30.4.07

Random matrix theory: lecture 13

Wigner's theorem

Preliminary: the semi-circle distribution and the Catalan numbers

Let μ be the distribution on $\mathbb{R}$ with pdf

$$p_\mu(x) = \frac{1}{2\pi}\sqrt{4-x^2}\cdot 1_{|x|\leq 2} \qquad x\in\mathbb{R}$$

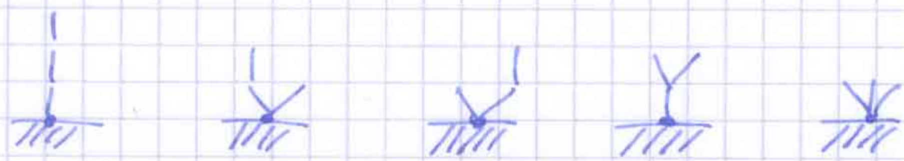
All moments of μ are finite and given by (homework)

$$m_{2k+1}=0\,,\quad m_{2k}=\frac{(2k)!}{k!\,(k+1)!}=\frac{1}{k+1}\binom{2k}{k}:=t_k\quad,\ k\geq 0 \qquad (*)$$

t_k are called the Catalan numbers; there are several combinatorial characterizations of these numbers. Let us mention two:

- t_k is the number of planar planted rooted trees with k branches

  illustration for $k=3$:



$$t_3=\frac{6!}{3!\,4!}=5$$

(*) Note moreover that

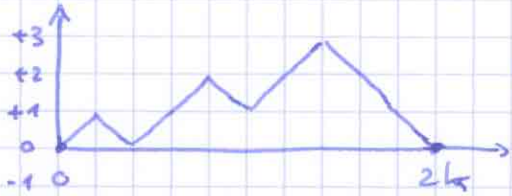
$$t_k\leq\frac{(2k)!}{(k!)^2}\leq\frac{\left((2k)(2k-2)\cdots 2\right)^2}{(k!)^2}\leq\left(\frac{2^k k!}{k!}\right)^2\leq 4^k$$

so the sequence (m_k) satisfies Carleman's condition.

30.4.07

Def: a Dyck path of length $2k$ is a path starting at zero and ending at zero in $2k$ steps, going up or down a each step, and conditioned to stay non-negative over the whole period.

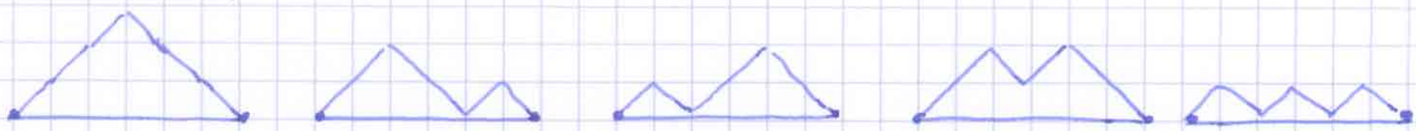
illustration:



a Dyck path

- t_k is the number of Dyck paths of length $2k$

illustration for $k=3$: $t_3 = 5$ again



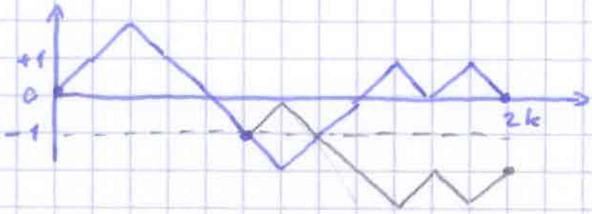
Identification between the two descriptions:

explore the tree from the root and go up and down along the branches; the record of these ups and downs is the corresponding Dyck path.

Counting Dyck paths: the reflexion principle

The number of Dyck paths of length $2k$ is the number of all paths from $(0,0)$ to $(2k, 0)$ minus the number of paths crossing the -1 horizontal line.

Note that to each of these paths crossing the -1 line corresponds a symmetric path reaching $(2k, -2)$ at the end:



reflexion principle

Note moreover that <u>all</u> paths from $(0,0)$ to $(2k,-2)$

× necessarily cross the horizontal line -1. Therefore, the number of Dyck paths of length $2k$ is:

×
$$\underbrace{\binom{2k}{k}}_{\substack{\#\text{ all paths}\\ \text{from 0 to 0}\\ (k \text{ downs})}} - \underbrace{\binom{2k}{k+1}}_{\substack{\#\text{ all paths}\\ \text{from 0 to }-2\\ (k+1 \text{ downs})}} = \frac{(2k)!}{k! \cdot k!} - \frac{(2k)!}{(k+1)!(k-1)!}$$

$$= \frac{(2k)!}{k!\, k!}\left(1 - \frac{k}{k+1}\right) = \binom{2k}{k}\frac{1}{k+1}$$

$$= t_k \quad \checkmark$$

We have therefore shown that the Catalan numbers are indeed the number of Dyck paths of length $2k$, or equivalently, the number of planar planted rooted trees with k branches.

We consider now the following random matrix ensemble: let H be a $n\times n$ real symmetric matrix such that

(i) $\{h_{jk}, j\leq k\}$ are i.i.d. random variables (and $h_{kj} := h_{jk}$)

× (ii) all moments of h_{11} are finite

(iii) $\mathbb{E}(h_{11}{}^{2\ell+1}) = 0 \quad \forall \ell \geq 0$

(iv) $\mathbb{E}(h_{11}{}^2) = 1$

Let moreover $H^{(n)} = \frac{1}{\sqrt{n}} H$ and $\lambda_1^{(n)} .. \lambda_n^{(n)}$ be the eigenvalues of $H^{(n)}$.

Theorem (Wigner, 1955)

$$F_n(t) := \frac{1}{n} \# \{ j : \lambda_j^{(n)} \leq t \} \xrightarrow[n\to\infty]{} \int_{-\infty}^{t} p_\mu(x)\, dx \quad \text{a.s. } \forall t\in\mathbb{R}$$

where

$$p_\mu(x) := \frac{1}{2\pi} \sqrt{4-x^2}\, 1_{|x|\leq 2} \quad \text{semi-circle distribution}$$

Remarks

- examples of distributions of h_{11} satisfying the assumptions are: $h_{11} \sim N_{\mathbb{R}}(0,1)$, $h_{11} = \pm 1$ w.p. $\frac{1}{2}$, $h_{11} \sim \mathcal{U}([-\sqrt{3}, \sqrt{3}])$, but the limit does not depend on this distribution; it actually only depends on its variance ($\mathbb{E}(h_{11}^2)$), so the result is universal (c.f. central limit theorem)

- the assumptions do not include the GOE, because of (i), that assumes iid diagonal and off-diagonal entries. It can actually be shown that the distribution of the diagonal entries does not influence the limit (as soon as these entries have bounded variance).
- there is another qualitative difference between the above theorem and the result that one obtains from finite-size analysis. The present theorem says that for any function $f : \mathbb{R} \to \mathbb{R}$ bounded and continuous

X $$\frac{1}{n}\sum_{j=1}^{n} f(\lambda_j^{(n)}) \xrightarrow[n\to\infty]{} \underbrace{\int_{\mathbb{R}} f(x)\, p_\mu(x)\, dx}_{\underline{NB}:\text{ deterministic limit!}} \quad a.s.$$

(c.f. definition of weak convergence), whereas the result obtained from finite-size analysis only says that

$$\mathbb{E}\left(\frac{1}{n}\sum_{j=1}^{n} f(\lambda_j^{(n)})\right) = \int_{\mathbb{R}} f(x)\, \rho^{(n)}(x)\, dx$$
$$\xrightarrow[n\to\infty]{} \int_{\mathbb{R}} f(x)\, p_\mu(x)\, dx.$$

- in general, the iid assumption may be relaxed in various ways; the important fact is that entries are independent with the same variance.

Proof of the theorem

- For convenience, we will prove the theorem under the slightly stronger assumption:

× (ii)' h_{11} is a bounded random variable

(i.e. $\exists C > 0$ such that $|h_{11}| \leq C$ a.s.)

- The technique used for proving weak convergence is via moments, i.e., we will show that $\forall \ell \geq 0$,

× $$m_\ell^{(n)} := \frac{1}{n} \sum_{j=1}^{n} \left(\lambda_j^{(n)}\right)^\ell = \frac{1}{n} \operatorname{Tr}\left(\left(H^{(n)}\right)^\ell\right) \xrightarrow[n\to\infty]{} m_\ell \text{ a.s.}$$

where m_ℓ are the moments of the semi-circle distribution.

- Given the criterion of the last lecture for a.s. convergence, it is sufficient to show that $\forall \ell \geq 0$,

1) $\left| \mathbb{E}\left(m_\ell^{(n)}\right) - m_\ell \right| = O\left(\frac{1}{n}\right)$
2) $\operatorname{Var}\left(m_\ell^{(n)}\right) = O\left(\frac{1}{n^2}\right)$

We will focus on the proof of (1) in the following, as the technique for proving (2) is similar.

- $\mathbb{E}(m_\ell^{(n)}) = \frac{1}{n}\,\mathbb{E}\left(\mathrm{Tr}\left((H^{(n)})^\ell\right)\right) = \frac{1}{n^{1+\ell/2}}\,\mathbb{E}\left(\mathrm{Tr}(H^\ell)\right)$

$$= \frac{1}{n^{1+\ell/2}} \sum_{j_1 \cdots j_\ell = 1}^{n} \mathbb{E}\left(h_{j_1 j_2} h_{j_2 j_3} \cdots h_{j_\ell j_1}\right)$$

- First note that $\mathbb{E}(m_{2\ell+1}^{(n)}) = 0 = m_{2\ell+1}$ $\forall \ell \geq 0$:

Indeed, $m_{2\ell+1} = 0$ since the semi-circle distribution is symmetric around zero.

Moreover, in the expectation of any of the products $\mathbb{E}(h_{j_1 j_2} h_{j_2 j_3} \cdots h_{j_{2\ell+1} j_1})$, at least one of the entries h_{jk} appears an odd number of times (we identify here h_{jk} and h_{kj}), so by the independence assumption and the assumption that $\mathbb{E}(h_{jk}^{2\ell+1}) = 0$ $\forall \ell \geq 0$, we obtain that the overall expectation is zero.

- Let us therefore focus our attention on

$$\mathbb{E}(m_{2\ell}^{(n)}) = \frac{1}{n^{1+\ell}} \sum_{j_1 \cdots j_{2\ell} = 1}^{n} \mathbb{E}\left(h_{j_1 j_2} h_{j_2 j_3} \cdots h_{j_{2\ell} j_1}\right)$$

We will see next time that $|\mathbb{E}(m_{2k}^{(n)}) - t_k| = O(\frac{1}{n})$, where t_k are the Catalan numbers ($= m_{2k}$). This will conclude the proof of (1).