Eigenvalue distributions for some correlated complex sample covariance matrices

2007 J. Phys. A: Math. Theor. 40 11093

(http://iopscience.iop.org/1751-8121/40/36/009)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 11093–11103 doi:10.1088/1751-8113/40/36/009

Eigenvalue distributions for some correlated complex sample covariance matrices

P J Forrester

Department of Mathematics and Statistics, University of Melbourne, Victoria 3010, Australia

Received 5 April 2007, in final form 1 August 2007
Published 21 August 2007
Online at stacks.iop.org/JPhysA/40/11093

Abstract
The distributions of the smallest and largest eigenvalues for the matrix product $Z^\dagger Z$, where Z is an $n \times m$ complex Gaussian matrix with correlations both along rows and down columns, are expressed as $m \times m$ determinants. In the case of correlation along rows, these expressions are computationally more efficient for the purpose of tabulation than those involving sums over partitions and Schur polynomials reported recently for the same distributions.

PACS numbers: 02.10.Yn, 02.50.Sk
Mathematics Subject Classification: 15A52, 62HXX

1. Introduction

A typical setting in multivariate statistics is to measure each of m variables $x_1, \dots, x_m$ a total of N times. For example, the variable x_k may denote the wind speed at weather station k at a specific time of day; recording the value on successive days gives a sequence of values $x_k^{(j)}$, $j = 1, \dots, N$ for the variable x_k, which forms a column vector $\vec{x}_k = \big[x_k^{(j)}\big]_{j=1,\dots,N}$. Collecting together the column vectors for each of the variables x_k gives the data matrix $X = [\vec{x}_k]_{k=1,\dots,m}$. Let the average of the readings of variable x_k be denoted $\bar{x}_k$, so that

$$\bar{x}_k = \frac{1}{N}\sum_{j=1}^{N} x_k^{(j)}.$$

Let $\vec{\bar{x}}_k = [\bar{x}_k]_{j=1,\dots,N}$ be the corresponding (constant) column vector, and set $\bar{X} := [\vec{\bar{x}}_k]_{k=1,\dots,m}$. Forming now

$$\frac{1}{n}A := \frac{1}{n}(X - \bar{X})^T (X - \bar{X}) = \left[\frac{1}{n}\sum_{j=1}^{N}\big(x_{k_1}^{(j)} - \bar{x}_{k_1}^{(j)}\big)\big(x_{k_2}^{(j)} - \bar{x}_{k_2}^{(j)}\big)\right]_{k_1,k_2=1,\dots,m},$$

$n = N - 1$, gives an empirical approximation to the covariance matrix $[\langle (x_{k_1} - \langle x_{k_1}\rangle)(x_{k_2} - \langle x_{k_2}\rangle)\rangle]_{k_1,k_2=1,\dots,m}$ for the variables $\{x_k\}_{k=1,\dots,m}$.

Analytic studies of the matrix A can be carried out in the case that the variables $x_1, \ldots, x_m$ relating to the data matrix X are chosen from a multivariate Gaussian distribution with variance matrix Σ and mean $\vec{\mu}$. Then it is known (see, e.g., [11]) that the distribution of A is the same as that for the matrix product $Y^T Y$, where Y is an $m \times n, n = N - 1$, Gaussian matrix in which each row is drawn from a multivariate Gaussian distribution with covariance matrix Σ and mean zero. Thus the joint probability density function (pdf) of the elements of Y is

$$\frac{1}{C} \, \mathrm{e}^{-\mathrm{Tr}(\Sigma^{-1} Y^T Y/2)}, \tag{1.1}$$

where here and throughout (unless otherwise stated) C represents *some* constant (i.e. quantity independent of the main variables of the equation, which here are the elements of Y).

A standard practice in studying the empirical covariance matrix is to form the eigenvalue–eigenvector decomposition. This comes under the name of principal component analysis (see, e.g., [13]). On a theoretical front one seeks analytic forms for eigenvalue distributions of the matrix $A = Y^T Y$ when Y is distributed according to (1.1). In fact the eigenvalue pdf can be written down in terms of a multivariable generalized hypergeometric function based on zonal polynomials (see, e.g., [15]). This function is inherently difficult to compute, but there have been some recent advances [14]. It is also possible to integrate over this pdf to express the distribution of the largest eigenvalue as another generalized hypergeometric function [12].

In a recent work [17] a study of the pdf for the smallest and largest eigenvalues of the matrix $A = Z^\dagger Z$ for Z an $n \times m$, $(n \geqslant m)$ complex Gaussian matrix with pdf

$$\frac{1}{C} \, \mathrm{e}^{-\mathrm{Tr}(\Sigma^{-1} Z^\dagger Z)} \tag{1.2}$$

has been undertaken. Earlier studies had considered these pdfs in the case $\Sigma = I$ [8, 9], while the case of Σ equal to the identity plus a finite rank matrix has received much recent attention [1, 4, 6]. The setting of complex data matrices is of great importance in recent quantitative studies of wireless communication (see, e.g., [19, 20]). A significant feature of the pdf (1.2) is that the corresponding joint eigenvalue pdf of A can be written as a determinant [1, 2, 10, 19]. Moreover, as to be shown in section 2 below, the marginal distributions by way of the pdf of the smallest and largest eigenvalues can also be evaluated as determinants. In relation to the largest eigenvalue this is given by substituting (2.12) in (2.5), while for the smallest eigenvalue it is given by substituting (2.16) in (2.13). In contrast, these same distributions where evaluated in [17] as a sum over Schur polynomials and as a generalized hypergeometric function based on the Schur polynomials respectively (see (2.17) and (2.10) below).

Suppose more generally that the complex data matrix Z has pdf

$$\frac{1}{C} \, \mathrm{e}^{-\mathrm{Tr}(\Sigma_1^{-1} Z^\dagger \Sigma_2^{-1} Z)}. \tag{1.3}$$

Here Σ_2 can be interpreted as the covariance coupling the measurements of a single variable z_k. Very recently [18, 19], it has been shown that for this distribution the canonical average

$$\langle \det(1 + u Z^\dagger Z)^p \rangle \tag{1.4}$$

can be expressed as an $n \times n$ determinant, even though the joint eigenvalue pdf of (1.3) cannot itself be written in a determinant form. In section 3 we will use the method of [19] to similarly express the pdf for the smallest and largest eigenvalues of $Z^\dagger Z$ with Z distributed as (1.3) in terms of determinants. In the case $m = n$, the pdf for the smallest eigenvalue is given by substituting (3.9) in (2.13), while the pdf of the largest eigenvalue is given by substituting (3.11) in (2.5).

2. Case of a single covariance matrix

2.1. Correlation across rows of Z

Consider the pdf (1.2). Introduce the singular value decomposition

$$Z = U\mathrm{diag}(\mu_1, \ldots, \mu_m)V, \tag{2.1}$$

where U (V) is a $m \times m$ $(n \times n)$ unitary matrix and the $\mu_j^2 =: \lambda_j$ are the eigenvalues of the positive definite matrix $Z^\dagger Z$.

We seek the joint distribution of the $\{\lambda_j\}_{j=1,\ldots,m}$, $p(\lambda_1, \ldots, \lambda_m)$ say. Firstly, with $A = Z^\dagger Z$, we know (see, e.g., [7])

$$\begin{aligned}(dA) &= \frac{1}{C}\det(Z^\dagger Z)^{n-m}(\mathrm{d}Z)\\ &= \frac{1}{C}\prod_{j=1}^m \lambda_j^{n-m}\prod_{1\leqslant j<k\leqslant m}(\lambda_k-\lambda_j)^2\,\mathrm{d}\lambda_1\cdots\mathrm{d}\lambda_m(V^\dagger\,\mathrm{d}Z),\end{aligned} \tag{2.2}$$

where $(V^\dagger\,\mathrm{d}Z)$ is the Haar measure (uniform distribution) on the space of $m \times m$ unitary matrices $U(m)$. Thus

$$p(\lambda_1, \ldots, \lambda_m) = \frac{1}{C}\prod_{j=1}^m \lambda_j^{n-m}\prod_{1\leqslant j<k\leqslant m}(\lambda_k-\lambda_j)^2\int_{V\in U(m)} \mathrm{e}^{-\mathrm{Tr}(\Sigma^{-1}V^\dagger\mathrm{diag}(\lambda_1,\ldots,\lambda_m)V)}(V^\dagger\,\mathrm{d}Z). \tag{2.3}$$

This is the well known Harish–Chandra/Itzykson–Zuber matrix integral (see, e.g., [16]). It has a closed form determinantal evaluation, which when substituted in (2.3), implies

$$p(\lambda_1, \ldots, \lambda_m) = \frac{1}{C}\prod_{j=1}^m \lambda_j^{n-m}\prod_{1\leqslant j<k\leqslant m}\frac{(\lambda_k-\lambda_j)}{(s_k-s_j)}\det[\mathrm{e}^{-s_j\lambda_k}]_{j,k=1,\ldots,m} \tag{2.4}$$

where $\{s_1, \ldots, s_m\}$ are the eigenvalues of Σ^{-1}. As referenced in the third sentence below (1.2), the result (2.4) has been made explicit in a number of recent works.

Consider now the probability $E((\lambda,\infty))$ that the interval (λ,∞) is free of eigenvalues. This is related to the pdf of the largest eigenvalue, $p^{\max}(\lambda)$ say, by

$$p^{\max}(\lambda) = -\frac{\mathrm{d}}{\mathrm{d}\lambda}E((\lambda,\infty)). \tag{2.5}$$

We have

$$\begin{aligned}E((\lambda,\infty)) &:= \int_0^\lambda \mathrm{d}\lambda_1\cdots\int_0^\lambda \mathrm{d}\lambda_m\, p(\lambda_1,\ldots,\lambda_m)\\ &= \frac{1}{C}\frac{1}{\prod_{j<k}^m(s_k-s_j)}\int_0^\lambda \mathrm{d}\lambda_1\cdots\int_0^\lambda \mathrm{d}\lambda_m\prod_{j=1}^m\lambda_j^{n-m}\prod_{j<k}^m(\lambda_k-\lambda_j)\det[\mathrm{e}^{-s_j\lambda_k}]_{j,k=1,\ldots,m}.\end{aligned}$$

Because both factors in the integrand are anti-symmetric in $\{\lambda_j\}_{j=1,\ldots,m}$, and noting too that with Asym denoting the operation of anti-symmetrization

$$\prod_{j<k}^m(\lambda_k-\lambda_j) = \mathrm{Asym}\,\lambda_1^0\lambda_2\cdots\lambda_m^{m-1},$$

the product can be replaced by $\lambda_1^0\lambda_2\cdots\lambda_m^{m-1}$ provided we multiply by $m!$. Doing this we see the integrations over $\{\lambda_k\}$ can be performed column by column, to give

$$E((\lambda,\infty)) = \frac{m!}{C}\frac{1}{\prod_{j<k}^m(s_k-s_j)}\det\left[\int_0^\lambda t^{n-m+k-1}\,\mathrm{e}^{-s_jt}\,\mathrm{d}t\right]_{j,k=1,\ldots,m}. \tag{2.6}$$

To evaluate C, we note $\lim_{\lambda\to\infty} E((\lambda,\infty)) = 1$. The integral in (2.6) can be evaluated in this limit to give

$$1 = \frac{m!}{C} \frac{1}{\prod_{j<k}^m (s_k - s_j)} \det\left[s_j^{-(n-m+k)}(n-m+k-1)!\right]_{j,k=1,\ldots,m}.$$

Factoring the factorials from the determinant and then making use of the Vandermonde determinant formula shows

$$C = (-1)^{m(m-1)/2} m! \prod_{k=1}^m (n-m+k-1)! \prod_{j=1}^m s_j^{-n}. \tag{2.7}$$

Substituting this in (2.6), and changing variables $s_j \mapsto \lambda s_j$ in the integral therein, we obtain for our final expression

$$E((\lambda,\infty)) = \frac{1}{\prod_{k=1}^m (n-m+k-1)!} \frac{\prod_{j=1}^m (\lambda s_j)^n}{\prod_{j<k}^m (-\lambda)(s_k - s_j)} \times \det\left[\int_0^1 t^{n-m+k-1} \mathrm{e}^{-\lambda s_j t}\, \mathrm{d}t\right]_{j,k=1,\ldots,m}. \tag{2.8}$$

We remark that in the case that $s_j = 1 (j = r+1,\ldots,m)$, the $m \to \infty$ limit of $E((\lambda,\infty))$, with $\lambda, s_1, \ldots, s_r$ appropriately scaled, is studied in [1]. We remark too that in [17, Corollary 3.3] $E((\lambda,\infty))$ is expressed in terms of the generalized multi-variable hypergeometric function

$${}_1F_1(a,b;x_1,\ldots,x_m) := \sum_\kappa \frac{[a]_\kappa}{d'_\kappa [b]_\kappa} s_\kappa(x_1,\ldots,x_m). \tag{2.9}$$

In (2.9) s_κ denotes the Schur polynomial labelled by a partition $\kappa = (\kappa_1,\ldots,\kappa_m)$, $\kappa_1 \geqslant \cdots \geqslant \kappa_m$,

$$[a]_\kappa := \prod_{j=1}^m \frac{\Gamma(a-j+1+\kappa_j)}{\Gamma(a-j+1)},$$

while

$$d'_\kappa = \frac{[m]_\kappa}{\bar{f}_m(\kappa)}, \qquad \bar{f}_m(\kappa) := \prod_{1\leqslant i<j\leqslant m} \frac{(j-i+\kappa_i-\kappa_j)}{j-i}.$$

Thus from [17, equation (3.5)]

$$E((\lambda,\infty)) = \prod_{k=1}^m \frac{\Gamma(k)}{\Gamma(n+k)} \prod_{j=1}^m (\lambda s_j)^n {}_1F_1(n; n+m; -\lambda s_1, \ldots, -\lambda s_m). \tag{2.10}$$

Comparing (2.10) and (2.6) gives the determinant formula

$$F_1(n; n+m; x_1,\ldots,x_m) = \prod_{k=1}^m \frac{\Gamma(n+k)}{\Gamma(k)\Gamma(n-m+k)} \frac{1}{\prod_{j<k}^m (x_k - x_j)} \times \det\left[\int_0^1 t^{n-m+k-1} \mathrm{e}^{x_j t}\, \mathrm{d}t\right]_{j,k=1,\ldots,m}. \tag{2.11}$$

The integral in (2.8) is itself a special case of a one variable confluent hypergeometric function ${}_1F_1$, allowing us to write

$$E((\lambda,\infty)) = \frac{1}{\prod_{k=1}^m (n-m+k)!} \frac{\prod_{j=1}^m (\lambda s_j)^n}{\prod_{j<k}^m (-\lambda)(s_k - s_j)} \times \det[{}_1F_1(n-m+k; n-m+k+1; -\lambda s_j)]_{j,k=1,\ldots,m}. \tag{2.12}$$

Note that (2.12) and (2.10) are identical in the case $m = 1$.

We draw attention to a limiting feature of (2.6) which is of relevance in the study of $E((\lambda,\infty))$ for fully correlated matrices (1.3). Suppose then that $n=m$ in (2.6), and consider the limit $s_n \to \infty$. Integrating the final row of integrals by parts, we see that the dominant term is that in the first column. Expanding by this term shows

$$\lim_{s_n\to\infty} E((\lambda,\infty))|_{m=n} = E((\lambda,\infty))|_{m=n-1},$$

and iterating this we have

$$\lim_{s_{n-m},\ldots,s_n\to\infty} E((\lambda,\infty))|_{m=n} = E((\lambda,\infty))$$

where on the right-hand side $E((\lambda,\infty))$ is for the general m case, as given by (2.8). The understanding of this result is that with $s_{n-m},\ldots,s_n \to \infty$, the final $n-m$ rows of Z become zero and so the eigenvalues of $Z^\dagger Z$ are those of $W^\dagger W$ for W the restriction of Z to its first m rows, together with m zero eigenvalues.

For the probability $E((0,\lambda))$ that the interval $(0,\lambda)$ is free of eigenvalues, related to the pdf of the smallest eigenvalue, $p^{\min}(\lambda)$ say, by

$$p^{\min}(\lambda) = \frac{d}{d\lambda}E((0,\lambda)), \tag{2.13}$$

we have

$$\begin{aligned}E((0,\lambda)) &:= \int_\lambda^\infty d\lambda_1\cdots\int_\lambda^\infty d\lambda_m\, p(\lambda_1,\ldots,\lambda_m)\\ &= \frac{1}{C}\frac{1}{\prod_{j<k}(s_k-s_j)}\int_\lambda^\infty d\lambda_1\cdots d\lambda_m \prod_{j=1}^m \lambda_j^{n-m}\prod_{1\leqslant j<k\leqslant m}(\lambda_k-\lambda_j)\det[e^{-s_j\lambda_k}]_{j,k=1,\ldots,m}.\end{aligned}$$

Proceeding now as in the derivation of (2.8) shows

$$\begin{aligned}E((0,\lambda)) &= \frac{m!}{C}\frac{1}{\prod_{j<k}^m(s_k-s_j)}\det\left[\int_\lambda^\infty t^{n-m+k-1}\,e^{-s_jt}\,dt\right]_{j,k=1,\ldots,m}\\ &= \frac{m!}{C}\frac{e^{-\lambda\sum_{j=1}^m s_j}}{\prod_{j<k}^m(s_k-s_j)}\det\left[\int_0^\infty (t+\lambda)^{n-m+k-1}\,e^{-s_jt}\,dt\right]_{j,k=1,\ldots,m}\end{aligned} \tag{2.14}$$

where C is given by (2.7). Note that for $m=n$ the determinant only contributes a constant (i.e. term independent of λ) and we have

$$E((0,\lambda))|_{m=n} = e^{-\lambda\sum_{j=1}^n s_j}, \tag{2.15}$$

which generalizes the same result known for $s_1=\cdots=s_n=1$ [5, 8].

In terms of the confluent hypergeometric function

$$U(a,b,z) := \frac{1}{\Gamma(a)}\int_0^\infty e^{-zt}t^{a-1}(1+t)^{b-a-1}\,dt,$$

which, unlike ${}_1F_1(a,b;z)$, is singular at $z=0$, but like ${}_1F_1$ is available as an inbuilt function on a number of mathematical computing packages, (2.14) reads

$$\begin{aligned}E((0,\lambda)) &= (-1)^{m(m-1)/2}\prod_{k=1}^m\frac{(\lambda s_k)^n}{\Gamma(n-m+k)}\frac{e^{-\lambda\sum_{j=1}^m s_j}}{\prod_{j<k}^m(s_k-s_j)}\\ &\quad\times\det[U(1,n-m+k+1,\lambda s_j)]_{j,k=1,\ldots,m}.\end{aligned} \tag{2.16}$$

The computationally more complex evaluation (due to the sum over partitions, which from an asymptotic formula of Hardy and Ramanujan for large k involves of order $e^{\pi\sqrt{2k/3}}$ terms)

$$E((0,\lambda)) = e^{-\lambda\sum_{j=1}^m s_j}\sum_{k=0}^{m(n-m)}\lambda^k\sum_{\substack{\kappa:|\kappa|=k\\ \kappa_1\leqslant n-m}}\frac{s_\kappa(s_1,\ldots,s_m)}{d'_\kappa} \tag{2.17}$$

is given in [17, equation (3.8)].

Also given in [17] is an expression involving sums over partitions and Schur polynomials for the probability $E((0,a)\cup(b,\infty))$ that there are no eigenvalues in either of the intervals $(0,a)$ or (b,∞). In terms of this quantity the joint pdf for the smallest and largest eigenvalues, $p(a,b)$ say, is given by

$$p(a,b) = -\frac{\partial^2}{\partial a\partial b}E((0,a)\cup(b,\infty)).$$

From $p(a,b)$ one can deduce the distribution of b/a, which is the square of the condition number and so is of relevance in numerical analysis. The method of derivation of (2.8) yields the determinant evaluation

$$\begin{aligned} E((0,a)\cup(b,\infty)) &= \frac{1}{C}\frac{1}{\prod_{j<k}^m(s_k-s_j)}\int_a^b \mathrm{d}\lambda_1\cdots\int_a^b \mathrm{d}\lambda_m \prod_{j=1}^m \lambda_j^{n-m}\prod_{j<k}^m(\lambda_k-\lambda_j)\det[\mathrm{e}^{-s_j\lambda_k}]_{j,k=1,\dots,m} \\ &= \frac{m!}{C}\frac{1}{\prod_{j<k}^m(s_k-s_j)}\det\left[\int_a^b t^{n-m+k-1}\,\mathrm{e}^{-s_j t}\,\mathrm{d}t\right]_{j,k=1,\dots,m} \end{aligned} \tag{2.18}$$

where C is given by (2.7).

2.2. *Correlation down columns of Z*

Consider next the case that Z has distribution

$$\frac{1}{C}\,\mathrm{e}^{-\mathrm{Tr}(Z^\dagger\Sigma_2^{-1}Z)} \tag{2.19}$$

((1.3) with $\Sigma_1 = I$). Now

$$\mathrm{e}^{-\mathrm{Tr}(Z^\dagger\Sigma_2^{-1}Z)} = \mathrm{e}^{-\mathrm{Tr}(\Sigma_2^{-1}ZZ^\dagger)} \tag{2.20}$$

and the non-zero eigenvalues of $ZZ^\dagger$ agree with the eigenvalues of $Z^\dagger Z$, which we again denote $\{\lambda_j\}_{j=1,\dots,m}$. With $A = ZZ^\dagger$ (2.2) again applies but with V replaced by U (recall (2.1)), and thus

$$\begin{aligned} p(\lambda_1,\dots,\lambda_m) &= \frac{1}{C}\prod_{j=1}^m \lambda_j^{n-m}\prod_{1\leqslant j<k\leqslant m}(\lambda_k-\lambda_j)^2 \lim_{\lambda_{m+1},\dots,\lambda_n\to 0} \\ &\quad\times \int_{U\in U(n)} \mathrm{e}^{-\mathrm{Tr}(\Sigma_2^{-1}U\mathrm{diag}(\lambda_1,\dots,\lambda_n)U^\dagger)}(U^\dagger\,\mathrm{d}U). \end{aligned} \tag{2.21}$$

(Here the limit could have been taken immediately, but as noted in [19] there are computational advantages in delaying this step.) Proceeding as in the derivation of (2.4) gives

$$p(\lambda_1,\dots,\lambda_m) = \frac{1}{C}\lim_{\lambda_{m+1},\dots,\lambda_n\to 0}\frac{\prod_{j=1}^m\lambda_j^{n-m}\prod_{j<k}^m(\lambda_k-\lambda_j)^2}{\prod_{j<k}^n(\lambda_k-\lambda_j)(s_k-s_j)}\det[\mathrm{e}^{-s_j\lambda_k}]_{j,k=1,\dots,n}$$

where here $\{s_1,\dots,s_n\}$ denotes the eigenvalues of Σ_2^{-1}. Now taking the limit this reads

$$p(\lambda_1,\dots,\lambda_m) = \frac{1}{C}\frac{\prod_{j<k}^m(\lambda_k-\lambda_j)}{\prod_{j<k}^n(s_k-s_j)}\det\left[[\mathrm{e}^{-s_j\lambda_k}]_{\substack{j=1,\dots,n\\k=1,\dots,m}}\quad [s_j^{k-1}]_{\substack{j=1,\dots,n\\k=1,\dots,n-m}}\right]. \tag{2.22}$$

The joint pdf (2.22) has been derived previously in [3, 19].

The derivations of (2.8) and (2.14) can be applied to (2.22) to deduce determinant formulas for $E((\lambda,\infty))$ and $E((0,\lambda))$. Thus we find

$$E((\lambda,\infty)) = \frac{m!}{C}\frac{1}{\prod_{j<k}^n(s_k-s_j)}\det\left[\left[\int_0^\lambda t^{k-1}\,\mathrm{e}^{-s_j t}\,\mathrm{d}t\right]_{\substack{j=1,\dots,n\\k=1,\dots,m}}\quad [s_j^{k-1}]_{\substack{j=1,\dots,n\\k=1,\dots,n-m}}\right] \tag{2.23}$$

and

$$E((0,\lambda)) = \frac{\prod_{k=1}^{m} k!}{C} \frac{e^{-\lambda \sum_{j=1}^{n} s_j}}{\prod_{j<k}^{n}(s_k - s_j)} \det \left[[s_j^{-k}]_{\substack{j=1,\dots,n \\ k=1,\dots,m}} \quad [e^{\lambda s_j} s_j^{k-1}]_{\substack{j=1,\dots,n \\ k=1,\dots,n-m}} \right] \tag{2.24}$$

where

$$C = \prod_{k=1}^{m} k! \prod_{j=1}^{n} s_j^{-m}. \tag{2.25}$$

3. Fully correlated case

We turn our attention now to the case (1.3), in which the data matrix Z is correlated both across rows and down columns. Here it does not appear possible to write the joint eigenvalue pdf of $Z^\dagger Z$ in determinant form. Nonetheless, it has been shown recently by Simon and Moustakos [18] (see [19] for a detailed presentation) that it is possible to give a determinant formula for the average (1.4). Here we will show that their calculation can be adopted to give determinant formulas for $E((\lambda,\infty))$ and $E((0,\lambda))$ (the latter being restricted to the case $m = n$).

To begin we suppose $m = n$. In the case of $E((\lambda,\infty))$, by using the limiting procedure discussed in the paragraph below (2.11), a formula can be deduced from this for general $m \leqslant n$. Our starting point is the formula [19]

$$p(\lambda_1,\dots,\lambda_n) = \prod_{j=1}^{n}(r_j s_j)^n \prod_{j<k}^{n}(\lambda_k - \lambda_j)\mathcal{I}(\{r_i\},\{s_i\},\{\lambda_i\}) \tag{3.1}$$

where, with the eigenvalues of Σ_1^{-1}, Σ_2^{-1} given by $\{r_i\}$, $\{s_i\}$ respectively,

$$\mathcal{I}(\{r_i\},\{s_i\},\{\lambda_i\}) = \frac{1}{n!} \sum_{k_1>k_2>\cdots>k_n\geqslant 0} \prod_{j=1}^{n} \frac{(-1)^{k_j}}{k_j!} \frac{\det\left[r_j^{k_l}\right]\det\left[\lambda_j^{k_l}\right]\det\left[s_j^{k_l}\right]}{\prod_{j<l}^{n}(k_l-k_j)(r_l-r_j)(s_l-s_j)}. \tag{3.2}$$

In [19, lemma 5] it is proved that $\mathcal{I}$ is bounded by an exponentially decaying function in λ_j for each $j = 1,\dots,n$.

Consider first $E((0,\lambda))$. We thus seek to integrate each λ_i in (3.1) over (λ,∞). For this we take inspiration from [19] and note from the Vandermonde determinant evaluation

$$\det\left[z_j^{k-1}\right]_{j,k=1,\dots,n} = \prod_{j<k}^{n}(z_k - z_j)$$

with $z_j = \lambda_j/(\lambda_j - \lambda)$ that

$$\prod_{j<k}^{n}(\lambda_k-\lambda_j) = \prod_{j=1}^{n}\left(\frac{\lambda_j-\lambda}{\lambda^{1/2}}\right)^{n-1} \det\left[\left(\frac{\lambda_j}{\lambda_j-\lambda}\right)^{k-1}\right]_{j,k=1,\dots,n}. \tag{3.3}$$

Expanding out the determinant according to its definition, substituting in (3.1) and integrating gives

$$E((0,\lambda))|_{m=n} = \prod_{j=1}^{n}(r_j s_j)^n \lambda^{-n(n-1)/2} \sum_{P\in S_N} \varepsilon(P) \int_\lambda^\infty d\lambda_1 \cdots \int_\lambda^\infty d\lambda_N$$

$$\times \prod_{j=1}^{n}(\lambda_j-\lambda)^{n-P(j)}\lambda_j^{P(j)-1}\mathcal{I}(\{r_i\},\{s_i\},\{\lambda_i\}). \tag{3.4}$$

For each integration variable, we integrate by parts $P(j) - 1$ times, making use of the simple formula

$$(\lambda_j - \lambda)^m = \frac{1}{m+1} \frac{\partial}{\partial \lambda_j} (\lambda_j - \lambda)^{m+1}, \qquad m \neq -1.$$

For $m > -1$, $(\lambda_j - \lambda)^{m+1}$ vanishes at $\lambda_j = \lambda$ (this is part of the motivation for the manipulation (3.3)), while from the remark below (3.2) the factor involving $\mathcal{I}$ vanishes at $\lambda_j = \infty$. Hence in the integration by parts there is no contribution from the end points. We must compute the partial derivatives with respect to λ_j of $\lambda_j^{P(j)-1}\mathcal{I}$. For this note that the only term dependent of λ_j in $\mathcal{I}$ is $\det[\lambda_j^{k_l}]$, and

$$\prod_{j=1}^{n} \lambda_j^{P(j)-1} \det\left[\lambda_j^{k_l}\right] = \det\left[\lambda_j^{k_l+P(j)-1}\right].$$

Performing $P(j) - 1$ integration by parts in each variable λ_j thus gives

$$\begin{aligned} E((0,\lambda))|_{m=n} = {} & \prod_{j=1}^{n} (r_j s_j)^n \frac{\lambda^{-n(n-1)/2}}{n!} \sum_{P \in S_N} \varepsilon(P) \int_\lambda^\infty \mathrm{d}\lambda_1 \cdots \int_\lambda^\infty \mathrm{d}\lambda_N \\ & \times \prod_{j=1}^{n} (\lambda_j - \lambda)^{(n-1)} \sum_{k_1 > \cdots > k_n \geqslant 0} \prod_{j=1}^{n} \frac{(-1)^{k_j}}{k_j!} \\ & \times \frac{\det\left[r_j^{k_l}\right] \det\left[\prod_{p=1}^{P(j)-1}\left(-\frac{k_l+p}{n-p}\right)\lambda_j^{k_l}\right] \det\left[s_j^{k_l}\right]}{\prod_{j<l}^{n}(k_l - k_j)(r_l - r_j)(s_l - s_j)}. \end{aligned} \tag{3.5}$$

Next we want to integrate row by row in the determinant. Although the integrand decays exponentially at infinity, this gives divergent integrals, as a result of interchanging the order of summation and integration. To overcome this, write

$$\int_\lambda^\infty \mathrm{d}\lambda_j = \lim_{L\to\infty} \int_\lambda^L \mathrm{d}\lambda_j \quad (j = 1, \ldots, n),$$

and so interchange only the finite range integrals with the summation of $\{h_j\}$. We see the resulting one dimensional integrals are the same down each column of the determinant and so can be factored. Furthermore, the sum over $P \in S_N$ then simply interchanges rows in the determinant, which is compensated for by $\varepsilon(P)$, thus contributing an overall factor of $n!$. Hence

$$\begin{aligned} E((0,\lambda))|_{m=n} = {} & \prod_{j=1}^{n} (r_j s_j)^n \lambda^{-n(n-1)/2} \lim_{L\to\infty} \sum_{k_1 > \cdots > k_n \geqslant 0} \\ & \times \prod_{j=1}^{n} \frac{(-1)^{k_j}}{k_j!} \frac{\det\left[r_j^{k_l}\right] \det\left[s_j^{k_l}\right]}{\prod_{j<l}^{n}(k_l - k_j)(r_l - r_j)(s_l - s_j)} \\ & \times \prod_{l=1}^{n} \int_\lambda^L (t-\lambda)^{n-1} t^{k_l}\, \mathrm{d}t \det\left[\prod_{p=1}^{j-1}\left(-\frac{k_l+p}{n-p}\right)\right]_{j,l=1,\ldots,n}. \end{aligned} \tag{3.6}$$

As noted in [19], it is straightforward to verify that

$$\det\left[\prod_{p=1}^{j-1}\left(-\frac{k_l+p}{n-p}\right)\right]_{j,l=1,\ldots,n} = (-1)^{n(n-1)/2} \prod_{j=1}^{n-1} \frac{1}{j^j} \prod_{j<l}^{n} (k_l - k_j),$$

thus cancelling $\prod_{j<l}^{n}(k_l - k_j)$ and reducing (3.6) to

$$E((0,\lambda))|_{m=n} = \prod_{j=1}^{n}(r_j s_j)^n (-\lambda)^{-n(n-1)/2} \prod_{j=1}^{n-1} \frac{1}{j^j} \lim_{L\to\infty} \sum_{k_1>\cdots k_n\geqslant 0} \times \prod_{j=1}^{n}\left(\int_{\lambda}^{L}(t-\lambda)^{n-1}\frac{(-t)^{k_j}}{k_j!}\right) \frac{\det\left[r_j^{k_l}\right]\det\left[s_j^{k_l}\right]}{\prod_{j<l}^{n}(r_l - r_j)(s_l - s_j)} \tag{3.7}$$

The lattice version of the well known Heine formula from random matrix theory (see, e.g., [7]),

$$\int_I \mathrm{d}\mu(x_1)\cdots\int_I \mathrm{d}\mu(x_N)\det[\phi_j(x_k)]_{j,k=1,\ldots,N}\det[\psi_j(x_k)]_{j,k=1,\ldots,N} = N!\det\left[\int_I \phi_j(x)\psi_k(x)\,\mathrm{d}\mu(x)\right]_{j,k=1,\ldots,N}$$

namely

$$\sum_{k_1>\cdots>k_n\geqslant 0}\det\left[a_i^{k_j}\right]\det\left[b_i^{k_j}\right]\prod_{i=1}^{n}w(k_i) = \det\left[\sum_{p=0}^{\infty}w(p)(a_i b_j)^p\right]_{i,j=1,\ldots,n},$$

referred to in [19] as the Cauchy–Binet formula, allows the sum in (3.7) to be computed. Taking then the limit $L \to \infty$ gives the sought determinant formula

$$E((0,\lambda))|_{m=n} = \prod_{j=1}^{n}(r_j s_j)^n(-\lambda)^{-n(n-1)/2}\prod_{j=1}^{n-1}\frac{1}{j^j}\frac{1}{\prod_{j<l}^{n}(r_l - r_j)(s_l - s_j)} \times \det\left[\int_{\lambda}^{\infty}(t-\lambda)^{n-1}\,\mathrm{e}^{-tr_j s_l}\,\mathrm{d}t\right]_{j,l=1,\ldots,n}. \tag{3.8}$$

And changing variables $t \mapsto t+\lambda$ in the integral allows (3.8) to be simplified further, giving

$$E((0,\lambda))|_{m=n} = \prod_{j=1}^{n-1} j!\frac{1}{\prod_{j<l}^{n}(-\lambda)(r_l - r_j)(s_l - s_j)}\det[\mathrm{e}^{-\lambda r_j s_l}]_{j,l=1,\ldots,n}. \tag{3.9}$$

Curiously this is the Harish–Chandra/Itzykson–Zuber matrix integral evaluation used in going from (2.3) to (2.4) and so we have the matrix integral representation

$$E((0,\lambda))|_{m=n} = \int \mathrm{e}^{-\mathrm{Tr}(\lambda R V^{\dagger} S V)}[V^{\dagger}\,\mathrm{d}Z], \tag{3.10}$$

where R, S are Hermitian matrices has eigenvalues $\{r_i\}, \{s_i\}$ respectively, and $[V^{\dagger}\,\mathrm{d}Z]$ denotes the normalized Haar measure, $\int[V^{\dagger}\,\mathrm{d}Z] = 1$. With $m = n$, in the limit $s_1, \ldots, s_n \to 1$ the pdf (1.3) reduces to (1.2). In keeping with this we can check that (3.9) reduces to (2.15) (with $s_j \mapsto r_j (j = 1, \ldots, n)$ in the latter).

We turn our attention now to $E((\lambda, \infty))$. For this we use a minor rewrite of (3.3),

$$\prod_{j<k}^{n}(\lambda_k - \lambda_j) = \prod_{j=1}^{n}\left(\frac{\lambda-\lambda_j}{\lambda^{1/2}}\right)^{n-1}\det\left[\left(\frac{\lambda_j}{\lambda-\lambda_j}\right)^{k-1}\right],$$

in (3.1) so that the analogue of (3.4) reads

$$E((\lambda,\infty))|_{m=n} = \prod_{j=1}^{n}(r_j s_j)^n\lambda^{-n(n-1)/2}\sum_{P\in S_N}\varepsilon(P)\int_0^{\lambda}\mathrm{d}\lambda_1\cdots\int_0^{\lambda}\mathrm{d}\lambda_N \times \prod_{j=1}^{n}(\lambda-\lambda_j)^{n-P(j)}\lambda_j^{P(j)-1}\mathcal{I}(\{r_i\},\{s_i\},\{\lambda_i\}).$$

The procedure of going from (3.4) to (3.8) can now be enacted. (Note that in the integration by parts the factor

$$\prod_{j=1}^{n}(\lambda-\lambda_j)^{n-P(j)}\lambda_j^{P(j)-1}$$

ensures that the integrand vanishes at the end points.) We thus arrive at the determinant evaluation

$$E((\lambda,\infty))|_{m=n}=\prod_{j=1}^{n-1}\frac{1}{j^j}\frac{\prod_{j=1}^{n}(\lambda r_j s_j)^n}{\prod_{j<l}^{n}(-\lambda)(r_l-r_j)(s_l-s_j)}\det\left[\int_0^1(1-t)^{n-1}\,\mathrm{e}^{-\lambda r_j s_l t}\,\mathrm{d}t\right]_{j,l=1,\dots,n}. \tag{3.11}$$

Analogous to the remark in the sentence below (3.10), we must have that for $s_1,\dots,s_n\to 1$ (3.11) coincides with (2.8) (after setting $s_j\mapsto r_j (j=1,\dots,n)$ in the latter). Now, taking the limit $s_1,\dots,s_n\to 1$ in (3.11) gives

$$\left(\frac{1}{(n-1)!}\right)^{n-1}\frac{\prod_{j=1}^{n}(\lambda r_j)^n}{\prod_{j<l}(-\lambda)(r_l-r_j)}\det\left[\int_0^1 \mathrm{e}^{-\lambda r_j t}\frac{d^{k-1}}{\mathrm{d}t^{k-1}}(t^{k-1}(1-t)^{n-1})\,\mathrm{d}t\right]_{j,k=1,\dots,n} \tag{3.12}$$

Expanding the derivatives using the product rule, we see after elementary column operations that the determinant in (3.12) is equal to

$$\det\left[\int_0^1 \mathrm{e}^{-\lambda r_j t}t^{k-1}\frac{d^{k-1}}{\mathrm{d}t^{k-1}}(t^{k-1}(1-t)^{n-1})\,\mathrm{d}t\right]_{j,k=1,\dots,n}$$
$$=\prod_{k=1}^{n-1}\frac{(n-1)!}{(n-k)!}\det\left[\int_0^1 \mathrm{e}^{-\lambda r_j t}(1-t)^{n-1}\left(\frac{t}{1-t}\right)^{k-1}\mathrm{d}t\right]_{j,k=1,\dots,n}. \tag{3.13}$$

But the determinant in (3.13) can be written

$$\int_0^1\mathrm{d}t_1\cdots\int_0^1\mathrm{d}t_n\prod_{j=1}^{n}\mathrm{e}^{-\lambda r_j t_j}(1-t_j)^{n-1}\det\left[\left(\frac{t_j}{1-t_j}\right)^{k-1}\right]_{j,k=1,\dots,n}$$
$$=\int_0^1\mathrm{d}t_1\cdots\int_0^1\mathrm{d}t_n\prod_{j=1}^{n}\mathrm{e}^{-\lambda r_j t_j}\prod_{j<k}(t_k-t_j)$$
$$=\det\left[\int_0^1\mathrm{e}^{-\lambda r_j t}t^{k-1}\,\mathrm{d}t\right]_{j,k=1,\dots,n} \tag{3.14}$$

Substituting (3.14) in (3.13), then substituting the result in (3.12) reclaims (2.8) in the case $m=n$.

It remains to apply the limiting procedure discussed in the paragraph below (2.11) to deduce from (3.11) the evaluation for general $m\leqslant n$. Noting the asymptotic expansion

$$\int_0^1(1-t)^{n-1}\,\mathrm{e}^{-\lambda r_j s_l t}\,\mathrm{d}t\sim\sum_{p=1}^{n}(-1)^{p-1}\frac{(n-1)\cdots(n-p+1)}{(\lambda r_j s_l)^p}$$

the required limits can be taken to give

$$E((\lambda,\infty))=(-1)^{(n-m)(n-m-1)/2}\prod_{j=1}^{n-1}\frac{1}{j^j}\prod_{p=1}^{n-m-1}\frac{\Gamma(n)}{\Gamma(n-p)}\frac{\left(\prod_{j=1}^{m}r_j\right)^n\left(\prod_{j=1}^{n}\lambda s_j\right)^n}{\prod_{j<l}^{m}(r_l-r_j)\prod_{j<l}^{n}\lambda(s_l-s_j)}$$
$$\times\det\begin{bmatrix}\left[\int_0^1(1-t)^{n-1}\,\mathrm{e}^{-\lambda r_j s_l t}\,\mathrm{d}t\right]_{\substack{j=1,\dots,m\\ l=1,\dots,n}}\\ \left[(\lambda s_l)^{-j}\right]_{\substack{j=1,\dots,n-m\\ l=1,\dots,n}}\end{bmatrix}. \tag{3.15}$$

Acknowledgment

This work was supported by the Australian Research Council.

References

[1] Baik J, Ben Arous G and Péché S 2005 Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices *Ann. Probab.* **33** 1643–97
[2] Bleher P M and Kuijlaars A 2004 Integral representations for multiple Hermite and Laguerre polynomials *Preprint* math.CA/0406616
[3] Chiani M, Win M Z and Zanella A 2003 On the capacity of spatially correlated MIMO Rayleigh-fading channels *IEEE Trans. Inform. Theory* **49** 2363–71
[4] Desrosiers P and Forrester P J 2006 Asymptotic correlations for Gaussian and Wishart matrices with external source *Int. Math. Res. Not.* **2006** 43
[5] Edelman A 1988 Eigenvalues and condition numbers of random matrices *SIAM J. Matrix Anal. Appl.* **9** 543–60
[6] Karoui N El 2006 A rate of convergence result for the largest eigenvalue of complex white Wishart matrices *Ann. Probab.* **34** 2077–117
[7] Forrester P J Log-gases and random matrices (www.ms.unimelb.edu.au/∼matpjf/matpjf.html)
[8] Forrester P J 1993 Exact results and universal asymptotics in the Laguerre random matrix ensemble *J. Math. Phys.* **35** 2539–51
[9] Forrester P J and Hughes T D 1994 Complex Wishart matrices and conductance in mesoscopic systems: exact results *J. Math. Phys.* **35** 6736–47
[10] Gao H and Smith P J 2000 A determinant representation for the distribution of quadratic forms in complex normal vectors *J. Multivariate Anal.* **73** 155–65
[11] Gupta A K and Nagar D K 1999 *Matrix Variate Distributions* (Boca Raton, FL: Chapman and Hall)
[12] James A T 1964 Distributions of matrix variate and latent roots derived from normal samples *Ann. Math. Stat.* **35** 475–501
[13] Johnstone I M 2001 On the distribution of the largest principal component *Ann. Math. Stat.* **29** 295–327
[14] Koev P and Edelman A 2005 The efficient evaluation of the hypergeometric function of a matrix argument *Preprint* math.PR/0505344
[15] Muirhead R J 1982 *Aspects of Multivariable Statistical Theory* (New York: Wiley)
[16] Orlov A Y 2004 New solvable matrix integrals *Int. J. Mod. Phys.* A **19** 276–93
[17] Ratnarajah T, Vaillancourt R and Alvo M 2005 Eigenvalues and condition numbers of complex random matrices *SIAM J. Matrix Anal. Appl.* **26** 441–56
[18] Simon S H and Moustakas A L 2004 Eigenvalue density of correlated random Wishart matrices *Phys. Rev.* E **69** 065101
[19] Simon S H, Moustakas A L and Marinelli L 2005 Capacity and character expansions: moment generating function and other exact results for MIMO correlated channels *Preprint* cs.IT/0509080
[20] Tulino A M and Verdú S 2004 *Random Matrix Theory and Wireless Communications (Foundations and Trends in Communications and Information Theory* vol 1*)* pp 1–182