Some applications of moments in RMT

1. Preliminaries

Let $\mu$ be a measure on $\mathbb{R}$; denote

$$M_k(\mu) = \int x^k \, d\mu, \quad k = 0, 1, 2, \ldots$$

We assume that these numbers (the moments of $\mu$) are finite.

Problem of moments:

given $m_0, m_1, \ldots, m_n$ ($n$ may be finite or $+\infty$),

(a) does there exist $\mu$ so that $\int x^k \, d\mu = m_k$?

(b) what can one say about $\mu$, e.g.

$$\max_{\mu} \mu(-\infty, x], \quad \min_{\mu} \mu(-\infty, x].$$

(a): "easy": $\iff$ for every positive polynomial $P(x) = \sum_{k=0}^{n} a_k x^k$,

$$\sum a_k M_k > 0 \quad (\iff \text{some determinants are positive}).$$

(b): harder and more useful. Typically, we want to extract information about $\mu$ from $m_0, \ldots, m_n$ (which are precisely or approximately known). [Origin: Chebyshev's approach to the CLT; Markov, …] Shelkys, …

Prop. 1 If $\mu$ is supported on a finite interval and

$$M_k(\nu) = M_k(\mu), \quad k = 0, 1, 2, \ldots$$

then $\nu = \mu$.

Proof (a) $M_k(\mu) \leq C^k M_k(\mu)$, hence $\nu$ is also supported on $[-C, C]$. 

(b) Any continuous $f$ on $[-C, C]$ can be approximated by pols (in uniform topology). Hence

$$\int f \, d\nu = \int f \, d\mu, \quad f \in C[-C, C] \implies \nu = \mu.$$

$\square$
Prop. 2. If \( \mu \) is such that \( \int e^{\varepsilon x} \, d\mu(x) < \infty \) for some \( \varepsilon > 0 \), and \( \mu_k(n) = \mu(n) \), \( k=0,1,2, \ldots \), then \( \nu = \mu \).

Hint: if \( \int e^{i\xi x} \, d\mu(x) = \int e^{i\xi x} \, d\nu(x) \), \( \xi \in \mathbb{R} \), then \( \mu = \nu \).

More precise criteria are available, e.g.: Carleman's condition in the positive, and Krein's condition in the negative. See e.g. Akhiezer's book.

Let \( \mu \) be a measure with finite moments. Orthogonalize \( 1, x, x^2, x^3, \ldots \) in \( L^2(\mu) \). We obtain the sequence \( P_0(x), P_1(x), \ldots \) of orthogonal polynomials w.r.t. to \( \mu \).

Ex. 1. \( \mu = \delta' \) (the standard Gaussian measure)
   \( \rightarrow \) Hermite polynomials

2. Wigner's measure \( d\mu_w(x) = \frac{2}{\pi} (1-x^2)^{1/2} \, dx \)
   \( \rightarrow \) Chebyshev polynomials of the second kind,
   \( U_k(x) = \sin((k+1)x)/\sin x \)

Hint: Check that \( \int U_k(x) U_l(x) \, d\mu = 0 \) using the change of variables \( x = \cos \theta \).

3. The Kasten-McKay measure \( d\mu_{KM}(x) = \frac{2d(d+1)}{\pi} \frac{(1-x^2)^{1/2} \, dx}{d^2 - 4(d-1)x^2} \)
   \( \rightarrow \) \( P_k(x) = \sqrt{\frac{d}{d-1}} \left[ U_k(x) - \frac{1}{d-1} U_{k-2}(x) \right] \).

Prop. 3. \( x \, P_k(x) = A_k \, P_{k+1}(x) + B_k \, P_k(x) + C_k \, P_{k-1}(x) \)

Hint: by induction.

Remark. For \( \mu_w \) (and hence also for \( \mu_{KM} \)), \( A_k = \frac{1}{2}, B_k = 0, C_k = 0 \).
Digression \( A_\xi, B_\xi, C_\xi \) can be written in a matrix (Jacobi matrix):

\[
\Sigma = \begin{pmatrix}
B_1 & C_1 \\
A_1 & B_2 & C_2 \\
& A_2 & B_3 & \ddots
\end{pmatrix}
\]

\( \Sigma \) represents multiplication by \( \xi \) in \( L^2(\mu) \).

(Direct) spectral problem: given \( A_\xi, B_\xi, C_\xi \), find \( \mu \).

(Inverse) spectral problem: given \( \mu \), find \( A_\xi, B_\xi, C_\xi \).

Sometimes (in RMT), one wishes to find the spectral measure of an operator \( H \), and one manages to find a tridiagonal matrix isospectral to it. Then one needs to solve the direct spectral problem.

See e.g. Dumitriu & Edelman, and further works.
Random matrices: the global regime

Given an $N \times N$ matrix $H$ and $\mu_H = \frac{1}{N} \sum \delta_{\lambda_i}$, where $\lambda_i$ are the eigenvalues of $H$. If $H$ is Hermitian, this is a measure on $\mathbb{R}$.

Now let $H_N$ be a sequence of matrices of increasing size. Set $\mu_N = \mu_{H_N}$. Question: $\mu_N \xrightarrow{w} ?$

Next, assume that $H_N$ are random. Then $\mu_N$ are also random (ie prob. distributions on the space of measures).

$\Omega \xrightarrow{w} ?$ Is the limit random or deterministic?

A Wigner matrices

- $\{H_N(u,v) \mid u < v\}$ are independent
- $\mathbb{E} H_N(u,v) = 0$; if $u < v$, $\mathbb{E} H_N(u,v)^2 = 1$
- $\mathbb{E} |H_N(u,v)|^k \leq A_k < \infty$.

We shall consider the spectral measures $\mu_N$ of $\frac{H_N}{\sqrt{N}}$ (this is just scaling).

Thus (Wigner) $\mu_N \xrightarrow{w} \mu_W$ (the limit is not random).

Plan of proof:

1) Consider $|E|_{\mu_N}$ and prove that $\int x^k d|E|_{\mu_N} \rightarrow \int x^k d\mu_W$ for $k = 0, 1, 2, \ldots$

In particular, the prob. measures $|E|_{\mu_N}$ have uniform exp. law and variance $\Rightarrow \{ |E|_{\mu_N} \}$ is precompact. By Prop. 1, $\mu_N \xrightarrow{w} \mu_W$. The measure $|E|_{\mu_W}$ is equal to $\mu_W$. Therefore, $\mu_W \xrightarrow{w} \mu_W$. 
(ii) Var $\sum x^k d\mu_N \to 0$, $k=0,1,2,\ldots$

Therefore $\mu_N$ are concentrated around the mean $\mathbb{E}\mu_N$, and hence also $\frac{\mu_N}{V_N} \to \mu$.

Now to business.

Proof: $\int x^k d\mu_N = \mathbb{E} \sum x^k d\mu_N = \mathbb{E} \frac{1}{N} \sum (\frac{H_N}{N})^k = 

= \frac{1}{2^k} \cdot \frac{1}{N^{k/2}} \sum \mathbb{E} H_N(u_0, u_1) H_N(u_1, u_2) \cdots H_N(u_k, u_0).

Every addend corresponds to a path $p_k$ in the complete graph $K_N$:

\[ u_0 \xrightarrow{u_1} u_2 \xrightarrow{u_3} u_4 \]

\[ \quad = p_k \quad \text{e.g.} \; 7 4 4 2 8 4 7 \]

We group the paths into equiv. classes with respect to the action of $S_N$.

NB The number of equivalence classes depends only on $k$

(and not on $N$).

On every edge $e = (u, v)$, we have a rv. $H_N(e) = H_N(u, v)$

$= H_N(u_i, u_i)$. If and edge $e$ appears $k(e)$ times,

\[ \mathbb{E} H_N(u_0, u_1) \cdots H_N(u_k, u_0) = \prod_{e} \mathbb{E} H_N(e)^{k(e)} \]

In particular, if $k(e) = 1$, the expectation is zero.

Consider the graph $G = (V, E)$ induced by $p_k$; it is connected, hence $\#V < \#E + 1 < \frac{k}{2} + 1$, with equality for trees.
Example:

Thus the contr. of the equiv. class of $G$ is

$$N \cdot \frac{1}{2^k N^{1+\frac{k}{2}}} \cdot A_k \leq A'_k;$$

if $G$ is not a tree: $\frac{A'_k}{N} \xrightarrow{N \to \infty} 0$. Thus we only consider trees.

Fact

# tree-like paths

of length $k = \begin{cases} 0, & k \text{ is odd} \\ \frac{k!}{(\frac{k}{2})!(\frac{k}{2}+1)!}, & \text{otherwise} \end{cases} = C_k$

(Catalan's number).

Hint

$E_g \leftrightarrow$ arrangements of brackets $(((\cdot)(\cdot))\cdot(\cdot))$

"(•" : new vertex, "\cdot\)" : back to prev. vertex.

Hence

$$\sum_k x^k E_d \mu_N \xrightarrow{\frac{C_k}{2^k}} \sum_k x^k d_j \mu_N(x).$$

Elementary computation by induction

(ii): Similar comb. argument for pairs of paths

the contribution to

$$\mathbb{E} \left[ \frac{1}{N} \text{tr}(H_i^{1/2}N) \cdot \frac{1}{N} \text{tr}(H_i^{1/2}N)^e \right]$$

comes from pair of disjoint trees.)
Consider a sequence of graphs \( \{G_N = (V_N, E_N)\} \), \( \#V_N = N \).

Denote \( \mu_N \) the spectral measure of \( A_N^{d-1} \), where \( A_N \) is the adjacency matrix of \( G_N \).

Then \( \mu_N \xrightarrow{w} \mu_{KM}^{d} \).

One can use moments as before and show that only trees contribute. Let us show a modification of this argument.

Let \( B_{k,N} \) be an \( N \times N \) matrix,

\[ B_{k,N}(u,v) = \# \text{ paths of length } k \text{ from } u \text{ to } v \text{ in } G_N \text{ without backtracking} \]

(eg \( u \xrightarrow{1} v \) is ok, but not \( u \xrightarrow{1} v \) )

Lemma: \( B_{k/N} \cdot A_N = B_{k+1,N} + (d-1) B_{k-1,N} \)

Proof: 

\( u \xrightarrow{w} v \) \( \Rightarrow \) last step back \( \Rightarrow B_{k+1,N} \)

\( u \xrightarrow{w} v \) \( \Rightarrow \) no last step \( \Rightarrow B_{k-1,N} \)
Therefore $B_{k,N}$ is a polynomial of $A_N$. Observe that, for every $k>0$, $tv B_{k,N}$ is eventually zero.

Now we are almost done:

$$\frac{B_{k,N}}{V d(\delta+1)^{k-1}} = P_k \left( \frac{A'}{2\sqrt{d-1}} \right), \quad \text{and}$$

$$Qx P_k(x) = P_{k+1}(x) + P_{k-1}(x).$$

Checking the cases $k=0,1,2$, one may see that there are exactly the orthogonal pols w.r. to $\mu_{kM}^d$. Thus:

$$\frac{1}{N} \text{tr} \frac{B_{k,N}}{V d(\delta+1)^{k-1}} = \int \sum_{k=1}^{\infty} \frac{d\mu_{kM}^d}{\mu_{kM}^d} = \int P_k \cdot 1 \ d\mu_{kM}^d, \quad k=1,2,\ldots$$

Therefore $\mu_N \rightarrow \mu_{kM}^d$.

Remarks:

1) One can put $\pm 1$-s on the edges

"weighted" adj. matrix

2) also, one may allow cycles, as long as there are "not many" of them.

E.g.: a random $d$-regular graph is fine.

3) The argument can be also used to prove Wigner's thm. this is easy for

$$H_N(u,v) = \begin{cases} \pm 1, & u < v \\ 0, & u = v \end{cases}.$$

4) (unlike moments), one can deduce a reasonable estimate on the rate of convergence.