Some work of Mario Wschebor on condition number of random matrices

Clapem Nov.5 2009

Jean-Marc Azaïs

IMT,LSP,University of Toulouse

Condition number

Let A be an n, n invertible real matrix, the condition number of A is defined as

$$\kappa(A) := \frac{\lambda_n}{\lambda_1},$$

λ_i ith singular value in decreasing order, Turing (1948).

Condition number

Let A be an n, n invertible real matrix, the condition number of A is defined as

$$\kappa(A) := \frac{\lambda_n}{\lambda_1},$$

λ_i ith singular value in decreasing order, Turing (1948).
It is related with the propagation of errors in the solution of the system

$$Ax = b.$$

Condition number

Let A be an n, n invertible real matrix, the condition number of A is defined as

$$\kappa(A) := \frac{\lambda_n}{\lambda_1},$$

λ_i ith singular value in decreasing order, Turing (1948).
It is related with the propagation of errors in the solution of the system

$$Ax = b.$$

Mesures the Frobenius distance between A and the set Σ of singular matrices

Theorem (Eckart-Young, 1936)

$$\kappa(A) = \frac{\lambda_n}{d_F(A, \Sigma)}$$

Non-Gaussian symmetric entries

Theorem

(Cuesta-Alberto and Wschebor, 2003)
We assume that the entries of A are i.i.d. with a symmetric distribution with non increasing density on $\mathbb{R}^+$ *such that*
$\mathrm{E}\left[|a_{1,1}|^r\right] = 1,$ *for some* $r > 0$.
(for example) $\mathbb{P}\{|a_{1,1}| \leq \alpha) \leq C\alpha,$ *for all* $\alpha > 0$. *Then*

$$\mathrm{E}[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + [3 \log n + \log C]^+ + 1,$$

where $x^+ = \max(x, 0)$ *for real* x.

Proof

$$\log(\kappa(A)) = \log(\lambda_n) - \log(\lambda_1).$$

The two terms are studied separately using elementary ad-hoc inequalities as Markov inequality. □

Proof

$$\log(\kappa(A)) = \log(\lambda_n) - \log(\lambda_1).$$

The two terms are studied separately using elementary ad-hoc inequalities as Markov inequality. □

Example : Uniform distribution on [-H,H] then

$$\mathrm{E}[\log \kappa(A)] \leq 4 \log n + 1.$$

Numerically the constant is 1 and not 4. No others results in the literature.

Proof

$$\log(\kappa(A)) = \log(\lambda_n) - \log(\lambda_1).$$

The two terms are studied separately using elementary ad-hoc inequalities as Markov inequality. □

Example : Uniform distribution on [-H,H] then

$$\mathrm{E}[\log \kappa(A)] \leq 4 \log n + 1.$$

Numerically the constant is 1 and not 4. No others results in the literature.

In the case of Gaussian distribution, this does not provide the best result due to Edelman (1988)

$$\mathrm{E}[\log \kappa(A)] \leq \log n + 1.537... + \varepsilon_n$$

Gaussian centered matrices

Theorem (Azaïs-Wschebor, 2005)

Assume that $A = ((a_{ij}))_{i,j=1,\dots,n}$, $n \geq 3$, and that the a_{ij}'s are i.i.d. standard normal random variables.
Then, there exist universal positive constants c, C such that for $x > 1$:

$$\frac{0.13}{x} \leq \frac{c}{x} < \mathbb{P}(\kappa(A) > n.x) < \frac{C}{x} \leq \frac{5.6}{x} \quad (1)$$

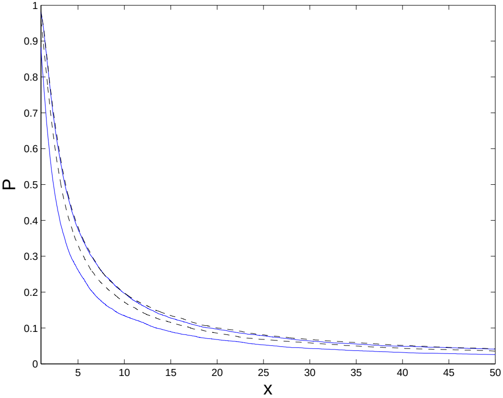
Gaussian centered matrices

Theorem (Azaïs-Wschebor, 2005)

Assume that $A = ((a_{ij}))_{i,j=1,\dots,n}$, $n \geq 3$, and that the a_{ij}'s are i.i.d. standard normal random variables.
Then, there exist universal positive constants c, C such that for $x > 1$:

$$\frac{0.13}{x} \leq \frac{c}{x} < \mathbb{P}(\kappa(A) > n.x) < \frac{C}{x} \leq \frac{5.6}{x} \qquad (1)$$

Extension to rectangular matrices and other types of "ensemble" have been presented by Edelman and Sutton (2005) using classical tools (orthogonal polynomials). But in fact the Azaïs and Wschebor method provides the best known result for non centered matrices.



Simulation suggests the values 0.88 and 2.00 for c and C.

Non-centered matrices, Smoothed analysis

Theorem (Wschebor, 2004)

Suppose that $A = M + D$ *where* D *is as before. For* $x > 0$:

$$\mathbb{P}(\kappa(A) > n.x) < \frac{1}{x}\left(\frac{1}{4\sqrt{2\pi n}} + C(M,n)\right) \tag{2}$$

where

$$C(M,n) = 7\left(5 + \frac{4\,\|M\|^2\,(1+\log n)}{n}\right)^{\frac{1}{2}},$$

where $\|M\|$: *greatest singular value of* M.

No other results in the literature

For example if $\|M\| \leq 1$, for $x > 0$

$$\mathbb{P}(\kappa(A) > n.x) < \frac{20}{x}$$

Proof

The proof is based on the study of the random field

$$X(t) = t^T B t = t^T A^T A t$$

on the unit sphere.

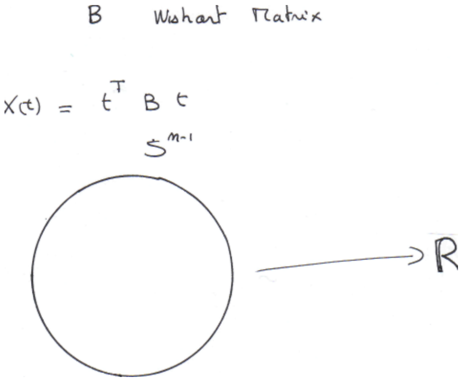
Proof

The proof is based on the study of the random field

$$X(t) = t^T B t = t^T A^T A t$$

on the unit sphere.
The critical values ($X'(t) = 0$) of this random field are the eigenvalues $\nu_1, \dots, \nu_n$ of B i.e. the squares of the singular values of A.

To compute the number of critical values that lies in an interval we use the Rice formula.



The process $X(t)$ on the sphere .

Theorem (Azaïs & Wschebor 2005, 2008)

Consider a random field $Z : S \to \mathbb{R}^d$, S open set $\mathbb{R}^d$ and $v \in \mathbb{R}^d$ a fixed point of the image (0 in our case) . Suppose that Z is Gaussian with $\mathcal{C}^1$ plus some technical hypotheses. Define $N_v^Z(B)$ as the number of roots in B of the equation $Z(t) = v$
Then,

$$\mathrm{E}\left(N_v^Z(B)\right) = \int_B \mathrm{E}\left(|\det(Z'(t))|/Z(t) = v\right)\ p_{Z(t)}(v)dt.$$

When B is compact, both sides are finite.

Can be extended to a process defined on a manifold as a sphere with standard extension arguments

How can we use this result

Joint density of (ν_1, ν_n)

$$\{\nu_n \in (a, a+da), \nu_1 \in (b, b+db)\}$$
$$= \left\{ \begin{array}{c} \exists\ s, t \in S^{n-1}, < s, t >= 0, X(s) \in (a, a+da), X(t) \in (b, b+db), \\ X'(s) = 0, X'(t) = 0, X''(s) \prec 0, X''(t) \succ 0 \end{array} \right.$$

$N_{a,b,da,db}$ of pairs (s, t) as above is equal to 0 or 4, so that:

$$\mathbb{P}\left(\nu_n \in (a, a+da), \nu_1 \in (b, b+db)\right) = \frac{1}{4}\mathrm{E}\left(N_{a,b,da,db}\right) \qquad (4)$$

We have to consider the number of zeros of the process $Z = (X'(s), X'(t))$ with parameter (s, t) lying in the $2n - 3$ manifold

$$V = \left\{(s, t) : s, t \in S^{n-1}, < s, t >= 0\right\}.$$

The additional conditions on $X(t)$, $X(s)$ $X''(t)$, $X''(s)$ correspond to direct generalization of Rice formulas.

Applying the theorem and letting *da* and *db* tend to zero. Using rotation invariance, we obtain the joint density of ν_1, ν_n

$$g(a,b) = \frac{1}{4}\sigma(V)$$
$$\mathrm{E}\Big(\Delta(e_1,e_2)\ \mathbb{I}_{\{X''(e_1)\prec 0, X''(e_2)\succ 0\}}\Big|X(e_1)=a, X(e_2)=b, Y(e_1,e_2)=0\Big)$$
$$.p_{X(e_1),X(e_2),Y(e_1,e_2)}(a,b,0),$$

with $\Delta(e_1,e_2) = \left[\det\left[\left(Y'(e_1,e_2)\right)^T Y'(e_1,e_2)\right]\right]^{\frac{1}{2}}$
The proof consists of computing all the ingredients of this formula.

An example of geometrical lemma

Lemma

The geometrical volume of the manifold V is equal to

$$\sigma(V) = \sqrt{2}\sigma_{n-1}.\sigma_{n-2}$$

Where $\sigma_{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}$ is the volume of the sphere.

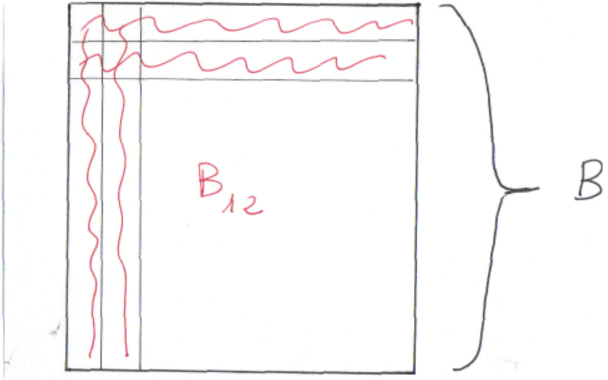
We have also

$$\left[\det\left[\left(Y'(e_1,e_2)\right)^T Y'(e_1,e_2)\right]\right]^{\frac{1}{2}}$$
$$= |\det(B_{12} - aI_{n-2})||\det(B_{12} - bI_{n-2})|(a-b), .$$

B_{12} is obtained from B by suppressing the first and second rows and columns. One of this term is bounded : $B_{12} \succ 0$ implies

$$|\det(B_{12} - aI_{n-2})| \ \mathbb{I}_{B_{12} - aI_{n-2} \prec 0} \leq a^{n-2}$$

The other is computed using known result on characteristic polynomials of Wishart matrices (Mehta)


B_{12}
B

Conclusion

Gaussian fields methods : new tool
to study random matrices
sometimes powerfull

THANK-YOU

GRACIAS

References

Wschebor, M.(2004). Smoothed analysis of K(A), Journal of Complexity, Vol. 20, 1, 97-107.

Cuesta-Albertos, J. and Wschebor M. (2003). Some Remarks on the Condition Number of a Real Random Square Matrix, *Journal of Complexity*, Vol. 19, 4, 548-554.

Azaïs, J-M. and Wschebor, M. (2005). Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix, *SIAM J. Matrix Anal. Appl.*, Vol. 26, No.2, 426-440.

Azaïs, J-M. and Wschebor, M. (2008). Level sets and extrema of random processes and fields. Wiley.

Edelman, A. and Sutton, B.D. (2005). Tails of condition number distributions, *SIAM J. Matrix Anal. Appl.*, Vol. 27, No. 2, 547-560.