Fluctuations from the Semicircle Law
Lecture 3

Ioana Dumitriu

University of Washington

Women and Math, IAS 2014

May 22, 2014

1. Higher Moments

Higher Moments

We have calculated the variance; now we move on to higher moments.

Examine

$$\mathbb{E}(X_{n,k}^l) = \mathbb{E}\left(\left(\text{tr}\left(W_n^k\right) - \mathbb{E}\left(\text{tr}(W_n^k)\right)\right)^l\right) .$$

This is the same as

$$\mathbb{E}(X_{n,k}^l) = \sum_{I_1,I_2,\ldots,I_l \in \mathcal{I}} \mathbb{E}\left(\prod_{j=1}^{l}(w_{I_j} - \mathbb{E}(w_{I_j}))\right) .$$

The same ideas apply.

The lth Moment

Consider again the graphs $G_{I_1}, \ldots, G_{I_l}$ as well as the union graph $\mathcal{G}$.

- If any of the edges in $\mathcal{G}$ appears only once in the union of the walks given by $I_1, \ldots, I_l$, the term has 0 contribution.
- Thus every edge has multiplicity 2 or more in the union of the walks.
- $\forall\, 1 \leq j \leq l$, $\exists\, 1 \leq \tilde{j} \leq l$ such that I_j and $I_{\tilde{j}}$ overlap, otherwise independence, 0 contribution.

The lth Moment

- Thus $\mathcal{G}$ has $c \leq \lfloor l/2 \rfloor$ connected components.
- What is the maximum number of vertices?
- Claim: $v \leq \lfloor \frac{(k-1)l}{2} \rfloor + c \leq \frac{kl}{2}$.

  Equality is achieved only when l even and $c = l/2$.

The lth Moment

Assume Claim is true.

Write

$$\begin{aligned} \mathbb{E}(X_{n,k}^l) &= \sum_{I_1, I_2, \ldots, I_l \in \mathcal{I}} \mathbb{E}\left(\prod_{j=1}^{l} (w_{I_j} - \mathbb{E}(w_{I_j})) \right) \\ &= \frac{1}{n^{kl/2}} \sum_{G_{I_1}, \ldots, G_{I_l}, \mathcal{G}} n^v (1 + o(1)) \cdot Q(e, v, k, l) \ , \end{aligned}$$

where the RHS sum is a polynomial in n, for which we must obtain the highest power coefficient.

For l odd, $v < kl/2$ (given Claim), so $\mathbb{E}(X_{n,k}^l) \to 0$.

The lth Moment

For l even, $v = kl/2$ if there are $l/2$ components, each being the overlap of 2 graphs G_{I_j} with a total of k vertices.

We already analyzed that! This is the same count as for the variance, σ_k^2, but now we do it $l/2$ times.

Moreover, the way we pair the graphs G_{I_j} is important: there are precisely $(l-1)!!$ possible matchings.

This yields that the asymptotics for l even are

$$\mathbb{E}((X_{n,k})^l) \to \sigma_k^l (l-1)!! \,.$$

The lth Moment

Thus, for all l,

$$\mathbb{E}\left(\left(\frac{X_{n,k}}{\sigma_k}\right)^l\right) \to \begin{cases} 0\,, & \text{if } l \text{ odd,} \\ (l-1)!!\,, & \text{if } l \text{ even.} \end{cases}$$

These are the moments of the standard normal variable.

Hence

$$\frac{X_{n,k}}{\sigma_k} \to N(0,1)\,,$$

with convergence in distribution. □

All that remains is to prove the Claim.

Claim

<u>Claim:</u> $v \leq \lfloor \frac{(k-1)l}{2} \rfloor + c \leq \frac{kl}{2}$. Equality is achieved only when l even, $c = k/2$.

Proof. Consider the graphs $G_{I_1}, \ldots, G_{I_l}, \mathcal{G}$, and the walks corresponding to $I_1, \ldots, I_l$.

We determined that the number of components of $\mathcal{G}$ is $c \leq \lfloor l/2 \rfloor$.

Let F be a *spanning forrest* for $\mathcal{G}$. Then F uses all v vertices and $e' \leq e$ edges. Moreover, $v = e' + c$.

We need to bound e'.

Claim

Claim: $v \le \lfloor \frac{(k-1)l}{2} \rfloor + c \le \frac{kl}{2}$. Equality is achieved only when l even, $c = k/2$.

Proof. Let $X = (X_{ij})$ be a $l \times k$ table of 0s and 1s, with the properties that

- each walk determined by I_j defines the jth row of X;
- an entry in row j is 1 only if the edge corresponding to it in the walk of I_j is an edge in F;
- every edge in F receives a 1 at least 2 times;
- every edge in F receives a 1 in each walk it appears.

Then it follows that $2e' \le \sum_{i,j} X_{ij}$.

Claim

Claim: $v \leq \lfloor \frac{(k-1)l}{2} \rfloor + c \leq \frac{kl}{2}$. Equality is achieved only when l even, $c = k/2$.

Proof. Note that such a table can be obtaied by simply putting a 1 in for every instance of a spanning tree edge (and 0 everywhere else).

We show each row of X can be made to have a 0.

If there is a row with no 0, all edges in the respective walk are in F. Since every component of F is a tree, the walk is a closed walk on a tree, every edge must appear twice.

Call this row j. There must be an edge in I_j which overlaps with another $I_{\tilde{j}}$ and thus receives a 1 in the row $\tilde{j}$. Replace one of the 1s corresponding to this edge in row j by a 0.

Claim

Claim: $v \leq \lfloor \frac{(k-1)l}{2} \rfloor + c \leq \frac{kl}{2}$. Equality is achieved only when l even, $c = k/2$.

Proof. Thus every row can be made to have a 0, so

$$\sum_{i,j} X_{ij} \leq \sum_{i=1}^{k} \sum_{j=1}^{l} X_{ij} \leq (k-1)l \,.$$

Hence $2e' \leq (k-1)l$, $e' \leq \lfloor \frac{(k-1)l}{2} \rfloor$, and

$$v = e' + c \leq \lfloor \frac{(k-1)l}{2} \rfloor + \lfloor \frac{l}{2} \rfloor \,,$$

with equality only if the number of components is $l/2$ and l is even. □

Putting it all together: CLT

We have shown:

- $\frac{X_{n,k}}{\sigma_k} \to N(0,1)$ in distribution, for all $k \geq 1$.
- Computed the covariances in Review Session.
- So we know all about the Gaussian Process defined on the monomials.
- Therefore we can immediately extend to all polynomials.

But the variances, covariances look nasty.

A Different Basis

The reason why the variances look nasty is because the monomial basis is NOT the “nice” one....

A Different Basis

... the right one is the Chebyshev basis!

A Different Basis

Let $T_0(x) = 1; T_1(x) = x; \ldots; T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$.

These are orthogonal with respect to the weight $\frac{1}{\sqrt{1-x^2}}$ on $[-1, 1]$.

If one renormalizes the matrices so that the eigenvalues are asymptotically on $[-1, 1]$ instead of $[-2, 2]$, one finds that

$$\frac{\mathrm{tr}(T_k(\tilde{W}_n)) - \mathbb{E}\left(\mathrm{tr}(T_k(\tilde{W}_n))\right)}{\sqrt{k}} \to N(0, 1) \ ,$$

for any $k \geq 1$.

A Special Basis

Moreover,

$$\mathrm{Cov}(T_k(W_n), T_l(W_n)) = \mathbb{E}(T_k(W_n)T_l(W_n)) - \mathbb{E}(T_k(W_n))\mathbb{E}(T_l(W_n)) \to k\delta_{k,l} \ .$$

This means that the Chebyshev basis diagonalizes the covariance matrix.

This phenomenon turns out to be *universal*, in that it applies to a multitude of other matrix ensembles, with other limiting laws (Marčenko-Pastur, Jacobi, regular graphs, etc.).