation in each direction is numerically characterized by the corresponding eigenvalue of Σ . In the context of the PCA assume that the diagonal entries of Λ are ordered, that is

$$\lambda_1 \geq \dots \lambda_p \geq 0.$$

If \mathbf{X} is a random vector with mean μ and covariance matrix Σ then the *principal component transformation* is the transformation

$$\mathbf{X} \mapsto \mathbf{Y} = (\gamma_1 \ \gamma_2 \ \dots \ \gamma_p)' (\mathbf{X} - \mu).$$

The i^{th} component of \mathbf{Y} is said to be the i^{th} *principal component* of \mathbf{X} .

Often it may not be desirable to use all original p variables – a smaller set of linear combinations of the p variables may capture most of the information. Naturally, the object of study becomes then a bunch of largest eigenvalues. Dependently on the *proportion of variation* explained by first few principal components, only a number of new variables can be retained. There are different measures of the proportions of variation and different rules of the decision: how many of the components to retain?

However, all such criterions are based on the behaviour of the **largest eigenvalues of the covariance matrix**. In practice, one works with estimators of the sample covariance matrix and its eigenvalues rather than with the true objects. Thus, the study of eigenvalue distribution, and particularly of the distributions of extreme eigenvalues, becomes a very important task with the practical consequences in the technique of the PCA. This is especially true in the age of high-dimensional data. See the thoughtful discussion on this in Donoho (2000).

Yet another motivating, although very general example can be given. Suppose we are interested in the question whether a given data matrix has a certain covariance structure. In line with the standard statistical methodology we could think as follows: for a certain adopted type of population distribution find the distribution of some statistics which is a function of the sample covariance matrix; then construct a test based on this distribution and use it whenever the data satisfies the conditions in which the test has been derived. There are tests following such a methodology, e.g. a test of *sphericity* in investigations of covariance structure of the data; see Kendall and Stuart (1968), Mauchly (1940) and Korin (1968) for references. Of course, this is a very general set up. However, in virtue of the recent results in the field of our interest, which are to be discussed where appropriate in further sections, one may hope that such a statistics for a covariance structure's testing may be based on the largest sample eigenvalue. Indeed, such tests have seen the development and are referred in literature to as the *largest root tests of $\Sigma = I$* , see Roy (1953).

1.3. Research question. As we saw, covariance matrices and their spectral behaviour play the crucial role in many techniques of the statistical analysis of high-dimensional data.

Suppose X is a data matrix, which we view as a realization of a random matrix $\mathcal{X} \sim N_{n \times p}(\mu, \Sigma)$. Here, by $N_{n \times p}(\mu, \Sigma)$ denote the distribution of $n \times p$ random matrices comprising n independent p -variate multivariate normal random rows with the mean μ and covariance matrix Σ .

Both for theoretical and practical needs it suffices to restrict ourselves by the zero mean case with the diagonal population covariance matrix Σ . This can be justified by the simple facts: the covariance structure is invariant under linear transformations and any linear transformation of a normal vector has the normal distribution.

Consider therefore a data matrix X as a realization of a random matrix $\mathcal{X} \sim N_{n \times p}(\mathbf{0}, \text{diag } \lambda_i)$. Let $S = (n-1)^{-1}X'X$ be the sample covariance matrix' estimator obtained from the fixed data matrix X . Let

$$\widehat{l}_1 \geq \widehat{l}_2 \geq \dots \geq \widehat{l}_p \geq 0$$

be the eigenvalues of S . This estimator S may be viewed as a realization of the sample covariance matrix $\mathcal{S} = (n-1)^{-1}\mathcal{X}'\mathcal{X}$ with the eigenvalues

$$l_1 \geq l_2 \geq \dots \geq l_p \geq 0$$

for which $\widehat{l}_1, \widehat{l}_2, \dots, \widehat{l}_p$ are also estimators.

In these settings the following problems are posed to be considered in this paper:

- **To study the distribution of the largest eigenvalues of the sample covariance matrix \mathcal{S} by reviewing the recent results on this topic as well as those preceding that led to these results. A special emphasis should be put on the study of the (marginal univariate) distribution of l_1 in the case of large data matrices' size.**
- **To take a further look at the practical aspects of these results by considering examples where the theory can be applied.**
- **To provide publicly available effective *S-Plus* routines implementing studied results and thus providing their use in statistical applications.**

1.4. Literature and historical review. Perhaps, the earliest papers devoted to the study of the behaviour of random eigenvalues and eigenvectors were those of Girshick (1939) and Hsu (1939). However, these followed the pioneering works in *random matrix theory* emerged in the late 1920's, largely due to Wishart (1928). Since then the random matrix theory has seen an impetuous development with applications in many fields of mathematics and natural science knowledge. One can mention (selectively) developments in:

- (1) **nuclear physics** – e.g. works of Wigner (1951, 1957), Gurevich *et al.* (1956), Dyson (1962), Liou *et al.* (1973), Brody *et al.* (1981), Bohigas *et al.* (1984a,b);
- (2) **multivariate statistical analysis** – e.g. works of Constantine (1963), James (1960, 1964), Muirhead (1982);
- (3) **combinatorics** – e.g. works of Schensted (1961), Fulton (1997), Tracy and Widom (1999), Aldous and Diaconis (1999), Deift (2000), Okounkov (2000);
- (4) **theory of algorithms** – e.g. work of Smale (1985);
- (5) **random graph theory** – e.g. works of Erdős and Rényi in late 1950's, Bollobás (1985), McKay (1981), Farkas *et al.* (2001);
- (6) **numerical analysis** – e.g. works of Andrew (1990), Demmel (1988), Edelman (1992);
- (7) **engineering: wireless communications and signal detection, information theory** – e.g. works of Silverstein and Combettes (1992), Telatar (1999), Tse (1999), Vismanath *et al.* (2001), Zheng and Tse (2002), Khan (2005);
- (8) **computational biology and genomics** – e.g. work of Eisen *et al.* (1998).

Spectrum properties (and thus distributions of eigenvalues) of sample covariance matrices are studied in this paper in the framework of *Wishart ensembles* (see § 2.1.4). Wishart (1928) proposed a matrix model which is known nowadays as the *Wishart real model*. He was also the first who computed the joint element distribution of this model.

Being introduced earlier than *Gaussian ensembles* (see §§ 2.1.2, 2.1.3), the Wishart models have seen the intensive study in the second half of the last century. The papers of Wigner (1955, 1958), Marčenko and Pastur (1967) were a great breakthrough in the study of the empirical distribution of eigenvalues and became classical. The limiting distributions which were found in these works became to be known as the *Wigner semicircle law* and the *Marčenko–Pastur distribution*, correspondingly.

The study of eigenvalue statistics took more clear forms after the works of James (1960), and Constantine (1963). The former defined and described zonal polynomials, and thus started the study of eigenvalue statistics in terms of special functions, the latter generalized the univariate hypergeometric functions in terms of zonal polynomial series and obtained a general result which permitted to derive the exact distribution of the largest eigenvalue [Muirhead (1982), p.420]. The work of Anderson (1963) is also among standard references on the topic.

Surveys of Pillai (1976, 1977) contain discussions on difficulties related to obtaining the marginal distributions of eigenvalues in Wishart ensembles. To extract the marginal distribution of an eigenvalue from the joint density was a cumbersome problem since it involved complicated integrations. On the other hand, the asymptotic behaviour of data matrices' eigenvalues as principal component variances under the condition of normality (and in the sense of distributional convergence) when p is fixed were known [see Anderson (2003), Flury (1988)].

However, modern applications require efficient methods for the large size matrices' processing. As Debashis (2004) notices, in fixed dimension scenario much of the study of the eigenstructure of a sample covariance matrix utilizes the fact that it is a good approximation of the population covariance matrix when sample size is large (comparatively to the number of variates). This is no longer the case when $\frac{n}{p} \rightarrow \gamma \in (0, \infty)$.

It was known that when the true covariance matrix is an identity matrix, the largest and the smallest eigenvalues in the corresponding Wishart matrix converges *almost surely* to the respective boundaries of the *support* of the *Marčenko–Pastur distribution*. However, no results regarding the variability information for this convergence were known until the work of Johnstone (2001), in which the asymptotic distribution for largest sample eigenvalue has been derived. Johnstone showed that the asymptotic distribution of the properly rescaled largest eigenvalue of the *white Wishart* population covariance matrix¹ when $\frac{n}{p} \rightarrow \gamma \in (0, \infty)$ is the *Tracy–Widom distribution* F_β , where $\beta = 1$. This was not a surprise for specialists in random matrix theory – the *Tracy–Widom law* F_β appeared as the limiting distribution of the first, second, etc. rescaled eigenvalues for *Gaussian Orthogonal* ($\beta = 1$, real case), *Gaussian Unitary* ($\beta = 2$, complex case) and *Gaussian Symplectic Ensemble* ($\beta = 4$,

¹*White Wishart distribution* corresponds to the case of n i.i.d. standard Gaussian p -variate vectors (see Definition 1).

quaternion case), correspondingly. The distribution was discovered by Tracy and Widom (1994, 1996). Dumitriu (2003) mentions that general β -ensembles were considered and studied as theoretical distributions with applications in lattice gas theory and statistical mechanics: see works of Baker and Forrester (1997), Johansson (1997, 1998). General β -ensembles were studied by a tridiagonal matrix constructions in the papers of Dumitriu and Edelman (2002) and Dumitriu (2003).

It should be noted here that the work of Johnstone followed the work of Johansson (2000) in which a limit theorem for the largest eigenvalue of a *complex Wishart matrix* has been proved. The limiting distribution was found to be the *Tracy-Widom distribution* F_β , where $\beta = 2$. However, particularly because of the principal difference in the construction of the real and complex models, these two different models require independent approach in investigation.

Soshnikov (2002) extended the results of Johnstone and Johansson by showing that the same limiting laws remain true for the covariance matrices of *subgaussian* real (complex) populations in the following mode of convergence: $n - p = O(n^{1/3})$.

Bilodeau (2002) studies the asymptotic distribution ($n \rightarrow \infty$, p is fixed) of the largest eigenvalue of a sample covariance matrix when the distribution for the population is *elliptical* with a finite kurtosis. The main feature of the results of Bilodeau's paper is that they are true regardless the multiplicity of the population's largest eigenvalue. Note that the class of elliptical distributions contains the multivariate normal law as a subclass. Definition of the elliptical distribution see in Muirhead (1982, p.34).

The limiting distributions of eigenvalues of sample correlation matrices in the setting of "large sample data matrices" are studied in the work of Jiang (2004). Although, the empirical distribution of eigenvalues is only considered in the last mentioned paper. Under two conditions on the ratio n/p , it is shown that the limiting laws are the Marčenko-Pastur distribution and the semicircular law, respectively.

Diaconis (2003) provides a discussion on fascinating connections between study of the distribution of the eigenvalues of large unitary and Wishart matrices and many problems and topics of statistics, physics and number theory: PCA, telephone encryption, the zeros of Riemann's zeta function; a variety of physical problems; selective topics from the theory of Toeplitz operators.

Finally, refer to a paper of Bai (1999) as a review of the methodologies and mathematical tools existing in spectral analysis of large dimensional random matrices.

2. THE LARGEST EIGENVALUES IN SAMPLE COVARIANCE MATRICES

In this section review of the results regarding the distribution of the largest eigenvalues (in bulk and individually) in sample covariance matrices is given. The Wishart ensemble is positioned as a family which will mostly attract our attention. However, an analogy with other Gaussian ensembles such as *Gaussian Orthogonal Ensemble (GOE) - O_n* , *Gaussian Unitary Ensemble (GUE) - U_n* and *Gaussian Symplectic Ensemble (GSE) - S_n* is discussed. These lead to the *Tracy-Widom laws* with parameter's values $\beta = 1, 2, 4$ as limiting distributions of the largest eigenvalue in GOE, GUE and GSE, correspondingly. The case where $\beta = 1$ corresponds also to the Wishart ensemble.

2.1. Random matrix models and ensembles. Depending on the type of studied random matrices, different families called *ensembles* can be distinguished.

2.1.1. *General methodology: basic notions.* In the spirit of the classical approach of the probability theory a *random matrix model* can be defined as a probability triple object $(\Omega, \mathcal{P}, \mathcal{F})$ where Ω is a set of matrices of interest, \mathcal{P} is a measure defined on this set, and \mathcal{F} is a σ -algebra on Ω , i.e. a family of measurable subsets of Ω , $\mathcal{F} \subseteq 2^\Omega$.

An ensemble of random matrices of interest usually forms a group (in algebraic sense). This permits to introduce a measure with the help of which the ensemble's "typical elements" are studied. The *Haar measure* is common in this context when the group is compact or locally compact, and then it is a probability measure \mathcal{P}_H on a group Ω which is translation invariant: for any measurable set $B \in \mathcal{F}$ and any element $\mathcal{M} \in \Omega$

$$\mathcal{P}_H(\mathcal{M}B) = \mathcal{P}_H(B).$$

Here, by $\mathcal{M}B$ the following set is meant

$$\mathcal{M}B \stackrel{\text{def}}{=} \{\mathcal{M}N \mid N \in B\}.$$

See Conway (1990) and Eaton (1983) for more information on Haar measure and Diaconis and Shahshahani (1994) for a survey of constructions of Haar measures.

2.1.2. *Wigner matrices.* Wigner matrices were introduced in mathematical physics by Eugene Wigner (1955) to study the statistics of the excited energy levels of heavy nuclei.

A *complex Wigner random matrix* is defined as a square $n \times n$ Hermitian matrix $\mathcal{A} = (A_{lm})$ with i.i.d. entries above the diagonal:

$$A_{lm} = \overline{A_{ml}}, \quad 1 \leq l \leq m \leq n, \quad \{A_{lm}\}_{l < m} \text{ - i.i.d. complex random variables,}$$

and the diagonal elements $\{A_{ll}\}_{l=1}^n$ are i.i.d. real random variables.

A Hermitian matrix whose elements are real is just a symmetric matrix, and, by analogy, a *real Wigner random matrix* is defined as a square $n \times n$ symmetric matrix $\mathcal{A} = (A_{lm})$ with i.i.d. entries above the diagonal:

$$A_{lm} = A_{ml}, \quad 1 \leq l \leq m \leq n, \quad \{A_{lm}\}_{l < m} \text{ - i.i.d. real random variables.}$$

By specifying the type of the elements' randomness, one can extract different sub-ensembles from the ensemble of Wigner matrices.

2.1.3. *Gaussian Unitary, Orthogonal and Symplectic Ensembles.* Assumption of normality in Wigner matrices leads to a couple of important ensembles: *Gaussian Unitary* (GUE) and *Gaussian Orthogonal* (GOE) ensembles.

The GUE is defined as the ensemble of complex $n \times n$ Wigner matrices with the Gaussian entries

$$\Re A_{lm}, \Im A_{lm} \sim \text{i.i.d. } N\left(0, \frac{1}{2}\right), \quad 1 \leq l < m \leq n,$$

$$A_{kk} \sim N(0, 1), \quad 1 \leq k \leq n.$$

The GOE is defined as the ensemble of real $n \times n$ Wigner matrices with the Gaussian entries

$$A_{lm} \sim \text{i.i.d. } N\left(0, \frac{1 + \delta_{lm}}{2}\right), \quad 1 \leq l \leq m \leq n,$$

where, as before, δ_{lm} is the Kronecker symbol.

For completeness we introduce a *Gaussian Symplectic Ensemble* (GSE).

The GSE is defined as the ensemble of Wigner-like random matrices with the quaternion Gaussian entries

$$\begin{aligned} A_{lm} &= X_{lm} + iY_{lm} + jZ_{lm} + kW_{lm}, \\ \text{where } X_{lm}, Y_{lm}, Z_{lm}, W_{lm} &\sim \text{i.i.d. } N\left(0, \frac{1}{2}\right), \quad 1 \leq l \leq m \leq n, \\ A_{kk} &\sim N(0, 1), \quad 1 \leq k \leq n. \end{aligned}$$

Note that all eigenvalues of a hermitian (symmetric) matrix are real.

For references on the relationship of GOE, GUE and GSE to quantum physics and their relevance to this field of science see Forrester (2005).

2.1.4. *Wishart ensemble.* As we saw, many of the random matrix models have (standard) normal entries. For an obvious reason, in statistics the symmetrization of such a matrix, say \mathcal{X} , by multiplying by the transpose \mathcal{X}' leads to the ensemble of positive definite matrices $\mathcal{X}'\mathcal{X}$, which coincides up to a constant with a family of sample covariance matrices. This ensemble was named after Wishart who first computed the joint element density of $\mathcal{X}'\mathcal{X}$.

Let $\mathcal{X} = (X_{ij})$ be a $n \times p$ real rectangular random matrix with independent identically distributed entries. The case where

$$X_{ij} \sim N(0, 1), \quad 1 \leq i \leq n, 1 \leq j \leq p,$$

leads to a family $\{\mathcal{X}'\mathcal{X}\}$ of positive definite matrices, which is known in the literature as the *Wishart ensemble of real sample covariance matrices*.

The following general definition is accepted.

Definition 1. A $p \times p$ random matrix \mathcal{M} is said to have a **Wishart distribution with scale matrix Σ and n degrees of freedom** if $\mathcal{M} = \mathcal{X}'\mathcal{X}$ where $\mathcal{X} \sim N_{n \times p}(\mu, \Sigma)$.

This is denoted by

$$\mathcal{M} \sim W_p(n, \Sigma).$$

The case when $\Sigma \equiv I$ is referred to as the "null" case, and then the distribution of \mathcal{M} is called a *white Wishart distribution*.

Analogously, the Wishart ensemble of complex sample covariance matrices can be defined. The spectral properties of such matrices are of a long-standing interest in nuclear physics. We restrict ourselves to the real case only.

The rest of this paragraph aims to demonstrate the difficulties arising with (i) representation of the joint density function and (ii) extraction of the marginal density of the Wishart matrices' largest eigenvalue.

To obtain an explicit form for the joint density of the eigenvalues is not, generally, an easy task. The following theorem can be found in Muirhead (1982, Th.3.2.17).

Theorem 1. *If \mathcal{A} is a $p \times p$ positive definite random matrix with density function $f(\mathcal{A})$ then the joint density function of the eigenvalues $l_1 > l_2 > \dots > l_p > 0$ of \mathcal{A} is*

$$\frac{\pi^{p^2/2}}{\Gamma_p(p/2)} \prod_{1 \leq i < j \leq p} (l_i - l_j) \int_{\mathcal{O}_p} f(HLH')(dH).$$

Here dH stands for the Haar invariant probability measure on the orthogonal group \mathcal{O}_p , normalized so that

$$\int_{\mathcal{O}_p} (dH) = 1.$$

For an explicit expression of dH as a differential form expressed in terms of exterior products see Muirhead (1982, p.72).

This general theorem applies to the Wishart case in the following way.

Theorem 2. *If $\mathcal{A} \sim W_p(n, \Sigma)$ with $n > p - 1$, the joint density function of the eigenvalues $l_1 > l_2 > \dots > l_p > 0$ of \mathcal{A} is*

$$\frac{\pi^{p^2/2} 2^{-np/2} (\det \Sigma)^{-n/2}}{\Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2})} \prod_{i=1}^p l_i^{(n-p-1)/2} \prod_{j=i+1}^p (l_i - l_j) \int_{\mathcal{O}_p} \text{etr}(-\frac{1}{2} \Sigma^{-1} HLH')(dH).$$

The latter theorem can be proven by substituting $f(\mathcal{A})$ in the Theorem 1 by the Wishart density function

$$(2.1) \quad \frac{2^{-np/2}}{\Gamma_p(\frac{n}{2}) (\det \Sigma)^{n/2}} \text{etr}(-\frac{1}{2} \Sigma^{-1} \mathcal{A}) (\det \mathcal{A})^{(n-p-1)/2},$$

and noticing that $\det \mathcal{A} = \prod_{i=1}^p l_i$. Here etr stands for the exponential of the trace of a matrix. The function Γ_p is a *multivariate gamma function*:

$$\Gamma_p(z) = \pi^{p(p-1)/4} \prod_{k=1}^p \Gamma(z - \frac{1}{2}(k-1)), \quad \Re z > \frac{1}{2}(p-1).$$

Generally, it is not easy to simplify the form of the integral appearing in the expression for the joint density of the eigenvalues or to evaluate that integral.

In the null case when $\Sigma = \lambda I$, using that

$$\begin{aligned} \int_{\mathcal{O}_p} \text{etr}(-\frac{1}{2} \Sigma^{-1} HLH')(dH) &= \int_{\mathcal{O}_p} \text{etr}(-\frac{1}{2\lambda} HLH')(dH) \\ &= \text{etr}\left(-\frac{1}{2\lambda} L\right) \int_{\mathcal{O}_p} (dH) \\ &= \exp\left(-\frac{1}{2\lambda} \sum_{i=1}^p l_i\right), \end{aligned}$$

one gets that the joint density distribution of the eigenvalues of the null Wishart matrix $\mathcal{A} \sim W_p(n, \lambda I_p)$ is

$$\frac{\pi^{p^2/2}}{(2\lambda)^{np/2} \Gamma_p(\frac{n}{2}) \Gamma_p(\frac{p}{2})} \exp\left(-\frac{1}{2\lambda} \sum_{i=1}^p l_i\right) \prod_{i=1}^p l_i^{(n-p-1)/2} \prod_{j=i+1}^p (l_i - l_j).$$

The case when $\Sigma = \text{diag}(\lambda_i)$ does not lead longer to such an easy evaluation.

2.2. Distribution of the largest eigenvalue: sample covariance matrices. If $\mathcal{X} \sim N_{n \times p}(\mu, \Sigma)$, then the sample covariance matrix $\mathcal{S} = \frac{1}{n} \mathcal{X}' H \mathcal{X}$ has the Wishart distribution $W_p(n-1, \frac{1}{n} \Sigma)$. Hence, any general result regarding the eigenvalue(s) of matrices in $W_p(n, \Sigma)$ can be easily applied to the sample covariance matrices' eigenvalue(s). Notice that the eigenvalues of $\mathcal{A} = n\mathcal{S} \sim W_p(n-1, \Sigma)$ are n times greater than those of \mathcal{S} .

Start with an example, concerning, again, the joint distribution of the eigenvalues. As a consequence of the observations made above, the Theorem 2 can be formulated in the following special form for sample covariance matrices.

Proposition 1. *The joint density function of the eigenvalues $l_1 > l_2 > \dots > l_p > 0$ of a sample covariance matrix $\mathcal{S} \sim W_p(n, \Sigma)$ ($n > p$) is of the following form*

$$\frac{\pi^{p^2/2} (\det \Sigma)^{-\tilde{n}/2} \tilde{n}^{-\tilde{n}p/2}}{\Gamma_p(\frac{\tilde{n}}{2}) \Gamma_p(\frac{p}{2})} \frac{1}{2} \prod_{i=1}^p l_i^{(\tilde{n}-p-1)/2} \prod_{j=i+1}^p (l_i - l_j) \int_{\mathcal{O}_p} \text{etr}\left(-\frac{1}{2} n \Sigma^{-1} H L H'\right) (dH),$$

where $\tilde{n} = n - 1$.

We have simply made a correction for n and substituted all l_i by $\tilde{n}l_i$.

The participating integral can be expressed in terms of the *multivariate hypergeometric function* due to James (1960). The fact is that

$$\begin{aligned} \text{etr}\left(-\frac{1}{2} \tilde{n} \Sigma^{-1} H L H'\right) &= {}_0F_0\left(-\frac{1}{2} \tilde{n} \Sigma^{-1} H L H'\right) \\ &\quad \text{averaging over the group } \mathcal{O}_p, \\ &\quad \text{see Muirhead (1982, Th. 7.2.12, p.256)} \\ &= {}_0F_0\left(-\frac{1}{2} \tilde{n} L, \Sigma^{-1}\right), \end{aligned}$$

where ${}_0F_0(\cdot)$ (${}_0F_0(\cdot, \cdot)$) is the (*two-matrix*) *multivariate hypergeometric function* – the function of a matrix argument(s) expressible in the form of zonal polynomial series.

Thus, the joint density function of the eigenvalues of \mathcal{S} is

$$\frac{\pi^{p^2/2} (\det \Sigma)^{-\tilde{n}/2} \tilde{n}^{-\tilde{n}p/2}}{\Gamma_p(\frac{\tilde{n}}{2}) \Gamma_p(\frac{p}{2})} \frac{1}{2} \prod_{i=1}^p l_i^{(\tilde{n}-p-1)/2} \prod_{j=i+1}^p (l_i - l_j) {}_0F_0\left(-\frac{1}{2} \tilde{n} L, \Sigma^{-1}\right).$$

The distribution of the largest eigenvalue can also be expressed in terms of the hypergeometric function of a matrix argument. The following theorem is formulated in Muirhead (1982) as a corollary of a more general result regarding the positive definite matrices.

Theorem 3. *If l_1 is the largest eigenvalue of \mathcal{S} , then the cumulative distribution function of l_1 can be expressed in the form*

$$(2.2) \quad \mathbb{P}(l_1 < x) = \frac{\Gamma_p(\frac{p+1}{2})}{\Gamma_p(\frac{n+p}{2})} \det(\frac{\tilde{n}}{2}\Sigma^{-1})^{\tilde{n}/2} {}_1F_1(\frac{\tilde{n}}{2}; \frac{n+p}{2}; -\frac{n}{2}x\Sigma^{-1}),$$

where, as before, $\tilde{n} = n - 1$.

The hypergeometric function ${}_1F_1$ in (2.2) represents the alternating series, which converges very slowly, even for small \tilde{n} and p . Sugiyama (1972) gives explicit evaluations for $p = 2, 3$. There is a stable interest in the methods of efficient evaluation of the hypergeometric function.

2.3. Convergence of eigenvalues. Being interested mostly in the behaviour of the largest eigenvalue of sample covariance matrix, we formulate the Johnstone's asymptotic result in the large sample case. The empirical distribution function of eigenvalues is also studied. A parallel between the Marčenko-Pastur distribution and the Wigner "semicircle"-type law is drawn.

2.3.1. Introduction. There are different modes of eigenvalues' convergence in sample covariance matrices. These depend not only on the convergence' type (weak convergence, almost surely convergence, convergence in probability), but also of the conditions imposed on the dimensions of the data matrix.

One of the classical results regarding the asymptotic distribution of eigenvalues of a covariance matrix in the case of a multinormal population is given by the following theorem [see Anderson (2003)].

Theorem 4. *Suppose \mathcal{S} is a $p \times p$ sample covariance matrix corresponding to a data matrix drawn from $N_{n \times p}(\mu, \Sigma)$ Asymptotically, the eigenvalues l_1, \dots, l_p of \mathcal{S} are distributed as follows:*

$$\sqrt{n}(l_i - \lambda_i) \xrightarrow{\text{dist}} N(0, 2\lambda_i^2), \text{ for } i = 1, \dots, p,$$

where $\{\lambda_i\}$ are the (distinct) eigenvalues of the population covariance matrix Σ .

Notice that here p is fixed. The convergence is meant in the weak sense, i.e., the pointwise convergence of c.d.f.'s of r.v.'s $\sqrt{n}(l_i - \lambda_i)$ to $\frac{1}{2\sqrt{\pi}\lambda_i} e^{-\frac{x^2}{4\lambda_i}}$ takes place.

Next, introduce the following notation. Let $\chi(C)$ be the event indicator function, i.e.

$$\chi(C) = \begin{cases} 1 & \text{if } C \text{ is true,} \\ 0 & \text{if } C \text{ is not true} \end{cases} .$$

Definition 2. *Let A be a $p \times p$ matrix with eigenvalues l_1, \dots, l_p . The **empirical distribution function** for the eigenvalues of A is the distribution*

$$l_p(x) \stackrel{\text{def}}{=} \frac{1}{p} \sum_{i=1}^p \chi(l_i \leq x).$$

The p.d.f. of the empirical distribution function can be represented in the terms of the Dirac delta function $\delta(x)$ (which is the derivative of the Heaviside step

function)

$$l'_p(x) = \frac{1}{p} \sum_{i=1}^p \delta(x - l_i).$$

Dependently on the type of random matrix, the empirical distribution function can converge to a certain non-random law: *semicircle law*, *full circle law*, *quarter circle law*, *deformed quarter law* and others [see Müller (2000)].

For example, the celebrated *Wigner semicircle law* can be formulated in the following way. Let $\mathcal{A}_p = (A_{ij}^{(p)})$ be a $p \times p$ Wigner matrix (either complex or real) satisfying the following conditions: (i) the laws of r.v.'s $\{A_{ij}^{(p)}\}_{1 \leq i \leq j \leq p}$ are symmetric, (ii) $\mathbb{E}[(A_{ij}^{(p)})^{2k}] \leq (\text{const} \cdot k)^k$, $k \in \mathbb{N}$, (iii) $\mathbb{E}[(\sqrt{p}A_{ij}^{(p)})^2] = \frac{1}{4}$, $1 \leq i < j \leq p$, $\mathbb{E}[\sqrt{p}A_{ii}^{(p)}] \leq \text{const}$. Then, the empirical eigenvalue density function $l'_p(x)$ converges in probability to

$$(2.3) \quad p(x) = \begin{cases} \frac{2}{\pi} \sqrt{1 - x^2} & |x| \leq 1 \\ 0 & |x| > 1 \end{cases}.$$

Convergence in probability means here that if $\{L_p\}_{p=1}^\infty$ are r.v.'s having the p.d.f.'s $l'_p(x)$, corresp., then

$$\forall \epsilon > 0 \quad \lim_{p \rightarrow \infty} \mathbb{P}(|L_p - L| \geq \epsilon) = 0,$$

where L is a r.v. with p.d.f. given by (2.3).

There is a kind of analog to the Wigner semicircle law for sample covariance in the case of Gaussian samples – Marčenko-Pastur distribution.

2.3.2. Marčenko-Pastur distribution: the "semicircle"-type law. The Marčenko-Pastur (1967) theorem states that the empirical distribution function $l_p(x)$ of the eigenvalues $l_1^{(p)} \geq l_2^{(p)} \geq \dots \geq l_p^{(p)}$ of a $p \times p$ sample covariance matrix \mathcal{S} of n normalized i.i.d. Gaussian samples satisfies the following statement:

$$l_p(x) \rightarrow G(x), \text{ as } n = n(p) \rightarrow \infty, \text{ s.t. } \frac{n}{p} \rightarrow \gamma,$$

almost surely where

$$(2.4) \quad G'(x) = \frac{\gamma}{2\pi x} \sqrt{(b-x)(x-a)}, \quad a \leq x \leq b,$$

and $a = (1 - \gamma^{-1/2})^2$ when $\gamma \geq 1$. When $\gamma < 1$, there is an additional mass point at $x = 0$ of weight $(1 - \gamma)$.

Figure 1 shows how the density curves of the Marčenko-Pastur distribution vary dependently on the value of the parameter γ .

For the purpose of illustration of the Marčenko-Pastur theorem, place two plots in the same figure: the plot of a realization of the empirical eigenvalue distribution ² for large n and p and the plot of the "limiting" Marčenko-Pastur distribution (for $\gamma = n/p$). Figure 2 represents such a comparison for $n = 80$, $p = 30$, $\gamma = 8/3$.

²Note that the empirical eigenvalue distribution function as a function of random variables (sample eigenvalues) is a random function!

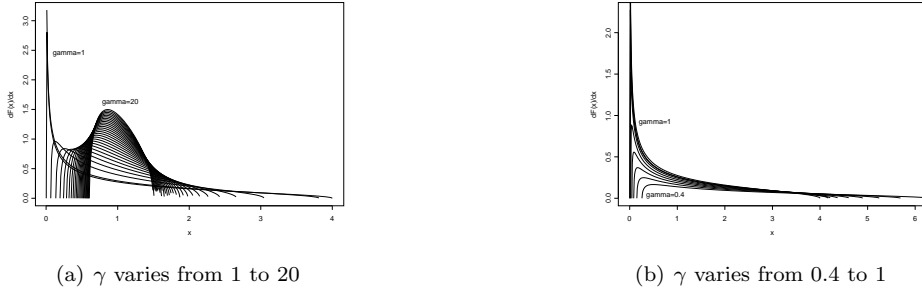


FIGURE 1. Densities of the Marčenko-Pastur distribution, corresponding to the different values of the parameter γ .

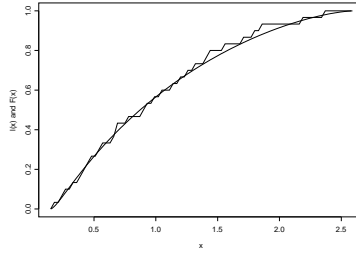


FIGURE 2. Realization of the empirical eigenvalue distribution ($n = 80, p = 30$) and the Marčenko-Pastur cdf with $\gamma = 8/3$.

The top and the bottom eigenvalues $l_1^{(p)}$ and $l_{\min\{n,p\}}^{(p)}$ converge almost surely to the edges of the support ³ $[a, b]$ of G [see Geman (1980) and Silverstein (1985)]:

$$(2.5) \quad \begin{aligned} l_1^{(p)} &\rightarrow (1 + \gamma^{-1/2})^2, \text{ a.s.}, \\ l_{\min\{n,p\}}^{(p)} &\rightarrow (1 - \gamma^{-1/2})^2, \text{ a.s.} \end{aligned}$$

Note that if $n < p$, then $l_{n+1}^{(p)}, \dots, l_p^{(p)}$ are zero. The almost sure convergence means here, that, for instance, the following holds for the largest eigenvalue:

$$\mathbb{P} \left(\lim_{n/p \rightarrow \gamma} l_1^{(p)} = (1 + \gamma^{-1/2})^2 \right) = 1.$$

Furthermore, the results of Bai *et al.* (1988) and Yin *et al.* (1988) state that (almost surely) $l_{\max}(n\mathcal{S}) = (\sqrt{n} + \sqrt{p})^2 + o(n + p)$, where $l_{\max}(n\mathcal{S})$ is the largest eigenvalues of the matrix $n\mathcal{S}$. Notice, that from this, the result given by (2.5) follows immediately.

However, the rate of convergence was unknown until the appearance of the papers of Johansson (2000) and Johnstone (2001). The former treated the complex case,

³By the support of a distribution it is meant the set closure of the set of arguments of the distribution's density function for which this density function is distinct from zero.

whereas the latter considered the asymptotic distribution of the largest eigenvalue of real sample covariance matrix – the case of our interest.

2.3.3. *Asymptotic distribution function for the largest eigenvalue.* The asymptotic distribution for the largest eigenvalue in Wishart sample covariance matrices was found by Johstone (2001). He also reported that the approximation is satisfactory for n and p as small as 10. The main result from his work can be formulated as follows.

Theorem 5. *Let \mathcal{W} be a white Wishart matrix and l_1 be its largest eigenvalue. Then*

$$\frac{l_1 - \mu_{np}}{\sigma_{np}} \xrightarrow{\text{dist}} F_1,$$

where the center and scaling constants are

$$\mu_{np} = (\sqrt{n-1} + \sqrt{p})^2, \quad \sigma_{np} = \mu_{np} \left((n-1)^{-\frac{1}{2}} + p^{-\frac{1}{2}} \right)^{1/3},$$

and F_1 stands for the distribution function of the Tracy-Widom law of order 1.

The limiting distribution function F_1 is a particular distribution from a family of distributions F_β . For $\beta = 1, 2, 4$ functions F_β appear as the limiting distributions for the largest eigenvalues in the ensembles GOE, GUE and GSE, correspondingly. It was shown by Tracy and Widom (1993, 1994, 1996). Their results state that for the largest eigenvalue $l_{\max}(\mathcal{A})$ of the random matrix \mathcal{A} (whenever this comes from GOE ($\beta = 1$), GUE ($\beta = 2$) or GSE ($\beta = 4$)) its distribution function $F_{N,\beta}(s) \stackrel{\text{def}}{=} \mathbb{P}(l_{\max}(A) < s)$, $\beta = 1, 2, 4$ satisfies the following limit law:

$$F_\beta(s) = \lim_{N \rightarrow \infty} F_{N,\beta}(2\sigma\sqrt{N} + \sigma N^{-1/6}s),$$

where F_β are given explicitly by

$$(2.6) \quad F_2(s) = \exp \left(- \int_s^\infty (x-s)q^2(x)dx \right),$$

$$(2.7) \quad F_1(s) = \exp \left(- \frac{1}{2} \int_s^\infty q(x)dx \right) [F_2(s)]^{1/2},$$

$$(2.8) \quad F_4(2^{-2/3}s) = \cosh \left(- \frac{1}{2} \int_s^\infty q(x)dx \right) [F_2(s)]^{1/2}.$$

Here $q(s)$ is the unique solution to the Painlevé II equation ⁴

$$(2.9) \quad q'' = sq + 2q^3 + \alpha \text{ with } \alpha = 0,$$

satisfying the boundary condition

$$(2.10) \quad q(s) \sim \text{Ai}(s), \quad s \rightarrow +\infty,$$

where $\text{Ai}(s)$ denotes the Airy special function – a participant of one of the pairs of linearly independent solutions to the following differential equation:

$$\omega'' - z\omega = 0.$$

⁴referred further to simply as the Painlevé II.

The boundary condition (2.10) means that

$$(2.11) \quad \lim_{s \rightarrow \infty} \frac{q(s)}{\text{Ai}(s)} = 1.$$

The Painlevé II (2.9) is one of the six exceptional second-order ordinary differential equations of the form

$$(2.12) \quad \frac{d^2 w}{dz^2} = F(z, w, dw/dz),$$

where F is locally analytic in z , algebraic in w , and rational in dw/dz . Painlevé (1902a,b) and Gambier (1909) studied the equations of the form (2.12) and identified all solutions to such equations which have no movable points, i.e. the branch points or singularities whose locations do not depend on the constant of integration of (2.12). From the fifty canonical types of such solutions, forty four are either integrable in terms of previously known functions, or can be reduced to one of the rest six new nonlinear ordinary differential equations known nowadays as the *Painlevé equations*, whose solutions became to be called the *Painlevé transcendents*. The Painlevé II is *irreducible* – it cannot be reduced to a simpler ordinary differential equation or combinations thereof, see Ablowitz and Segur (1977).

Hastings and McLeod (1980) proved that there is a unique solution to (2.9) satisfying the boundary conditions:

$$q(s) \sim \begin{cases} \text{Ai}(s), & s \rightarrow +\infty \\ \sqrt{-\frac{1}{2}s}, & s \rightarrow -\infty \end{cases}.$$

Moreover, these conditions are independent of each other and correspond to the same solution.

For the mathematics beyond the connection between the Tracy-Widom laws and the Painlevé II see the work of Tracy and Widom (2000).

3. TRACY-WIDOM AND PAINLEVÉ II: COMPUTATIONAL ASPECTS OF REALIZATION IN *S-Plus*

A description of the numerical work on the Tracy-Widom distributions and Painlevé II equation is given in this section.

3.1. Painlevé II and Tracy-Widom laws. An approach of the representation of the Painlevé II equation (2.9) by a system of ODE's is discussed. We describe an approximation algorithm of its solving in *S-Plus*, supporting it by all analytical and algebraic manipulations needed for this purpose. The motivation is to evaluate numerically the Tracy-Widom distribution and to provide appropriate statistical tools for using in *S-Plus*.

3.1.1. From ordinary differential equation to a system of equations. To solve numerically the Painlevé II and evaluate the Tracy-Widom distributions we exploit heavily the idea of Per-Olof Persson (2002) [see also Edelman and Per-Olof Persson (2005)]. His approach for implementation in *MATLAB* is adapted for a numerical work in *S-Plus*.

Since the Tracy-Widom distributions F_β ($\beta = 1, 2, 4$) are expressible in terms of the Painlevé II whose solutions are transcendental, consider the problem of numerical evaluation of the particular solution to the Painlevé II satisfying (2.10):

$$(3.1) \quad q'' = sq + 2q^3$$

$$(3.2) \quad q(s) \sim \text{Ai}(s), s \rightarrow \infty.$$

To obtain a numeric solution to this problem in *S-Plus*, first rewrite (3.1) as a system of the first order differential equations (in the vector form):

$$(3.3) \quad \frac{d}{ds} \begin{pmatrix} q \\ q' \end{pmatrix} = \begin{pmatrix} q' \\ sq + 2q^3 \end{pmatrix},$$

and by virtue of (2.11) substitute the condition (3.2) by the condition $q(s_0) = \text{Ai}(s_0)$, where s_0 is a sufficiently large positive number. This conditions, being added to the system (3.3), have the form

$$(3.4) \quad \left. \begin{pmatrix} q(s) \\ q'(s) \end{pmatrix} \right|_{s=s_0} = \begin{pmatrix} \text{Ai}(s_0) \\ \text{Ai}'(s_0) \end{pmatrix}.$$

Now, the problem (3.3)+(3.4) can be solved in *S-Plus* as initial-value problem using the function `ivp.ab` – the initial value solver for systems of ODE's, which finds the solution of a system of ordinary differential equations by an adaptive Adams-Bashforth predictor-corrector method at a point, given solution values at another point. However, before applying this, notice that for the evaluation of the Tracy-Widom functions $F_1(s)$, $F_2(s)$ and $F_4(s)$ some integrals of $q(s)$ should be found, in our situation – numerically estimated. This can be done using the tools for numerical integration, but this would lead to very slow calculations since such numerical integration would require a huge number of calls of the function `ivp.ab`, which is inadmissible for efficient calculations. Instead, represent F_1 , F_2 and F_4 in the following form

$$(3.5) \quad F_2(s) = e^{-I(s; q(s))},$$

$$(3.6) \quad F_1(s) = e^{-\frac{1}{2}J(s; q(s))} [F_2(s)]^{1/2},$$

$$(3.7) \quad F_4(2^{-2/3}s) = \cosh\left(-\frac{1}{2}J(s)\right) [F_2(s)]^{1/2},$$

by introducing the following notation

$$(3.8) \quad I(s; h(s)) \stackrel{\text{def}}{=} \int_s^\infty (x-s)h(s)^2 dx,$$

$$(3.9) \quad J(s; h(s)) \stackrel{\text{def}}{=} \int_s^\infty h(x) dx,$$

for some function $h(s)$.

Proposition 2. *The following holds*

$$(3.10) \quad \frac{d^2}{ds^2} I(s; q(s)) = q(s)^2$$

$$(3.11) \quad \frac{d}{ds} J(s; q(s)) = -q(s)$$

Proof. First, consider the function

$$W(s) = \int_s^\infty R(x, s) dx, \quad s \in \mathbb{R},$$

where $R(x, s)$ is some function.

Recall [e.g. Korn and Korn (1968, pp 114-115)] that the *Leibnitz rule* of the differentiation under the sign of integral can be applied in the following two cases as follows:

$$(3.12) \quad \frac{d}{ds} \int_a^b f(x, s) dx = \int_a^b \frac{\partial}{\partial s} f(x, s) dx,$$

$$(3.13) \quad \frac{d}{ds} \int_{\alpha(s)}^{\beta(s)} f(x, s) dx = \int_{\alpha(s)}^{\beta(s)} \frac{\partial}{\partial s} f(x, s) dx + f(\alpha(s), s) \frac{d\alpha}{ds} - f(\beta(s), s) \frac{d\beta}{ds},$$

under conditions that the function $f(x, s)$ and its partial derivative $\frac{\partial}{\partial s} f(x, s)$ are continuous in some rectangular $[a, b] \times [s_1, s_2]$ and the functions $\alpha(s)$ and $\beta(s)$ are differentiable on $[s_1, s_2]$ and are bounded on this segment. Furthermore, the formula (3.12) is also true for improper integrals under condition that the integral $\int_a^b f(x, s) dx$ converges, and the integral $\int_a^b \frac{\partial}{\partial s} f(x, s)$ converges uniformly on the segment $[s_1, s_2]$. In this case the function $f(x, s)$ and its partial derivative $\frac{\partial}{\partial s} f(x, s)$ are only supposed to be continuous on the set $[a, b] \times [s_1, s_2]$, or $(a, b] \times [s_1, s_2]$, depending on which point makes the integral improper.

Now represent $W(s)$ as follows

$$W(s) = \int_s^b R(x, s) dx + \int_b^\infty R(x, s) dx, \text{ for some } b \in [s, \infty),$$

and apply the Leibnitz rule for each of the summands under the suitable conditions imposed on $R(x, s)$:

$$\begin{aligned} \frac{d}{ds} \int_s^b R(x, s) dx &= \int_s^b \frac{\partial}{\partial s} R(x, s) dx + R(b, s) \frac{db}{ds} - R(s, s) \frac{ds}{ds} \\ &= \int_s^b \frac{\partial}{\partial s} R(x, s) dx - R(s, s), \text{ and} \end{aligned}$$

$$\frac{d}{ds} \int_b^\infty R(x, s) dx = \int_b^\infty \frac{\partial}{\partial s} R(x, s) dx,$$

where for the second differentiation we have used the rule (3.13) for improper integrals as exposed above.

Finally, derive that

$$(3.14) \quad \frac{d}{ds} \int_s^\infty R(x, s) dx = \int_s^\infty \frac{\partial}{\partial s} R(x, s) dx - R(s, s).$$

This particularly gives for $R(x, s) \equiv I(s; q(s))$ the expression

$$\frac{d}{ds} I(s; q(s)) = - \int_s^\infty q(x)^2 dx,$$

which after the repeated differentiation using the same rule becomes

$$\frac{d^2}{ds^2}I(s; q(s)) = q(s)^2.$$

Similarly, for $J(s; q(s))$ one gets

$$\frac{d}{ds}J(s; q(s)) = -q(s).$$

The conditions under which the differentiation takes place can be easily verified knowing the properties of the Painlevé II transcendent $q(s)$ and its asymptotic behaviour at $+\infty$. \square

Write further $[\cdot]$ for a vector and $[\cdot]^T$ for a vector transpose.

Define the following vector function

$$(3.15) \quad \mathbf{V}(s) \stackrel{\text{def}}{=} [q(s), q'(s), I(s; q(s)), I'(s; q(s)), J(s)]^T.$$

The results of the Proposition 2 can be used together with the idea of artificial representation of an ODE with a system of ODE's. Namely, the following system with the initial condition is a base of using the solver `ivp.ab`, which will perform the numerical integration automatically:

$$(3.16) \quad \mathbf{V}'(s) = [q'(s), sq + 2q^3(s), I'(s; q(s)), q^2(s), -q(s)]^T,$$

$$(3.17) \quad \mathbf{V}(s_0) = [\text{Ai}(s_0), \text{Ai}'(s_0), I(s_0; \text{Ai}(s)), \text{Ai}(s_0)^2, J(s_0; \text{Ai}(s))]^T.$$

The initial values in (3.17) should be computed to be passed to the function `ivp.ab`. The Airy function $\text{Ai}(s)$ can be represented in terms of other common special functions, such as the Bessel function of a fractional order, for instance. However, there is no default support neither for the Airy function, nor for the Bessel functions in *S-Plus*. Therefore, the values from (3.17) at some "large enough" point s_0 can, again, be approximated using the asymptotics of the Airy function ($s \rightarrow \infty$). The general asymptotic expansions for large **complex** s of the Airy function $\text{Ai}(s)$ and its derivative $\text{Ai}'(s)$ are as follows [see, e.g., Antosiewicz (1972)]:

$$(3.18) \quad \text{Ai}(s) \sim \frac{1}{2}\pi^{-1/2}s^{-1/4}e^{-\zeta} \sum_{k=0}^{\infty} (-1)^k c_k \zeta^{-k}, \quad |\arg s| < \pi,$$

where

$$(3.19) \quad c_0 = 1, \quad c_k = \frac{\Gamma(3k + \frac{1}{2})}{54^k k! \Gamma(k + \frac{1}{2})} = \frac{(2k+1)(2k+3)\dots(6k-1)}{216^k k!}, \quad \zeta = \frac{2}{3}s^{3/2};$$

and

$$(3.20) \quad \text{Ai}'(s) \sim -\frac{1}{2}\pi^{-1/2}s^{1/4}e^{-\zeta} \sum_{k=0}^{\infty} (-1)^k d_k \zeta^{-k}, \quad |\arg s| < \pi,$$

where

$$(3.21) \quad d_0 = 1, \quad d_k = -\frac{6k+1}{6k-1}c_k, \quad \zeta = \frac{2}{3}s^{3/2}.$$

Setting

$$\tilde{c}_k \stackrel{\text{def}}{=} \ln c_k = \ln(2k+1) + \ln(2k+3) + \dots + \ln(6k-1) - k \ln 216 - \sum_{i=1}^k \ln i, \text{ and}$$

$$\tilde{d}_k \stackrel{\text{def}}{=} \ln(-d_k) = \ln \frac{6k+1}{6k-1} + \tilde{c}_k,$$

one gets the following recurrences:

$$\tilde{c}_k = \tilde{c}_{k-1} + \ln(3 - (k-1/2)^{-1}),$$

$$\tilde{d}_k = \tilde{c}_{k-1} + \ln(3 + (k/2 - 1/4)^{-1}),$$

which can be efficiently used for the calculations of the asymptotics (3.19) and (3.20) in the following form:

$$(3.22) \quad \text{Ai}(s) \sim \frac{1}{2} \pi^{-1/2} s^{-1/4} e^{-\zeta} \sum_{k=0}^{\infty} (-1)^k e^{\tilde{c}_k} \zeta^{-k},$$

$$(3.23) \quad \text{Ai}'(s) \sim \frac{1}{2} \pi^{-1/2} s^{1/4} e^{-\zeta} \sum_{k=0}^{\infty} (-1)^{k+1} e^{\tilde{d}_k} \zeta^{-k}.$$

The function `AiryAsymptotic` can be found in *Listing*. For example, its calls with 200 terms of the expansion at the points 4.5, 5 and 6 return:

```
> AiryAsymptotic(c(4.5,5,6),200)
[1] 3.324132e-004 1.089590e-004 9.991516e-006
```

which is precise up to a sixth digit (compare with the example in Antosiewicz (1972, p.454) and with MATLAB's output).

However, if the value for s_0 is set to be fixed appropriately, there is no need in asymptotic expansions of $\text{Ai}(s)$ and $\text{Ai}'(s)$ for solving the problem⁵ (3.16)+(3.17). From the experimental work I have found that the value $s_0 = 5$ would be an enough "good" to perform the calculations. The initial values are as follows:

$$\mathbf{V}(5) = [1.0834e^{-4}, -2.4741e^{-4}, 5.3178e^{-10}, 1.1739e^{-8}, 4.5743e^{-5}]^T.$$

From now on let $I(s) \equiv I(s; q(s))$ and $J(s) \equiv J(s; q(s))$. Further we show how to express F_β, f_β in terms of $I(s)$ and $J(s)$ and their derivatives. This will permit to evaluate approximatively the Tracy-Widom distribution for $\beta = 1, 2, 4$ by solving the initial value problem (3.16)+(3.17).

From (3.5)-(3.7) the expressions for F_1 and F_4 follows immediately:

$$(3.24) \quad F_1(s) = e^{-\frac{1}{2}[I(s)+J(s)]}$$

$$(3.25) \quad F_4(s) = \cosh\left(-\frac{1}{2}J(\gamma s)\right)e^{-I(\gamma s)},$$

where $\gamma = 2^{2/3}$.

Next, find expressions for f_β . From (3.5) it follows that

$$(3.26) \quad f_2(s) = -I'(s)e^{-I(s)}.$$

⁵Note also that the using of (3.19) and (3.20) would add an additional error while solving the Painlevé II with the initial boundary condition (2.10) and thus while evaluating F_β .

$\beta \setminus p$ -points	0.995	0.975	0.95	0.05	0.025	0.01	0.005	0.001
1	-4.1505	-3.5166	-3.1808	0.9793	1.4538	2.0234	2.4224	3.2724
2	-3.9139	-3.4428	-3.1945	-0.2325	0.0915	0.4776	0.7462	1.3141
4	-4.0531	-3.6608	-3.4556	-1.0904	-0.8405	-0.5447	-0.3400	0.0906

TABLE 1. Values of the Tracy-Widom distributions ($\beta=1, 2, 4$) exceeded with probability p .

The expressions for f_1, f_4 follows from (3.24) and (3.25):

$$(3.27) \quad f_1(s) = -\frac{1}{2}[I'(s) - q(s)]e^{-\frac{1}{2}[I(s)+J(s)]},$$

and

$$(3.28) \quad f_4(s) = -\frac{\gamma}{2}e^{-\frac{I(\gamma s)}{2}} \left[\sinh\left(\frac{J(\gamma s)}{2}\right) q(\gamma s) + I'(\gamma s) \cosh\left(\frac{J(\gamma s)}{2}\right) \right].$$

Note that $J'(s) = -q(s)$ as shown in the Proposition 2.

Implementation notes. As already mentioned, the problem (3.16)+(3.17) can be solved in *S-plus* using the initial value solver `ivp.ab`.

For instance, the call

```
> out <- ivp.ab(fin=s,init=c(5,c(1.0834e-4,-2.4741e-4, 5.3178e-10,
1.1739e-8,4.5743e-5)), deriv=fun,tolerance=1e-11),
```

where the function `fun` is defined as follows

```
> fun<-function(s,y) c(y[2],s*y[1]+2*y[1]^3,y[4],y[1]^2,-y[1]),
```

evaluates the vector \mathbf{V} , defined in (3.15), at some point \mathbf{s} , and hence evaluates the functions $q(s)$, $I(s)$ and $J(s)$. Further, F_β , f_β can be evaluated using (3.5), (3.24), (3.25), (3.27)-(3.28). The corresponding functions are `FTWb(s,beta)` and `fTWb(s,beta)`, and can be found in *Listing*.

The Tracy-Widom quantile function `qTWb(p,beta)` uses the dichotomy method applied for the cdf of the corresponding Tracy-Widom distribution. Alternatively, the Newton-Raphson method can be used, since we know how to evaluate the derivative of the distribution function of the Tracy-Widom law, i.e. how to calculate the Tracy-Widom density.

Finally, given a quantile returning function it is easy now to generate Tracy-Widom random variables using the *Inverse Transformation Method* (e.g. see Ross (1997)). The corresponding function `rTWb(beta)` can be found in *Listing*.

Using the written *S-Plus* functions which are mentioned above, the statistical tables of the Tracy-Widom distributions ($\beta = 1, 2, 4$) have been constructed. These can be found in Appendix A. Table 1 comprises some standard tail p -values of the Tracy-Widom laws.

Some examples of using the described functions follow:

0.05% p -value of TW_2 is

```
> qTWb(1-0.95,2)
```

```
[1] -3.194467
```

Conversely, check the 0.005% tail p -value of TW_1 :

```
> 1-FTWb(2.4224,1)
```

```
val 5
```

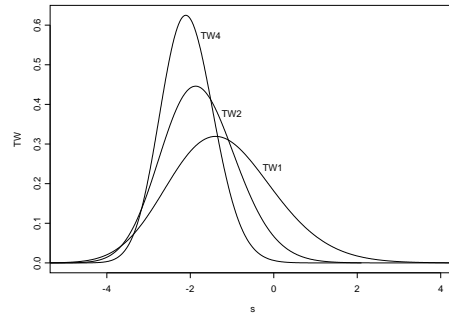


FIGURE 3. Tracy-Widom density plots, corresponding to the values of β : 1, 2 and 4.

```
0.004999185
```

(compare with the corresponding entries of Table 1)

Next, evaluate the density of TW_4 at the point $s = -1$:

```
> fTWb(-1,4)
```

```
val 3
```

```
0.1576701
```

Generate a realization of a random variable distributed as TW_1 :

```
> rTWb(1)
```

```
[1] -1.47115
```

Finally, for the density plots of the Tracy-Widom distributions ($\beta = 1, 2, 4$) appeal to Figure 3.

Performance and stability issues. There are two important questions: how fast the routines evaluating the Tracy-Widom distributions related functions are, and whether the calculations provided by these routines are enough accurate to use them in statistical practice.

The answer to the former question is hardly not to be expected – the "Tracy-Widom" routines in the form as they are presented above exhibit an extremely slow performance. This reflects dramatically on the quantile returning functions which use the dichotomy method and therefore need several calls of corresponding cumulative functions. The slow performance can be well explained by a necessity to solve a system of ordinary differential equations given an initial boundary condition. However, if the calculations provided by these routines are "enough" precise we could tabulate the Tracy-Widom distributions on a certain grid of points and then proceed with an approximation of these distributions using, for instance, *smoothing* or *interpolating splines*.

Unfortunately, there is not much to say on the former question, although some insight can be gained. The main difficulty here is the evaluation of the error appearing while substituting the Painlevé II problem with a boundary condition at $+\infty$ by a problem with an initial value at some finite point, i.e., while substituting (3.1)+(3.2) by the problem (3.3)+(3.4). However, it would be really good to have any idea about how sensible such substitution is. For this, appeal to Figure 4,

where the plots of three different Painlevé II transcendents evaluated on a uniform grid from the segment $[-15, 5]$ using the calls of `ivp.ab` with different initial values are given, such that these transcendents "correspond" to the following boundary conditions: $q(s) \sim k \operatorname{Ai}(s)$, $k = 1 - 10^{-4}, 1, 1 + 10^4$. The plots were produced by evaluating the transcendents using the substitution (3.4) of the initial boundary condition. The value for s_0 has been chosen the same as in the calculations for the Tracy-Widom distribution: $s_0 = 5$. The question is how the output obtained with the help of `ivp.ab` after a substitution of the boundary condition at $+\infty$ with the local one agrees with the theory?

Consider the null parameter Painlevé II equation

$$(3.29) \quad q'' = sq + 2q^3,$$

and a boundary condition

$$(3.30) \quad q(s) \sim k \operatorname{Ai}(s), \quad s \rightarrow +\infty.$$

The asymptotic behaviour on the negative axis of the solutions to (3.29) satisfying the condition (3.30) is as follows [Clarkson and McLeod (1988), Ablowitz and Segur (1977)]:

- if $|k| < 1$, then as $z \rightarrow -\infty$

$$(3.31) \quad q(s) \sim d(-s)^{-1/4} \sin\left(\frac{2}{3}(-s)^{3/2} - \frac{3}{4}d^2 \ln(-s) - \theta\right),$$

where the so called *connection formulae* for d and θ are

$$\begin{aligned} d^2(k) &= -\pi^{-1} \ln(1 - k^2), \\ \theta(k) &= \frac{3}{2}d^2 \ln 2 + \arg\left(\Gamma\left(1 - \frac{1}{2}id^2\right)\right); \end{aligned}$$

- if $|k| = 1$, then as $z \rightarrow -\infty$

$$(3.32) \quad q(s) \sim \operatorname{sign}(k) \sqrt{-\frac{1}{2}s};$$

- if $|k| > 1$, then $q(s)$ blows up at a finite s^* :

$$(3.33) \quad q(s) \sim \operatorname{sign}(k) \frac{1}{s - s^*}.$$

From Figure 4 one can see that the cases when $k < 1$ and $k > 1$ are sensitively different, whereas the case with $k = 1$ appeared to have the asymptotics of the solutions with $k < 1$. However, all the three solutions are hardly distinguished to the right from $s = -2.5$.

Next, compare the plot of the solution corresponding to the case $k = 1$ with the curve corresponding to the theoretical asymptotic behaviour given by (3.32). The plot is presented in Figure 5. Here we see that the plot of our solution congruently follows the asymptotic curve up to the point $s \approx -4$. In this connection an important observation should be made – to the left from this point the Tracy-Widom densities' values ($\beta = 1, 2, 4$) become smaller and smaller, starting with $\max_{\beta=1,2,4} f_\beta(-4) = f_1(-4) \approx 0.0076$.

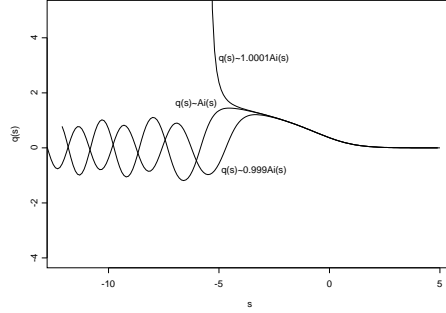


FIGURE 4. Comparison of the evaluated Painlevé II transcendent $q(s) \sim k \text{Ai}(s)$, $s \rightarrow \infty$, corresponding to the different values of k : $1 - 10^{-4}$, 1 , $1 + 10^{-4}$, as they were passed to `ivp.ab`. Practically, the solutions are hardly distinguished to the right from $s = -2.5$.

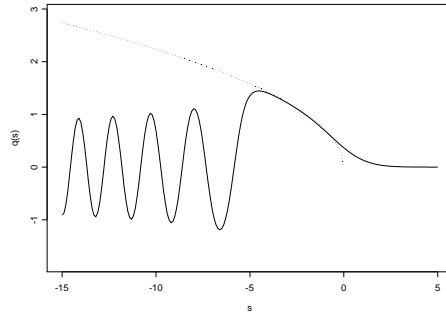


FIGURE 5. The evaluated Painlevé II transcendent $q(s) \sim \text{Ai}(s)$, $s \rightarrow \infty$, and the parabola $\sqrt{-\frac{1}{2}s}$.

Finally, Figure 6 represents a graphical comparison of the evaluated Painlevé II transcendent, for which k is less but close to one, with the predicted theoretical behaviour of the decay oscillations type, given by (3.31).

All these graphical comparisons permit us to conclude that the using of the function `ivp.ab` for solving the initial value problem for a system of ordinary differential equations, which was a result of some artificial transformations, delivers quite sensible calculations, according with the theoretical prediction based on the asymptotic behaviour of the desired objects. Unfortunately, there is not much information on the numerical estimation of the accuracy of the computations presented here.

It is believed by the author of this paper that the values of the tabulated Tracy-Widom distributions from Tables (2)-(7) are precise up to the fourth digit.

The problem is now to struggle with the speed of performance. We move to the spline approximation.

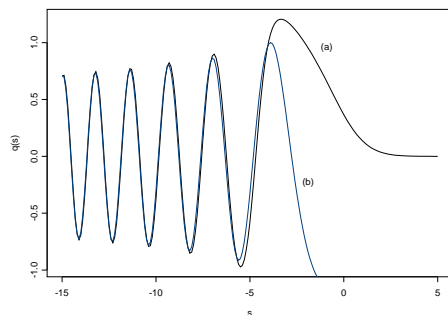


FIGURE 6. (a) The evaluated Painlevé II transcendent $q(s) \sim 0.999 \text{Ai}(s)$, $s \rightarrow +\infty$, and (b) the asymptotic ($s \rightarrow -\infty$) curve (3.31), corresponding to $k = 1 - 10^{-4}$.

3.1.2. *Spline approximation.* As we found, the functions calculating the Tracy-Widom distributions, described above, have a very slow performance, which makes their use almost ineffective in applications. On the other hand, our *heuristic* analysis has shown that the accuracy of those functions is relatively high. How to use this latter property and eliminate of the former one?

We use the cubic spline interpolation for representing the functions related with the Tracy-Widom (cumulative distribution functions, density functions and quantile functions).

The main idea is in the following: the Tracy-Widom distributions are tabulated on a uniform grid from the segment of the maximal concentration of a respective distribution ($\beta = 1, 2, 4$) using "precise" functions from *Listing* and described in § 3.1.1. Then, this data set is used for interpolating by using the cubic splines.

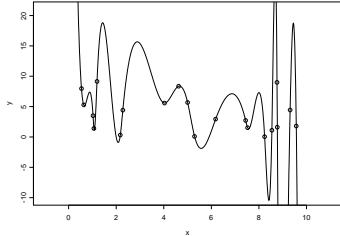
The *S-Plus* function `spline`, which interpolates through data points by means of a cubic spline, is used rather for visualizing purposes, since this functions can only interpolate the initial data set at evenly situated new points. The function `bs` can be used for a generation of a basis matrix for polynomial splines and can be used together with linear model fitting functions. Its use is rather tricky and, again, inappropriate in our situation. What we want is the function which would permit us to interpolate through a given set of tabulated points of a desired distribution. Preferably, it should be a vector-supporting function.

Such function have been written for using in *S-Plus*. The function `my.spline.eval` is designed for the coefficients' calculation of a cubic spline given a data set, and the function `my.spline` returns a value of the spline with the coefficients passed to this function. Practically, the both functions can be used efficiently together. See their description and body texts in *Listing*.

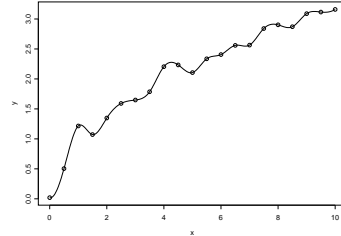
The following examples are to demonstrate the use of the functions `my.spline.eval` and `my.spline`, as well as some peculiarities of the spline interpolation depending on the data set's features and character.

Figures 7(a) and 7(b) contain the graphical outputs of the following two *S-Plus* sessions:

```
> x<-runif(20,0,1);y<-runif(20,0,1);x<-sort(x);reg<-seq(0,10,0.01)
> plot(x,y,xlim=c(-1,11),ylim=c(-10,21))
> lines(reg,my.spline.eval(reg,x,y,my.spline(x,y)))
```



(a) first example: phenomenon of a sudden oscillation in spline interpolation of data



(b) second example: no sudden oscillations

FIGURE 7. Characteristic example of using the functions `my.spline` and `my.spline.eval` in *S-Plus*.

```
> x<-seq(0,10,0.5);y<-sqrt(x)+rnorm(21,0,0.1);reg<-seq(0,10,0.01)
> plot(x,y);lines(reg,my.spline.eval(reg,x,y,my.spline(x,y)))
```

Listing contains the text bodies of the functions `dtw`, `ptw` and `qtw` which implements the evaluation of the Tracy-Widom density, cumulative distribution and quantile, correspondingly, using the cubic spline interpolation. The use of these functions is standard, e.g.

```
> dtw(runif(4,-6,7),1)
[1] 0.0002247643 0.2891297912 0.0918758337 0.2393426156
```

Interestingly, it has been found that the *S-Plus* function `smooth.spline` provides even faster calculations, though relatively as good as the functions based on `my.spline` and `my.spline.eval`. Analogs of `dtw`, `ptw` and `qtw` based on the using of `smooth.spline` can be found in Listing 5b.

```
> reg<-runif(4,-6,7); dtw(reg,1);dtwsmooth1(reg)
> [1] 0.00738347825 0.06450056711 0.00009679960 0.00004254274
> 0.007383475 0.06450054 0.00009685025 0.00004254317
```

3.2. Summary. To summarize, the work relating to the numerical evaluation of the Tracy-Widom density, cumulative distribution functions as well as quantile and random number generation functions for the using in *S-Plus* was described in this chapter. The approach based on the representation of an ordinary differential equation of the order higher than one with the system of ordinary differential equations led to extremely slow, though "sensible" computations. Therefore, the method of spline approximation on a uniform grid has been used. The implementation of this method provided significantly faster calculations which permit an efficient statistical work. Although, there are some principal difficulties with the estimation of the accuracy of the calculations.

4. APPLICATIONS

This section is designed for the discussion on some applications of the results relating to the distribution of the largest eigenvalue(s) in covariance matrices.

4.1. Tests on covariance matrices and sphericity tests. Sphericity tests represent inferential tools for answering the question: is there an evidence that a population matrix is not proportional to the identity one? Such tests can be used while studying the question whether a certain sample covariance matrix (its realization) corresponds to a population with a given matrix describing the population covariance structure.

Put it more formally. Suppose the matrix S is a sample covariance matrix calculated from some data obtained in an experiment. Suppose also that we want to check the following hypothesis: does S correspond to a population with true covariance matrix Σ ?

Kendall and Stuart (1968) reduce this hypothesis to a *sphericity hypothesis*. Since Σ is known, they point, it can be linearly transformed to a unity matrix (of course, this linear transformation is nothing but Σ^{-1}). Now it suffices to check whether the matrix $C = \Sigma^{-1}S$ does correspond to the theoretic matrix $\sigma^2 I$, where σ is unknown. This is called the *sphericity test*.

The statistics of such test under the assumption of multinormality is due to Mauchly (1940) and it is given by

$$(4.1) \quad l = \left[\frac{\det C}{((\text{tr } C)/p)^p} \right]^{n/2},$$

where n and p are the data dimensions as usual. The quantity $-n \log l^{2/n}$ is distributed as χ^2 with $f = p(p+1)/2 - 1$ degrees of freedom.

The Johnstone's result (Theorem 5) leads straightforwardly to a test of sphericity. Indeed, the result is stated for white Wishart matrices, hence a sphericity test for covariance matrices based on the largest eigenvalue statistics can be constructed. Namely, consider the following null hypothesis

$$(4.2) \quad H_0 : \Sigma = \sigma^2 I,$$

against the alternative one

$$H_1 : \Sigma \neq \sigma^2 I.$$

Under the assumption of multinormality, the properly rescaled multiple of the largest eigenvalue of the sample covariance matrix has approximately the Tracy-Widom ($\beta = 1$) distribution. Thus, approximate $p \cdot 100\%$ significance values for the sphericity test will be the corresponding $p/2 \cdot 100\%$ and $(1 - p/2) \cdot 100\%$ quantiles of this distribution. The test is double-sided.

Consider an illustrative example, which is rather artificial, nevertheless sufficient to demonstrating the technique.

Let `SIGMA` be an *S-Plus* variable representing some positive-definite square matrix of size 10×10 . Simulate 40 samples from the multivariate normal population with the true covariance matrix $\Sigma \equiv \text{SIGMA}$:

```
> X<-rmvnorm(40, mean=rep(0,10), cov=SIGMA);
then derive a sample covariance matrix and its inverse
```

```

> S<-var(X)
> L<-ginverse(SIGMA);
which after the appropriate transformation
> C<-L%*%S

```

can be tested on sphericity ($C = \sigma^2 I$?) This can be done by using the statistics l from (4.1)

```

> s<-sum(eigen(C)$values)
> p<-prod(eigen(C)$values)
> l<-( p/((s/10)^10) )^(20)
> -40*log(l^(2/40))
> [1] 52.31979

```

and this value is not significant, since the 99% quantile of the χ^2 distribution with 54 degrees of freedom is much greater than the mean of this distribution which is just 54 – the number of degrees of freedom. There is no evidence for rejecting the hypothesis that S is a sample covariance matrix corresponding to the population with true covariance matrix which is equal to $SIGMA$.

Consider now the verification of the same hypothesis using the largest eigenvalue statistics

```

> l<-eigen(C)$values[1]
> n<-40; p<-10; mu<-(sqrt(n-1)+sqrt(p))^2
> sigma<-sqrt(mu)*(1/sqrt(n-1)+1/sqrt(p))^(1/3)
> (40*l-mu)/sigma
> -2.939404

```

Now appeal to Table 1 to find that -2.94 does not fall in the rejection 99% double-sided significance region of the TW_1 . The null hypothesis can be accepted.

4.2. Principal Component Analysis. As already noted, the Principal Component Analysis is one of the most employed statistical techniques. This technique is based on the spectral decomposition of covariance matrices, and it is not surprisingly, that the variability and asymptotical results on the boundary of the spectra of such matrices significantly contribute to the development of the PCA technique.

4.2.1. Lizard spectral reflectance data. A data set used as an example in the further exposition is described here.

The data set we used contains the lizard spectral reflectance data. Such measurements of colours' variations within and/or between populations are often used in biology to study and analyze the sexual selection, species recognition and differential predation risk in the animal world.

The measurements of lizard front legs' spectra were made with the use of a spectrophotometer, normal to the surface. They were expressed relative to a certified 99% white reflectance standard. The range of measurements was from 320 to 700 nm, and the data set comprised 77 data points at 5 nm intervals of wavelength. Each measurement was an average of three consecutive readings. The graphical summary of the data is presented in Figure 8. There are 95 records of 77 wavelength points and this data set is a typical example of high dimensional data.

The data was kindly offered by Professor Firth and Mr. Caralambous. This assistance is gratefully acknowledged.

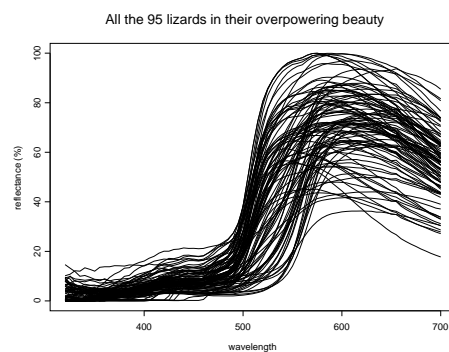
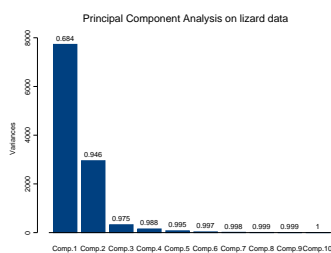
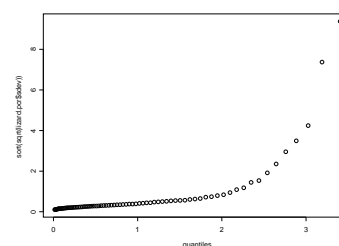


FIGURE 8. Reflectance samples measured from 95 lizards' front legs (wavelengths are measured in nanometers).

4.2.2. *Wachter plot.* There may be given many suggestions and rules regarding how many principal components to retain while proceeding with PCA. One of the heuristical ways of the graphical nature for doing this is the so called *screepplot*. This is a plot of the ordered principal component variances, i.e. a plot of the eigenvalues of a corresponding realization of the sample covariance matrix. The screeplot of eigenvalues in the lizard example is given in Figure 9(a).



(a) Scree plot from the PCA on the lizard data



(b) Wachter plot for the lizard spectral reflectance data

FIGURE 9. Comparison of the screeplot and Wachter plot for the lizard data.

There is a modification of a screeplot due to Wachter (1976) who proposed a graphical data-analytic use of the Marčenko-Pastur "semicircle"-type law behaviour of eigenvalues⁶: to make a probability plot of the ordered observed eigenvalues l_{p-i+1} against the quantiles $G^{-1}\left(\frac{i-1/2}{p}\right)$ of the Marčenko-Pastur distribution with the density given by (2.4). Let us call such a plot the *Wachter plot*.

The Wachter plot corresponding to the lizard data is presented in Figure 9(b). This plot shows an interesting feature – there is an uptick from the right, – a departure from the straight line displacement, which is predicted theoretically by

⁶See the discussion in § 2.3.2.

the Marčenko-Pastur law. Can one use the Wachter plot in the context of the lizard data? Recall that the Marčenko-Pastur theorem (§ 2.3.2) was formulated for the white case, i.e. the i.i.d. Gaussian data samples were assumed then.

Perform now a test of the departure from the multinormality on the lizard data. The test is based on the *Mahalanobis distance statistics* [e.g. Boik 2004]:

$$D_i = [(\mathbf{x}_i - \bar{\mathbf{x}})' S^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})]^{1/2},$$

where S is a sample covariance matrix corresponding to a data matrix consisting of samples $\{\mathbf{x}_i\}_{i=1}^n$. If the data have been sampled from a multivariate normal population, then

$$(4.3) \quad \frac{n}{(n-1)^2} D_i^2 \sim \text{Beta} \left(\frac{p}{2}, \frac{n-p-1}{2} \right),$$

where n is the number of p -variate data samples.

The QQ plot of the ordered D_i^2 statistics against the quantiles of the beta distribution with parameters $77/2$ and $(95 - 77 - 1)/2$ is reproduced in Figure 10. There is an evidence of the departure from the multivariate normal distribution⁷. More of this, it can be easily found that "neighbour" wavelengths are strongly correlated, which is natural to expect.

However, it is known that the Marčenko-Pastur result still holds for some special models. One of such models is a *spiked population models*. This model helps to understand the phenomenon occurring in the Wachter plot and to extract some useful information of the population covariance matrix even in the case when the sample covariance matrix is not a good approximate.

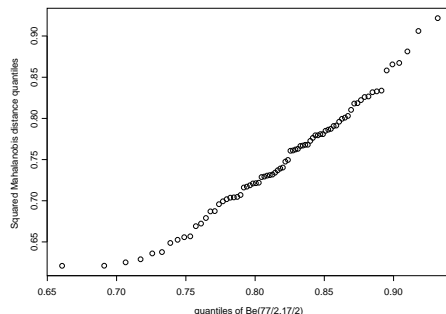


FIGURE 10. QQ plot of squared Mahalanobis distance corresponding to the lizard data: test on the departure from multinormality.

4.2.3. Spiked Population Model. It has been observed by many researches that for certain types of data only a few first (ordered) sample eigenvalues have the limiting behaviour differing from the behaviour that would be expected under identity covariance scenario. Debashis (2004) mentions that this phenomenon occurs in

⁷Notice also that the assumption of multinormality would lead to marginal normality for each wavelength value measured, which evidently is not the case. Recall that the multinormality implies marginal normality, the converse is generally not true.

speech recognition [Buja *et al* (1995)], wireless communication and engineering [Telatar (1999)], statistical learning [Hoyle and Rattray (2003, 2004)].

To study this phenomenon, Johnstone (2001) introduced the so called *spiked population model* – a model in which independently and identically distributed zero-mean real-valued observations $\mathbf{X}_1, \dots, \mathbf{X}_n$ from a p -variate normal distribution have the covariance matrix $\Sigma = \text{diag}(l_1, l_2, \dots, l_r, 1, \dots, 1)$ where $l_1 \geq l_2 \geq \dots \geq l_r > 1$.

Baik and Silverstein (2004) studied the problem of the sample eigenvalues' behaviour in the spiked population models for a more general class of samples. In their settings, the *spiked population model* comprises a class of p -variate populations with non-negative definite Hermitian covariance matrices, which have the following spectra (given in the form of a diagonal matrix extracted from the spectral decomposition of covariance matrix):

$$\Lambda_r = \text{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{k_1}, \underbrace{\lambda_2, \dots, \lambda_2}_{k_2}, \dots, \underbrace{\lambda_M, \dots, \lambda_M}_{k_M}, \underbrace{1, \dots, 1}_{p-r}),$$

where $r = k_1 + k_2 + \dots + k_M$, $k_i \geq 0$, $i = 1, \dots, M$, and $\lambda_1 > \dots > \lambda_M$. One can say that Λ defines a spiked model.

In such settings, following Johnstone (2001), denote the distribution of the k^{th} largest **sample** eigenvalue l_k in the spiked model defined by Λ_r by $\mathcal{L}(l_k|n, p, \Lambda_r)$. It was proved by Johnstone that

$$(4.4) \quad \mathcal{L}(l_{r+1}|n, p, \Lambda_r) \stackrel{\text{st}}{<} \mathcal{L}(l_1|n, p-r, I_{p-r}),$$

where $r < p$, s.t.

$$\mathbb{P}_{I_{p-r}}(l_1 < x) < \mathbb{P}_{\Lambda_r}(l_{r+1} < x),$$

where $\mathbb{P}_{I_{p-r}}(\cdot)$ and $\mathbb{P}_{\Lambda}(\cdot)$ are the distribution functions for the largest and the $(r+1)^{\text{th}}$ eigenvalues in the null and the spiked models, correspondingly.

This result provides a conservative p -value for testing the null hypothesis $H_0 : \lambda_{r+1} = 1$ in the spiked model. This means, that the null hypothesis must be rejected on the $p \cdot 100\%$ level of significance once the corresponding p -value for $\mathcal{L}(l_1|n, p-r, I_{p-r})$ happens to be significant⁸.

Consider for a moment the same *S-Plus* variable **SIGMA** introduced in § 4.1. The Wachter plot, corresponding to **SIGMA** is presented in Figure 11. It suggests departure from the sphericity at least in 3 directions. Here $n = 40$, $p = 10$, and the spiked model corresponding to the three separated eigenvalues is of our primary interest. Let Λ_3 to be a diagonal 10×10 matrix with the top values $\lambda_1, \lambda_2, \lambda_3$ equal to the observed ones, and the remaining diagonal elements all set to one. The following *S-Plus* code shows that by virtue of the Johnstone's result (4.4) the fourth eigenvalue in the data is significantly larger that would be expected under the corresponding null model.

```
> n<-40; p<-10; r<-3;
> mu<-(sqrt(n-1)+sqrt(p-r))^2;
> sigma<-sqrt(mu)*(1/sqrt(n-1)+1/sqrt(p-r))^(1/3)
> len<-length(eigen(SIGMA)$values);
> su<-sum(eigen(SIGMA)$values[(r+1):len])
```

⁸Note, that p in the word combination "p-value" is a number between 0 and 1, whereas p in $\mathcal{L}(l_{r+1}|n, p, \Lambda_r)$ and $\mathcal{L}(l_1|n, p-r, I_{p-r})$ stands for the dimension of the data.

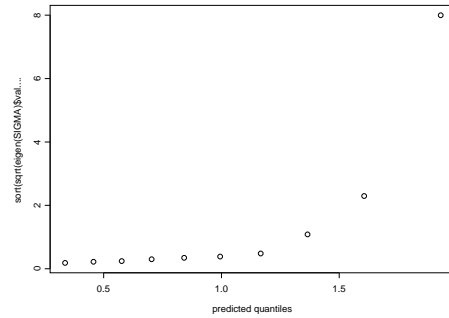


FIGURE 11. Wachter plot for the sample data simulated from a population with the true covariance matrix SIGMA .

```
> tauhatsquared<-su/(n*(p-r)); tauhatsquared*(mu+2.02*sigma)
> [1] 0.2447291
> eigen(SIGMA)$values[r]
> [1] 0.176367
```

The calculations were performed on the 99% level of significance: $0.2447291 > 0.176367$. The corresponding approximate percentile of TW_1 can be found from the first row of Table 1 to be 0.02. The case with $r = 4$ also gives significant result, whereas $r = 5$ does not lead us to the rejection of the null hypothesis. We can suggest the model with only 4 direction of the variability in our data.

```
> n<-40; p<-10; r<-5
> mu<-(sqrt(n-1)+sqrt(p-r))^2;
> sigma<-sqrt(mu)*(1/sqrt(n-1)+1/sqrt(p-r))^(1/3)
> len<-length(eigen(SIGMA)$values)
> su<-sum(eigen(SIGMA)$values[(r+1):len])
> tauhatsquared<-su/(n*(p-r)); tauhatsquared*(mu+2.02*sigma)
[1] 0.1522345
> eigen(SIGMA)$values[r]
[1] 0.1480558
```

Interestingly, it took some 70 analogous verifications (tests) for the lizard data to come up to the first non-significant result: $r = 73$ (recall, that $p = 77$). This is not in the accordance with the Wachter plot, Figure 9(b).

Some farther work should be done to explain this phenomenon. Notice, that the result (4.4) is valid for the null case, but the lizard data, as we saw, can even hardly be assumed to come from the multivariate normal distribution.

5. SUMMARY AND CONCLUSIONS

5.1. Conclusions. The behaviour of the largest eigenvalues of sample covariance matrices as random objects has been studied in this paper. A special accent was put on the study of the asymptotic results. The choice of the largest eigenvalues as the objects of interest was motivated by their importance in many techniques of

multivariate statistics, including the Principal Component Analysis, and possibility of their using in statistical tests as test statistics.

The study was done under the standard assumption of the data normality (Gaussianity). Such assumption lead us to the special class of the covariance matrices which have the Wishart distribution.

Even for the white Wishart case the exact distribution of the largest eigenvalue of a sample covariance matrix cannot be extracted explicitly from the joint density distribution of the eigenvalues, though the latter has an explicit form (see § 2.1.4).

The asymptotic result of Johnstone (§ 2.3.3) in the real white Wishart case followed the work of Johansson for the complex case, and established the convergence of the properly rescaled largest eigenvalue to the Tracy-Widom distribution TW_1 – one of the three limiting distributions for the largest eigenvalues in the Gaussian Orthogonal, Unitary and Symplectic Ensembles, found by Tracy and Widom just some ten years ago. These distributions have seen great generalizations, both theoretical and practical in the consecutive works of many authors.

The Johnstone's asymptotic result gives a good approximation for the data matrices' sizes as small as 10×10 . This is especially important for the today applications where one deals with data matrices of great, sometimes huge sizes. However, the main difficulty in using of the results of Johnstone, as well as similar and earlier results of Johansson, Tracy and Widom, is due to the fact the the limiting distributions, – the Tracy-Widom distributions – are stated in terms of the particular solution to the second order differential equation Painlevé II. This particular solution is specified with the boundary condition.

To evaluate the functions relating to the Tracy-Widom distributions in *S-Plus* (density, cumulative distribution, quantile, random number generation functions), the boundary condition which specified the solution to Painlevé II has been substituted by a local initial value condition. This permitted to reduce set of ordinary differential equations to a system of ODE's and use the initial value solver for such systems in *S-Plus*.

Using these, relatively precise calculations, the Tracy-Widom distributions ($\beta = 1, 2, 4$) have been tabulated on the intervals of the most concentration. The table of standard p -values is also presented. It is useful for the inferential statistical work involving these distributions.

However, the use of the written *S-Plus* functions for evaluating the Tracy-Widom distributions using the described above conception of solving the Painlevé II practically is hardly possible, since the performance of those functions is extremely slow. The cubic spline interpolation has been used to "recover" the distribution values from the set of pre-tabulated points. As a result, fast performance functions evaluating the Tracy-Widom distributions have been presented for the statistical work in *S-Plus*.

Interestingly, both the Tracy-Widom distribution as the limiting law and the exact largest eigenvalue distribution law required quite sophisticated computational algorithms to be efficiently evaluated using the standard mathematical software packages. We saw it for the asymptotic case, and the problem of the evaluation of the exact distribution was studied in the work of Koev and Edelman (2005), from which I draw the same conclusion.

Finally, a few practical examples have been considered to show how the studied distributions can be used in applications. These include the test on covariance matrices, and particularly the sphericity test, the spectral analysis of the high dimensional data with the use of the spiked model introduced recently by Johnstone (2001). However, one should acknowledge that in the last example (with the lizard data) many things remained not interpreted and justified at this stage of study. Here we see a necessity of some farther research.

5.2. Open problems and further work. In this paper many interesting problems from the area of the study have not been touched.

We did not study the recent developments relating to the efficient evaluation of the hypergeometric function of a matrix argument. This refers to the problem of the evaluation of the largest eigenvalue's exact distribution. The most recent results in this area are reported by Koev and Edelman (2005). A comparative analysis of the two different ways: the use of asymptotic distributions and evaluation of the exact ones, – could be done in perspective.

However, both approaches require methods of estimation of an error of approximation appearing therein. In our study it is the error appearing while substituting of the Painlevé II with a boundary condition with the Painlevé II with a local initial value condition. Although, it was shown on examples that the methods we implement by the means of *S-Plus* to solve the Painlevé II and evaluate the Tracy-Widom distributions are enough "sensible" to distinguish between what is predicted by theory even when we make such substitution of initial value problems. However, more rigorous justification should be found to use the methods we follow and to estimate the approximation error we make by using these methods. The use of splines and general theory of spline approximation would permit then to estimate the error appearing while recovering the Tracy-Widom distributions from the tabulated set of points, by considering these distributions as functions-members of a certain class of functions, and thus to estimate the final approximation error. Currently, the author cannot estimate exactly the accuracy of the numerical evaluation of the Tracy-Widom distributions presented in Tables 2-7, Appendix A. It is only believed from the experimental work the approximation is exact up to a fourth digit.

A general suggestion can be given to increase the performance of the written *S-Plus* routines: the calculation code should be written in C++ where possible and then compiled into *.dll file to be used in *S-Plus*.

There are some gaps in the consideration of using the studied asymptotic results in applications. Thus, the aspects related to the power of the sphericity test based on the largest eigenvalue statistics and the Tracy-Widom ($\beta = 1$) distribution were not touched here. Also, one needs to be more careful while using the spiked model. It also needs some further development and study. Particularly, it should be found out how robust it is in the sense of the departure of the normality of the data.

REFERENCES

- [1] Ablowitz, M.J., Segur, H. (1977) Exact linearization of a Painlevé transcendent. *Phys. Rev. Lett.* **38**(20), 1103–1106.

- [2] Aldous, D., Diaconis, P. (1999) Longest increasing subsequences: From patience sorting to the Baik-Deift-Johansson theorem. *Bull. Amer. Math. Soc.*, **36**, 413–432.
- [3] Anderson, T.W. (1963) Asymptotic theory of principal component analysis. *Ann. Math. Statist.*, **34**, 122–148.
- [4] Anderson, T.W. (2003) *An Introduction to Multivariate Statistical Analysis*. (Third edition.) New York: Wiley.
- [5] Andrew, A.L. (1990) Choosing tests matrices for numerical software. In *Computational Techniques and Applications Conference: CTAC-89*, W.L. Hogarth and B.J. Nege (eds.), 561–567.
- [6] Antosiewicz, H.A. (1972) Bessel functions of fractional order. In *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Abramowitz, M., Stegun, I. (eds.), 10th printing, 435–456.
- [7] Bai, Z.D., Silverstein, J.W., Yin, Y.Q. (1988) A note on the largest eigenvalue of a large dimensional sample covariance matrix. *J. Mult. Anal.*, **26**, 166–168.
- [8] Bai, Z.D. (1999) Methodologies in spectral analysis of large dimensional random matrices, a review. *Statist. Sinica*, **9** 611–677.
- [9] Baik, J., Silverstein, J.W. (2004) Eigenvalues of large sample covariance matrices of spiked population models. *Preprint*.
<http://arxiv.org/pdf/math.ST/0408165>
- [10] Baker, T., Forrester, P. (1997) The Calogero-Sutherland model and generalized classical polynomials. *Commun. Math. Phys.*, **188** 175–216.
- [11] Bilodeau, M. (2002) Asymptotic distribution of the largest eigenvalue. *Commun. Statist.-Simula.*, **31**(3), 357–373.
- [12] Bohigas, O., Giannoni, M.J., Schmit, C. (1984a) Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.*, **52**, 1–4.
- [13] Bohigas, O., Giannoni, M.J. (1984b) Chaotic motion and random matrix theories. In *Mathematical and Computational Methods in Nuclear Physics*, J.S. Dehesa, J.M.G. Gomez and A. Poll (eds.), pp. 1–99. Lecture Notes in Physics, vol. 209. New York: Springer Verlag.
- [14] Boik, R.J. (2004) Classical Multivariate Analysis. *Lecture Notes: Statistics 537*
<http://www.math.montana.edu/~rjboik/classes/537/notes.537.pdf>
- [15] Bollobás, B. (1985) *Random Graphs*. Academic, London.
- [16] Brody, T.A., Flores, J., French, J.P., Mello, P.A., Pandey, S., Wong, S.M. (1981) Random-matrix physics: spectrum and strength fluctuations. *Review of Modern Physics*, **53**, 385–479.
- [17] Buja, A., Hastie, T., Tibshirani, R. (1995) Penalized discriminant analysis. *Ann. Statist.*, **23**, 73–102.
- [18] Clarkson, P.A., McLeod, J.B. (1988) A connection formula for the second Painlevé transcendent. *Archive for Rational Mechanics and Analysis*, **103**, 97–138.
- [19] Constantine, A.G. (1963) Some noncentral distribution problems in multivariate analysis. *Ann. Math. Statist.*, **34**, 1270–1285.
- [20] Conway, J.A. (1990) *A course in Functional Analysis*. New York: Springer Verlag.
- [21] Debashis, P. (2004) Asymptotics of the leading sample eigenvalues for a spiked covariance model. *Preprint*.
<http://www-stat.stanford.edu/~debashis/eigenlimit.pdf>
- [22] Deift, P. (2000) Integrable systems and combinatorial theory. *Notices of the American Mathematical Society*, **47**, 631–640.
- [23] Demmel, J. (1988) The probability that a numerical analysis problem is difficult. *Math. of Comp.*, **50**, 449–480.
- [24] Diaconis, P., Shahshahani, M. (1994) On the eigenvalues of random matrices. In *Studies in Applied probability*; J. Gani (ed.), *Jour. Appl. Probab.*: Special vol.**31A**, 49–62.
- [25] Diaconis, P. (2003) Patterns in eigenvalues: The 70th Josiah Willard Gibbs lecture. *Bull. Amer. Math. Soc.*, **40**(2), 155–178.
- [26] Donoho, D.L. (2000) High-dimensional data analysis: The Curses and Blessings of dimensionality. *Technical Report*. Department of Statistics, Stanford University.
<http://www-stat.stanford.edu/~donoho/Lectures/AMS2000/Curses.pdf>
- [27] Dumitriu, I., Edelman, A. (2002) Matrix models for beta-ensembles. *J. Math. Phys.*, **43**, 5830–5847.
- [28] Dumitriu, I. (2003) Eigenvalue statistics for beta-ensembles. *PhD thesis*, Massachusetts Institute of Technology.

- [29] Dyson, F.G. (1962) Statistical theory of energy levels of complex systems, I, II and III. *Journal of Mathematical Physics*, **3**, 140–175. (Reprinted in *Selected Papers of Freeman Dyson, with Commentary*. Foreword by Elliot H. Lieb, 1996. Providence, R.I.: American Mathematical Society; Cambridge, Mass.: International Press.)
- [30] Eaton, M. (1983) *Multivariate Statistics*. Wiley: New York.
- [31] Edelman, A. (1992) On the distribution of a scaled condition number. *Math. of Comp.*, **58**, 185–190.
- [32] Edelman, A., Persson, Per-Olof (2005) Numerical methods for eigenvalue distributions of random matrices. *Preprint*.
<http://arxiv.org/pdf/math-ph/0501068>
- [33] Eisen, M.B., Spellman, P.T., Brown, P.O., Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci.*, **95**, 14863–14868.
- [34] Farkas, I.J., Derényi, I., Barabási, A.-L., Vicsek, T. (2001) Spectra of "real-world" graphs: Beyond the semicircle law *Phys. Rev. E*, **64**, 026704.
- [35] Favard, J. (1936) Application de la formule sommatoire d'Euler a la demonstration de quelques proprietes extremales des integrals des fonctions periodiques. *Math. Tidsskrift*, **4**, 81–94.
- [36] Favard, J. (1937) Sur les meilleurs procedes d'approximation de certaines classes de fonctions par des polynomes trigonometriques. *Bull. Sc. Math.*, **61**, 209–224.
- [37] Favard, J. (1940) Sur l'interpolation. *J. Math. Pures et Appl.*, **19**, 281–306.
- [38] Flury, B. (1988) *Common Principal Components and Related Multivariate Models*. New York: Wiley.
- [39] Forrester, P. (2005) *Log-gases and Random matrices*. Book in progress.
<http://www.ms.unimelb.edu.au/~matpjf/matpjf.html>
- [40] Fulton, W. (1997) *Young Tableaux: With Applications to Representation Theory and Geometry*. New York: Cambridge University Press.
- [41] Gambier, B. (1909) Sur les équations différentielles du second ordre et du premier degré dont l'intégrale générale est à points critiques fixes. *Acta Math.* **33**, 1–55.
- [42] Geman, S. (1980) A limit theorem for the norm of random matrices *Ann. Probab.* **8**(2), 252–261.
- [43] Gentle, J.E. (1998) Singular value factorization. In *Numerical Linear Algebra for Applications in Statistics*. Berlin: Springer-Verlag, 102–103.
- [44] Girshick, M.A. (1939) On the sampling theory of roots of determinantal equations. *Ann. Math. Statist.*, **10**, 203–224.
- [45] Golub, G.H., Van Loan, C.F. (1996) The singular value decomposition. In *Matrix Computations*. Baltimore, MD: John Hopkins University Press, 70–71.
- [46] Gurevich, I.I., Pevsner, M.I. (1956) Repulsion of nuclear levels. *Nuclear Physics*, **2**, 575–581. (Reprinted in *Statistical Theories of Spectra: Fluctuations. A Collection of Reprints and Original Papers, with an Introductory Review*, Charles E. Porter (ed), 1965, New York and London: Academic Press, pp. 181–187)
- [47] Hastings, S.P., McLeod, J.B. (1980) A boundary value problem associated with the second Painlevé transcendent and the Korteweg-de Vries equation. *Arch. Rational Mech. Anal.* **73**, 31–51.
- [48] Hoyle, D., Rattray, M. (2003) Limiting form of the sample covariance eigenspectrum in PCA and kernel PCA. *Advances in Neural Information Processing Systems (NIPS 16)*.
- [49] Hoyle, D., Rattray, M. (2004) Principal component analysis eigenvalue spectra from data with symmetry breaking structure. *Phys. Rev. E*, **69**.
- [50] Hsu, P.L. (1939) On the distribution of the roots of certain determinantal equations. *Ann. Eugenics.*, **9**, 250–258.
- [51] Jiang, T. (2004) The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā*, **66**(1), 35–48.
- [52] James, A.T. (1960) The distribution of the latent roots of the covariance matrix. *Ann. Math. Stat.*, **31**, 151–158.
- [53] James, A.T. (1964) Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Stat.*, **35**, 476–501.
- [54] Johansson, K. (1997) On random matrices from the compact classical groups. *Ann. of Math.*, **145**(3), 519–545.

- [55] Johansson, K. (1998) On fluctuations of random hermitian matrices. *Duke Math. J.*, **91**, 151–203.
- [56] Johansson, K. (2000) Shape fluctuations and random matrices. *Comm. Math. Phys.*, **209**, 437–476.
- [57] Johnstone, I.M. (2001) On the distribution of the largest eigenvalue in principal component analysis. *Ann. of Stat.*, **29**(2), 295–327.
- [58] Kendall, M.G., Stuart, A. (1968) Design and analysis, and time-series. Volume 3 of *The Advanced Theory of Statistics*. Second edition, Griffin, London.
- [59] Khan, E. (2005) Random matrices, information theory and physics: new results, new connections. *Informational Processes*, **5**(1), 87–99.
- [60] Koev, P., Edelman A. (2005) The efficient evaluation of the hypergeometric function of a matrix argument. *Preprint*.
<http://arxiv.org/pdf/math.PR/0505344>
- [61] Kolmogorov, A.N. (1936) Über die besste Annäherung von Funktionen einer gegebenen Funktionklassen. *Ann. of Math.*, **37**, 107–110.
- [62] Korin, B.P. (1968) On the distribution of a statistic used for testing a covariance matrix. *Biometrika*, **55**, 171.
- [63] Korn, G.A., Korn, T.M. (1968) *Mathematical Handbook for Scientists and Engineers. Definitions, Theorems and Formulas for Reference and review*. McGraw-Hill, New-York.
- [64] Liou, H.I., Camarda, H.S., Hacken, G., Rahn, F., Rainwater, J., Slagowitz, M., Wynchank, S. (1973) Neutron resonanse spectroscopy: XI: The separated isotopes of Yb. *Physival Review C*, **7**, 823–839.
- [65] Marčenko, V.A., Pastur, L.A. (1967) Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSSR - Sbornik*, **1**(4), 457–483.
- [66] Mardia, K.V., Kent, J.T., Bibby, J.M. (1979) *Multivariate Analysis*. Academic Press.
- [67] Mauchly, J.W. (1940) Significance test for sphericity of a normal n - variate distribution. *Ann. Math. Statist.*, **11**, 204.
- [68] McKay, B.D. (1981) The expected eigenvalue distribution of a large regular graph. *Lin. Alg. Appl.*, **40**, 203–216.
- [69] Mehta, M.L. (1991) *Random Matrices*. Academic Press, Boston (second edition).
- [70] Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*. Wiley, New York.
- [71] Müller, R.R. (2000) Random matrix theory. *Tutorial Overview*.
<http://lthiwww.epfl.ch/~leveque/Matrix/muller.pdf>
- [72] Okounkov, A. (2000) Random matrices and random permutations. *Mathematical Research Notices*, **20**, 1043–1095.
- [73] Painlevé, P. (1902a) Mémoires sur les équations différentielles dont l'intégrale générale est uniforme. *Bull. Soc. Math.*, **28**, 201–261.
- [74] Painlevé, P. (1902b) Sur les équations différentielles du second ordre supérieur dont l'intégrale générale est uniforme. *Acta Math.*, **25**, 1–85.
- [75] Persson, Per-Olof (2002) Random matrices. Numerical methods for random matrices. <http://www.mit.edu/~persson/numrand.report.pdf>
- [76] Pillai, K.C.S. (1976) Distribution of characteristic roots in multivariate analysis. Part I: Null distributions. *Can. J. Statist.*, **4**, 157–184.
- [77] Pillai, K.C.S. (1976) Distribution of characteristic roots in multivariate analysis. Part II: Non-null distributions. *Can. J. Statist.*, **5**, 1–62.
- [78] Ross, S.M. (1997) *Simulation*. Academic Press.
- [79] Roy, S.N. (1953) On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Statist.*, **24**, 220–238.
- [80] Samarskiy, A.A. (1982) *Introduction to numerical methods*. Moscow, Nauka (in Russian).
- [81] Schensted, C. (1961) Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics*, **13**, 179–191.
- [82] Silverstein, J.W. (1985) The smallest eigenvalue of a large dimensional Wishart matrix. *Ann. Probab.*, **13**(4), 1364–1368.
- [83] Silverstein, J.W., Combettes, P.L. (1992) Large dimensional random matrix theory for signal detection and estimation in array processing. *Preprint*.
<http://www4.ncsu.edu/~jack/workshop92.pdf>
- [84] Smale (1985) On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc.*, **13**, 87–121.

- [85] Soshnikov, A. (2002) A note on universality of the distribution of the largest eigenvalues in certain sample covariance matrices. *J. Statist. Phys.*, **108**, 1033–1056.
- [86] Stanley, R. (1989) Some combinatorial properties of Jack symmetric functions. *Adv. Math.*, **77**(1), 76–115.
- [87] Sugiyama, T. (1970) Joint distribution of the extreme roots of a covariance matrix. *Ann. Math. Statist.* **41**, 655–657.
- [88] Sugiyama, T. (1972) Percentile points of the largest latent root of a matrix and power calculation for testing $\Sigma = I$. *J. Japan. Statist. Soc.* **3**, 1–8.
- [89] Telatar, E. (1999) Capacity of multi-antenna Gaussian channels. *European Transactions on Telecommunications*, **10**, 585–595.
- [90] Tracy, C.A., Widom, H. (1993) Level-spacing distribution and Airy kernel. *Phys. Letts. B* **305**, 115–118.
- [91] Tracy, C.A., Widom, H. (1994) Level-spacing distribution and Airy kernel. *Comm. Math. Phys.* **159**, 151–174.
- [92] Tracy, C.A., Widom, H. (1996) On orthogonal and symplectic matrix ensembles. *Comm. Math. Phys.* **177**, 727–754.
- [93] Tracy, C.A., Widom, H. (1999) On the distribution of the lengths of the longest monotone subsequences in random words. <http://arXiv.org/abs/math/9904042>
- [94] Tracy, C.A., Widom, H. (2000) Airy Kernel and Painlevé II. [arXiv:solv-int/9901004](http://arXiv.org/abs/solv-int/9901004)
- [95] Tse, D.N.C (1999) Multiuser receivers, random matrices and free probability. In *Proc. of 37th Annual Allerton Conference on Communication, Control and Computing*, 1055–1064. degas.eecs.berkeley.edu/~tse/pub.html
- [96] Vismanath, P., Tse, D., Anantharam, V. (2001) Asymptotically optimal waterfilling in vector multiple access channels. *IEEE Trans. Inform. Theory*, **47**, 241–267.
- [97] Wachter, K.W. (1976) Probability plotting points for principal components. In *Ninth Interface Symposium Computer Science and Statistics*. D. Hoaglin and R. Welsch (eds.) Prindle, Weber and Schmidt, Boston, 299–308.
- [98] Weisstein, E.W. *et al.* (2005) MathWorld. A Wolfram Web Resource. <http://mathworld.wolfram.com>
- [99] Wigner, E.P. (1951) On the statistical distribution of the widths and spacings of nuclear resonance levels. *Proceedings of the Cambridge Philosophical Society*, **47**, 790–798. (Reprinted in *Statistical Theories of Spectra: Fluctuations. A Collection of Reprints and Original Papers, with an Introductory Review*, Charles E. Porter (ed), 1965, New York and London: Academic Press, pp. 120–128)
- [100] Wigner, E.P. (1955) Characteristic vectors of bordered matrices of infinite dimensions. *Ann. Math.* **62**, 548–564.
- [101] Wigner, E.P. (1957) Statistical properties of real symmetric matrices with many dimensions. *Canadian Mathematical Congress Proceedings*, 174–184. (Reprinted in *Statistical Theories of Spectra: Fluctuations. A Collection of Reprints and Original Papers, with an Introductory Review*, Charles E. Porter (ed), 1965, New York and London: Academic Press, pp. 188–198)
- [102] Wigner, E.P. (1958) On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, **67** 325–326. (Reprinted in *Statistical Theories of Spectra: Fluctuations. A Collection of Reprints and Original Papers, with an Introductory Review*, Charles E. Porter (ed), 1965, New York and London: Academic Press, pp. 226–227)
- [103] Wishart, J. (1928) The generalized product moment distribution in samples from a normal multivariate population. *Biometrika A.*, **20**, 32–43.
- [104] Yin, Y.Q, Bai, Z.D., Krishnaiah, P.R. (1988) On the limit of the largest eigenvalue of a large dimensional sample covariance matrix. *Probab. Theory Related Fields*, **78**, 509–521.
- [105] Zheng, L., Tse, S. (2001) Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels. *IEE Trans. Inf. Th.*

S	9	8	7	6	5	4	3	2	1	0
-3.8	0.010182	010449	010722	011001	011286	011577	011875	012179	012489	012806
-3.7	013130	013460	013798	014142	014494	014853	015219	015593	015974	016363
-3.6	016760	017165	017577	017998	018427	018865	019311	019765	020229	020701
-3.5	021182	021672	022171	022680	023198	023726	024263	024810	025367	025934
-3.4	026511	027098	027696	028304	028922	029552	030192	030844	031506	032180
-3.3	032865	033561	034269	034989	035720	036463	037219	037986	038766	039559
-3.2	040363	041181	042011	042854	043710	044579	045461	046356	047265	048188
-3.1	049124	050074	051037	052015	053006	054012	055032	056067	057116	058179
-3.0	059257	060350	061458	062581	063718	064871	066039	067223	068421	069636
-2.9	070865	072111	073372	074649	075942	077251	078576	079917	081274	082647
-2.8	084037	085443	086866	088305	089760	091232	092721	094227	095749	097288
-2.7	098844	100417	102006	103613	105237	106877	108535	110210	111902	113611
-2.6	115338	117081	118842	120620	122415	124227	126057	127904	129768	131649
-2.5	133547	135463	137396	139346	141314	143298	145300	147319	149354	151407
-2.4	153477	155564	157668	159789	161927	164081	166252	168440	170645	172866
-2.3	175104	177358	179629	181916	184220	186539	188875	191227	193594	195978
-2.2	198377	200793	203223	205670	208131	210608	213100	215608	218130	220667
-2.1	223219	225786	228367	230962	233572	236196	238834	241486	244152	246831
-2.0	249524	252231	254950	257683	260428	263186	265957	268740	271536	274344
-1.9	277163	279995	282838	285693	288559	291436	294324	297223	300133	303053
-1.8	305983	308924	311874	314835	317804	320783	323772	326769	329775	332790
-1.7	335813	338844	341884	344931	347986	351048	354118	357194	360278	363368
-1.6	366464	369567	372675	375790	378910	382036	385166	388302	391443	394588
-1.5	397737	100891	404048	407209	410374	413542	416713	419887	423064	426243
-1.4	429425	432608	435793	438980	442169	445358	448549	451740	454932	458124
-1.3	461317	464509	467701	470893	474083	477274	480462	483650	486836	490021
-1.2	493203	496384	499562	502738	505911	509081	512248	515411	518572	521728
-1.1	524881	528030	531174	534314	537449	540580	543705	546826	549941	553050
-1.0	556154	559252	562344	565429	568508	571581	574647	577706	580757	583802
-0.9	586839	589869	592891	595905	598911	601908	604898	607879	610851	613814
-0.8	616769	619714	622650	625577	628494	631402	634300	637188	640066	642934
-0.7	645791	648639	651475	654301	657116	659921	662714	665496	668267	671027
-0.6	673775	676512	679237	681951	684652	687342	690020	692685	695339	697980
-0.5	700609	703225	705829	708420	710999	713565	716118	718658	721185	723699
-0.4	726200	728688	731163	733624	736072	738507	740928	743336	745730	0748111
-0.3	750478	752832	755171	757497	75981	762108	764393	766663	76892	771163
-0.2	773392	775607	777808	779995	782168	784327	786472	788603	790719	792822
-0.1	794910	796985	799045	801091	803123	805141	807145	809134	811110	813071
-0.0	815019	816952	818871	820776	822668	824545	826408	828257	830092	831913

TABLE 2. Values of the TW_1 cumulative distribution function for the negative values of the argument.

S	0	1	2	3	4	5	6	7	8	9
0.0	0.831913	833720	835513	837293	839058	840810	842548	844272	845982	847679
0.1	849362	851031	852687	854329	855958	857573	859175	860763	862338	863899
0.2	865448	866983	868505	870013	871509	872992	874461	875918	877362	878793
0.3	880211	881616	883009	884389	885756	887111	888454	889784	891102	892407
0.4	893700	894981	896250	897507	898752	899985	901206	902416	903613	904799
0.5	905974	907137	908288	909428	910557	911674	912780	913875	914959	916032
0.6	917094	918146	919186	920216	921235	922243	923241	924229	925206	926173
0.7	927129	928076	929012	929938	930855	931761	932658	933545	934422	935290
0.8	936148	936997	937836	938666	939487	940299	941101	941895	942679	943455
0.9	944222	944980	945730	946471	947203	947927	948643	949350	950049	950740
1.0	951423	952097	952764	953423	954074	954717	955353	955981	956602	957215
1.1	957820	958418	959009	959593	960170	960739	961302	961858	962406	962948
1.2	963483	964012	964534	965049	965558	966061	966557	967046	967530	968007
1.3	968479	968944	969403	969856	970304	970746	971181	971612	972036	972455
1.4	972869	973277	973680	974077	974469	974856	975238	975615	975986	976353
1.5	976714	977071	977423	977770	978113	978450	978784	979112	979436	979756
1.6	980071	980382	980688	980991	981289	981582	981872	982158	982440	982717
1.7	982991	983261	983527	983789	984048	984303	984554	984802	985046	985286
1.8	985523	985757	985987	986214	986438	986658	986875	987089	987300	987507
1.9	987712	987914	988112	988308	988501	988690	988877	989062	989243	989422
2.0	989598	989771	989942	990110	990276	990439	990599	990758	990913	991067
2.1	991218	991366	991513	991657	991799	991938	992076	992211	992344	992476
2.2	992605	992732	992857	992980	993101	993220	993338	993453	993567	993679
2.3	993789	993897	994004	994109	994212	994313	994413	994511	994608	994703
2.4	994797	994889	994979	995068	995156	995242	995327	995410	995492	995573
2.5	995652	995730	995807	995882	995956	996029	996101	996172	996241	996309
2.6	996376	996442	996507	996570	996633	996695	996755	996815	996873	996931
2.7	996987	997043	997097	997151	997203	997255	997306	997356	997405	997454
2.8	997501	997548	997594	997639	997683	997726	997769	997811	997852	997893
2.9	997932	997972	998010	998048	998085	998121	998157	998192	998226	998260
3.0	998294	998326	998358	998390	998421	998451	998481	998510	998539	998567
3.1	998595	998622	998649	998675	998701	998726	998751	998775	998799	998823
3.2	998846	998868	998891	998912	998934	998955	998975	998996	999015	999035
3.3	999054	999073	999091	999109	999127	999144	999162	999178	999195	999211
3.4	999227	999242	999257	999272	999287	999301	999315	999329	999343	999356
3.5	999369	999382	999394	999407	999419	999431	999442	999454	999465	999476

TABLE 3. Values of the TW_1 cumulative distribution function for the positive values of the argument.

APPENDIX A. TRACY-WIDOM DISTRIBUTIONS: STATISTICAL TABLES

SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES, HERIOT-WATT UNIVERSITY, RICCAR-
TON, EDINBURGH EH14 4AS, UNITED KINGDOM

E-mail address: A.I.Bejan@ma.hw.ac.uk

S	9	8	7	6	5	4	3	2	1	0
-3.8	0.005479	005691	005910	006137	006370	006612	006861	007119	007384	007659
-3.7	007941	008233	008534	008844	009164	009494	009833	010183	010544	010915
-3.6	011297	011691	012096	012513	012941	013383	013836	014303	014782	015275
-3.5	015782	016302	016837	017386	017950	018529	019123	019733	020358	021000
-3.4	021659	022334	023026	023736	024463	025209	025972	026755	027556	028376
-3.3	029216	030076	030956	031856	032777	033719	034683	035668	036675	037704
-3.2	038756	039831	040929	042050	043195	044364	045558	046776	048019	049287
-3.1	050581	051901	053247	054619	056017	057443	058895	060375	061883	063419
-3.0	064982	066574	068195	069845	071524	073232	074969	076736	078534	080361
-2.9	082219	084107	086026	087976	089957	091969	094013	096088	098194	100333
-2.8	102503	104706	106940	109207	111506	113838	116202	118599	121028	123490
-2.7	125984	128511	131071	133664	136290	138948	141639	144363	147120	149909
-2.6	152730	155585	158471	161390	164342	167325	170341	173389	176468	179579
-2.5	182721	185895	189100	192336	195603	198901	202229	205587	208975	212392
-2.4	215839	219316	222821	226355	229917	233507	237125	240771	244443	248142
-2.3	251868	255620	259397	263200	267028	270880	274756	278657	282580	286527
-2.2	290496	294487	298500	302534	306589	310664	314759	318873	323006	327157
-2.1	331327	335513	339717	343937	348173	352424	356691	360971	365265	369573
-2.0	373893	378225	382569	386924	391290	395665	400050	404444	408846	413256
-1.9	417673	422097	426527	430962	435402	439847	444295	448747	453201	457657
-1.8	462115	466574	471033	475492	479950	484407	488862	493315	497765	502211
-1.7	506654	511091	515524	519951	524372	528785	533192	537591	541982	546363
-1.6	550736	555099	559451	563792	568123	572441	576747	581040	585320	589587
-1.5	593839	598077	602300	606507	610698	614873	619031	623172	627296	631401
-1.4	635489	639557	643607	647637	651647	655637	659606	663554	667482	671388
-1.3	675272	679133	682973	686789	690583	694353	698100	701823	705522	709196
-1.2	712846	716471	720072	723646	727196	730720	734218	737690	741136	744555
-1.1	747948	751315	754654	757967	761252	764511	767742	770946	774122	777271
-1.0	780392	783485	786550	789588	792597	795579	798533	801458	804356	807225
-0.9	810066	812880	815665	818422	821150	823851	826524	829168	831785	834374
-0.8	836935	839467	841973	844450	846900	849322	851716	854084	856423	858736
-0.7	861021	863280	865511	867716	869894	872045	874170	876268	878340	880387
-0.6	882407	884401	886370	888313	890231	892124	893992	895834	897652	899446
-0.5	901215	902960	904681	906378	908052	909702	911328	912932	914512	916070
-0.4	917605	919118	920609	922078	923524	924950	926354	927737	929098	930439
-0.3	931760	933060	934340	935599	936840	938060	939261	940443	941606	942751
-0.2	943876	944984	946073	947145	948198	949235	950253	951255	952240	953208
-0.1	954160	955095	956015	956918	957806	958678	959535	960377	961204	962016
-0.0	962814	963598	964367	965123	965865	966593	967308	968010	968699	969375

TABLE 4. Values of the TW_2 cumulative distribution function for the negative values of the argument.

S	0	1	2	3	4	5	6	7	8	9
0.0	0.969375	970038	970689	971328	971955	972570	973173	973764	974345	974914
0.1	975472	976019	976556	977082	977598	978103	978599	979085	979561	980027
0.2	980485	980933	981371	981801	982223	982635	983039	983435	983823	984202
0.3	984574	984938	985294	985643	985984	986318	986645	986965	987278	987585
0.4	987885	988178	988465	988746	989020	989289	989551	989808	990059	990305
0.5	990545	990780	991009	991234	991453	991667	991877	992081	992281	992477
0.6	992668	992854	993036	993214	993388	993558	993724	993886	994044	994198
0.7	994349	994496	994640	994780	994917	995050	995181	995308	995432	995553
0.8	995671	995787	995899	996009	996116	996220	996322	996421	996518	996612
0.9	996704	996794	996881	996967	997050	997131	997210	997286	997361	997434
1.0	997506	997575	997642	997708	997772	997835	997896	997955	998012	998069
1.1	998123	998177	998228	998279	998328	998376	998422	998468	998512	998554
1.2	998596	998637	998676	998715	998752	998789	998824	998858	998892	998924
1.3	998956	998987	999017	999046	999074	999102	999128	999154	999180	999204
1.4	999228	999251	999274	999296	999317	999338	999358	999377	999396	999414
1.5	999432	999450	999467	999483	999499	999514	999529	999544	999558	999572
1.6	999585	999598	999610	999622	999634	999646	999657	999668	999678	999688
1.7	999698	999708	999717	999726	999735	999743	999751	999759	999767	999774
1.8	999782	999789	999795	999802	999808	999815	999821	999827	999832	999838
1.9	999843	999848	999853	999858	999863	999867	999871	999876	999880	999884
2.0	999888	999891	999895	999898	999902	999905	999908	999911	999914	999917
2.1	999920	999923	999925	999928	999930	999933	999935	999937	999939	999941
2.2	999943	999945	999947	999949	999951	999952	999954	999956	999957	999959
2.3	999960	999961	999963	999964	999965	999967	999968	999969	999970	999971
2.4	999972	999973	999974	999975	999976	999977	999977	999978	999979	999980

TABLE 5. Values of the TW_2 cumulative distribution function for the positive values of the argument.

S	9	8	7	6	5	4	3	2	1	0
-3.9	0.006650	006950	007263	007588	007925	008275	008638	009015	009406	009812
-3.8	010232	010668	011119	011587	012071	012572	013091	013628	014184	014758
-3.7	015352	015966	016600	017255	017932	018631	019352	020096	020864	021655
-3.6	022472	023313	024181	025074	025994	026942	027918	028922	029955	031018
-3.5	032110	033234	034389	035576	036795	038047	039332	040652	042006	043396
-3.4	044821	046283	047781	049317	050891	052503	054154	055844	057575	059346
-3.3	061158	063012	064907	066845	068826	070850	072918	075030	077186	079388
-3.2	081635	083928	086267	088653	091085	093564	096091	098666	101289	103960
-3.1	106679	109448	112265	115131	118047	121012	124027	127092	130206	133370
-3.0	136584	139848	143162	146526	149940	153403	156916	160479	164092	167754
-2.9	171465	175225	179034	182892	186798	190752	194754	198803	202899	207042
-2.8	211231	215466	219746	224071	228441	232854	237312	241811	246354	250937
-2.7	255562	260228	264933	269677	274459	279279	284136	289029	293958	298921
-2.6	303918	308947	314009	319102	324225	329378	334559	339767	345003	350263
-2.5	355549	360858	366191	371544	376919	382313	387726	393156	398603	404066
-2.4	409543	415033	420536	426050	431574	437127	442649	448197	453751	459309
-2.3	464871	470436	476002	481569	487134	492698	498259	503816	509368	514914
-2.2	520452	525983	531504	537015	542514	548001	553475	558934	564378	569805
-2.1	575215	580607	585980	591332	596664	601973	607260	612523	617761	622974
-2.0	628160	633319	638451	643554	648627	653670	658683	663663	668612	673527
-1.9	678409	683256	688069	692846	697587	702291	706958	711588	716179	720731
-1.8	725245	729718	734152	738545	742898	747209	751478	755706	759892	764035
-1.7	768135	772193	776207	780177	784105	787988	791827	795622	799373	803079
-1.6	806741	810359	813932	817460	820944	824383	827777	831127	834432	837693
-1.5	840909	844081	847209	850293	853332	856328	859280	862189	865054	867876
-1.4	870655	873391	876085	878737	881346	883914	886440	888926	891370	893773
-1.3	896136	898460	900743	902987	905192	907359	909487	911577	913630	915645
-1.2	917623	919565	921471	923341	925176	926975	928741	930472	932169	933833
-1.1	935464	937063	938629	940164	941667	943140	944582	945995	947377	948731
-1.0	950055	951352	952620	953861	955075	956263	957424	958559	959669	960754
-0.9	961814	962850	963863	964852	965818	966761	967682	968582	969460	970317
-0.8	971154	971970	972766	973543	974301	975040	975761	976464	977149	977816
-0.7	978467	979101	979719	980321	980907	981478	982034	982576	983103	983616
-0.6	984115	984601	985074	985534	985981	986416	986840	987251	987651	988040
-0.5	988418	988785	989142	989489	989826	990153	990471	990779	991079	991370
-0.4	991652	991926	992192	992450	992701	992943	993179	993407	993629	993844
-0.3	994052	994254	994449	994639	994822	995000	995172	995339	995501	995658
-0.2	995809	995956	996098	996235	996368	996497	996622	996742	996858	996971
-0.1	997080	997185	997287	997386	997481	997573	997662	997748	997830	997911
-0.0	997988	998063	998135	998204	998272	998337	998399	998460	998518	998574

TABLE 6. Values of the TW_4 cumulative distribution function for the negative values of the argument.

S	0	1	2	3	4	5	6	7	8	9
0.0	0.998574	998629	998681	998731	998780	998827	998872	998916	998958	998999
0.1	999038	999075	999112	999146	999180	999212	999244	999274	999303	999330
0.2	999357	999383	999408	999432	999455	999477	999498	999519	999538	999557
0.3	999575	999593	999610	999626	999641	999656	999670	999684	999697	999710
0.4	999722	999734	999745	999756	999766	999776	999786	999795	999804	999812
0.5	999820	999828	999835	999842	999849	999856	999862	999868	999874	999879
0.6	999885	999890	999895	999899	999904	999908	999912	999916	999920	999923

TABLE 7. Values of the TW_4 cumulative distribution function for the positive values of the argument.