Random matrices and communication systems: WEEK 12

In this lecture, we reprove the theorem from last time using the Stieltjes transform method.

1 Marčenko-Pastur's theorem

Let H be an $n \times n$ random matrix with i.i.d. complex-valued entries such that for all $1 \le j, l \le n$:

(i) $\mathbb{E}(h_{jl}) = 0$, $\mathbb{E}\left(|h_{jl}|^2\right) = 1$;

(ii) the distribution of h_{jl} is compactly supported (this second assumption may be relaxed).

As last time, let us consider $W^{(n)} = \frac{1}{n} HH^*$, $\lambda_1^{(n)}, \dots, \lambda_n^{(n)}$ its (non-negative) eigenvalues and the empirical distribution $\mu_n = \frac{1}{n}\sum_{j=1}^n \delta_{\lambda_j^{(n)}}$. A particular instance of Marčenko-Pastur's theorem is the following.

Theorem 1.1. Under assumptions (i) and (ii), almost surely, the sequence $(\mu_n,\, n \ge 1)$ converges weakly towards the quarter-circle law μ, whose pdf is given by

$$p_\mu(x) = \frac{1}{\pi}\sqrt{\frac{1}{x} - \frac{1}{4}}\, 1_{\{0<x<4\}}$$

As a preliminary to the proof of the theorem, let us consider, for a given $z \in \mathbb{C}\backslash\mathbb{R}$, the quadratic equation:

$$z\, g(z)^2 + z\, g(z) + 1 = 0 \tag{1}$$

This quadratic equation has two solutions

$$g_\pm(z) = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{z}}$$

It turns out that $g_+(z)$ is the Stieltjes transform of the above distribution μ. The proof is left as an exercise in the homework.

Proof of Theorem 1.1 (sketch).
The strategy for today is to use the characterization of weak convergence via Stieltjes transform: a sequence of distributions converges weakly towards a limiting distribution if the corresponding sequence of Stieltjes transforms converges pointwise on $\mathbb{C}_+ = \{z \in \mathbb{C} : \text{Im } z > 0\}$ towards the limiting Stieltjes transform. Thus, we will prove that almost surely, $g_n(z)$ converges in the large n limit towards a solution of equaton (1). As all the g_n are by definition Stieltjes transforms, but only one of the two solutions of this equation is a Stieltjes transform, a continuity argument allows then to conclude that g_n can only converge to g_+ and not to g_-.

The rule of the game is therefore now to try writing $g_n(z)$ on both sides of an equality sign. To this end, let us compute

$$g_n(z) = \int_{\mathbb{R}} \frac{1}{x-z}\, d\mu_n(x) = \frac{1}{n}\sum_{j=1}^n \frac{1}{\lambda_j^{(n)} - z} = \frac{1}{n}\,\text{Tr}\left(\left(W^{(n)} - zI\right)^{-1}\right)$$

First notice that $W^{(n)} = \frac{1}{n} HH^* = \frac{1}{n}\sum_{k=1}^n h_k\, h_k^*$, where h_k is the k^{th} column of H (and is therefore $n \times 1$). This way, $W^{(n)}$ is expressed as a sum of rank one $n \times n$ matrices. For a given $1 \le k \le n$, let us define

$$W_k^{(n)} = W^{(n)} - \frac{1}{n} h_k\, h_k^* = \frac{1}{n}\sum_{l=1,\, l\neq k}^n h_l\, h_l^*$$

as well as the resolvents

$$G^{(n)}(z) = \left(W^{(n)} - zI\right)^{-1} \quad \text{and} \quad G_k^{(n)}(z) = \left(W_k^{(n)} - zI\right)^{-1}$$

Notice that the object we are interested in is $g_n(z) = \frac{1}{n} \operatorname{Tr}\left(G^{(n)}(z)\right)$. Let us now prove the following two lemmas.

Lemma 1.2.

$$\frac{1}{n}\, h_k^* \, G^{(n)}(z) \, h_k = \frac{\frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}{1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}$$

Proof. Let us compute

$$\begin{aligned} h_k^* \, G_k^{(n)}(z) \, \left(G^{(n)}(z)\right)^{-1} &= h_k^* \, G_k^{(n)}(z) \, \left(W^{(n)} - zI\right) = h_k^* \, G_k^{(n)}(z) \, \left(W_k^{(n)} - zI + \frac{1}{n}\, h_k \, h_k^*\right) \\ &= h_k^* + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k \, h_k^* = \left(1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k\right) \, h_k^* \end{aligned}$$

Therefore,

$$h_k^* \, G_k^{(n)}(z) = \left(1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k\right) \, h_k^* \, G^{(n)}(z)$$

and

$$h_k^* \, G_k^{(n)}(z) \, h_k = \left(1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k\right) \, h_k^* \, G^{(n)}(z) \, h_k$$

which concludes the proof. □

Lemma 1.3.

$$g_n(z) = \frac{1}{n} \operatorname{Tr}\left(G^{(n)}(z)\right) = -\frac{1}{nz} \sum_{k=1}^{n} \frac{1}{1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}$$

Proof. Let us compute

$$\begin{aligned} 1 &= \frac{1}{n} \operatorname{Tr}(I) = \frac{1}{n} \operatorname{Tr}\left(\left(W^{(n)} - zI\right) \, G^{(n)}(z)\right) = \frac{1}{n} \operatorname{Tr}\left(\frac{1}{n} \sum_{k=1}^{n} h_k \, h_k^* \, G^{(n)}(z) - z \, G^{(n)}(z)\right) \\ &= \frac{1}{n} \sum_{k=1}^{n} \left(\frac{1}{n}\, h_k^* \, G^{(n)}(z) \, h_k\right) - z \, g_n(z) = \frac{1}{n} \sum_{k=1}^{n} \left(\frac{\frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}{1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}\right) - z \, g_n(z) \end{aligned}$$

by Lemma 1.2. So

$$z \, g_n(z) = \frac{1}{n} \sum_{k=1}^{n} \left(\frac{\frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}{1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k} - 1\right) = -\frac{1}{n} \sum_{k=1}^{n} \frac{1}{1 + \frac{1}{n}\, h_k^* \, G_k^{(n)}(z) \, h_k}$$

which concludes the proof. □

Notice that so far, these formulas hold for *any* matrix of the form $W^{(n)} = \frac{1}{n} HH^*$, without any further assumption on the matrix H. On the contrary, the next lemma relies strongly on the assumptions (i) and (ii).

Lemma 1.4. Under assumptions (i) and (ii), for all $z \in \mathbb{C}\backslash\mathbb{R}$ and all $\varepsilon > 0$, there exists $C > 0$ independent of n such that

$$\mathbb{P}\left(\left|\frac{1}{n}\, h_k^* \, G_k^{(n)} \, h_k - \frac{1}{n} \operatorname{Tr}\left(G^{(n)}(z)\right)\right| \geq \varepsilon\right) \leq \frac{C}{n^2}$$

which implies by the Borel-Cantelli lemma that

$$\frac{1}{n}\, h_k^* \, G_k^{(n)} \, h_k - \frac{1}{n} \operatorname{Tr}\left(G_k^{(n)}(z)\right) \underset{n \to \infty}{\to} 0 \quad \text{almost surely}$$

We do not prove this lemma, but simply show in the following that for all $n \geq 1$,

$$\mathbb{E}\left(\frac{1}{n}\, h_k^*\, G_k^{(n)}\, h_k - \frac{1}{n}\mathrm{Tr}\left(G_k^{(n)}(z)\right)\right) = 0$$

which also requires the use of the assumptions made above:

$$\mathbb{E}\left(\frac{1}{n}\, h_k^*\, G_k^{(n)}\, h_k\right) = \frac{1}{n}\sum_{j,l=1}^{n} \mathbb{E}\left(\overline{h_{jk}}\, \left(W_k^{(n)} - zI\right)^{-1}_{jl}\, h_{lk}\right)$$

As the matrix $W_k^{(n)}$ does not "contain" h_k, the k^{th} column of H, it is independent of both $\overline{h_{jk}}$ and h_{lk}, so

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{n}\, h_k^*\, G_k^{(n)}\, h_k\right) &= \frac{1}{n}\sum_{j,l=1}^{n} \mathbb{E}\left(\overline{h_{jk}}\, h_{lk}\right)\, \mathbb{E}\left(\left(W_k^{(n)} - zI\right)^{-1}_{jl}\right) \overset{(*)}{=} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left(\left(W_k^{(n)} - zI\right)^{-1}_{jj}\right) \\
&= \mathbb{E}\left(\frac{1}{n}\,\mathrm{Tr}\left(W_k^{(n)} - zI\right)\right) = \mathbb{E}\left(\frac{1}{n}\,\mathrm{Tr}\left(G_k^{(n)}(z)\right)\right)
\end{aligned}$$

where $(*)$ follows from the fact that $\mathbb{E}(\overline{h_{jk}}\, h_{lk}) = \delta_{jl}$, according to assumption (i) and the independence assumption.

The actual proof of Lemma 1.4 relies on the use of Chebychev's inequality with $\phi(x) = x^4$ and a similar analysis of the expectation (except that one must consider the moment of order 4 instead of the first order moment).

The last lemma is a technicality, which holds again for any matrix of the form $W^{(n)} = \frac{1}{n}\, HH^*$.

Lemma 1.5. For all $z \in \mathbb{C}\backslash\mathbb{R}$,

$$\left|\frac{1}{n}\mathrm{Tr}\left(G_k^{(n)}(z)\right) - \frac{1}{n}\mathrm{Tr}\left(G^{(n)}(z)\right)\right| \leq \frac{1}{n\,|\mathrm{Im}\ z|}$$

The proof of this lemma is still rather long for a technicality and is therefore omitted.

Gathering together the results of Lemmas 1.3, 1.4 and 1.5, we obtain that for large values of n,

$$\begin{aligned}
g_n(z) &= \frac{1}{n}\,\mathrm{Tr}\left(G^{(n)}(z)\right) = -\frac{1}{nz}\sum_{k=1}^{n} \frac{1}{1 + \frac{1}{n}\, h_k^*\, G_k^{(n)}(z)\, h_k} \\
&\simeq -\frac{1}{nz}\sum_{k=1}^{n} \frac{1}{1 + \frac{1}{n}\,\mathrm{Tr}\left(G_k^{(n)}(z)\right)} \simeq -\frac{1}{z}\,\frac{1}{1 + \frac{1}{n}\,\mathrm{Tr}\left(G^{(n)}(z)\right)} = -\frac{1}{z\,(1 + g_n(z))}
\end{aligned}$$

which may be rewritten as

$$z\, g_n(z)^2 + z\, g_n(z) + 1 \simeq 0$$

Taking some more precautions, we can conclude that $g_n(z)$ converges almost surely towards a solution of the quadratic equation (1), which should be chosen as g_+ for the reasons explained above. This "completes" the proof of the theorem. $\square$