Smoothed analysis of $\kappa(A)$.

Mario Wschebor
Centro de Matemática. Facultad de Ciencias.
Universidad de la República. Calle Iguá 4225.
11400 Montevideo. Uruguay.
Tel: (5982)5252522, Fax: (5982)5258617
E mail: wschebor@cmat.edu.uy

May 3, 2009

Abstract

Let $A = ((a_{ij}))$ be an $m \times m$ $(m \geq 3)$ real random matrix, with independent Gaussian entries with a common variance σ^2. Denote by M the matrix of expected values of the entries of A. For $x > 0$ we prove that $P(\kappa(A) > m.x) < \frac{1}{x}\left(\frac{1}{4\sqrt{2\pi m}} + C(M,\sigma,m)\right)$with $C(M,\sigma,m) = 7\left(5 + \frac{4\|M\|^2(1+\log m)}{\sigma^2 m}\right)^{\frac{1}{2}}$. Here $\kappa(A) = \|A\| \left\|A^{-1}\right\|$ is the usual condition number of A, $\|.\|$ is Euclidean operator norm. This implies that if $0 < \sigma \leq 1$ and $\|M\| \leq 1$ then, for $x > 0$, $P(\kappa(A) > m.x) < \frac{K}{\sigma x}$ where K is a universal constant.

Mathematics Subject Classification (2000): Primary: 15A12, 15A52. Secondary: 60G15, 60G60, 65F35.

Key words and phrases: Random matrices, Condition Number, Smoothed Analysis, Rice formulae for random fields.

Let $A = ((a_{ij}))_{i,j=1,\dots,m}$ be an $m \times m$ real matrix. Denote by $\|A\| = \sup_{\|x\|=1} \|Ax\|$ its Euclidean operator norm. $\|x\|$ denotes Euclidean norm of x in $\mathbb{R}^m$. We assume throughout that $m \geq 3$.

If A is non-singular, its *condition number* $\kappa(A)$ is defined by

$$\kappa(A) = \|A\|\|A^{-1}\|$$

and if A is singular we put $\kappa(A) = +\infty$ ([11], [12]). The role of $\kappa(A)$ in numerical linear algebra is well-established (see for example [3], [5], [8], [13]) In [7] it was conjectured that if the entries a_{ij} are independent Gaussian random variables having a common variance σ^2, $0 < \sigma \leq 1$ and $\sup_{i,j} |m_{ij}| \leq 1$, where $m_{ij} = E(a_{ij})$ then

$$P(\kappa(A) > x) = O\left(\frac{m}{\sigma x}\right). \tag{1}$$

(see also [9] for some related questions).

In a recent paper this has been proved in the centered case. More precisely:

Theorem 1 (Azaïs & Wschebor 2003) *Assume that $A = ((a_{ij}))_{i,j=1,\dots,m}$, $m \geq 3$, and that the a_{ij}'s are i.i.d. Gaussian standard random variables.*

Then, there exist universal positive constants c, C (for example, $c = 0,13$ and $C = 5,60$) such that for $x > 1$:

$$\frac{c}{x} < P(\kappa(A) > m.x) < \frac{C}{x} \tag{2}$$

This implies (1) when $m_{ij} = 0$ for all i, j (notice that $\kappa(\sigma A) = \kappa(A)$ for any $\sigma > 0$). Moreover, the lower bound shows that in this case this is the precise behaviour up to a constant factor.

We will use the following notations. Given A, an $m \times m$ real matrix, we denote by $\lambda_1, ..., \lambda_m$, $0 \leq \lambda_1 \leq .... \leq \lambda_m$ the eigenvalues of $A^T A$. If $X : S^{m-1} \to \mathbb{R}$ is the quadratic polynomial $X(x) = x^T A^T A x$, then:

- $\lambda_m = \|A\|^2 = \max_{x \in S^{m-1}} X(x)$
- in case $\lambda_1 > 0$, $\lambda_1 = \frac{1}{\|A^{-1}\|^2} = \min_{x \in S^{m-1}} X(x)$.

Then,

$$\kappa(A) = \left(\frac{\lambda_m}{\lambda_1}\right)^{\frac{1}{2}}$$

when $\lambda_1 > 0$.

- $< .,. >$ is usual scalar product in $\mathbb{R}^m$.
- I_k denotes the $k \times k$ identity matrix.
- $B = A^T A = ((b_{ij}))_{i,j=1,\dots,m}$

- For $s \neq 0$ in $\mathbb{R}^m$, $\pi_s : \mathbb{R}^m \to \mathbb{R}^m$ is the orthogonal projection onto $\{s\}^{\perp}$, the orthogonal complement of s in $\mathbb{R}^m$.

- $C \succ 0$ (resp. $C \prec 0$) means that the symmetric matrix C is positive definite (resp. negative definite).

- If ξ is a random vector $p_\xi(.)$ is the density of its distribution whenever it exists.

- For a differentiable function F defined on a smooth manifold S embedded in some Euclidean space, $F'(s)$ and $F''(s)$ are the first and the second derivative of F that we will represent, in each case, with respect to an appropriate orthonormal basis of the tangent space.

- $M = ((m_{ij}))_{i,j=1,\dots,m}$ will denote the matrix of the expected values of the entries of A.

The next proposition exhibits an example which shows that in the non-centered case (1) does not hold with the hypotheses proposed by Sankar, Spielman & Teng.

Proposition 1 *Let $a_{ij} = 1+\sigma g_{ij}$ for all i, j, where the g_{ij}'s i.i.d. Gaussian, with mean zero and variance 1, 0¡$\sigma < \frac{1}{2}$.*

Then, there exists a positive constant K such that for each $h > \frac{3}{2}$, one can find an integer m_h so that for all $m \geq m_h$ one has

$$P(\kappa(A) > x) \geq K\frac{m^{3/2}}{\sigma x} \quad \text{for} \quad \frac{m^{3/2}}{\sigma} < x < \frac{m^h}{\sigma} \tag{3}$$

Proof.

We denote $G = ((g_{ij}))$.

Clearly $\|M\| = m$. Fix $\varepsilon > 0$ in such a way that

$$(2+\varepsilon)\sqrt{m} \leq m$$

for any $m \geq 5$.

We have the inclusions:

$$\{\kappa(A) > (1-\sigma)m\alpha\} \supset \{\|A\| \geq (1-\sigma)m\} \cap \left\{\left\|A^{-1}\right\| > \alpha\right\}$$

$$\{\|A\| \geq (1-\sigma)m\} \supset \left\{\|A\| \geq m - \sigma(2+\varepsilon)\sqrt{m}\right\} \supset \left\{\|G\| \leq (2+\varepsilon)\sqrt{m}\right\}.$$

Hence,

$$P\left(\kappa(A) > (1-\sigma)m\alpha\right) \geq P\left(\left\|A^{-1}\right\| > \alpha\right) - P\left(\|G\| > (2+\varepsilon)\sqrt{m}\right) \tag{4}$$

The second term in (4) is bounded by

$$C_1 e^{-C_2 m \varepsilon^2} \tag{5}$$

where C_1, C_2 are positive constants. This is a well-known bound on the largest eigenvalue of Wishart matrices (see for example [10], [2] or [6]).

As for the first term in the right-hand side of (4), one has

$$P\left(\left\|A^{-1}\right\| > \alpha\right) = P(\lambda_1 < \frac{1}{\alpha^2})$$

Let u be the point $u = (1, ..., 1)^T$ and L the subspace

$$L = \{x \in \mathcal{R}^m : \langle x, u \rangle = 0\}.$$

Then,

$$\lambda_1 = \min_{x \in S^{m-1}} x^T A^T A x \leq \min_{x \in S^{m-1}, x \in L} x^T A^T A x = \sigma^2 \min_{x \in S^{m-1}, x \in L} x^T G^T G x = \widetilde{\lambda}_1$$

Since the law of G is independent of the choice of an orthonormal basis in $\mathcal{R}^m$, one may choose a basis of the form $\left\{\frac{u}{\|u\|}, u_2, ..., u_m\right\}$ and the quadratic form to be minimized in the computation of $\widetilde{\lambda}_1$ is $\sum_{i,j=2}^m \widetilde{g}_{ij} y_i y_j$ where $((\widetilde{g}_{ij}))$ is the associated matrix to the linear transformation $x \rightsquigarrow Gx$ in the new basis and the y_i's are the new coordinates. In other words, $\widetilde{\lambda}_1$ is the smallest eigenvalue of a matrix of size $(m-1) \times (m-1)$ which is σ^2 times a standard Wishart. We can then use a known lower bound (see for example [10] or [1]) and obtain, if $\frac{(m-1)}{\sigma^2 \alpha^2} < 1$:

$$P(\lambda_1 < \frac{1}{\alpha^2}) \geq P(\widetilde{\lambda}_1 < \frac{1}{\alpha^2}) \geq \beta \frac{\sqrt{m-1}}{\sigma \alpha} \tag{6}$$

where β is a positive constant.

Putting together (5) and (6), since $0 < \sigma < \frac{1}{2}$ we get:

$$P\left(\kappa(A) > \frac{m\alpha}{2}\right) \geq P\left(\kappa(A) > (1-\sigma)m\alpha\right) \geq \beta \frac{\sqrt{m-1}}{\sigma \alpha} - C_1 e^{-C_2 m \varepsilon^2} \tag{7}$$

Replacing in (7) $x = \frac{m\alpha}{2}$ and noticing that $m^h C_1 e^{-C_2 m\varepsilon^2}$ is bounded by a constant $K(h)$ depending on h, we see that (7) implies that

$$\text{if } x > \frac{m^{\frac{3}{2}}}{2\sigma} \quad \text{then} \quad P\left(\kappa(A) > x\right) \geq \frac{\beta m\sqrt{m-1}}{2\sigma x} - \frac{K(h)}{m^h}.$$

To conclude, take

$$K < \frac{\beta}{2}$$

and m_h large enough to insure that $m \geq m_h$ implies $\beta m\sqrt{m-1} - K(h) \geq Km^{\frac{3}{2}}$. Then, under the additional condition $x < \frac{m^h}{\sigma}$, inequality (3) follows.□

The aim of the present paper is to prove Theorem 2 below, in which a positive result is given in the non-centered case, but using a different norm on the expected matrix, i.e., replacing sup norm by Euclidean norm.

Theorem 2 *Assume that the a_{ij}'s are independent with a common variance σ^2.*

Then, for $x > 0$ one has:

$$P(\kappa(A) > m.x) < \frac{1}{x}\left(\frac{1}{4\sqrt{2\pi m}} + C(M, \sigma, m)\right) \tag{8}$$

where

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

Remarks.

1. Theorem 2 implies a modified form of conjecture (1), namely if $0 < \sigma \leq 1$ and $\|M\| \leq 1$ then, for $x > 0$:

$$P(\kappa(A) > m.x) < \frac{20}{\sigma x} \tag{9}$$

   This is an immediate consequence of the statement in the Theorem.

2. With similar calculations than the ones we will perform for the proof of Theorem 2, one can improve somewhat the constants in (8) and (9).

Proof of Theorem 2.

Due to the homogeneity of $\kappa(A)$, with no loss of generality we may assume $\sigma = 1$, changing the expected matrix M by $\frac{1}{\sigma}M$ in the final result.

We follow closely the proof of Theorem 1 in [1], with some changes to adapt it to the present conditions. In exactly the same way as it is done in that paper, using a so-called Rice-type formula for the expectation of the number of critical points of a random field, one can prove that the joint density $g(a,b)$, $a > b$ of the random variables λ_m, λ_1 satisfies inequality:

$$\begin{aligned} & g(a,b) \qquad\qquad (10) \\ \leq \; & \int_V E\left(\Delta(s,t)\mathbb{1}_{\{X''(s)\prec 0, X''(t)\succ 0\}}/X(s)=a, X(t)=b, Y(s,t)=0\right). \\ & \cdot p_{X(s),X(t),Y(s,t)}(a,b,0) \;\; \sigma_V(d(s,t)). \end{aligned}$$

Here,

- $V = \left\{(s,t) : s,t \in S^{m-1}, < s,t >= 0\right\}$ is a C^∞-differentiable manifold without boundary, embedded in $\mathbb{R}^{2m}$, $\dim(V) = 2m-3$. $\tau = (s,t)$ denotes a generic point in V and $\sigma_V(d\tau)$ the geometric measure on V. It is not hard to check that $\sigma_V(V) = \sqrt{2}\sigma_{m-1}.\sigma_{m-2}$ where σ_{m-1} denotes the surface area of $S^{m-1} \subset R^m$, that is $\sigma_{m-1} = \frac{2\pi^{m/2}}{\Gamma(m/2)}$.

- $Y : V \to \mathbb{R}^{2m}$ is the random field defined by

$$Y(s,t) = \begin{pmatrix} \pi_s(Bs) \\ \pi_t(Bt) \end{pmatrix}.$$

- For $\tau = (s,t)$ a given point in V, we have that

$$Y(\tau) \in \{(t,-s)\}^\perp \cap \left[\{s\}^\perp \times \{t\}^\perp\right] = W_\tau$$

  for any value of the matrix B, where $\{(t,-s)\}^\perp$ is the orthogonal complement of the point $(t,-s)$ in $\mathbb{R}^{2m}$. Notice that $\dim(W_\tau) = 2m-3$.

- $p_{X(s),X(t),Y(s,t)}$ is the density of the triplet $X(s), X(t), Y(s,t)$ in $\mathcal{R} \times \mathcal{R} \times W_{(s,t)}$.

- $\Delta(\tau) = \left[\det\left[(Y'(\tau))^T Y'(\tau)\right]\right]^{\frac{1}{2}}$.

Next, we compute the ingredients in the right-hand member of (10). This has some differences with the centered case.

Put $a_{ij} = m_{ij} + g_{ij}$ with the g_{ij}'s i.i.d. Gaussian standard and $G = ((g_{ij}))$.

For each $(s,t) \in V$, we take an orthonormal basis of $\mathcal{R}^m$ so that its first two elements are respectively s and t, say $\{s, t, w_3, ..., w_m\}$. When expressing the linear transformation $x \rightsquigarrow A.x$ ($x \in \mathcal{R}^m$) in this new basis, we denote $A^{s,t}$ the associated matrix and by $a_{ij}^{s,t}$ its i,j entry. In a similar way we get $G^{s,t}, M^{s,t}, B^{s,t}$. Notice that $G^{s,t}$ has the same law than G, but the non-random part $M^{s,t}$ can vary with the point (s,t).

We denote by $B_1^{s,t}$ (respectively $B_2^{s,t}$) the $(m-1)\times(m-1)$ matrix obtained from $B^{s,t}$ by supressing the first (respectively the second) row and column. $B_{1,2}^{s,t}$ denotes the $(m-2)\times(m-2)$ matrix obtained from $B^{s,t}$ by supressing the first and second row and column.

To get an estimate for the right-hand member in (10) we start with the density $p_{X(s),X(t),Y(s,t)}(a,b,0)$.

We denote $B^{s,t} = ((b_{ij}^{s,t}))$ (and similarly for the other matrices).

We have:

$$\begin{aligned}
X(s) &= b_{11}^{s,t} \\
X(t) &= b_{22}^{s,t} \\
X''(s) &= B_1^{s,t} - b_{11}^{s,t} I_{m-1} \\
X''(t) &= B_2^{s,t} - b_{22}^{s,t} I_{m-1}.
\end{aligned}$$

Take the following orthonormal basis of the subspace $W_{(s,t)}$:

$$\{(w_3,0), ..., (w_m,0), (0,w_3), ..., (0,w_m), \frac{1}{\sqrt{2}}(t,s)\} = L_{s,t}$$

Since the expression of $Y(s,t)$ in the canonical basis of $\mathcal{R}^{2m}$ is:

$$Y(s,t) = (0, b_{21}^{s,t}, b_{31}^{s,t}, ..., b_{m1}^{s,t}, b_{12}^{s,t}, 0, b_{32}^{s,t}, ..., b_{m2}^{s,t}, b_{12}^{s,t})^T,$$

it is written in the orthonormal basis $L_{s,t}$ as the linear combination:

$$Y(s,t) = \sum_{i=3}^{m} \left[b_{i1}^{s,t}.(w_i,0) + b_{i2}^{s,t}.(0,w_i) \right] + \sqrt{2} b_{12}^{s,t}. \left[\frac{1}{\sqrt{2}}(t,s) \right]$$

It follows that the joint density of $X(s), X(t), Y(s,t)$ appearing in (10) in the space $\mathcal{R} \times \mathcal{R} \times W_{(s,t)}$ is the joint density of the r.v.'s

$$b_{11}^{s,t}, b_{22}^{s,t}, \sqrt{2} b_{12}^{s,t}, b_{31}^{s,t}, ..., b_{m1}^{s,t}, b_{32}^{s,t}, ..., b_{m2}^{s,t}$$

at the point $(a, b, 0)$. To compute this density, first compute the joint density q of

$$b_{31}^{s,t}, ..., b_{m1}^{s,t}, b_{32}^{s,t}, ..., b_{m2}^{s,t},$$

given $a_1^{s,t}, a_2^{s,t}$, where $a_j^{s,t}$ denotes the j-th column of $A^{s,t}$, with the additional conditions that

$$\left\| a_1^{s,t} \right\| = b_{11}^{s,t} = a, \ \left\| a_2^{s,t} \right\| = b_{22}^{s,t} = b, \ \left\langle a_1^{s,t}, a_2^{s,t} \right\rangle = b_{12}^{s,t} = 0$$

q is the normal density in $\mathbb{R}^{2(m-2)}$, with the same variance matrix as in the centered case, that is

$$\begin{pmatrix} a.I_{m-2} & 0 \\ 0 & b.I_{m-2} \end{pmatrix}.$$

but not necessarily centered.

So, the conditional density q is bounded above by

$$\frac{1}{(2\pi)^{m-2}} \frac{1}{(ab)^{\frac{m-2}{2}}}. \tag{11}$$

Our next task is to obtain an upper bound useful for our purposes for the density of the triplet

$$(b_{11}^{s,t}, b_{22}^{s,t}, b_{12}^{s,t}) = (\|a_1^{s,t}\|^2, \|a_2^{s,t}\|^2, < a_1^{s,t}, a_2^{s,t} >)$$

at the point $(a, b, 0)$ which together with (11) will provide an upper bound for $p_{X(s),X(t),Y(s,t)}(a, b, 0)$. We do this in the next Lemma, which we will apply afterwards with $\xi = a_1^{s,t}, \eta = a_2^{s,t}$.

Lemma 1 *Let ξ, η be to independent Gaussian vectors in $\mathcal{R}^m$ ($m \geq 2$), $E(\xi) = \mu$, $E(\eta) = \nu, Var(\xi) = Var(\eta) = I_m$.*

Then, the density p of the random triplet $\left(\|\xi\|^2, \|\eta\|^2, \langle \xi, \eta \rangle \right)$ satisfies the following inequality, for $a \geq 4 \|\mu\|^2$:

$$p(a, b, 0) \leq \frac{1}{4(2\pi)^m} \sigma_{m-1} \sigma_{m-2} (ab)^{\frac{m-3}{2}} \exp(-\frac{a}{8}) \quad (a, b > 0) \tag{12}$$

Proof. Let $F : \mathcal{R}^m \times \mathcal{R}^m \to \mathcal{R}^3$ be the function

$$F(x, y) = \left(\|x\|^2, \|y\|^2, \langle x, y \rangle \right)^T$$

According to the coarea formula, the density p at the point $(a,b,0)$ can be written as

$$\begin{aligned} & p(a,b,0) \qquad (13) \\ = & \int_{F^{-1}(a,b,0)} \left(\det\left[F'(x,y).\left(F'(x,y)\right)^T\right]\right)^{-\frac{1}{2}} \frac{1}{(2\pi)^m} e^{-\frac{1}{2}\left[\|x-\mu\|^2+\|y-\nu\|^2\right]} d\gamma(x,y) \end{aligned}$$

where γ denotes the geometric measure on $F^{-1}(a,b,0)$.

The C^∞-differentiable manifold $F^{-1}(a,b,0)$ is given by the set of equations

$$\|x\|^2 = a, \quad \|y\|^2 = b, \quad \langle x,y\rangle = 0$$

and has dimension $2m-3$. One can verify that

$$\gamma(F^{-1}(a,b,0)) = (a+b)^{\frac{1}{2}} \sigma_{m-1}\sigma_{m-2}(ab)^{\frac{m-2}{2}}.$$

(note that the manifold V considered above is $F^{-1}(1,1,0)$).

On the other hand,

$$F'(x,y) = \begin{pmatrix} 2.x^T & 0 \\ 0 & 2.y^T \\ y^T & x^T \end{pmatrix}$$

so that if $(x,y) \in F^{-1}(a,b,0)$, one gets:

$$\det\left[F'(x,y).\left(F'(x,y)\right)^T\right] = 16.ab(a+b)$$

Replacing into (13) and taking into account condition $a \geq 4\|\mu\|^2$, the result in the Lemma follows.□

Summing up this part, (11) plus (12) imply that

$$p_{X(s),X(t),Y(s,t)}(a,b,0) \leq \frac{1}{2^{2m-\frac{3}{2}}\pi^{m-2}} \frac{1}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m-1}{2}\right)} \frac{\exp(-\frac{a}{8})}{\sqrt{ab}} \qquad (14)$$

We now consider the conditional expectation in (10).

First, observe that the $(2m-3)$- dimensional tangent space to V at the point (s,t) is parallel to the orthogonal complement in $\mathcal{R}^m \times \mathcal{R}^m$ of the triplet of vectors $(s,0);(0,t);(t,s)$. This is immediate from the definition of V.

To compute the associated matrix for $Y'(s,t)$ take the set

$$\{(w_3,0),...,(w_m,0),(0,w_3),...,(0,w_m),\frac{1}{\sqrt{2}}(t,-s)\} = K_{s,t}.$$

as orthonormal basis in the tangent space. As for the codomain of Y, we take the canonical basis in $\mathcal{R}^{2m}$.

A direct calculation gives :

$$Y'(s,t) = \begin{pmatrix} -v^T & 0_{1,m-2} & -\frac{1}{\sqrt{2}} b_{21}^{s,t} \\ w^T & 0_{1,m-2} & \frac{1}{\sqrt{2}}(-b_{11}^{s,t} + b_{22}^{s,t}) \\ B_{12}^{s,t} - b_{11}^{s,t} I_{m-2} & 0_{m-2,m-2} & \frac{1}{\sqrt{2}} w \\ 0_{1,m-2} & -w^T & \frac{1}{\sqrt{2}}(-b_{11}^{s,t} + b_{22}^{s,t}) \\ 0_{1,m-2} & v^T & \frac{1}{\sqrt{2}} b_{21}^{s,t} \\ 0_{m-2,m-2} & B_{12}^{s,t} - b_{22}^{s,t} I_{m-2} & -\frac{1}{\sqrt{2}} v \end{pmatrix}$$

where $v^T = (b_{31}^{s,t}, ..., b_{m1}^{s,t})$, $w^T = (b_{32}^{s,t}, ..., b_{m2}^{s,t})$, $0_{i,j}$ is a null matrix with i rows and j columns. The columns represent the derivatives in the directions of $K_{s,t}$ at the point (s,t). The first m rows correspond to the components of $\pi_s(Bs)$, the last m ones to those of $\pi_t(Bt)$.

Thus, under the conditioning in (10),

$$Y'(s,t) = \begin{pmatrix} 0_{1,m-2} & 0_{1,m-2} & 0 \\ 0_{1,m-2} & 0_{1,m-2} & \frac{1}{\sqrt{2}}(b-a) \\ B_{12}^{s,t} - aI_{m-2} & 0_{m-2,m-2} & 0_{m-2,1} \\ 0_{1,m-2} & 0_{1,m-2} & \frac{1}{\sqrt{2}}(b-a) \\ 0_{1,m-2} & 0_{1,m-2} & 0 \\ 0_{m-2,m-2} & B_{12}^{s,t} - bI_{m-2} & 0_{m-2,1} \end{pmatrix}$$

and

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

Since $B_{12}^{s,t} \succ 0$ one has

$$|\det(B_{12}^{s,t} - aI_{m-2})| \mathbb{1}_{B_{12}^{s,t} - aI_{m-2} \prec 0} \leq a^{m-2}$$

and the conditional expectation in (10) is bounded by:

$$a^{m-1} E\left[\begin{array}{c} |\det(B_{12}^{s,t} - bI_{m-2})| \mathbb{1}_{B_{12}^{s,t} - bI_{m-2} \succ 0} / \\ b_{11}^{s,t} = a, b_{22}^{s,t} = b, b_{12}^{s,t} = 0, b_{i1}^{s,t} = b_{i2}^{s,t} = 0 \ (i = 3, ..., m). \end{array}\right] \tag{15}$$

We further condition on $a_1^{s,t}$ and $a_2^{s,t}$, with the additional requirement that $\left\|a_1^{s,t}\right\|^2 = a$, $\left\|a_2^{s,t}\right\|^2 = b$, $\left\langle a_1^{s,t}, a_2^{s,t}\right\rangle = 0$. Since unconditionally,

$a_3, ..., a_m$ are independent Gaussian vectors in $\mathcal{R}^m$ each having variance equal to 1 and mean smaller or equal to $\|M\|$, under the conditioning, their joint law becomes the law of $(m-2)$ Gaussian vectors in $\mathcal{R}^{m-2}$,independent of the condition and also having variance equal to 1 and mean with Euclidean norm smaller than or equal to $\|M\|$.

As a consequence, the conditional expectation in (15) is bounded by

$$E\left(\det(C^{s,t})\right)$$

where $C^{s,t}$ is an $(m-2)\times(m-2)$ random matrix, $C^{s,t} = ((c_{ij}^{s,t}))$, $c_{ij}^{s,t} =< u_i^{s,t}, u_j^{s,t} >$, $(i,j = 3, ..., m)$,

$$u_i^{s,t} = \zeta_i + \mu_i^{s,t} \qquad i = 3, ..., m,$$

$\zeta_3, ..., \zeta_m$ are i.i.d. standard normal in $\mathcal{R}^{m-2}$ and $\left\|\mu_i^{s,t}\right\| \leq \|M\|$ for $i = 3, ..., m$.

The usual argument to compute $\det(C^{s,t})$ as the square of the volume in $\mathcal{R}^{m-2}$ of the set of linear combinations of the form $\sum_{i=3}^{i=m} \lambda_i u_i^{s,t}$ with $0 \leq \lambda_i \leq 1$ $(i = 3, ..., m)$, shows that

$$\begin{aligned} E\left(\det(C^{s,t})\right) &\leq \left(1 + \|M\|^2\right)\left(2 + \|M\|^2\right)...\left(m - 2 + \|M\|^2\right) \\ &= (m-2)! \prod_{i=1}^{i=m-2}\left(1 + \frac{\|M\|^2}{i}\right) \\ &\leq (m-2)!\left[\left(1 + \|M\|^2 \frac{1+\log m}{m}\right)\right]^m \end{aligned}$$

where we have bounded the geometric mean by the arithmetic mean.

Replacing in (15) and on account of the bound (14) we get from (10) the following bound for the joint density, valid for $a \geq 4\|M\|^2$:

$$g(a,b) \leq \overline{C}_m \frac{e^{-\frac{a}{8}}}{\sqrt{ab}} a^{m-1} \tag{16}$$

where

$$\overline{C}_m = \frac{1}{4(m-2)!}\left[1 + \|M\|^2 \frac{1+\log m}{m}\right]^m$$

We now turn to the proof of (8).

One has, for $x > 1$:

$$P(\kappa(A) > x) = P(\frac{\lambda_m}{\lambda_1} > x^2) \leq P(\lambda_1 < \frac{L^2 m}{x^2}) + P(\frac{\lambda_m}{\lambda_1} > x^2, \lambda_1 \geq \frac{L^2 m}{x^2}) \tag{17}$$

where L is a positive number to be chosen later on.

For the first term in (17), we use Proposition 9 in [2], which is a slight modification of Theorem 3.2. in [7]; see also [4]:

$$P(\lambda_1 < \frac{L^2 m}{x^2}) = P(\|A^{-1}\| > \frac{x}{L\sqrt{m}}) \leq C_2 \frac{Lm}{x} \tag{18}$$

where C_2 is a constant, $C_2 \approx 2.35$.

Impose first on L the condition

$$L^2 m \geq 4 \|M\|^2$$

so that for the second term in (17) we can make use of the bound (16) on the joint density $g(a,b)$:

$$P(\frac{\lambda_m}{\lambda_1} > x^2, \lambda_1 \geq \frac{L^2 m}{x^2}) = \int_{L^2 m x^{-2}}^{+\infty} db \int_{bx^2}^{+\infty} g(a,b) da \leq H_m(x^2) \tag{19}$$

with

$$H_m(y) = \overline{C}_m \int_{L^2 m y^{-1}}^{+\infty} db \int_{by}^{+\infty} \frac{\exp(-\frac{a}{8})}{\sqrt{ab}} a^{m-1} da,$$

We have:

$$H_m'(y) = \overline{C}_m \left[\begin{array}{l} -\int_{L^2 m y^{-1}}^{+\infty} \exp(-\frac{by}{8})(by)^{m-1} \frac{db}{\sqrt{y}} \\ +\frac{Lm^{\frac{1}{2}}}{y^{\frac{3}{2}}} \int_{L^2 m}^{+\infty} \exp\left(-\frac{a}{4}\right) a^{m-3/2} da \end{array} \right]$$

which implies

$$\begin{aligned} -H_m'(y) &\leq \overline{C}_m y^{m-3/2} \int_{L^2 m y^{-1}}^{+\infty} \exp\left(-\frac{by}{8}\right) b^{m-1} db \\ &\leq \frac{\overline{C}_m}{y^{3/2}} 8^m \int_{\frac{L^2 m}{8}}^{+\infty} e^{-z} z^{m-1} dz \leq \frac{\overline{C}_m}{y^{3/2}} 8^m \frac{5}{3} e^{-\frac{L^2 m}{8}} \left(\frac{L^2 m}{8}\right)^{m-1} = \overline{D}_m \frac{1}{y^{3/2}} \end{aligned}$$

if we choose $L^2 > 20$.

So,

$$H_m(y) = -\int_y^{+\infty} H_m'(s) ds \leq \overline{D}_m \int_y^{+\infty} \frac{ds}{s^{\frac{3}{2}}} \leq 2\overline{D}_m \frac{1}{y^{1/2}} \tag{20}$$

where

$$\overline{D}_m \leq \frac{10}{3\sqrt{2\pi} L^2} \frac{m}{\sqrt{m-2}} \exp\left[\left(1 - \frac{L^2}{8} + \log L^2 + \log\theta\right) m\right]$$

where $\theta = 1 + \|M\|^2 \frac{1+\log m}{m}$.

Choosing

$$L = 2\sqrt{2}(1+4\theta)^{\frac{1}{2}}$$

conditions $L^2 > 20$ and $L^2 m \geq 4\|M\|^2$ are verified and $1 - \frac{L^2}{8} + \log L^2 + \log\theta < 0$.

Hence,

$$2\overline{D}_m \leq \frac{1}{4}\sqrt{\frac{m}{2\pi}}.$$

On account of (18), (19) and (20), replacing in the right-hand side of (17), inequality (8) in the statement of the Theorem follows.□

The author wants to thank Professor Jean-Marc Azaïs for frutiful discussions on the subject.

REFERENCES.

[1] Azaïs, J-M; Wschebor, M. Upper and lower bounds for the tails of the distribution of the condition number of a Gaussian matrix, 2003, *submitted.*

[2] Cuesta-Albertos, J.; Wschebor, M. Condition numbers and the extrema of random fields. In *Proc. IV Ascona Seminar, Birkhaüser*, 2003, to appear.

[3] Demmel, J.W. *Applied numerical linear algebra.* Philadelphia Pa. SIAM, 1997.

[4] Edelman, A. Eigenvalues and condition numbers of random matrices. *SIAM J. of Matrix Anal. and Applic., 1988, 9*, 543-556.

[5] Higham, N.J. *Accuracy and stability of numerical algorithms.* Philadelphia Pa. SIAM, 1996.

[6] Ledoux, M. A remark on hypercontractivity and tail inequalities for the largest eigenvalues of random matrices. *Séminaire de Probabilités XXXVII. Lecture Notes in Math. Springer-Verlag*, 2003, to appear.

[7] Sankar, A.; Spielman, D.A.;Teng, S.H. Smoothed analysis of the condition numbers and growth factors of matrices, 2002, *preprint.*

[8] Smale, S. On the efficiency of algorithms of analysis. *Bull. Amer. Math. Soc. (N.S.), 1985, 13 (2)*, 87-121.

[9] Spielman, D.A.; Teng, S.H. Smoothed Analysis of Algorithms. *ICM 2002, Beijing, Vol. I*, 597-606, 2002.

[10] Szarek, S.J. *Condition numbers of random matrices. J. Complexity, 1991, 7 (2)*, 131-149.

[11] Turing, A.M. Rounding-off errors in matrix processes. *Quart. J. Mech. Appl. Math. 1948, 1*, 287-308.

[12] von Neumann, J. ; Goldstine H.H. Numerical inverting of matrices of high order. *Bull. Amer. Math. Soc., 1947, 53*, 1021-1099.

[13] Wilkinson, J.H. *Rounding errors in algebraic processes.* Englewood Cliffs, N.J., Prentice Hall Inc., 1963.