1 Stieltjes transform

From last time, we saw that it suffices to study the last entry of \( \left( \frac{1}{\sqrt{n}}M_n - zI_n \right)^{-1} \).

Definition 1.1. Let \( M \) be a \((p + q) \times (p + q)\) matrix, and write

\[
M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},
\]

where \( A \) is \( p \times p \) and \( D \) is \( q \times q \). Then the Schur complement of \( A \) is the \( q \times q \) matrix \( M/A := D - CA^{-1}B \).

If \( A \) is invertible, then

\[
M^{-1} = \begin{pmatrix} A^{-1} + A^{-1}B(M/A)^{-1}CA^{-1} & -A^{-1}B(M/A)^{-1} \\ -(M/A)^{-1}CA^{-1} & (M/A)^{-1} \end{pmatrix}.
\]

Apply to our situation, if we write

\[
M_n = \begin{pmatrix} M_{n-1}^{-1} & X \\ X^* & \xi_{n,n} \end{pmatrix},
\]

we see that

\[
\left( \frac{1}{\sqrt{n}}M_n - zI_n \right)^{-1}_{n,n} \left( \frac{1}{\sqrt{n}}M_{n-1} - zI_{n-1} \right)^{-1}\left( \frac{1}{\sqrt{n}}M_{n-1} - zI_{n-1} \right)^{-1}. X \right)^{-1}.
\]

As we have seen before, to prove the semicircle law, we may assume that the diagonal entries are 0. Therefore,

\[
E \left[ \left( \frac{1}{\sqrt{n}}M_n - zI_n \right)^{-1} \right] = -E \left( z + \frac{1}{n}X^* \left( \frac{1}{\sqrt{n}}M_{n-1} - zI_{n-1} \right)^{-1} X \right)^{-1}. \quad (1.1)
\]

Write \( R = \left( \frac{1}{\sqrt{n}}M_{n-1} - zI_{n-1} \right)^{-1} \). We would like to know what \( X^*RX \) is. First, observe that \( X \) and \( R \) are independent. Therefore, we may condition on \( R \) and study \( X^*RX \) assuming \( R \) is some deterministic matrix. But before that, we also observe that \( M_{n-1} \) is Hermitian, and hence all its eigenvalues are real. In particular, this implies the operator norm of \( R, \|R\| \), is of order \( O(1) \), since we assume that the imaginary part of \( z \) is positive. So it suffices to understand what \( X^*RX \) is when \( R \) is deterministic and \( \|R\| = O(1) \).

Before study \( X^*RX \), let’s first study \( X^*AX \), where \( A \) is a positive semidefinite matrix with \( \|A\| = O(1) \). Also, we will further assume that the entries of \( M_n \) are uniformly bounded (which is fine by the reduction we saw before). In this case, the map \( X \mapsto (X^*AX)^{1/2} = \|A^{1/2}X\| \) is Lipschitz. Therefore, we can apply Talagrand’s inequality to see

\[
P \left( \| (X^*AX)^{1/2} - \mathbf{M}(X^*AX)^{1/2} \| \geq \lambda \right) \leq C e^{-c\lambda^2}.
\]
If \( A \) has \( k \) nonzero eigenvalues, then using Hoeffding’s inequality one can show that \( \| A^{1/2}X \| \geq \Omega(\sqrt{k}) \) with high probability, and this also implies \( M(X^*AX)^{1/2} \geq \Omega(\sqrt{k}) \). Moreover, observe that median satisfies \( (M(X^*AX)^{1/2})^2 = M(X^*AX) \). Therefore, multiplying both sides in the probability by \( \| (X^*AX)^{1/2} + M(X^*AX)^{1/2} \| \), we see that

\[
P \left( |X^*AX - M(X^*AX)| \geq \lambda \sqrt{n} \right) \leq Ce^{-c\lambda^2},
\]

for some possibly different \( C \) and \( c \). If \( A \) is Hermitian instead of positive definite, we may write \( A = A_+ + A_- \), where \( A_+ \) has only nonnegative eigenvalues and \( A_- \) has only nonpositive eigenvalues, and applying triangle inequality we obtain

\[
P \left( |X^*AX - M(X^*AX)| \geq \lambda \sqrt{n} \right) \leq Ce^{-c\lambda^2}
\]

for some different \( C \) and \( c \). Using the fact that any random variable \( Y \) with finite second moment satisfies \( |MY - EY| = O(\text{Var}(Y)^{1/2}) \), we can replace median by the mean. If \( R \) is a general matrix with \( \| R \| = O(1) \), we may write \( R = A + B \), where \( A \) is Hermitian and \( B \) is skew-Hermitian, and similarly we will obtain

\[
P \left( |X^*RX - EX^*RX| \geq \lambda \sqrt{n} \right) \leq Ce^{-c\lambda^2}.
\]

Now, since \( \xi_{i,j} \) has mean zero and variance 1, we have

\[
EX^*RX = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} E\xi_{i,n}r_{i,j}\xi_{j,n} = \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} r_{i,j}\delta_{ij} = \text{tr}(R).
\]

Thus, we have

\[
P \left( |X^*RX - \text{tr}(R)| \geq \lambda \sqrt{n} \right) \leq Ce^{-c\lambda^2}.
\]

The above bound holds for deterministic matrix \( R \) with \( \| R \| = O(1) \). If \( R \) is random, then

\[
P \left( |X^*RX - \text{tr}(R)| \geq \lambda \sqrt{n} \right) = E \left[ P \left( |X^*RX - \text{tr}(R)| \geq \lambda \sqrt{n} | R \right) \right] \leq Ce^{-c\lambda^2}.
\]

By some computations, we have

\[
\text{tr}(R) = \sqrt{n(n-1)}s_{n-1} \left( \frac{\sqrt{n}}{\sqrt{n-1}} z \right).
\]

This expression is exactly the same as what we saw last time, and hence by what we showed last time, we have

\[
\text{tr}(R) = n(s_n(z) + o(1)).
\]
Also, recall that almost surely, \( s_n(z) - E s_n(z) \to 0 \) as \( n \to \infty \), and hence
\[
\text{tr}(R) = n(E s_n(z) + o(1)).
\]

Finally, writing \( E_n = \{|X^*RX - \text{tr}(R)| \geq n^{1/3}\} \), and recalling the left hand side of (1.1) is \( E s_n(z) \), we obtain
\[
E s_n(z) = -E \left( z + \frac{1}{n} X^*RX \right)^{-1}
= -E \left[ \left( z + \frac{1}{n} X^*RX \right)^{-1} 1_{E_n} \right] + o(1)
= -E \left( z + \frac{1}{n} (n E s_n(z) + o(1)) \right)^{-1} + o(1)
= -\frac{1}{z + E s_n(z)} + o(1).
\]

It is not difficult to show that \( E s_n \) is locally uniformly equicontinuous and locally uniformly bounded away from the real line. By the Arzelà-Ascoli theorem, \( E s_n \) converges locally uniformly to a limit \( s \) along a subsequence. So we have
\[
s(z) = -\frac{1}{z + s(z)}.
\]
Solving for \( s(z) \), we have
\[
s(z) = \frac{-z \pm \sqrt{z^2 - 4}}{2}.
\]
Since the Stieltjes transform goes to 0 as \( z \to \infty \), we conclude that
\[
s(z) = \frac{-z + \sqrt{z^2 - 4}}{2}.
\]

As there is only one possible subsequential limit of \( E s_n \), we conclude that \( E s_n \) converges locally uniformly to \( s \), and thus \( s_n(z) \) converges to \( s(z) \) almost surely.

To finish the proof, it remains to find which distribution has the Stieltjes transform \( s \), but this can be found by observing
\[
\frac{s(\cdot + bi) - s(\cdot - bi)}{2\pi i} \Rightarrow \mu_{sc}
\]
as \( b \downarrow 0 \).

References
