Accuracy of the Tracy-Widom limit for extreme eigenvalues in white Wishart matrices

Zongming Ma

Department of Statistics
Stanford University

zongming@stanford.edu

May 15, 2008

Principal component analysis

Data: $X_{n\times p} \sim N_{n\times p}(\mathbf{1}'\mu, \Sigma \otimes I_n)$
Sample covariance matrix: $S = (n-1)^{-1}X'HX$ with $H = I - n^{-1}\mathbf{1}\mathbf{1}'$

PCA (sample version)

find $a_1, \cdots, a_p \in \mathbb{R}^p$ with $\|a_i\| = 1$, s.t.

$$a_i = \text{argmax}\{a'Sa : a'a_j = 0, j < i\}, \quad a_0 = 0.$$

Question: Is this procedure meaningful?

Testing isotropic variance: $\Sigma = I_p$

- Under the null, $(n-1)S \sim W_p(I_p, n-1)$
- Roy's union intersection principle (UIP):

$$\text{reject the null if} \quad \lambda_1(S) > c_1 \quad \text{or} \quad \lambda_p(S) < c_2.$$

Modern asymptotics

An alternative asymptotic setting

$$p \to \infty,\ n = n(p) \to \infty \text{ and } n/p \to c \in (0, \infty) \tag{1}$$

Theorem (Johnstone 2001)

Under the alternative asymptotic setting (1), *for*

$$\mu_p = (\sqrt{n-1} + \sqrt{p})^2,$$
$$\sigma_p = (\sqrt{n-1} + \sqrt{p})(1/\sqrt{n-1} + 1/\sqrt{p})^{1/3},$$

$(\lambda_1(A) - \mu_p)/\sigma_p \Rightarrow F_1$ *(Tracy-Widom law of order 1).*

Example

Patterson et al. (2006): testing the presence of population heterogeneity with SNP data.

The Tracy-Widom law

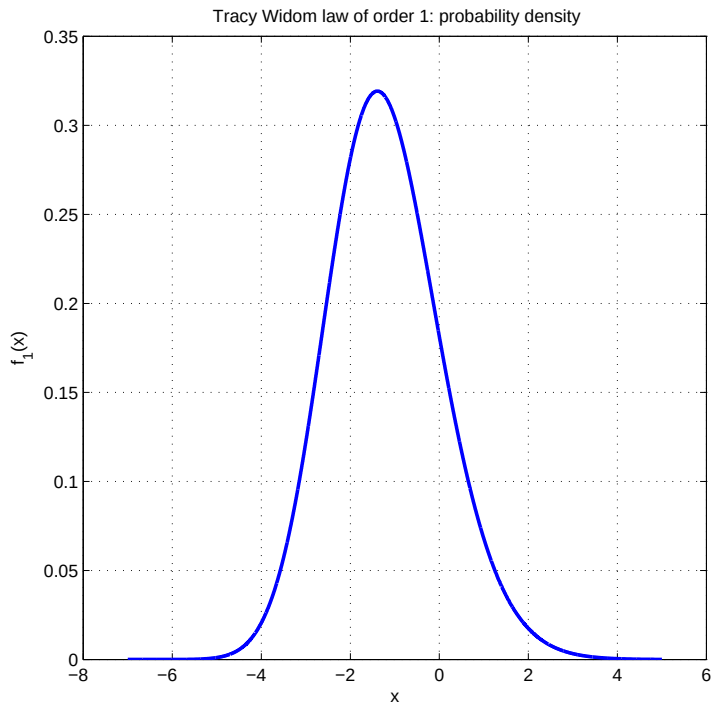
Definition

Let q satisfies: $q" = sq + 2q^3, \quad q(s) \sim \mathrm{Ai}(s), s \to \infty$, then

$$F_1(s) = \exp(-\frac{1}{2}\int_s^\infty q(x) + (x-s)q^2(x)dx).$$

History

Discovered in Tracy & Widom (1996) as the limiting law of the largest eigenvalue in an $n \times n$ Gaussian symmetric matrix;



The largest eigenvalue: second order accuracy

Theorem (ZM 2008)

Under (1), *changing the rescaling constants to*

$$\mu_{n,p} = (\sqrt{n-1/2} + \sqrt{p-1/2})^2$$
$$\sigma_{n,p} = (\sqrt{n-1/2} + \sqrt{p-1/2})(1/\sqrt{n-1/2} + 1/\sqrt{p-1/2})^{1/3}$$

one obtains second order accuracy

$$P((\lambda_1(A) - \mu_{n,p})/\sigma_{n,p} \leq s) = F_1(s) + O(\min(n,p)^{-2/3}).$$

Remarks

- Correction terms increase accuracy for small (n, p)
- Accuracy is $O(\min(n,p)^{-2/3})$ instead of $O(\min(n,p)^{-1/3})$

Finite sample approximation: probability table

Simulations for finite $n \times p$ vs. Tracy-Widom limit ($R = 40,000$)

TW1	500×5		20×5		5×5		2×SE
.01	.010	.020	.001	.002	.000	.000	.001
.05	.049	.083	.018	.028	.002	.002	.002
.10	.098	.150	.054	.073	.021	.020	.003
.30	.296	.385	.259	.303	.218	.213	.005
.50	.502	.589	.483	.531	.465	.460	.005
.70	.705	.772	.704	.737	.702	.698	.005
.90	.906	.933	.906	.921	.908	.907	.003
.95	.955	.969	.954	.962	.954	.953	.002
.99	.990	.994	.990	.992	.989	.989	.001

Conventional significance levels are shown in red fonts.
Results based on Johnstone (2001) are shown in gray fonts for comparison.

The smallest eigenvalue: convergence and accuracy

Some modifications

- Asymptotic regime: $n-1 \geq p$, and $n/p \to c \in (1, \infty)$
- Reflected Tracy-Widom law (Paul, 2006): $G_1(s) = 1 - F_1(-s)$

Rescaling constants

$$\mu_{n,p}^- = (\sqrt{n-1/2} - \sqrt{p-1/2})^2$$
$$\sigma_{n,p}^- = (\sqrt{n-1/2} - \sqrt{p-1/2})(1/\sqrt{p-1/2} - 1/\sqrt{n-1/2})^{1/3}$$
$$\nu_{n,p}^- = \log \mu_{n,p}^- + O(\min(n,p)^{-4/3}), \quad \tau_{n,p}^- = \sigma_{n,p}^- / \mu_{n,p}^-.$$

Theorem (ZM 2008)

$$P((\log \lambda_p(A) - \nu_p^-)/\tau_p^- \leq s) = G_1(s) + O(\min(n,p)^{-2/3}).$$

Finite sample approximation: probability table

Smallest eigenvalue: $\nu_{n,p}^- = \log \mu_{n,p}^- + \frac{1}{8}(\tau_{n,p}^-)^2, \tau_{n,p}^- = \sigma_{n,p}^-/\mu_{n,p}^-$

Simulations for finite $n \times p$ vs. Tracy-Widom limit ($R = 40,000$)

TW1	10×5	50×25	20×5	100×25	2×SE
.99	.999	.997	.999	.993	.001
.95	.995	.973	.977	.960	.002
.90	.976	.931	.939	.915	.003
.70	.798	.728	.740	.713	.005
.50	.555	.515	.522	.505	.005
.30	.310	.302	.301	.298	.005
.10	.095	.097	.096	.098	.003
.05	.047	.048	.047	.048	.002
.01	.011	.010	.009	.009	.001

Conventional significance levels are shown in red fonts.