Condition Numbers and Extrema of Random Fields.

J.A. Cuesta-Albertos*
Departamento de Matemáticas, Estadística y Computación,
Universidad de Cantabria. Santander. Spain
E-mail: cuestaj@unican.es

Mario Wschebor
Centro de Matemática, Facultad de Ciencias
Universidad de la Republica. Montevideo. Uruguay.
E-mail: wschebor@cmat.edu.uy

May 20, 2003

1 Introduction.

Let A be an $m \times m$ real matrix. We denote by

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

its Euclidean operator norm, where we denote by $\|v\|$ the Euclidean norm of $v \in \mathcal{R}^m$. If A is non-singular, its *condition number* $\kappa(A)$ is defined by

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

The role of $\kappa(A)$ in Numerical Linear Algebra has been recognized since a long time [11], [12], [13] as well as its importance in the evaluation of algorithm complexity [5], [7]. $\kappa(A)$ measures, to the first order of approximation, the largest expansion in the relative error of the solution of the $m \times m$ linear system of equations

$$Ax = b \tag{1}$$

when its input is measured with error.

In other words, $\log_2 \kappa(A)$ is the loss of precision in the solution x of (1) due to ill-conditioning of A, measured in number of places in the finite binary expansion of x.

A natural problem is trying to understand the behaviour of $\kappa(A)$ when the matrix A is chosen at random, that is, to estimate the tail

$$P[\kappa(A) > x], \text{ for each } x \in \mathcal{R}^+,$$

*This author has been partially supported by the Spanish Ministerio de Ciencia y Tecnología, grants PB98-0369-C02-02 and BFM2002-04430-C02-02.

(where P is the probability defined on the probability space in which A is defined) or the moments of the random variable $\kappa(A)$. Of course, a priori this will depend on the meaning of "choosing A at random", that is, which is the probability distribution of A. A typical result is the following:

Theorem 1 (Edelman, 1988) *Let* $A = (a_{i,j})_{i,j=1,\dots,m}$ *and assume that the* $a_{i,j}$*'s are i.i.d. Gaussian standard random variables. Then:*

$$E\{\log\kappa(A)\} = \log m + C_0 + \varepsilon_m, \tag{2}$$

where C_0 *is a konwn constant* ($C_0 \cong 1,537$) *and* $\varepsilon_m \to 0$ *as* $m \to +\infty$.

In [3] one can find some elementary inequalities for the moments of $\log\kappa(A)$ when the entries of A are i.i.d. but not necessarily Gaussian. In a recent paper [9] bounds for $P[\kappa(A) > x]$ are given when the $a_{i,j}$'s are i.i.d. Gaussian with a common variance but may be non-centered (this has been called "smoothed analysis"). More precisely:

Theorem 2 (Sankar, Spielman, Teng, 2002) *Assume* $a_{i,j} = m_{i,j} + g_{i,j}$ $(i,j = 1,...,m)$ *where the* $g_{i,j}$*'s are i.i.d. centered Gaussian with common variance* $\sigma^2 \le 1$ *and the (non random) matrix*

$$M = (m_{i,j})_{i,j=1,\dots,m}$$

verifies $\|M\| \le m^{\frac{1}{2}}$.

Then, there exists x_0 *such that, if* $x > x_0$, *then*

$$P[\kappa(A) > x] \le \frac{4.734m\left(1 + 4(\log x)^{\frac{1}{2}}\right)}{x\sigma}. \tag{3}$$

Remark 3 There are a few differences between this statement and the actual statement in [9]. The first one is that instead of 4.734 their constant is 3.646, apparently due to a mistake in the numerical evaluation. The second one, their hypothesis is $\sup_{i,j}|m_{i,j}| \le 1$ instead of $\|M\| \le m^{\frac{1}{2}}$, which they actually use in their proof and which is not implied by the previous one. Finally, the inequality from [10], which is applied in their proof, does not apply for every $x > 0$.

If one denotes $\lambda_1, ...\lambda_m$, $0 \le \lambda_1 \le .... \le \lambda_m$, the eigenvalues of the matrix A^tA (A^t stands for the transpose of A), then

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

where

$$M_A = \max_{\|x\|=1} f(x); \quad m_A = \min_{\|x\|=1} f(x); \quad f(x) = x^tA^tAx \ (x \in \mathcal{R}^m).$$

It is possible to study the random variable $\kappa(A)$ using techniques related to extrema of random fields. More precisely, if $a > 0$:

$$P[M_A > a] = P[M^+(X,a) \geq 2] \leq \frac{1}{2} E\left\{M^+(X,a)\right\}, \tag{4}$$

where, if S^{m-1} is the unit sphere in the m-dimensional euclidean space, then X is the real-valued random field

$$X = f \mid S^{m-1}$$

and

$$M^+(X,a) = \#\left\{x : x \in S^{m-1},\ X \ \text{ has a local maximum at the point } x \text{ and } X(x) > a\right\}$$

(note that since f is an even function, $\{M_A > a\}$ occurs if and only if $\{M^+(X,a) \geq 2\}$).

The main point in making inequality (4) a useful tool is that the expectation in the right-hand side member can be computed - or at least estimated - using Rice formula for the expectation of the number of critical points of the random field X' (the derivative of X).

In fact, we will only use an upper bound for $E\{M^+(X,a)\}$, as will be explained below. The upper bound thus obtained for $P[M_A > a]$ will be one of the tools to prove Theorem 11 which contains a variant of (3) that implies an improvement if x is large enough. However Conjecture 1 in [9] which states that $P[\kappa(A) > x] \leq O(\frac{m}{x\sigma})$ remains an open problem.

Inequality in Proposition 6 is a variant of results that have been known since a certain time (see for example Lemma 2.8 in [10]).

Our main point here is the connection between the spectrum of random matrices and the zeros of random fields which makes useful Rice formulae for the moments of the number of zeros. In our context, inequality (13) is interesting for large values of a for which the classical inequalities are of the same order. Note also that in this case, the constant $1/4$ in the exponent can be replaced by any constant strictly smaller than $1/2$, if a is large enough.

On the other hand, for the time being, this method does not provide the precise bounds on the distribution of the largest eigenvalue of a Wishart matrix for values of a close to $a = 2$ (c.f. [4] or [8]).

These inequalities permit to deduce inequalities for the moments of $\log \kappa(A)$, as in Corollary 13, which gives a bound for $E\{\log \kappa(A)\}$ for non-centered random matrices. This also leads to an alternative proof of a weak version of Edelman's Theorem, which instead of (2) states that

$$E\{\log \kappa(A)\} \leq \log m + C \tag{5}$$

for some constant C.

Rice formulae for the moments of the number of zeros of a random field can be applied in some other related problems, which are in fact more complicated than the one we are adressing here. In [2] this is the case for condition numbers in linear programming. We briefly sketch one of the results in this paper.

Consider the system of inequalities

$$Ax < 0 \tag{6}$$

where A is an $n \times m$ real matrix, $n > m$, and $y < 0$ denotes that all the coordinates of the vector y are negative. In [1] the following condition number was defined, for the (feasibility) problem of determining wheather the set of solutions of (6) is empty or not.

Denote by $a_1^t, ..., a_n^t$ the rows of A,

$$f_k(x) = \frac{a_k^t x}{\|a_k\|} \ (k = 1, ..., n), \quad D(A) = \min_{x \in S^{m-1}} \max_{1 \le k \le n} f_k(x).$$

Cheung-Cucker condition number is

$$\mathcal{C}(A) = |D(A)|^{-1},$$

with the convention $\mathcal{C}(A) = +\infty$ when $D(A) = 0$. [2] contains the following result:

Theorem 4 *Assume that*

$$\frac{m(1 + \log n)}{n} \le 1.$$

If $a_{i,j}, , i = 1, ..., n, j = 1, ..., m$ are i.i.d. Gaussian standard random variables, then

$$E\{\log \mathcal{C}(A)\} \le \max(\log m, \log\log n) + K, \tag{7}$$

where K is a constant.

To prove (7) one can also use a method based upon the formulae on extrema of random fields, since the problem consists in giving fine bounds for the probability

$$P\left[\left|\min_{x \in S^{m-1}} Z(x)\right| < b\right],$$

where $Z(x) = \max_{1 \le k \le n} f_k(x)$ and b is a (small) positive number.

The difference between the proof of Theorem 4 and the study of $\kappa(A)$ for square matrices is that in the latter case the random function X that is to be considered is the restriction of a quadratic form to the unit sphere, hence a nice regular function, while the study of the local extrema of Z is complicated. This is due to the fact that $x \to Z(x)$ is a non-differentiable piecewise affine function.

The plan of the paper is as follows. In Section 2 we include some technical results which are required to state main results in this paper. Those results appear in Section 3.

2 Technical preliminaries.

In this section, $B_{m-1}(0, \delta)$ is the Euclidean ball centered at the origin with radius δ in $\mathcal{R}^{m-1}$, $|B_{m-1}(0, \delta)|$ is its Lebesgue measure, σ_{m-1} the $(m-1)$-dimensional geometric measure in S^{m-1} and $T \prec 0$ denotes that the bilinear form T is negative definite.

Proposition 5 (Kac's formula) *Let $F : S^{m-1} \to \mathcal{R}$ be a C^2 function. Denote:*

$$\mathcal{M}^+(F,a) = \left\{x : x \in S^{m-1},\ F \ \text{ has a local maximum at the point } x \text{ and } F(x) > a\right\},$$

$$M^+(F,a) = \#\left[\mathcal{M}^+(F,a)\right].$$

We assume that

$$\left\{x : x \in S^{m-1},\ F'(x) = 0,\ \det(F''(x)) = 0\right\} = \phi \tag{8}$$

i.e. that there are no critical points of F in which F'' is singular.

Then,

$$M^+(F,a) = \lim_{\delta \downarrow 0} \frac{1}{|B_{m-1}(0,\delta)|} \int_{S^{m-1}} |\det(F''(x))|\, 1_{\{\|F'(x)\|<\delta, F''(x)\prec 0, F(x)>a\}} \sigma_{m-1}(dx) \tag{9}$$

Proof. The hypothesis implies that the points of $\mathcal{M}^+(F,a)$ are isolated, hence, that $M^+(F,a)$ is finite. Put

$$\mathcal{M}^+(F,a) = \{x_1, ..., x_N\}.$$

Then, for $j = 1, ..., N$:

$$F'(x_j) = 0; \quad F''(x_j) \prec 0.$$

If δ_0 is small enough, using the inverse function theorem, there exist pairwise disjoint open neighbourhoods $U_1, ..., U_N$ in S^{m-1} of the points $x_1, ..., x_N$ respectively, such that for each $j = 1, ..., N$ the map $x \to F'(x)$ is a diffeomorphism between U_j and $B_{m-1}(0,\delta_0)$ and

$$\bigcup_{j=1}^{N} U_j = \left\{x : x \in S^{m-1}, \|F'(x)\| < \delta_0, F''(x) \prec 0, F(x) > a\right\}.$$

Using the change of variable formula,

$$\int_{U_j} |\det(F''(x))|\, \sigma_{m-1}(dx) = |B_{m-1}(0,\delta_0)|,$$

it follows that

$$\begin{aligned} N\,|B_{m-1}(0,\delta_0)| &= \int_{\bigcup_{j=1}^N U_j} |\det(F''(x))|\, \sigma_{m-1}(dx) \\ &= \int_{S^{m-1}} |\det(F''(x))|_{\{\|F'(x)\|<\delta_0, F''(x)\prec 0, F(x)>a\}}\, \sigma_{m-1}(dx). \end{aligned}$$

This proves (9). ■

Suppose now that $X(x) = x^t A^t A x$, $A = (a_{i,j})_{i,j=1,...,m}$, $x \in \mathcal{R}^m$ and introduce the following notations: for $x \in S^{m-1}$, let $\{e_1, ..., e_m\}$ be an orthonormal basis of $\mathcal{R}^m$ such that $e_1 = x$. We denote $A^x = (a^x_{i,j})_{i,j=1,...,m}$ the matrix associated to the linear transformation in $\mathcal{R}^m$ defined by $y \to Ay$, when one takes $\{e_1, ..., e_m\}$ as reference basis.

Put also $B^x = (A^x)^t A^x = (b^x_{i,j})_{i,j=1,...,m}$, $b^x_{i,j} = \sum_{h=1}^m a^x_{h,i} a^x_{h,j}$.

Direct computations show that:

$$X(x) = b^x_{1,1} \tag{10}$$

$$X'(x) = 2\left(b^x_{2,1}, ..., b^x_{m,1}\right)^t \tag{11}$$

$$X''(x) = 2\begin{pmatrix} b^x_{2,2} - b^x_{1,1} & .... & b^x_{2,m} \\ ... & .... & .... \\ b^x_{m,2} & ... & b^x_{m,m} - b^x_{1,1} \end{pmatrix} = 2(B^x_{2,2} - b^x_{1,1} I_{m-1}) \tag{12}$$

(I_{m-1} is the $(m-1) \times (m-1)$ identity matrix).

In the rest of this paper, $G = (g_{i,j})_{i,j=1,...,m}$ will denote a random matrix with i.i.d. centered Gaussian entries and common variance σ^2 and $A = (a_{i,j})_{i,j=1,...,m}$, where $a_{i,j} = g_{i,j} + m_{i,j}$ and $M = (m_{i,j})_{i,j=1,...,m}$ is a nonrandom matrix.

Since our interest is in studying $\kappa(A)$, the fact that $\kappa(\frac{1}{\sigma}A) = \kappa(A)$, for every $\sigma \neq 0$, implies that we may assume that $\sigma = 1$ if we replace the expected matrix M by $\frac{1}{\sigma}M$.

The next proposition is a variant of a known inequality (Lemma 2.8 in [10] and references therein). We give here an independent proof.

Proposition 6 *Assume $G = (g_{i,j})_{i,j=1,...,m}$ where the $g_{i,j}$'s are i.i.d. random variables, each one having standard Gaussian probability distribution. Assume $m \geq 3$.*

Then, for each $a \geq 4$ one has

$$P\left[\|G\| \geq a\sqrt{m}\right] \leq \frac{C_1(a)}{\sqrt{m}} \exp\left[-\frac{1}{4}a^2 m\right], \tag{13}$$

where $C_1(a) = \frac{36\sqrt{2}e^{-2}}{7a^3\sqrt{\pi}} \leq \frac{36\sqrt{2}e^{-2}}{4^3 7\sqrt{\pi}} = C_1 \approx 0.008677...$

Proof. *Step 1.*

We consider the quadratic form defined on $\mathcal{R}^m$:

$$f_G(x) = x^t G^t G x.$$

We have, for $t > 0$:

$$P\left[\|G\|^2 > t\right] \leq \frac{1}{2} E\left\{M^+(f_G, t)\right\}. \tag{14}$$

To be able to apply Proposition 5 to $M^+(f_G, t)$ we need to check condition (8). One way to do this is to use Proposition 4 in [2], applying it to the random vector field $V = f'_G$ since the random variable $f'_G(x)$ has a bounded density in $\mathcal{R}^{m-1}$. One can conclude that almost surely formula (9) holds true for $F = f_G$.

Step 2.

For each $x \in S^{m-1}$ let us compute the joint distribution of $f_G(x)$ and $f'_G(x)$ in $\mathcal{R} \times \mathcal{R}^{m-1}$.

Note first that due to the invariance under linear isometries, this joint distribution is the same for all $x \in S^{m-1}$. We compute it for $x = w = (1, 0, ..., 0)^t$. Notice that, in this case $A^w = A$, $B^w = B, ...$

Conditionally on $(g_{1,1},...,g_{m,1})$, $f_G(w)$ is constant and equal to $b_{1,1}=\sum_{h=1}^m g_{h,1}^2$ and the random variables

$$b_{i,1}=\sum_{h=1}^m g_{h,i}g_{h,1} \quad (i=2,...,m)$$

are independent, each one being Gaussian centered with variance $b_{1,1}$.

So, since the distribution of $b_{1,1}$ is χ^2 with m degrees of freedom and on account of (10) and (11), the joint density of $f_G(w)$ and $f'_G(w)$ is equal to:

$$\begin{aligned} p_{f_G(w),f'_G(w)}(y,z) &= \chi^2_m(y)\frac{\exp\left[-\frac{1}{2}\frac{\|z\|^2}{4y}\right]}{(2\pi)^{\frac{m-1}{2}}2^{m-1}y^{\frac{m-1}{2}}} \\ &= \frac{1}{2^{\frac{3m}{2}-1}\Gamma(\frac{m}{2})(2\pi)^{\frac{m-1}{2}}}\frac{\exp\left[-\frac{1}{2}\left(y+\frac{\|z\|^2}{4y}\right)\right]}{\sqrt{y}}. \end{aligned} \tag{15}$$

On the other hand, using Step 1 and Proposition 5, we obtain:

$$\begin{aligned} &2P(\|G\|^2>t) \\ &\leq E\left\{\lim_{\delta\downarrow 0}\frac{1}{|B_{m-1}(0,\delta)|}\int_{S^{m-1}}|\det(f''_G(x))|\,1_{\{\|f'_G(x)\|<\delta,f''_G(x)\prec 0,f_G(x)>t\}}\sigma_{m-1}(dx)\right\} \\ &\leq \int_t^{+\infty}dy\int_{S^{m-1}}E\left\{|\det(f''_G(x))|\,1_{f''_G(x)\prec 0}/f_G(x)=y,f'_G(x)=0\right\}p_{f_G(x),f'_G(x)}(y,0)\ \sigma_{m-1}(dx) \\ &= \sigma_{m-1}(S^{m-1})\int_t^{+\infty}E\left\{|\det(f''_G(w))|\,1_{f''_G(w)\prec 0}/f_G(w)=y,f'_G(w)=0\right\}p_{f_G(w),f'_G(w)}(y,0)dy. \end{aligned}$$

In the last equality we have used again the fact that the law of the random field $\{f_G(x):x\in S^{m-1}\}$ is invariant under a linear isometry of $\mathcal{R}^m$.

Substituting the density from (15) and taking into account that

$$\sigma_{m-1}(S^{m-1})=\frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})},$$

we obtain:

$$\begin{aligned} &P(\|G\|^2>t) \\ &\leq \frac{2\sqrt{2\pi}}{\left[2^m\Gamma(\frac{m}{2})\right]^2}\int_t^{+\infty}E\left\{|\det(f''_G(w))|\,1_{f''_G(w)\prec 0}/(b_{1,1},...,b_{m,1})=(y,0,...,0)\right\}\frac{\exp\left[-\frac{y}{2}\right]}{\sqrt{y}}\ dy. \end{aligned} \tag{16}$$

Step 3.

From the expression (12) for $f''_G(w)$, since $B_{2,2}$ is positive definite we have that:

$$|\det\left(f''_G(w)\right)|\,1_{\left\{f''_G(w)\prec 0\right\}}\leq(2b_{1,1})^{m-1}\,1_{\left\{f''_G(w)\prec 0\right\}}\leq(2b_{1,1})^{m-1}.$$

Replacing in the integrand in (16) we get:

$$P\left[\|G\|^2>t\right]\leq\frac{\sqrt{2\pi}}{2^m\left[\Gamma(\frac{m}{2})\right]^2}\int_t^{+\infty}y^{m-\frac{3}{2}}\exp\left[-\frac{y}{2}\right]\ dy=\frac{\sqrt{2\pi}}{2^m\left[\Gamma(\frac{m}{2})\right]^2}J_m(t).$$

For the remaining, we use the inequalities:

$$\Gamma\left(\frac{m}{2}\right) \geq \left(\frac{m}{2}-1\right)^{\frac{m}{2}-1} \exp\left[-\left(\frac{m}{2}-1\right)\right]\sqrt{2\pi\left(\frac{m}{2}-1\right)}$$

and

$$J_m(t) \leq 2t^{m-\frac{3}{2}}\left(1+\frac{1}{8}+...+\left(\frac{1}{8}\right)^{m-1}\right)\exp\left[-\frac{t}{2}\right] \leq \frac{16}{7}t^{m-\frac{3}{2}}\exp\left[-\frac{t}{2}\right],$$

the second one valid for $t \geq 16m$.

So, if $a \geq 4$:

$$\begin{aligned}
P\left[\|G\|^2 > a^2 m\right] &\leq \frac{e^{-2}}{4}\left(\frac{2}{\pi}\right)^{1/2}\frac{e^m}{(m-2)^{m-1}}\frac{16}{7}(a^2m)^{m-\frac{3}{2}}\exp\left[-\frac{a^2m}{2}\right] \\
&= \frac{4\sqrt{2}e^{-2}}{7a^3\sqrt{\pi m}}(ea^2)^m\left(1+\frac{2}{m-2}\right)^{m-1}\exp\left[-\frac{a^2m}{2}\right] \\
&\leq \frac{36\sqrt{2}e^{-2}}{7a^3\sqrt{\pi m}}(ea^2)^m\exp\left[-\frac{a^2m}{2}\right] \\
&\leq \frac{36\sqrt{2}e^{-2}}{7a^3\sqrt{\pi m}}\exp\left[-m\left(\frac{a^2}{2}-1-\log(a^2)\right)\right] \\
&\leq \frac{36\sqrt{2}e^{-2}}{7a^3\sqrt{\pi}}\frac{1}{\sqrt{m}}\exp\left[-\frac{a^2m}{4}\right],
\end{aligned}$$

which is the inequality in the statement. ■

Next we obtain an upper bound for the tail probabilities $P\left[\|A^{-1}\| > x\right]$. This was done in Theorem 3.2 in [9] .We include here a proof that in fact uses their technique and also provides a slight improvement in the numerical constant.

We will employ the following lemma.

Lemma 7 (Lemma 3.1, [9]) *Assume that* $A = (a_{i,j})_{i,j=1,...,m}$, $a_{i,j} = m_{i,j} + g_{i,j}$ $(i,j = 1,...,m)$*, where the* $g_{i,j}$*'s are i.i.d. standard Gaussian r.v.'s. Let* $v \in S^{m-1}$*. Then*

$$P\left[\left\|A^{-1}v\right\| > x\right] < \left(\frac{2}{\pi}\right)^{1/2}\frac{1}{x}.$$

Lemma 8 *Let* $U = (U_1,...,U_m)$ *be an* m*-dimensional vector chosen uniformly on* S^{m-1} *and let* t_{m-1} *be a real valued r.v. with a Student distributon with* $m-1$ *degrees of freedom.*

Then, if $c \in (0,m)$*, we have that*

$$P\left[U_1^2 > \frac{c}{m}\right] = P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right].$$

Proof. Let $V = (V_1, ..., V_m)$ be a m-dimensional random vector with standard Gaussian distribution. We can assume that

$$U = \frac{V}{\|V\|}.$$

Let us denote, to simplify the notation $K = V_2^2 + ... + V_m^2$. Then the statement

$$\frac{V_1^2}{V_1^2 + K} > \frac{c}{m}$$

is equivalent to that

$$\frac{V_1^2}{K} > \frac{c}{m-c},$$

and we have that

$$P\left[U_1^2 > \frac{c}{m}\right] = P\left[\frac{(m-1)V_1^2}{K} > \frac{m-1}{m-c}c\right] = P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right],$$

where t_{m-1} is a real valued r.v. having Student's distribution with $m-1$ degrees of freedom. ■

Proposition 9 *Assume that* $A = (a_{i,j})_{i,j=1,...,m}$, $a_{i,j} = m_{i,j} + g_{i,j}$ $(i, j = 1, ..., m)$, *where the* $g_{i,j}$*'s are i.i.d. standard Gaussian r.v.'s and* $M = (m_{i,j})_{i,j=1,...,m}$ *is non random.*

Then, for each $x > 0$:

$$P[\|A^{-1}\| \geq x] \leq C_2(m)\frac{m^{1/2}}{x}, \tag{17}$$

where

$$C_2(m) = \left(\frac{2}{\pi}\right)^{1/2}\left(\sup_{c\in(0,m)} \sqrt{c}P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right]\right)^{-1} \leq C_2(\infty) = C_2 \approx 2.34737...$$

Proof. Let U be an n-dimensional random vector, independent of A with uniform distribution on S^{m-1}.

Aplying Lemma 7 we have that

$$P\left[\|A^{-1}U\| > x\right] = E\left\{P\left[\|A^{-1}U\| > x\big/ U\right]\right\} \leq \left(\frac{2}{\pi}\right)^{1/2}\frac{1}{x}. \tag{18}$$

Now, since if w_A satisfies that $\|A^{-1}w_A\| = \|A^{-1}\|$, and $\|u\| = 1$, then,

$$\|A^{-1}u\| \geq \|A^{-1}\| \times |< w_A, u >|,$$

we have that, if $c \in (0, m)$, then

$$P\left[\|A^{-1}U\| \geq x\left(\frac{c}{m}\right)^{1/2}\right] \geq P\left[\left\{\|A^{-1}\| \geq x\right\} \text{ and } \left\{|< w_A, U >| \geq \left(\frac{c}{m}\right)^{1/2}\right\}\right]$$

$$
\begin{aligned}
&= E\left\{P\left[\left\{\|A^{-1}\| \geq x\right\} \text{ and } \left\{| < w_A, U > | \geq \left(\frac{c}{m}\right)^{1/2}\right\} \Big/ A\right]\right\} \\
&= E\left\{I_{\{\|A^{-1}\| \geq x\}} P\left[| < w_A, U > | \geq \left(\frac{c}{m}\right)^{1/2} \Big/ A\right]\right\} \\
&= E\left\{I_{\{\|A^{-1}\| \geq x\}} P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right]\right\} \\
&= P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right] P[\|A^{-1}\| \geq x].
\end{aligned}
$$

where we have applied Lemma 8. From here and (18) we have that

$$
P[\|A^{-1}\| \geq x] \;\leq\; \frac{1}{P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right]} \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x} \left(\frac{m}{c}\right)^{1/2}.
$$

To end the proof notice that, if g is a standard Gaussian random variable, then

$$
\begin{aligned}
\sup_{c\in(0,m)} c^{1/2} P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right] &\geq \sup_{c\in(0,1)} c^{1/2} P\left[t_{m-1}^2 > \frac{m-1}{m-c}c\right] \qquad (19)\\
&\geq \sup_{c\in(0,1)} c^{1/2} P\left[t_{m-1}^2 > c\right] \\
&\geq \sup_{c\in(0,1)} c^{1/2} P\left[g^2 > c\right] \\
&\geq 0.565^{1/2} P\left[g^2 > 0.565\right] = 0.3399.
\end{aligned}
$$

■

Remark 10 Explicit expressions for $C_2(m)$ don't seem to be easy to obtain. Therefore, we have carried out some numerical computations with MatLab in order to have approximations to this value.

In the following table we include the results.

Table 1. Optimal values for $C_2(m)$ and values of c in which they are reached.

m	3	4	5	10	25	50	100	∞
$C_2(m)$	1.879	2.038	2.086	2.244	2.309	2.328	2.338	2.347
c	1.146	0.923	0.823	0.672	0.604	.584	0.574	0.565

Notice from the table that restriction in (19) to that $c \in (0,1)$ is not important as long as $m \geq 4$.

3 Main results.

Theorem 11 *Assume that* $A = (a_{i,j})_{i,j=1,\dots,m}$, $a_{i,j} = m_{i,j} + g_{i,j}$ $(i,j = 1,\dots,m)$, *where the* $g_{i,j}$*'s are i.i.d. centered Gaussian with common variance* σ^2 *and* $M = (m_{i,j})_{i,j=1,\dots,m}$

is non random. Let $m \geq 3$. If $\log x \geq 4m$ one has:

$$P[\kappa(A) > x] \leq \frac{1}{x}\left[\frac{C_1}{\sqrt{m}} + C_2(m)\sqrt{m}\frac{\|M\|}{\sigma} + C_2(m)\sqrt{4m}\,(\log x)^{\frac{1}{2}}\right], \tag{20}$$

where C_1 and $C_2(m)$ were defined in Propositions 6 and 9 respectively.

Proof. As we noticed above, we may assume that $\sigma = 1$ and replace the matrix M by $\frac{1}{\sigma}M$. Put $G = (g_{i,j})_{i,j=1,\ldots,m}$. From Proposition 6, if $a \geq 4$:

$$P\left[\|A\| > \frac{1}{\sigma}\|M\| + a\sqrt{m}\right] \leq P\left[\|G\| > a\sqrt{m}\right] \leq \frac{C_1}{\sqrt{m}}\exp\left[-\frac{a^2 m}{4}\right].$$

Using also Proposition 9:

$$\begin{aligned} P[\kappa(A) > x] &\leq P\left[\|A\| > \frac{1}{\sigma}\|M\| + a\sqrt{m}\right] + P\left[\left\|A^{-1}\right\| > \frac{x}{\sigma^{-1}\|M\| + a\sqrt{m}}\right] \\ &\leq \frac{C_1}{\sqrt{m}}\exp\left[-\frac{a^2 m}{4}\right] + \frac{C_2(m)\sqrt{m}}{x}\left(\frac{\|M\|}{\sigma} + a\sqrt{m}\right). \end{aligned}$$

Putting

$$a = \sqrt{\frac{4\log x}{m}}$$

the result follows. ■

Corollary 12 *With the notations and hypotheses of Theorem 11, $m \geq 3$, for any x large enough*

$$P(\kappa(A) > x) \leq H\frac{\sqrt{m}}{x}\left[\frac{1}{m} + \frac{\|M\|}{\sigma} + (\log x)^{\frac{1}{2}}\right].$$

where H is a constant.

Proof. Apply Theorem 11. ■

One can also use Propositions 6 and 9 to get bounds for the moments of $\log\kappa(A)$. For example we can obtain the following corollary:

Corollary 13 *With the notations and hypotheses of Theorem 11. If $m \geq 3$, then*

$$E\{\log\kappa(A)\} \leq \log(m) + 1 + \log C_2 + \log\left(\frac{\|M\|}{\sigma\sqrt{m}} + 4\right) + \frac{C_1}{2m}\exp\left[-4m\right].$$

Proof. We may assume that $\sigma = 1$ and replace the matrix M by $\frac{1}{\sigma}M$. Let $\beta = \log(C_2\sqrt{m})$. Applying Proposition 9, we have that

$$\begin{aligned} E\left\{\log\left\|A^{-1}\right\|\right\} &\leq \beta + \int_{\beta}^{\infty} P\left[\left\|A^{-1}\right\| > e^x\right] \\ &\leq \beta + C_2\sqrt{m}e^{-\beta} = \log\left(C_2\sqrt{m}\right) + 1. \end{aligned} \tag{21}$$

Now, let $\gamma = \log\left(\frac{\|M\|}{\sigma} + 4\sqrt{m}\right)$. Notice that, if $x \geq \gamma$, then $\left(e^x - \frac{\|M\|}{\sigma}\right) \geq 4\sqrt{m}$. Therefore, applying Proposition 6 we obtain

$$\begin{aligned} E\left\{\log\|A\|\right\} &\leq \gamma + \int_{\gamma}^{\infty} P\left[\|A\| > e^x\right] dx \\ &\leq \gamma + \int_{\gamma}^{\infty} P\left[\|G\| > e^x - \frac{\|M\|}{\sigma}\right] dx \\ &\leq \gamma + \frac{C_1}{\sqrt{m}}\int_{\gamma}^{\infty} \exp\left(-\frac{1}{4}\left(e^x - \frac{\|M\|}{\sigma}\right)^2\right) dx. \end{aligned}$$

From here, if we make the change of variable $y = e^x - \frac{\|M\|}{\sigma}$, we obtain that

$$\begin{aligned} E\left\{\log\|A\|\right\} &\leq \gamma + \frac{C_1}{\sqrt{m}}\int_{4\sqrt{m}}^{\infty} \exp\left(-\frac{1}{2}y^2\right) dy \\ &\leq \gamma + \frac{C_1}{2m}\exp\left(-4m\right). \end{aligned}$$

And the corollary follows from here and (21). ■

Putting $M = 0$, $\sigma = 1$, the last Corollary provides a weak version of Edelman's Theorem of the form (5).

Acknowledgment. Authors want to thank an anonymous referee whose comments and suggestions have improved the paper.

References

[1] Cheung, D.; Cucker, F. (2001). A new Condition Number for Linear Programming. *Math. Programming,* **91,** 163-174.

[2] Cucker, F.; Wschebor, M. (2002). On the Expected Condition Number of Linear Programming Problems. *Numer. Mathem.* To appear.

[3] Cuesta-Albertos, J.; Wschebor, M. (2002). Some Remarks on the Condition Number of a Real Random Square Matrix. Submitted.

[4] Davidson, K.R.; Szarek, S.J. (2001) Local Operator Theory, Random Matrices and Banach Spaces, in *Handbook of the Geometry of Banach Spaces*, Vol. 1, Ch. 8, Eds. W.B. Johnson and J. Lindenstrauss, Elsevier, pp. 317-366.

[5] Demmel, J. (1997). *Applied Numerical Linear Algebra.* Ed. SIAM.

[6] Edelman, A. (1988). Eigenvalues and condition numbers of random matrices. *SIAM J. of Matrix Anal. and Appl.,* **9,** 543-556.

[7] Higham, N. (1996). *Accuracy and Stability of Numerical Algorithms.* Ed. SIAM.

[8] Ledoux, M. (2002) A Remark on Hypercontractivity and Tail Inequalities for the Largest Eigenvalues of Random Matrices, *preprint.*

[9] Sankar, A.; Spielman, D.A.; Teng, S.H. (2002). Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices. Preprint.

[10] Szarek, S.J. (1990). Spaces with large distance to l_∞^n and random matrices. *Amer. J. Math.* **112**, 899-942.

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

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

[13] Wilkinson, J.H. (1963). *Rounding Errors in Algebraic Processes.* Ed. Prentice-Hall.