Strong approximation of eigenvalues of large dimensional Wishart matrices by roots of generalized Laguerre polynomials

Holger Dette
Ruhr-Universität Bochum
Fakultät für Mathematik
44780 Bochum
Germany

email: holger.dette@ruhr-uni-bochum.de

October 9, 2001

Abstract

The purpose of this note is to establish a link between recent results on asymptotics for classical orthogonal polynomials and random matrix theory. Roughly speaking it is demonstrated that the ith eigenvalue of a Wishart matrix $W(I_n, s)$ is close to the ith zero of an appropriately scaled Laguerre polynomial, when

$$\lim_{n,s\to\infty} n/s = y \in [0,\infty).$$

As a by-product we obtain an elemantary proof of the Marčenko-Pastur and the semicircle law without relying on combinatorical arguments. Moreover, our approach also allows a simple treatment of the case $y=\infty$, where a new semicircle law can be established for the s largest eigenvalues of the Wishart matrix.

AMS 1991 Subject Classifications: Primary 60F15; Secondary 33C45

Key words and phrases: Random matrix theory, Marčenko-Pastur law, semi-circle law, Laguerre polynomials, roots of orthogonal polynomials, strong approximation

1 Introduction

The study of sample covariance matrices is important in multivariate statistics and since the pioneering work of Marčenko and Pastur (1967) much effort has been devoted to this subject [see e.g. Silverstein (1985), Bai and Yin (1988a,b, 1993), Johnstone (2001) among many others]. In this note we present a new approach for the derivation of the asymptotic sprectral distribution of a Wishart matrix $W(I_n, s)$, when the parameters n and s both converge to infinity at appropriate rates. This method relies on a close connection between the eigenvalues of the Wishart matrix and the zeros of classical orthogonal polynomials. To be precise, let $V_s \in \mathbb{R}^{n\times s}$ denote a random matrix with i.i.d. standard normally distributed entries, define

$$M_s = \frac{1}{s} V_s V_s^T \in \mathbb{R}^{n\times n} \tag{1.1}$$

as the sample covariance matrix and let $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ denote the ordered eigenvalues of the matrix M_s, where the double index has been omitted for the sake of simplicity; i.e. $\lambda_i = \lambda_i^{(n)}$. It is well known that the joint densitiy of the eigenvalues is proportional to the function

$$\prod_{1\leq i<j\leq n}^{n} |\lambda_i - \lambda_j| \prod_{i=1}^{n} \lambda_i^{(s-n-1)/2} e^{-\lambda_i/2}$$

and a *typical* vector of ordered eigenvalues should be close to the mode of this density. By classical results of Stieltjes [see e.g. Szegö (1959)] the above density becomes maximal for the zeros of the Laguerre polynomial. The asymptotic properties of these polynomials have been recently investigated independently from the random matrix literature in the context of approximation theory. We refer to Gawronski (1993) and Bosbach and Gawronski (1998) for some results on strong asymptotics for Laguerre polynomials with varying coefficients and to Faldey and Gawronski (1995), Dette and Studden (1995), Kuijlaars and Van Assche (1999) for recent results on the asymptotic zero distribution of these polynomials.

It is the purpose of the present paper to provide a link between the results in random matrix theory and the theory of orthogonal polynomials. To this end we derive an almost sure approximation of the eigenvalues of the Wishart matrix M_s defined in (1.1) by the zeros of appropriately scaled generalized Laguerre polynomials, when

$$\lim_{n,s\to\infty} n/s = y \in [0, \infty]. \tag{1.2}$$

This generalizes recent work of Silverstein (1985), who established almost sure convergence of the smallest eigenvalue of the Wishart matrix $W(I_n, s)$, when

$$\lim_{n,s\to\infty} n/s = y \in (0, 1).$$

Note that our results include the case $y = \infty$, which was not considered so far. As a by-product we obtain a simple proof of the Marčenko-Pastur law for the empirical spectral distribution function

$$F_{M_s}(x) = \frac{1}{n} \#\{i \mid \lambda_i \leq x\} \tag{1.3}$$

[note that this function has already been appropriately standardized] by using recent results on weak asymptotics for classical orthogonal polynomials. Additionally, we provide a new proof of the classical semicircle law when $n/s \to 0$ and our method allows the derivation of a new semicircle law for the largest s eigenvalues of an approriately scaled Wishart matrix in the case $n, s \to \infty$, $n/s \to \infty$.

2 Eigenvalues of Wishart matrices and zeros of Laguerre polynomials

Throughout this paper let for $k = 0, 1, \ldots, n$ $L_k^{(\alpha_n)}(x)$ denote the kth generalized Laguerre polynomial orthogonal with respect to the weight function $x^{\alpha_n} \exp(-x) I_{(0,\infty)}(x)$. We note that the orthogonalizing measure is varying with the degree n and that we are interested in a comparison of the roots $x_1 < \ldots < x_n$ of appropriately scaled versions of the polynomial $L_n^{(\alpha_n)}(x)$ with the ordered eigenvalues $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_n$ of the matrix M_s (or an appropriately scaled version) defined in (1.1). The scaling of the polynomial and the Wishart matrix depends on the limit y in (1.2) and we use the roots of the polynomial

$$L_n^{(s-n+1)}(sx)$$

in the case $y \in (0, \infty)$, the zeros of the polynomial

$$L_n^{(s)}(2\sqrt{ns}x + s + n)$$

in the case $y = 0$ and the roots of the polynomial

$$L_n^{(s-n)}(2\sqrt{ns}x + n)$$

in the case $y = \infty$ for a comparison. The scaling of the Laguerre polynomials is motivated by weak asymptotic properties of their zeros [see Theorem 2.4], which were recently obtained by Gawronski (1993), Bosbach and Gawronski (1998), Faldey and Gawronski (1995), Dette and Studden (1995), Kuijlaars and Van Assche (1999). Throughout this paper I_k denotes the $k \times k$ identity matrix. The main result of this paper is the following.

Theorem 2.1.
a) *Let $\lambda_1 \le \ldots \le \lambda_n$ denote the ordered eigenvalues of the sample covariance matrix M_s defined in (1.1) and $x_1 < \ldots < x_n$ denote the zeros of the Laguerre polynomial $L_n^{(s-n+1)}(sx)$. If $n, s \to \infty$, $n/s \to y \in (0, \infty)$, then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |\lambda_j - x_j|^2 \;=\; 0 \qquad a.s.$$

b) *Let $\lambda_1 \le \ldots \le \lambda_n$ denote the ordered eigenvalues of the sample covariance matrix*

$$N_s = \frac{1}{2\sqrt{ns}} \{V_s V_s^T - sI_n\} \tag{2.1}$$

and $x_1 < \ldots < x_n$ denote the zeros of the Laguerre polynomial $L_n^{(s)}(2(\sqrt{ns})x + s + n)$. If $n, s \to \infty$, $n/s \to 0$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} |\lambda_j - x_j|^2 \;=\; 0 \qquad a.s.$$

c) *Let* $\lambda_1 \leq \ldots \leq \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix*

$$P_s = \frac{1}{2\sqrt{ns}}\{V_s V_s^T - nI_n\} \tag{2.2}$$

and $-\frac{1}{2}\sqrt{n/s} = x_1 = \ldots = x_{n-s} < x_{n-s+1} < \ldots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n)}(2\sqrt{ns}x + n)$*. If* $n, s \to \infty$, $n/s \to \infty$, *then*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} |\lambda_j - x_j|^2 \;=\; \lim_{n\to\infty} \frac{1}{n}\sum_{j=n-s+1}^{n} |\lambda_j - x_j|^2 \;=\; 0 \qquad a.s.$$

Proof:
a) For a proof of part (a) we assume at first that $y \in (0,1)$, that is $s \geq n$ for sufficiently large n. According to Silverstein (1985; p. 1366) the matrix M_s is orthogonally similar to a triangular matrix $\tilde{A} = (\tilde{a}_{i,j})_{i,j=1}^n$ with entries

$$\begin{aligned}
\tilde{a}_{i,i} &= \frac{1}{s}(Y_{n-i+1}^2 + X_{s-i+1}^2) \quad\; i = 1,\ldots,n,\\
\tilde{a}_{i,i+1} &= \frac{1}{s}X_{s-i+1}Y_{n-i}, \quad i = 1,\ldots,n-1,\\
\tilde{a}_{i+1,i} &= \frac{1}{s}X_{s-i+1}Y_{n-i}, \quad i = 1,\ldots,n-1,
\end{aligned}$$

where $Y_n^2 = 0$, $X_i^2 \sim \chi_i^2, Y_i^2 \sim \chi_i^2$ are independent chi-square distributed random variables $(X_i \geq 0, Y_i \geq 0)$. Therefore it is easy to see that the matrix M_s has the same eigenvalues as the matrix $A = (a_{i,j})_{i,j=1}^n$ defined by

$$\begin{aligned}
a_{i,i} &= \tilde{a}_{n-i+1,n-i+1} = \frac{1}{s}(Y_i^2 + X_{s-n+i}^2) \quad i = 1,\ldots,n,\\
a_{i,i+1} &= \tilde{a}_{n-i,n-i+1} = \frac{1}{s}X_{s-n+i+1}Y_i \qquad i = 1,\ldots,n-1,\\
a_{i+1,i} &= \tilde{a}_{n-i+1,n-i} = \frac{1}{s}X_{s-n+i+1}Y_i \qquad i = 1,\ldots,n-1.
\end{aligned}$$

Now consider the kth Laguerre polynomial

$$\hat{L}_k^{(s-n+1)}(x)$$

orthogonal with respect to the weight function $x^{s-n+1}\exp(-x)I_{(0,\infty)}(x)$ with leading coefficient 1. According to Chihara (1978) we have the recursion $(\alpha_n = s - n + 1)$

$$\hat{L}_{k+1}^{(\alpha_n)}(x) = (x - \{2k+1+\alpha_n\})\hat{L}_k^{(\alpha_n)}(x) - k(k+\alpha_n)\hat{L}_{k-1}^{(\alpha_n)}(x) \tag{2.3}$$

with initial conditions $\hat{L}_{-1}^{(\alpha_n)}(x) = 0, \hat{L}_0^{(\alpha_n)}(x) = 1$. It is now straightforward to see that the zeros of the polynomial $\hat{L}_n^{(\alpha_n)}(sx)$ are precisely the eigenvalues of the triangular matrix

$B = (b_{i,j})_{i,j}^n$, where

$$\begin{aligned} b_{i,i} &= \frac{1}{s}(\alpha_n + 2i - 1) \qquad & i = 1, \ldots, n, \\ b_{i,i+1} &= \frac{1}{s}\sqrt{i(i+\alpha_n)} \qquad & i = 1, \ldots, n-1, \\ b_{i+1,i} &= \frac{1}{s}\sqrt{i(i+\alpha_n)} \qquad & i = 1, \ldots, n-1. \end{aligned}$$

This follows by factorizing $(-1)^n$, $(\frac{1}{s})^n$ in the determinant equation

$$\det(B - \lambda I) = 0,$$

and identifying the recursion (2.3) for the polynomial $\hat{L}_n^{(\alpha_n)}(s\lambda)$.

Now the discussion following Lemma 2.3 in Bai (1999) yields for the distance between the eigenvalues of the matrix M_s and the zeros of the polynomial $\hat{L}_n^{(s-n+1)}(sx)$

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^n |\lambda_j - x_j|^2 &\leq \frac{1}{n}\operatorname{tr}(A-B)^2 = \frac{1}{n}\sum_{i,j=1}^n (a_{i,j} - b_{i,j})^2 \\ &= \frac{1}{n}\sum_{i=1}^n (a_{i,i} - b_{i,i})^2 + \frac{2}{n}\sum_{i=1}^{n-1}(a_{i,i+1} - b_{i,i+1})^2 \ , \end{aligned} \tag{2.4}$$

where the first equality follows from the symmetry of the matrices A and B.
The two terms in (2.4) are estimated seperately. For the first term we have with some finite constant $c > 0$ (observing $Y_n^2 = 0$)

$$\begin{aligned} c \cdot \sum_{i=1}^n (a_{i,i} - b_{i,i})^2 &\leq \sum_{i=1}^{n-1}\Big(\frac{Y_i^2 - i}{s}\Big)^2 + \Big(\frac{n}{s}\Big)^2 + \sum_{i=1}^n \Big(\frac{X_{s-n+i}^2 - s + n - i}{s}\Big)^2 \\ &\leq 2n \cdot M_n^2 + \Big(\frac{n}{s}\Big)^2 \ , \end{aligned}$$

where the the random variable M_n is defined by

$$M_n = \max\Big\{\max_{1\leq i\leq n-1}|\frac{Y_i^2 - i}{s}|, \max_{s-(n-1)\leq i\leq s}|\frac{X_i^2 - i}{s}|\Big\} \ . \tag{2.5}$$

From Silverstein (1985, p. 1367) it follows that $M_n \to 0$ a.s. and we obtain

$$\frac{1}{n}\sum_{i=1}^n (a_{i,i} - b_{i,i})^2 \to 0 \quad \text{a.s.} \tag{2.6}$$

For the remaining term in (2.4) we have

$$\frac{1}{n}\sum_{i=1}^{n-1}(a_{i,i+1} - b_{i,i+1})^2 = \frac{1}{n}\sum_{i=1}^{n-1}\Big|\frac{X_{s-n+i+1}Y_i}{s} - \frac{\sqrt{i(i+\alpha_n)}}{s}\Big|^2$$

$$
\begin{aligned}
&\leq \frac{1}{n}\sum_{i=1}^{n-1}\Big\{\Big|\frac{X^2_{s-n+i+1}-(s-n+i+1)}{s}\Big|^{1/2}\Big|\frac{Y_i^2-i}{s}\Big|^{1/2} \\
&\qquad +\sqrt{\frac{i+s-n+1}{s}}\Big|\frac{Y_i^2-i}{s}\Big|^{1/2}+\sqrt{\frac{i}{s}}\Big|\frac{X^2_{s-n+i+1}-(s-n+i+1)}{s}\Big|^{1/2}\Big\}^2 \\
&\leq (M_n+2\sqrt{M_n})^2 \to 0 \text{ a.s. },
\end{aligned}
$$

where the random variable M_n is defined in (2.5) and we have used the inequality

$$
\begin{aligned}
|\underline{a}\,\underline{b}-ab| &\leq |\underline{a}^2-a^2|^{1/2}|\underline{b}^2-b^2|^{1/2} \\
&+ |b||\underline{a}^2-a^2|^{1/2}+|a||\underline{b}^2-b^2|^{1/2}
\end{aligned}
$$

for nonnegative $\underline{a}, \underline{b}, a, b$ [see Silverstein (1985)]. Observing (2.4) the assertion (a) of the Theorem 2.1 follows in the case $y \in (0,1)$.

In the case $y > 1$ (which means $n > s$ for suffiently large n) the result is established by interchanging the roles of s and n and from a representation for generalized Laguerre polynomials with negative parameter. To be precise, we note that in this case the matrix M_s is orthogonally similar to an $n \times n$ matrix A with principal $s \times s$ block containing the (s-dimensional) rows

$$
\frac{1}{s}(X_n^2+Y_{s-1}^2, Y_{s-1}X_{n-1}, 0, \dots, 0),
$$

$$
\frac{1}{s}(Y_{s-i+1}X_{n-i+1}, X^2_{n-i+1}+Y^2_{s-i}, Y_{s-i}X_{n-i}, 0, \dots, 0),
$$

($i = 2, \dots, s-1$) and

$$
\frac{1}{s}(0, \dots, 0, Y_1X_{n-s+1}, X^2_{n-s+1}) ,
$$

where all other entries in the matrix A are 0 and the meaning of the random variables X_i, Y_i, X_i^2, Y_i^2 is the same as in the prevoius paragraph. Observing the identitiy

$$
L_n^{(-k)}(x) = (-x)^k\frac{(n-k)!}{n!}L_{n-k}^{(k)}(x) \tag{2.7}
$$

[see Szegö (1959), Section 5.2)] the assertion now follows by similar arguments as given for the case $y \in (0,1)$.

The remaining case $y = 1$ is proved by considering two subsequences corresponding to the cases $s \geq n$ and $s < n$, respectively.

b) By the same argument as given in the proof of part a) the eigenvalues of the matrix N_s defined in (2.1) are obtained as the eigenvalues of the tridiagonal matrix A defined by

$$
\begin{aligned}
a_{i,i} &= \frac{1}{2\sqrt{sn}}(Y_i^2+X^2_{s-n+i}-s) \\
a_{i,i+1} &= a_{i+1,i} = \frac{1}{2\sqrt{sn}}X_{s-n+i+1}Y_i.
\end{aligned}
$$

Now consider the Laguerre polynomials with leading coefficient 1 and parameter $\alpha_n = s$ and define polynomials

$$p_k(x) = \hat{L}_k^{(s)}(2\sqrt{ns}x + s + n).$$

The zeros of the polynomial $p_n(x)$ are given by the eigenvalues of the matrix B defined by

$$\begin{aligned} b_{i,i} &= \frac{1}{2\sqrt{ns}}(\alpha_n + 2i - 1 - n - s) = \frac{1}{2\sqrt{ns}}(2i - 1 - n) , \qquad i = 1, \dots, n \\ b_{i,i+1} &= b_{i+1,i} = \frac{1}{2\sqrt{ns}}\sqrt{i(i + \alpha_n)} = \frac{1}{2\sqrt{ns}}\sqrt{i(i+s)} , \qquad i = 1, \dots, n-1 \end{aligned}$$

where $\alpha_n = s$. The assertion now follows by similar arguments as given in the proof of part a).

c) The asymptotic properties in the case $y = \infty$ follow from a combination of the arguments given in the proof of part a) for the case $y > 1$ and the proof of part b). The matrix P_s defined in (2.2) is orthogonally similar to an $n \times n$ matrix A with principal $s \times s$ block containing the (s-dimensional) rows

$$\frac{1}{2\sqrt{ns}}(X_n^2 + Y_{s-1}^2 - n, Y_{s-1}X_{n-1}, 0, \dots, 0),$$

$$\frac{1}{2\sqrt{ns}}(Y_{s-i+1}X_{n-i+1}, X_{n-i+1}^2 + Y_{s-i}^2 - n, Y_{s-i}X_{n-i}, 0, \dots, 0),$$

$(i = 2, \dots, s-1)$ and

$$\frac{1}{2\sqrt{ns}}(0, \dots, 0, Y_1X_{n-s+1}, X_{n-s+1}^2 - n) ,$$

where all other entries in the matrix A are 0 and the meaning of the random variables X_i, Y_i, X_i^2, Y_i^2 is the same as in the proof of part a). From (2.7) we have for some constant $c \neq 0$

$$L_n^{(s-n)}(2\sqrt{ns}x + n) \; = \; c \cdot (2\sqrt{ns}x + n)^{n-s} L_s^{(n-s)}(2\sqrt{ns}x + n) ,$$

where the s positive zeros of the polynomial on the right hand side are obtained as the eigenvalues of the tridiagonal matrix B with elements

$$\begin{aligned} b_{i,i} &= \frac{1}{2\sqrt{ns}}(2i - 1 - s) , \qquad i = 1, \dots, s \\ b_{i,i+1} &= b_{i+1,i} = \frac{1}{2\sqrt{ns}}\sqrt{i(i+n-s)} , \qquad i = 1, \dots, s-1 . \end{aligned}$$

The assertion now follows by similar arguments as given in the proof of part a). $\square$

The following results is an immediate consequence of Theorem 2.1 and recent results on the location of the zeros of classical orthogonal polynomials.

Corollary 2.2.

a) *Let* $\lambda_1 \le \ldots \le \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* M_s *defined in (1.1) and* $x_1 < \ldots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n+1)}(sx)$. *If* $d_1 \le d_2 \le \ldots \le d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty$, $n/s \to y \in (0,\infty)$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} d_{\lfloor nt \rfloor} \;=\; 0 \qquad a.s.$$

for all $t \in (0,1)$. *In particular we obtain for the smallest and largest eigenvalue of the matrix* M_s *and for the smallest and largest zero of the polynomial* $L_n^{(s-n+1)}(sx)$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,1]}} x_1 \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,1]}} \lambda_1 \;=\; (1-\sqrt{y})^2 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} x_n \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to y\in(0,\infty)}} \lambda_n \;=\; (1+\sqrt{y})^2 \qquad a.s.$$

and in the case $y \ge 1$

$$\lim_{\substack{n,s\to\infty \\ n/s\to y\in[1,\infty)}} x_{n-s+1} \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to y\in[1,\infty)}} \lambda_{n-s+1} \;=\; (1-\sqrt{y})^2 \qquad a.s.$$

b) *Let* $\lambda_1 \le \ldots \le \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* N_s *defined in (2.1) and* $x_1 < \ldots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s)}(2\sqrt{ns}x + s + n)$. *If* $d_1 \le d_2 \le \ldots \le d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty$, $n/s \to 0$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} d_{\lfloor nt \rfloor} \;=\; 0 \qquad a.s.$$

for all $t \in (0,1)$. *In particular we obtain for the largest and smallest eigenvalue of the matrix* N_s *and for the smallest and largest zero of the polynomial* $L_n^{(s)}(2\sqrt{ns}x + s + n)$

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} x_1 \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to 0}} \lambda_1 \;=\; 1 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty \\ n/s\to 0}} x_n \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to 0}} \lambda_n \;=\; -1 \qquad a.s.$$

c) *Let* $\lambda_1 \le \ldots \le \lambda_n$ *denote the ordered eigenvalues of the sample covariance matrix* P_s *defined in (2.2) and* $x_1 < \ldots < x_n$ *denote the zeros of the Laguerre polynomial* $L_n^{(s-n)}(2\sqrt{ns}x + n)$. *If* $0 = d_1 = \ldots = d_{n-s} \le d_{n-s+1} \le d_2 \le \ldots \le d_n$ *denote the ordered differences* $|\lambda_i - x_i|$ *and* $n, s \to \infty$, $n/s \to \infty$, *then*

$$\lim_{\substack{n,s\to\infty \\ n/s\to \infty}} d_{\lfloor nt \rfloor} \;=\; 0 \qquad a.s.$$

for all $t \in (0,1)$. *In particular we obtain for the largest and* $(n-s+1)$*th smallest eigenvalue of the matrix* P_s *and for the* $(n-s+1)$*th smallest and largest zero of the polynomial* $L_n^{(s-n)}(2\sqrt{ns}x + n)$

$$\lim_{\substack{n,s\to\infty \\ n/s\to \infty}} x_{n-s+1} \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to \infty}} \lambda_{n-s+1} \;=\; -1 \qquad a.s.$$

$$\lim_{\substack{n,s\to\infty \\ n/s\to \infty}} x_n \;=\; \lim_{\substack{n,s\to\infty \\ n/s\to \infty}} \lambda_n \;=\; 1 \qquad a.s.$$

Proof: The first part is an immediate consequence of Theorem 2.1. The assertion regarding the largest and smallest eigenvalue follows similary to the proof of Theorem 2.1 by an application of the Theorem of Geršgorin (1931) [see Silverstein (1985)]. The results for the largest and smallest zero of the Laguerre polynomial can be obtained from Theorem 4.4 in Dette and Studden (1985) and formula (2.7). □

The asymptotic properties of the largest and smallest eigenvalue in part a) of Corollary 2.2 were already observed by Silverstein (1985), but we did not find the result for sample covariance matrices for the case $n/s \to 0$ or $n/s \to \infty$ in the literature [for a proof of the analogue for $n/s \to 0$ in the case of Wigner matrices see Bai and Yin (1988b)]. The following example illustrates the quality of approximation in Corollary 2.2.

n	10		15		20	
	λ_j	x_j	λ_j	x_j	λ_j	x_j
	-0.74099	-0.72672	-0.76857	-0.76658	-0.78275	-0.78866
	-0.57733	-0.57065	-0.64970	-0.65404	-0.68759	-0.69867
	-0.43048	-0.42409	-0.54791	-0.55201	-0.60781	-0.61862
	-0.28494	-0.27741	-0.45075	-0.45334	-0.53318	-0.54254
	-0.13478	-0.12609	-0.35486	-0.35519	-0.46089	-0.46816
	0.02480	0.03340	-0.25757	-0.25592	-0.38926	-0.39427
	0.19766	0.20499	-0.15823	-0.15433	-0.31717	-0.32007
	0.39190	0.39425	-0.05522	-0.04934	-0.24446	-0.24499
	0.62150	0.61121	0.05261	0.06015	-0.17025	-0.16852
	0.93142	0.88111	0.16661	0.17546	-0.09418	-0.09021

Table 2.1. *The* 10 *smallest zeros of the scaled Laguerre polynomials $p_n(x)$ defined in (2.8) and the n smallest eigenvalues of the standardized Wishart matrix N_{10n} defined in (2.9) for various values of n.*

Example 2.3. Consider the case $s = 10n$ and note that for finite samples the limits $n/s \to y \in (0, \infty)$ and $n/s \to 0$ cannot be distinguished. Therefore both approximations of part a) and b) in Theorem 2.1 could be used in prinicple. For the sake of brevity we use only case b). Table 2.1 shows the zeros $x_1, \ldots, x_n$ of the Laguerre polynomial

$$p_n(x) = L_n^{(10n)}(2n\sqrt{10}x + 11n) \tag{2.8}$$

and the eigenvalues $\lambda_1, \ldots, \lambda_n$ of the standardized matrix

$$N_{10n} = \frac{1}{2n\sqrt{10}}(V_{10n}V_{10n}^T - 10nI_n) \tag{2.9}$$

for $n = 10, 15, 20$. These eigenvalues have been obtained by simulations based on 100.000 runs. For the sake of brevity only the ten smallest eigenvalues and zeros are displayed.

We will use Theorem 2.1 for an alternative proof of the famous Marčenko-Pastur and semicircle law in the normal case using recent results for the asymptotic zero distribution of classical orthogonal polynomials. Additionally we provide a semicircle law for the s largest eigenvalues of the appropriately scaled Wishart matrix, when $n, s \to \infty$ $n/s \to \infty$, which seems to be unknown in the literature. Conversely the arguments given in this paper show that the Marčenko-Pastur and semicircle law could also be used to provide an alternative proof for the asymptotic zero distribution of the Laguerre polynomials with varying integer valued parameters. For the sake of completeness we recall a result on the asymptotic zero distribution for the zeros of the Laguerre polynomials with varying (not necessarily integer valued) parameters. A proof can be found in Dette and Studden (1995) [see also Faldey and Gawronski (1995), Dette and Wong (1995) or Kuijlaars and Van Assche (1999)]. For a real sequence $(\alpha_n)_{n\in\mathbb{N}}$ with elements > -1 let

$$N^{(\alpha_n)}(\xi) \; := \; \#\left\{ \, x \mid L_n^{(\alpha_n)}(x) = 0, \; x \le \xi \right\} \tag{2.10}$$

denote the number of zeros of the generalized Laguerre polynomial $L_n^{(\alpha_n)}(x)$ less or equal than ξ, then we have the following result.

Theorem 2.4. (Dette and Studden (1995))

a) *If* $\lim\limits_{n\to\infty} \frac{\alpha_n}{n} = a \ge 0$, *then*

$$\lim_{n\to\infty} \frac{1}{n} N^{(\alpha_n)}(n\xi) \; = \; \frac{1}{2\pi} \int_{r_1}^{\xi} \frac{\sqrt{(r_2 - x)(x - r_1)}}{x} dx \qquad \text{for all } \; \xi \in [r_1, r_2] \; ,$$

where $r_{1,2} = 2 + a \pm 2\sqrt{1+a}$

b) *If* $\lim\limits_{n\to\infty} \frac{\alpha_n}{n} = \infty$, *then*

$$\lim_{n\to\infty} \frac{1}{n} N^{(\alpha_n)}(\sqrt{n\alpha_n}\xi + \alpha_n) \; = \; \frac{1}{2\pi} \int_{-2}^{\xi} \sqrt{4 - x^2} dx \qquad \text{for all } \; |\xi| \le 2 \; .$$

Theorem 2.5. (Marčenko-Pastur and extended semicircle law)

a) *If* $n \to \infty, n/s \to y \in (0, \infty)$ *and* F_{M_s} *denotes the empirical spectral distribution function of the matrix* M_s *defined in (1.1), then for all* $\xi \in \mathbb{R}$

$$F_{M_s}(\xi) \to F_M(\xi) \qquad a.s. \tag{2.11}$$

where the distribution function F_M *has densitiy*

$$f_M(x) \; := \; \frac{1}{2\pi y} \frac{\sqrt{(b-x)(x-a)}}{x} I_{[a,b]}(x) \; ,$$

the quantities a and b are given by $a = (1-\sqrt{y})^2, b = (1+\sqrt{y})^2$, respectively, and there is an additional jump of size $1 - 1/y$ in the case $y > 1$.

b) *If $n/s \to 0$ and F_{N_s} denotes the empirical spectral distribution function of the matrix (2.1), then we have for any $x \in [-1, 1]$*

$$F_{N_s}(x) \to F_N(x) \ := \ \frac{2}{\pi}\int_{-1}^{x} \sqrt{1-t^2}dt \qquad a.s. \tag{2.12}$$

($F_{N_s}(x) \to 1$ if $x > 1, F_{N_s}(x) \to 0$ if $x < 1$).

c) *If $n, s \to \infty$, $n/s \to \infty$ and F_{P_s} denotes the empirical distribution function of the s largest eigenvalues of the matrix P_s defined (2.2), then we have for any $x \in [-1, 1]$*

$$F_{P_s}(x) \to F_N(x) \ = \ \frac{2}{\pi}\int_{-1}^{x} \sqrt{1-t^2}dt \qquad a.s. \tag{2.13}$$

($F_{P_s}(x) \to 1$ if $x > 1, F_{P_s}(x) \to 0$ if $x < 1$).

Proof.

a) Consider at first the case (a) with $y \in (0, 1]$. From Bai (1999) and Theorem 2.1 it follows for the Levy distance L between the distribution functions F_{M_s} and F_B that

$$L^3(F_{M_s}, F_B) \ \le \ \frac{1}{n}\sum_{i=1}^{n} |\lambda_j - x_j|^2 \to 0 \ \text{ a.s. },$$

where

$$F_B(\xi) \ = \ \frac{1}{n}\#\{x \mid \hat{L}_n^{(\alpha_n)}(sx) = 0, x \le \xi\}$$

denotes the empirical distribution function of the zeros of the Laguerre polynomial $L_n^{(\alpha_n)}(sx)$ with parameter $\alpha_n = s - n + 1$. From the first part of Theorem 2.4 we therefore have for any $\xi \in [r_1, r_2]$

$$\begin{aligned} F_B(\xi) \ &= \ \frac{1}{n}\#\{x \mid \hat{L}_n^{(\alpha_n)}(n\frac{s}{n}x) = 0, x \le \xi\} \\ &= \ \frac{1}{n}N^{(\alpha_n)}(n\frac{s}{n}\xi) \underset{\substack{n\to\infty \\ \frac{n}{s}\to y}}{\longrightarrow} \frac{1}{2\pi}\int_{r_1}^{\xi/y} \frac{\sqrt{(r_2-x)(x-r_1)}}{x}dx \end{aligned} \tag{2.14}$$

where $r_{1,2} = (1 \pm \frac{1}{\sqrt{y}})^2$. Substitution and differentiation yields for the density of the limiting distribution

$$\frac{1}{2\pi y} \cdot \frac{\sqrt{(b-x)(x-a)}}{x} I_{[a,b]}(x) \ ,$$

where

$$\begin{aligned} a \ &= \ (1-\sqrt{y})^2 \\ b \ &= \ (1+\sqrt{y})^2. \end{aligned}$$

The argument for the case $y > 1$ follows exactly in the same way using at first the identity (2.7).

b) Again we obtain from Theorem 2.1

$$L^3(F_{N_s}, F_B) \to 0 \quad \text{a.s.}$$

where F_B denotes the empirical distribution function of the roots of the polynomial $L_n^{(\alpha_n)}(2\sqrt{n\alpha_n}x + s + n)$ with $\alpha_n = s$, that is

$$\begin{aligned} F_B(\xi) &= \frac{1}{n}\#\{x \mid L_n^{(\alpha_n)}(2\sqrt{n\alpha_n}x + s + n) = 0, x \le \xi\} \\ &= \frac{1}{n}N^{(\alpha_n)}(2\sqrt{n\alpha_n}\xi + s + n) \ . \end{aligned}$$

Observing $\frac{n}{\sqrt{n\alpha_n}} = \sqrt{\frac{n}{s}} = o(1)$ and Example 2.7 in Dette and Studden (1995) the second part of Theorem 2.4 now gives

$$\begin{aligned} \lim_{n\to\infty} F_B(\xi) &= \lim_{n\to\infty} \frac{1}{n}N^{(\alpha_n)}(2\sqrt{n\alpha_n}\xi + \alpha_n) \\ &= \frac{1}{2\pi}\int_{-2}^{2\xi} \sqrt{4 - x^2}dx = \frac{2}{\pi}\int_{-1}^{\xi} \sqrt{1 - t^2}dt \end{aligned}$$

whenever $|\xi| \le 1$, which proves the assertion of Theorem 2.5 b).

c) This is proved in the same way using the identity (2.7). □

Acknowledgements. The author would like to thank I. Gottschlich, who typed most parts of the paper with considerable technical expertise and I. Johnstone for some help with the references. The work of H. Dette was supported by the Deutsche Forschungsgemeinschaft (SFB 475, Komplexitätsreduktion in multivariaten Datenstrukturen).

References

Z.D. Bai (1999). Methodologies in spectral analysis of large dimensional random matrices, a review. Statist. Sinica 9, 611-677.

Z.D. Bai, Y.Q. Yin (1988a). Convergence to the semicircle law. Ann. Prob. 16, 863-875.

Z.D. Bai, Y.Q. Yin (1988b). Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a Wigner matrix. Ann. Prob. 16, 1729-1741.

Z.D. Bai, Y.Q. Yin (1993). Limit of the smallest eigenvalue of a large dimensional covariance matrix. Ann. Prob. 21, 1275-1294.

C. Bosbach, W. Gawronski (1998). Strong asymptotics for Laguerre polynomials with varying weights. J. Comput. Appl. Math. 99, 77-89.

T. Chihara (1978). An Introduction to Orthogonal Polynomials. Gordon and Breach.

H. Dette, W.J. Studden (1995). Some new asymptotic properties for the zeros of Jacobi, Laguerre and Hermite polynomials. Constr. Approx. 11, 227-238.

H. Dette, W.K. Wong (1995). Recurrence moment formulas for D-optimal designs. Scand. J. Statist. 22, 505-512.

J. Faldey, W. Gawronski (1995). On the limit distribution of the zeros of Jonquiere polynomials and generalized classical orthogonal polynomials. J. Approx. Theory 81, 231-249.

W. Gawronski (1993). Strong asymptotics and the asymptotic zero distributions of Laguerre polynomials $L_n^{(an+\alpha)}$ and Hermite polynomials $H_n^{(an+\alpha)}$. Analysis 13, 29-67.

S.A. Geršgorin (1931). Über die Abgrenzung der Eigenwerte einer Matrix. Ivz. Akad. Nauk. SSSR Ser. Fiz.-Mat. 6, 749 - 754.

I.M. Johnstone (2001). On the distribution of the largest principal component. Ann. Statist. to appear.

A.B.J. Kuijlaars, W. Van Assche (1999). The asymptotic zero distribution of orthogonal polynomials with varying recurrence coefficients. J. Approx. Theory 99, 167-197.

V.A. Marčenko, L.A. Pastur (1967). Distributions of eigenvalues of some sets of random matrices. Math. USSR-Sb. 1, 507-536.

G. Szegö (1959). Orthogonal Polynomials. American Mathematical Society.

J. Silverstein (1985). The smallest eigenvalue of a large dimensional Wishart matrix. Ann. Prob. 13, 1364-1368.

E.P. Wigner (1955). On the distribution of the roots of certain symmetric matrices. Ann. Math. 62, 548-564.