Infinite random matrix theory, tridiagonal bordered Toeplitz matrices, and the moment problem

Alexander Dubbs [*], Alan Edelman

Mathematics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, United States

A R T I C L E   I N F O

A B S T R A C T

The four major asymptotic level density laws of random matrix theory may all be showcased through their Jacobi parameter representation as having a bordered Toeplitz form. We compare and contrast these laws, completing and exploring their representations in one place. Inspired by the bordered Toeplitz form, we propose an algorithm for the finite moment problem by proposing a solution whose density has a bordered Toeplitz form.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Consider the "big" laws for asymptotic level densities for various random matrices:

Wigner semicircle law	[19]
Marchenko–Pastur law	[13]
Kesten–McKay law	[10,14]
Wachter law	[18]

---

* Corresponding author.
   E-mail addresses: alex.dubbs@gmail.com (A. Dubbs), edelman@math.mit.edu (A. Edelman).

Table 1

Random matrix laws in raw form. The Kesten–McKay and Wachter laws are related by the linear transform $(2x_{\text{Wachter}} - 1)v = x_{\text{Kesten–McKay}}$ and $a = b = v/2$.

Measure	Support	Parameters
Wigner semicircle $d\mu_{WS} = \frac{\sqrt{4-x^2}}{2\pi}dx$	$I_{WS} = [\pm 2]$	N/A
Marchenko–Pastur $d\mu_{MP} = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi x}dx$	$I_{MP} = [\lambda_-, \lambda_+]$	$\lambda_\pm = (1 \pm \sqrt{\lambda})^2$, $\lambda \geq 1$
Kesten–McKay $d\mu_{KM} = \frac{v\sqrt{4(v-1)-x^2}}{2\pi(v^2-x^2)}dx$	$I_M = [\pm 2\sqrt{v-1}]$	$v \geq 2$
Wachter $d\mu_W = \frac{(a+b)\sqrt{(\mu_+ - x)(x - \mu_-)}}{2\pi x(1-x)}dx$	$I_W = [\mu_-, \mu_+]$	$\mu_\pm = (\frac{\sqrt{b} \pm \sqrt{a(a+b-1)}}{a+b})^2$, $a, b \geq 1$

In raw form, these laws (Table 1) appear as somewhat complicated expressions involving square roots. This paper highlights a unifying principle that underlies these four laws, namely the laws may be encoded as Jacobi symmetric tridiagonal matrices that are Toeplitz with a length 1 boundary.

This suggests that some of the nice properties of the big laws are connected to this property, and further suggests the importance of the larger family of laws encoded as Toeplitz with length k boundary, known as "nearly Toeplitz" matrices. This motivates the two parts of this paper:

(1) We tabulate in one place key properties of the four laws, not all of which can be found in the literature. These sections are expository, with the exception of the as-of-yet unpublished Wachter moments, and the Kesten–McKay and Wachter law Jacobi parameters and free cumulants.

(2) We describe a new algorithm to exploit the Toeplitz-with-length-k boundary structure. In particular, we show how practical it is to approximate distributions with incomplete information using distributions having nearly-Toeplitz encodings.

Studies of nearly Toeplitz matrices in random matrix theory have been pioneered by Anshelevich [1,2].

Historically, the Wigner semicircle law is the most famous. The weight function is classical, and corresponds to Chebychev polynomials of the second kind. It is the equilibrium measure [3] for Hermite polynomials and the asymptotic distribution for Gaussian or Hermite ensembles (GOE, GUE, GSE, etc.). None of the other weight functions are classical, but they are all equilibrium measures for classical polynomials. The second most famous law is the Marchenko–Pastur law. It is the equilibrium measure for Laguerre polynomials and is the asymptotic distribution for Wishart matrices or Laguerre ensembles. The Kesten–McKay law, described in [9], is the equilibrium measure for Gegenbauer polynomials. It is not commonly included among the Wigner semicircle, Marchenko–Pastur, and Wachter laws, but we believe that it merits inclusion on account of its place in the

Table 2

Jacobi parameter encodings for the big level density laws. Upper left: symmetric Toeplitz tridiagonal with 1-boundary. Upper right: laws organized by Toeplitz property. Below: specific parameter values.

$$\begin{bmatrix} \alpha_0 & \beta_0 & & & & \\ \beta_0 & \alpha_1 & \beta_1 & & & \\ & \beta_1 & \alpha_1 & \beta_1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & \beta_1 & \alpha_1 & \beta_1 \\ & & & & \beta_1 & \alpha_1 \end{bmatrix}$$

	$\alpha_0 = \alpha_1$	$\alpha_0 \neq \alpha_1$
$\beta_0 = \beta_1$	Wigner semicircle	Marchenko–Pastur
$\beta_0 \neq \beta_1$	Kesten–McKay	Wachter

Measure	α_0	$\alpha_n \ (n \geq 1)$	β_0	$\beta_n \ (n \geq 1)$
Wigner semicircle	0	0	1	1
Marchenko–Pastur	λ	$\lambda + 1$	$\sqrt{\lambda}$	$\sqrt{\lambda}$
Kesten–McKay	0	0	$\sqrt{v}$	$\sqrt{v-1}$
Wachter	$\frac{a}{a+b}$	$\frac{a^2-a+ab+b}{(a+b)^2}$	$\frac{\sqrt{ab}}{(a+b)^{3/2}}$	$\frac{\sqrt{ab(a+b-1)}}{(a+b)^2}$

Table 3

Cauchy transforms.

Measure	Cauchy transform
Wigner semicircle	$\frac{z-\sqrt{z^2-4}}{2}$
Marchenko–Pastur	$\frac{1-\lambda+z-\sqrt{(1-\lambda+z)^2-4z}}{2z}$
Kesten–McKay	$\frac{(v-2)z-v\sqrt{4(1-v)+z^2}}{2(v^2-z^2)}$
Wachter	$\frac{1-a+(a+b-2)z-\sqrt{(a+1-(a+b)z)^2-4a(1-z)}}{2z(1-z)}$

upper-right box in Table 2. The Wigner and arcsine distributions are special cases of the Kesten–McKay distribution. The Wachter law generalizes to Jacobi polynomials. They describe sections of random unitary matrices, MANOVA matrices, and general Jacobi ensembles.

The laws have been encoded in many formats over the years. Wigner's earliest work encoded the semicircle law through its exponential generating function, the Bessel function. For many decades the Stieltjes or Cauchy transforms (Table 3) have been valuable. Free probability has proposed the R-transform and S-transform as efficient encodings for these laws (Table 5). In this note, we list each representation of each distribution, but we focus on the Jacobi parameter representation.

Other laws may be characterized as being asymptotically Toeplitz or numerically Toeplitz fairly quickly, such as the limiting histogram of the eigenvalues of $(X/\sqrt{m} + \mu I)^t(X/\sqrt{m} + \mu I)$, where X is $m \times n$, n is $O(m)$, and $m \longrightarrow \infty$ (Fig. 3).

This nearly Toeplitz property inspires an underlying approximation concept. Instead of simply truncating a Jacobi matrix, we can construct an infinite Toeplitz matrix that is Toeplitz on all but an initial finite set of rows and columns. In turn, we apply this approximation idea to the moment problem in Section 5: given a finite set of moments,

we can use the Lanczos iteration to find corresponding Jacobi parameters, put those parameters in a continued fraction, and take its inverse Stieltjes transform to find a smooth, compactly-supported distribution with the correct moments. Instead of using the moments to find the Jacobi parameters we can also use a discretization of the measure. In [16], the Cauchy transform and its relationship to continued fractions of depth 2 are discussed, an idea that is generalized here in Table 9.

Mathematical investigations into weight functions with Toeplitz or asymptotically Toeplitz Jacobi parameters may be found in the work of two mathematicians with coincidentally similar names: Geronimus [7] and Geronimo [5], [6]. It seems known that the algorithm may have issues, perhaps reminiscent of the Gibbs phenomenon of Fourier analysis, in that atoms may emerge.

In retrospect, a good part of this paper may be found explicitly or implicitly in the work of Anshelevich [2]. Nonetheless, as we began to form the various tables and noticed the nice Catalan and Narayana properties, and especially the little box in the upper right in Table 2, we realized that in addition to the algorithm, there were enough ideas that were below the surface that we worked out for ourselves and wished to share.

2. The Jacobi symmetric tridiagonal encoding of probability distributions

All distributions have corresponding tridiagonal matrices of Jacobi parameters. They may be computed, for example, by the continuous Lanczos iteration, described in [17, p. 286] and reproduced in Table 8.

We computed the Jacobi representations of the four laws providing the results in Table 2. The Jacobi parameters (α_i and β_i for $i = 0, 1, 2, \ldots$) are elements of an infinite Toeplitz tridiagonal representations bordered by the first row and column, which may have different values from the Toeplitz part of the matrix.

Anshelovich [2] provides a complete table of six distributions that have Toeplitz Jacobi structure. The first three of which are semicircle, Marchenko–Pastur, and Wachter. The other three distributions occupy the same box as Wachter in Table 2. Anshelovich casts the problem as the description of all distributions whose orthogonal polynomials have generating functions of the form

$$\sum_{n=0}^{\infty} P_n(x)z^n = \frac{1}{1 - xu(z) + tv(z)},$$

which he calls Free Meixner distributions.

He includes the one and two atom forms of the Marchenko–Pastur and Wachter laws which correspond in random matrix theory to the choices of tall-and-skinny vs. short-and-fat matrices in the SVD or CS decompositions, respectively.

Table 4
Moments.

Measure	Moment n
Wigner semicircle	$C_{n/2}$ if n is even, 0 otherwise
Marchenko–Pastur	$N_n(\lambda)$
Kesten–McKay	$\sum_{j=1}^{n/2} \binom{n-j}{n/2}\left(\frac{j}{n-j}\right)v^j(v-1)^{n/2-j}$ if n is even, 0 otherwise
Wachter	$\frac{a}{a+b} - (a+b)\sum_{j=0}^{n-2}[(\frac{\sqrt{a(a+b-1)}}{a+b})^{2j+4}N_{j+1}(\frac{b}{a(a+b-1)})]$

Table 5
R-transforms and S-transforms computed as $S(z) = R^{-1}(z)/z$.

Measure	R-transform	S-transform
Wigner semicircle	w	1
Marchenko–Pastur	$\frac{\lambda}{1-w}$	$\frac{z-\lambda}{z^2}$
Kesten–McKay	$\frac{-v+v\sqrt{1+4w^2}}{2w}$	$\frac{v}{v^2-z^2}$
Wachter	$\frac{-a-b+w+\sqrt{(a+b)^2+2(a-b)w+w^2}}{2w}$	$\frac{a-az-bz}{z^2(z-1)}$

3. Infinite RMT laws

This section compares the properties of all four major infinite random matrix theory laws, the Wigner semicircle law, the Marchenko–Pastur law, the Kesten–McKay law, and the Wachter law.

We state the four laws of infinite dimensional random matrix theory and their intervals of support. μ_{WS} on I_{WS}, μ_{MP} on I_{MP}, μ_{KM} on I_{KM}, and μ_W on I_W correspond to the Wigner semicircle, Marchenko–Pastur, Kesten–McKay, and Wachter laws. These laws are originally credited to [19,13,14,18]. See Table 1.

We can also write down the moments for each measure in Table 4, for Wigner and Marchenko–Pastur see [4], for Kesten–McKay see [14], and for Wachter see Theorem 6.1 in Section 6. Remember the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ and the Narayana polynomial $N_n(r) = \sum_{j=1}^{n} N_{n,j} r^j$, where $N_{n,j} = \frac{1}{n}\binom{n}{j}\binom{n}{j-1}$, excepting $N_0(r) = 1$. The coefficients of $v^j(1-v)^{n/2-j}$ in the Kesten–McKay moments form the Catalan triangle. We discuss the pyramid created by the Wachter moments in Section 4.

Inverting the Cauchy transforms and subtracting $1/w$, computes the R-transform, see Table 5. If there are multiple roots, we pick one with a series expansion with no pole at $w = 0$.

The free cumulants κ_n for each measure appear in Table 6 by expanding the R-transform above (the generating function for the Narayana polynomials is given by [12], the generating function for the Catalan numbers is well known).

It is widely known that the Catalan numbers are the moments of the semicircle law, but we have not seen any mention that the same numbers figure prominently as the free cumulants of the Kesten–McKay law. The Narayana polynomials are prominent

Table 6
Free cumulants.

Measure	κ_n
Wigner semicircle	$\delta_{n,2}$
Marchenko–Pastur	λ
Kesten–McKay	$(-1)^{(n-2)/2} v C_{(n-2)/2}$ if n is even, 0 otherwise
Wachter	$-N_n(-\frac{b}{a})\frac{(-a)^{n+1}}{(a+b)^{2n+1}}$

Table 7
Sequences of polynomials orthogonal over of the four major laws.

Measure	$q_n(x)$, $n \geq 1$.
Wigner semicircle	$U_n(\frac{x}{2})$
Marchenko–Pastur	$\lambda^{(n-1)/2}(x-\lambda)U_{n-1}(\frac{x-\lambda-1}{2\sqrt{\lambda}}) - \lambda^{n/2}U_{n-2}(\frac{x-\lambda-1}{2\sqrt{\lambda}})$
Kesten–McKay	$(v-1)^{(n-1)/2}xU_{n-1}(\frac{x}{2\sqrt{v-1}}) - v(v-1)^{(n-2)/2}U_{n-2}(\frac{x}{2\sqrt{v-1}})$
Wachter	$(x-\frac{a}{a+b})(\frac{\sqrt{ab(a+b-1)}}{(a+b)^2})^{n-1}U_{n-1}(\frac{-b-a(a+b-1)+(a+b)^2x}{2\sqrt{ab(a+b-1)}})$
	$\quad - \frac{a+b}{a+b-1}(\frac{\sqrt{ab(a+b-1)}}{(a+b)^2})^n U_{n-2}(\frac{-b-a(a+b-1)+(a+b)^2x}{2\sqrt{ab(a+b-1)}})$

as the moments of the Marchenko–Pastur law, but they also figure clearly as the free cumulants of the Wachter law. There are well known relationships, involving Catalan numbers, between the moments and free cumulants of any law [15], but we do not know if the pattern is general enough to take the moments of one law, transform it somewhat, and have them show up in the free cumulants in another law.

We compute an S-transform as $S(z) = R^{-1}(z)/z$. See Table 5.

Each measure has a corresponding three-term recurrence for its orthonormal polynomial basis, with $q_{-1}(x) = 0$, $q_0(x) = 1$, $\beta_{-1} = 0$, and for $n \geq 0$, $q_{n+1}(x) = ((x - \alpha_n)q_n(x) - \beta_{n-1}q_{n-1}(x))/\beta_n$. In the case of the Wigner semicircle, Marchenko–Pastur, Kesten–McKay, and Wachter laws, the Jacobi parameters α_n and β_n are constant for $n \geq 1$ because they are all versions of the Meixner law [2] (a linear transformation may be needed). The Wigner semicircle case is given by simplifying the Meixner law in [1], and the Marchenko–Pastur, Kesten–McKay, and Wachter cases are given by taking two iterations Lanczos algorithm symbolically to get α_1 and β_1. See Table 2.

Each measure also has an infinite sequence of monic polynomials $q_n(x)$ which are orthogonal with respect to that measure. They can be written as sums of Chebyshev polynomials of the second kind, $U_n(x)$, which satisfy $U_{-1} = 0$, $U_0(x) = 1$, and $U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x)$ for $n \geq 1$, [11]. See Table 7. For $n = 0$, $q_0(x) = 1$, and in general for $n \geq 1$,

$$q_n(x) = \beta_1^{n-1}(x - \alpha_0)U_{n-1}\big((x - \alpha_1)/(2\beta_1)\big) - \beta_0^2\beta_1^{n-2}U_{n-2}\big((x - \alpha_1)/(2\beta_1)\big).$$

In the Wigner semicircle case the polynomials can be combined using the recursion rule for Chebyshev polynomials.

Table 8
The Lanczos iteration produces the Jacobi parameters in α and β.

<u>Lanczos on Measure μ</u>

$\beta_{-1} = 0$, $q_{-1}(x) = 0$, $q_0(x) = 1$
for $n = 0, 1, 2, \ldots$ **do**
    $v(x) = x q_n(x)$
    $\alpha_n = (q_n(x), v(x))$
    $v(x) = v(x) - \beta_{n-1} q_{n-1}(x) - \alpha_n q_n(x)$
    $\beta_n = \|v(x)\|$
    $q_{n+1}(x) = v(x)/\beta_n$
end for

Table 9
Algorithm recovering or approximating an analytic measure by Toeplitz matrices with boundary.

Algorithm: Compute Measure from Nearly Toeplitz Jacobi Matrix.

(1) Nearly Jacobi Toeplitz Representation: Run the continuous Lanczos algorithm up to step k, after which all α_i are equal and all β_i are equal, or very nearly so. If they are equal, this algorithm will recover $d\mu$ exactly, otherwise it will find it approximately. The Lanczos algorithm may be run using a discretization of the measure μ, or its initial moments.

$$(\alpha_{0:\infty}, \beta_{0:\infty}) = \text{Lanczos}(d\mu(x)).$$

(2) Cauchy transform: evaluate the finite continued fraction below on the interval of x where it is imaginary.

$$g(x) = \cfrac{1}{x - \alpha_0 - \cfrac{\beta_0^2}{x - \alpha_1 - \cfrac{\beta_1^2}{\ddots - \cfrac{\beta_{k-2}^2}{\alpha_{k-1} - \cfrac{2\beta_{k-1}^2}{x - \alpha_k + \sqrt{(\alpha_k - x)^2 - 4\beta_k^2}}}}}}.$$

(3) Inverse Cauchy transform: divide the imaginary part by $-\pi$, to compute the desired measure.

$$d\mu(x) = -\frac{1}{\pi}\text{Im}(g(x)).$$

4. The Wachter law moment pyramid

Using Mathematica we can extract an interesting number pyramid from the Wachter moments, see Fig. 1. Each triangle in the pyramid is formed by taking the coefficients of a and b in the i-th Wachter moment, with the row number within the pyramid determined by the degree of the corresponding monomial in a and b. All factors of $(a+b)$ are removed from the numerator and denominator beforehand and alternating signs are ignored.

Furthermore, there are many patterns within the pyramid. The top row of each triangle is a list of Narayana numbers, which sum to Catalan numbers. The bottom entries of each pyramid are triangular numbers. The second-to-bottom entry on the right of every pyramid is a sum of consecutive triangular numbers. The second to both the left and right on the top row of every triangle are also triangular numbers.
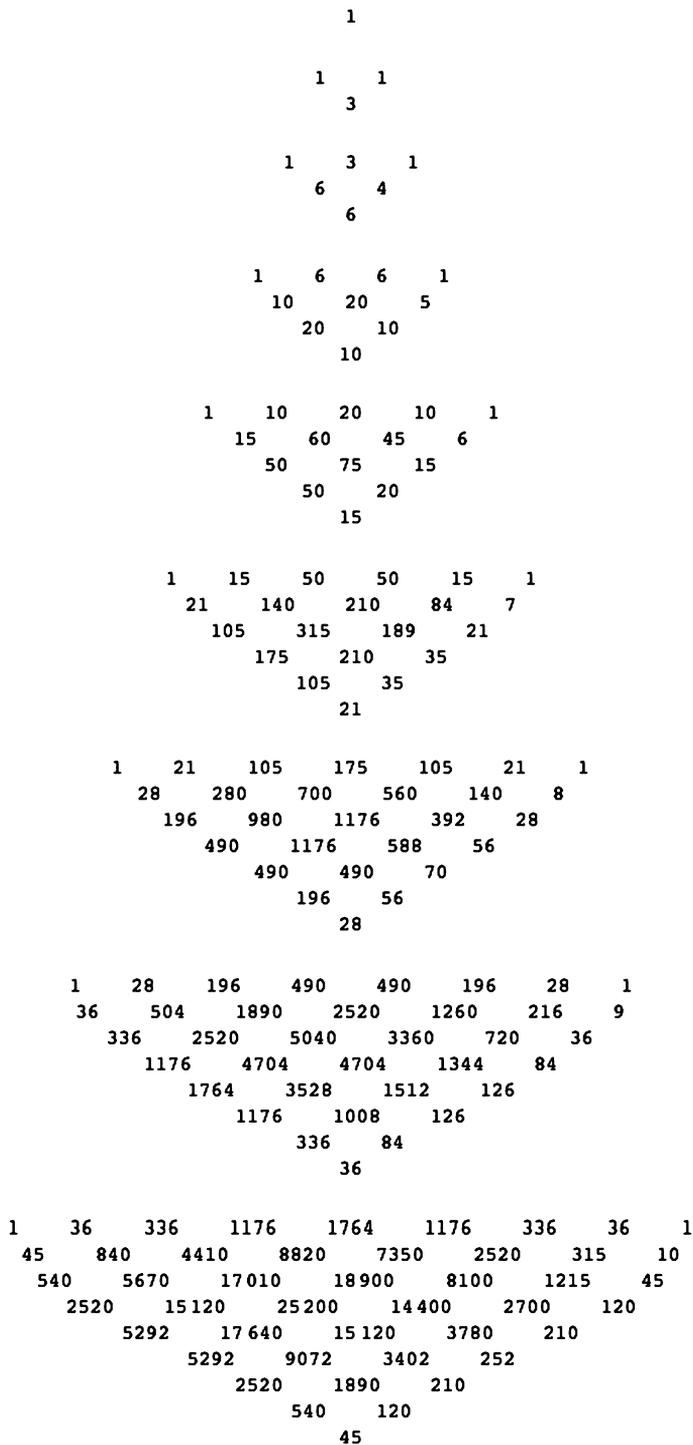
```
                              1

                         1        1
                              3

                     1        3        1
                          6        4
                          6

                 1        6        6        1
                     10       20       5
                         20       10
                         10

             1       10       20       10       1
                 15       60       45       6
                     50       75       15
                         50       20
                         15

         1       15       50       50       15       1
             21      140      210      84       7
                 105      315      189      21
                     175      210      35
                         105      35
                         21

     1       21      105      175      105      21       1
         28      280      700      560      140      8
             196      980      1176      392      28
                 490      1176      588      56
                     490      490      70
                         196      56
                         28

   1       28      196      490      490      196      28      1
       36      504      1890      2520      1260      216      9
           336      2520      5040      3360      720      36
               1176      4704      4704      1344      84
                   1764      3528      1512      126
                       1176      1008      126
                           336      84
                           36

 1      36      336      1176      1764      1176      336      36      1
    45      840      4410      8820      7350      2520      315      10
      540      5670      17010      18900      8100      1215      45
         2520      15120      25200      14400      2700      120
            5292      17640      15120      3780      210
               5292      9072      3402      252
                   2520      1890      210
                       540      120
                       45
```

Fig. 1. A number pyramid from the coefficients of the Wachter law moments.

5. Moments build nearly-Toeplitz Jacobi matrices

This section is concerned with recovering a probability distribution from its Jacobi parameters, α_i and β_i such that they are "nearly Toeplitz," i.e. there exists a k such that for $i \geq k$ all α_i are equal and all β_i are equal. Note that i ranges from k to ∞. The Jacobi parameters are found from a distribution by the Lanczos iteration.

We now state the continuous Lanczos iteration, replacing the matrix A by the variable x and using $\mu = \mu_{WS}, \mu_{MP}, \mu_M, \mu_W$ to compute dot products. A good source is [17]. For a given measure μ on an interval I, let

$$\big(p(x), q(x)\big) = \int_I p(x)q(x)d\mu,$$

and $\|p(x)\| = \sqrt{(p(x), p(x))}$. Then the Lanczos iteration is described by Table 8.

There are two ways to compute the integrals numerically. The first is to sample x and $q_n(x)$ at many points on the interval of support for $q_0(x) = 1$ and discretize the integrals on that grid. The second can be done if you know the moments of μ. If $r(x)$ and $s(x)$ are polynomials, $(r(x), s(x))$ can be computed given μ's moments. Since the $q_n(x)$ are polynomials, every integral in the Lanczos iteration can be done in this way. In that case, the $q_n(x)$ are stored by their coefficients of powers of x instead of on a grid. Once we have reached k iterations, we have fully constructed the infinite Jacobi matrix using the first batch of μ's moments, or a discretization of μ.

Step 1 can start with a general measure in which case Step 3 finds an approximate measure with a nearly Toeplitz representation. Step 1 could also start with a sequence of moments. It should be noted that the standard way to go from moments to Lanczos coefficients uses a Hankel matrix of moments and its Cholesky factorization [8, (4.3)].

As an example, we apply the algorithm to the histogram of the eigenvalues of $(X/\sqrt{m} + \mu I)^t(X/\sqrt{m} + \mu I)$, where X is $m \times n$, which has Jacobi parameters α_i and β_i that converge asymptotically and quickly. We smooth the histogram using a Gaussian kernel and then compute its Jacobi parameters. The reconstruction of the histogram is in Fig. 3. We also use the above algorithm to reconstruct a normal distribution from its first sixty moments, see Fig. 4.

The following theorem concerning continued fractions allows one to stably recover a distribution from its Lanczos coefficients α_i and β_i. As we have said, if the first batch of μ's moments are known, we can find all α_i and β_i from $i = 0$ to ∞ using the continuous Lanczos iteration.

Theorem 5.1. *Let μ be a measure on interval $I \subset \mathbb{R}$ with Lanczos coefficients α_i and β_i, with the property that all α_i are equal for $i \geq k$ and all β_i are equal for $i \geq k$. We can recover $I = [\alpha_k - 2\beta_k, \alpha_k + 2\beta_k]$, and we can recover $d\mu(x)$ using a continued fraction. This theorem combines Theorems 1.97 and 1.102 of [9].*
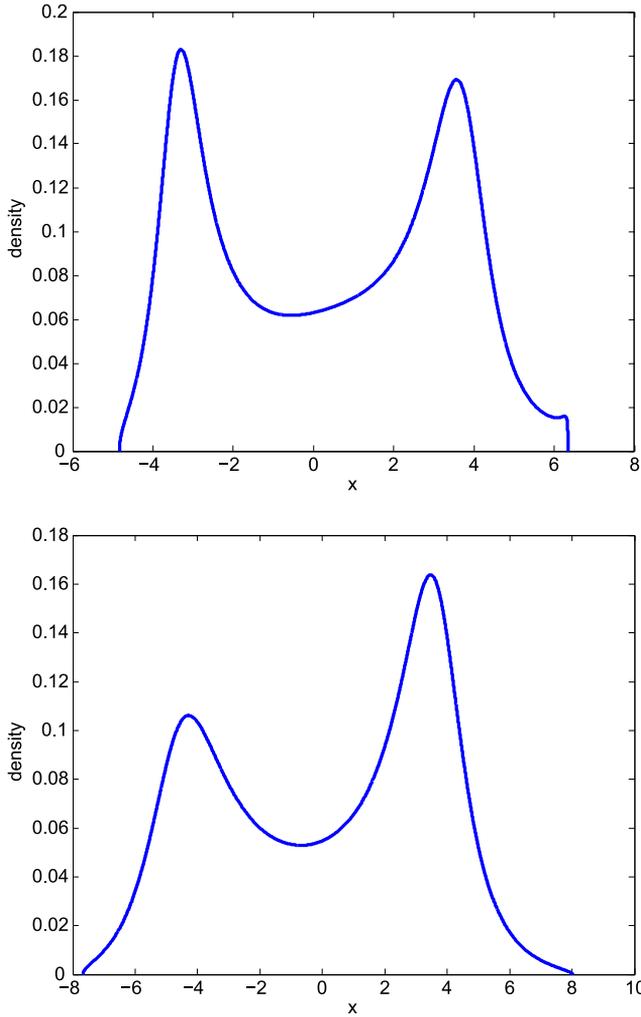
Fig. 2. Recovery from of a distribution from random α_i and β_i using Theorem 5.1. On top we use $k = 5$, on bottom we use $k = 3$.

$$g(x) = \cfrac{1}{x - \alpha_0 - \cfrac{\beta_0^2}{x - \alpha_1 - \cfrac{\beta_1^2}{\ddots - \cfrac{2\beta_{k-1}^2}{x - \alpha_k + \sqrt{(\alpha_k - x)^2 - 4\beta_k^2}}}}}$$

$$d\mu(x) = -\frac{1}{\pi}\mathrm{Im}\big(g(x)\big).$$

Fig. 2 illustrates curves recovered from random terminating continued fractions $g(x)$ such that the β_i are positive and greater in magnitude than the α_i. In both cases, the above theorem allows correct recovery of the α_i and β_i (which is not always numerically possible). In the first one, $k = 5$, in the second, $k = 3$.

Fig. 3. Eigenvalues taken from $(X/\sqrt{m} + \mu I)^t(X/\sqrt{m} + \mu I)$, where X is $m \times n$, $m = 10^4$, $n = 3m$, $\mu = 5$. The blue bar histogram is taken using hist.m, a better one was taken by convolving the data with Gaussian kernel. That convolution histogram was used to initialize the continuous Lanczos algorithm which produced five α's and β's. They were put into a continued fraction as described above, assuming α_i and β_i to be constant after $i = 5$. The continued fraction recreated the histogram, which is the thick red line.

If X is an $m \times n$, $m < n$ matrix of normals for m and n very large, $(X/\sqrt{m} + \mu I)^t(X/\sqrt{m} + \mu I)$ has α_i and β_i which converge to a constant, making its eigenvalue distribution recoverable up to a very small approximation. See Fig. 3.

We also tried to reconstruct the normal distribution, whose Jacobi parameterization is not at all Toeplitz, and which is not compactly supported. Fig. 4 plots the approximations using 10 and 20 moments.

6. Direct computation of the Wachter law moments

While the moments of the Wachter law may be obtained in a number of ways, including expanding the Cauchy transform, or applying the mobius inverse formula to the free cumulants, in this section we show that a direct computation of the integral is possible.

Theorem 6.1. *We find the moments of the Wachter law, m_k.*

$$m_k = \frac{a}{a+b} - (a+b) \sum_{j=0}^{k-2} \left[\left(\frac{\sqrt{a(a+b-1)}}{a+b} \right)^{2j+4} N_{j+1} \left(\frac{b}{a(a+b-1)} \right) \right].$$

Proof. We start by integrating the following expression by comparing it to the Marchenko–Pastur law.
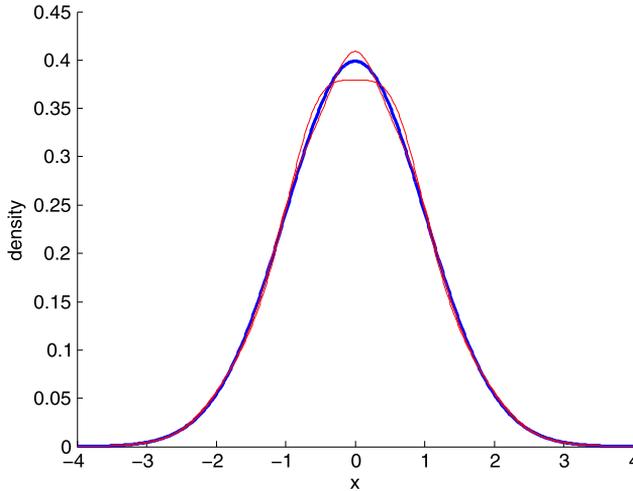
Fig. 4. The normal distribution's Jacobi matrix is not well approximated by Toeplitz plus boundary, but with sufficiently many moments good approximations are possible. The above graph shows the normal distribution recovered by the method in this paper using 10 and 20 moments. The thick line is the normal computed by $e^{-x^2/2}\sqrt{2\pi}$, and the thin lines on top of it use our algorithm.

$$J_1 = \frac{1}{2\pi} \int\limits_{\mu_-}^{\mu_+} x^k \sqrt{(\mu_+ - x)(x - \mu_-)}dx.$$

If $x = su$, $dx = sdu$ and this integral becomes

$$J_1 = \frac{s^{k+2}}{2\pi} \int\limits_{\frac{\mu_-}{s}}^{\frac{\mu_+}{s}} u^k \sqrt{\left(\frac{\mu_+}{s} - u\right)\left(u - \frac{\mu_-}{s}\right)}du.$$

To compare this expression to the Marchenko–Pastur law, we need to pick s and λ such that $\frac{\mu_+}{s} = (1 + \sqrt{\lambda})^2$ and $\frac{\mu_-}{s} = (1 - \sqrt{\lambda})^2$ for $\lambda \geq 1$. There are more than one choices of each parameter, but we pick $\sqrt{s} = \frac{1}{2}(\sqrt{\mu_+} - \sqrt{\mu_-})$ and

$$\sqrt{\lambda} = \frac{\sqrt{\mu_+} + \sqrt{\mu_-}}{\sqrt{\mu_+} - \sqrt{\mu_-}}.$$

Using the Narayana numbers, and the formula for the moments of the Marchenko–Pastur law, the integral equals

$$J_1 = \left(\frac{1}{2}(\sqrt{\mu_+} - \sqrt{\mu_-})\right)^{2k+4} N_{k+1}\left(\left(\frac{\sqrt{\mu_+} + \sqrt{\mu_-}}{\sqrt{\mu_+} - \sqrt{\mu_-}}\right)^2\right).$$

Using a and b, this becomes

$$J_1 = \left(\frac{\sqrt{a(a+b-1)}}{a+b}\right)^{2k+4} N_{k+1}\left(\frac{b}{a(a+b-1)}\right).$$

We also need to integrate

$$J_2 = \frac{1}{2\pi} \int\limits_{\mu_-}^{\mu_+} \frac{\sqrt{(\mu_+ - x)(x - \mu_-)}}{1 - x} dx$$

Let $su = x - 1$, $sdu = dx$. This becomes

$$J_2 = -\frac{s}{2\pi} \int\limits_{\frac{\mu_- - 1}{s}}^{\frac{\mu_+ - 1}{s}} \frac{\sqrt{(\frac{\mu_+ - 1}{s} - u)(u - \frac{\mu_- - 1}{s})}}{u} du,$$

which by symmetry is

$$J_2 = \frac{s}{2\pi} \int\limits_{\frac{1 - \mu_+}{s}}^{\frac{1 - \mu_-}{s}} \frac{\sqrt{(\frac{1 - \mu_-}{s} - u)(u - \frac{1 - \mu_+}{s})}}{u} du.$$

Using the same technique as previously, $\sqrt{s} = \frac{1}{2}(\sqrt{1 - \mu_-} - \sqrt{1 - \mu_+})$ and

$$\sqrt{\lambda} = \frac{\sqrt{1 - \mu_+} + \sqrt{1 - \mu_-}}{\sqrt{1 - \mu_-} - \sqrt{1 - \mu_+}}.$$

Using the fact that the Marchenko–Pastur law is normalized, the answer is

$$J_2 = s = \frac{1}{4}(\sqrt{1 - \mu_-} - \sqrt{1 - \mu_+})^2 = \frac{a}{(a + b)^2}.$$

Now we are ready to find the moments of the Wachter law. Using the geometric series formula,

$$m_k = \frac{a + b}{2\pi} \int\limits_{\mu_-}^{\mu_+} \frac{x^{k-1}\sqrt{(\mu_+ - x)(x - \mu_-)}}{1 - x} dx$$

$$= \frac{a + b}{2\pi} \sum_{j=k-1}^{\infty} \int\limits_{\mu_-}^{\mu_+} x^j \sqrt{(\mu_+ - x)(x - \mu_-)} dx$$

$$= \frac{a + b}{2\pi} \int\limits_{\mu_-}^{\mu_+} \frac{\sqrt{(\mu_+ - x)(x - \mu_-)}}{1 - x} dx - \frac{a + b}{2\pi} \sum_{j=0}^{k-2} \int\limits_{\mu_-}^{\mu_+} x^j \sqrt{(\mu_+ - x)(x - \mu_-)} dx$$

$$= \frac{a}{a + b} - (a + b) \sum_{j=0}^{k-2} \left[\left(\frac{\sqrt{a(a + b - 1)}}{a + b} \right)^{2j+4} N_{j+1} \left(\frac{b}{a(a + b - 1)} \right) \right]. \qquad \square$$

Acknowledgements

References

[1] Michael Anshelevich, Wojciech Młotkowski, The free Meixner class for pairs of measures, http://arxiv.org/abs/1003.4025, 2011.
[2] Michael Anshelevich, Bochner–Pearson-type characterization of the free Meixner class, Adv. in Appl. Math. 46 (2011) 25–45.
[3] Percy Deift, Orthogonal Polynomials and Random Matrices: A Riemann–Hilbert Approach, Courant Lecture Notes in Mathematics, 1998.
[4] Alan Edelman, Random matrix theory, in preparation.
[5] J. Geronimo, K.M. Case, Scattering theory and polynomials orthogonal on the real line, Trans. Amer. Math. Soc. 258 (2) (April 1980).
[6] J. Geronimo, P. Iliev, Bernstein–Szegő measures, Banach algebras, and scattering theory, arXiv.
[7] Ja.L. Geronimus, Orthogonal polynomials, Amer. Math. Soc. Transl. Ser. 2 108 (1977).
[8] G.H. Golub, J.A. Welsch, Calculation of Gauss quadrature rules, Math. Comp. 23 (1969) 221.
[9] Akihito Hora, Nobuaki Obata, Quantum Probability and Spectral Analysis of Graphs, Theoretical and Mathematical Physics, Springer, Berlin, Heidelberg, 2007.
[10] Harry Kesten, Symmetric random walks on groups, Trans. Amer. Math. Soc. 92 (1959) 336–354.
[11] Timothy Kusalik, James A. Mingo, Roland Speicher, Orthogonal polynomials and fluctuations of random matrices, 2005, arXiv.
[12] T. Mansour, Y. Sun, Identities involving Narayana polynomials and Catalan numbers, Discrete Math. 309 (2009) 4079–4088.
[13] V.A. Marchenko, L.A. Pastur, Distribution of the eigenvalues in certain sets of random matrices, Mat. Sb. 72 (114) (1967).
[14] Brendan D. McKay, The expected eigenvalue distribution of a large regular graph, Linear Algebra Appl. 40 (1981) 203–216.
[15] Alexandru Nica, Roland Speicher, Lectures on the Combinatorics of Free Probability, Cambridge University Press, 2006.
[16] N. Raj Rao, Alan Edelman, The polynomial method for random matrices, Found. Comput. Math. (2007), in press.
[17] Lloyd N. Trefethen, David Bau III, Numerical Linear Algebra, SIAM, 1997.
[18] Kenneth W. Wachter, The strong limits of random matrix spectra for sample matrices of independent elements, Ann. Probab. 6 (1978).
[19] Eugene P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions, Ann. of Math. 62 (1955).