

CONDITION NUMBERS OF RANDOM TRIANGULAR MATRICES*

D. VISWANATH[†] AND L. N. TREFETHEN[‡]

Abstract. Let L_n be a lower triangular matrix of dimension n each of whose nonzero entries is an independent $N(0, 1)$ variable, i.e., a random normal variable of mean 0 and variance 1. It is shown that κ_n , the 2-norm condition number of L_n , satisfies

$$\sqrt[n]{\kappa_n} \rightarrow 2 \text{ almost surely}$$

as $n \rightarrow \infty$. This exponential growth of κ_n with n is in striking contrast to the linear growth of the condition numbers of random *dense* matrices with n that is already known. This phenomenon is not due to small entries on the diagonal (i.e., small eigenvalues) of L_n . Indeed, it is shown that a lower triangular matrix of dimension n whose diagonal entries are fixed at 1 with the subdiagonal entries taken as independent $N(0, 1)$ variables is also exponentially ill conditioned with the 2-norm condition number κ_n of such a matrix satisfying

$$\sqrt[n]{\kappa_n} \rightarrow 1.305683410\dots \text{ almost surely}$$

as $n \rightarrow \infty$. A similar pair of results about complex random triangular matrices is established. The results for real triangular matrices are generalized to triangular matrices with entries from any symmetric, strictly stable distribution.

Key words. random triangular matrices, matrix condition numbers, exponentially nonnormal matrices, strong limit theorems

AMS subject classifications. 15A52, 15A12, 65F35, 60F15

PII. S0895479896312869

1. Introduction. Random dense matrices are well conditioned. Edelman has shown that if each of the n^2 entries of a matrix of dimension n is an independent $N(0, 1)$ variable, the probability density function (PDF) of κ_n/n , where κ_n is the 2-norm condition number of such a matrix, converges pointwise to the function

$$\frac{2x + 4}{x^3} \exp(-2x^{-1} - 2x^{-2})$$

as $n \rightarrow \infty$ [5]. Since the distribution of κ_n/n is independent of n in the limit $n \rightarrow \infty$, we can say that the condition numbers of random dense matrices grow only linearly with n . Using this PDF, it can be shown, for example, that $E(\log(\kappa_n)) = \log(n) + 1.537\dots + o(1)$ [5].

In striking contrast, the condition number of a random lower triangular matrix L_n , a matrix of dimension n all of whose diagonal and subdiagonal entries are independent $N(0, 1)$ variables, grows exponentially with n . If κ_n is the 2-norm condition number of L_n (defined as $\|L_n\|_2 \|L_n^{-1}\|_2$), we show that

$$\sqrt[n]{\kappa_n} \rightarrow 2 \text{ almost surely}$$

as $n \rightarrow \infty$ (Theorem 4.3). Figure 1.1a illustrates this result.

*Received by the editors December 2, 1996; accepted for publication by A. Edelman April 21, 1997. This work was supported by NSF grant DMS-9500975CS and DOE grant DE-FG02-94ER25199.

<http://www.siam.org/journals/simax/19-2/31286.html>

[†]Department of Computer Science, Cornell University, Ithaca, NY 14853 (divakar@cs.cornell.edu).

[‡]Computing Laboratory, Oxford University, Oxford, OX1 3QD, UK (lnt@comlab.ox.ox.ac.uk).

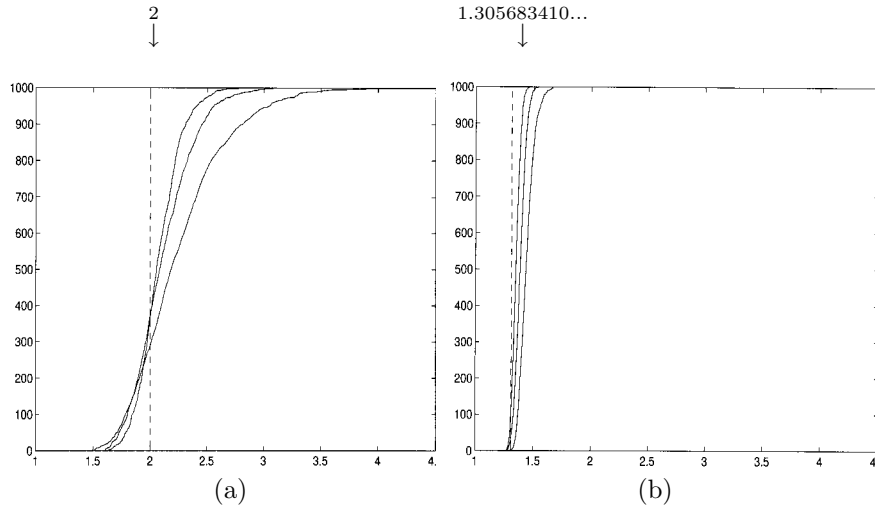


FIG. 1.1. Empirical cumulative density functions of $\sqrt[3]{\kappa_n}$, for triangular and unit triangular matrices, respectively, with $n = 25, 50, 100$ obtained from 1000 random matrices for each n . The random entries are $N(0, 1)$ variables. The higher values of n correspond to the steeper curves. In the limit $n \rightarrow \infty$, the cumulative density functions converge to Heaviside step functions with jumps at the dashed lines.

The matrices that arise in the experiments reported in Figure 1.1 are so ill conditioned that the standard, normwise stable method of finding the condition number using the SVD [10] fails owing to rounding errors. The method used to generate the figures finds the inverse of the triangular matrix explicitly using the standard algorithm for triangular inversion, and then computes the norms of the matrix and its inverse independently. This works because the computation of each column of the inverse by the standard triangular inversion algorithm is componentwise backward stable [12].

The exponential growth of κ_n with n is not due to small entries on the diagonal since the probability of a diagonal entry being exponentially small is also exponentially small. For a further demonstration that the diagonal entries do not cause the exponential growth in κ_n , we consider condition numbers of *unit* triangular matrices, i.e., triangular matrices with ones on the diagonal. If κ_n is the condition number of a unit lower triangular matrix of dimension n with subdiagonal entries taken as independent $N(0, 1)$ variables, then

$$\sqrt[3]{\kappa_n} \rightarrow 1.305683410\dots \text{ almost surely}$$

as $n \rightarrow \infty$ (Theorem 5.3). Obviously, in this case the ill conditioning has nothing to do with the diagonal entries (i.e., the eigenvalues) since they are all equal to 1. The relationship of the exponential ill conditioning of random unit triangular matrices to the stability of Gaussian elimination with partial pivoting is discussed in section 7.

We will use L_n to refer to triangular matrices of various kinds — real or complex, with or without a unit diagonal. But L_n always denotes a lower triangular matrix of dimension n . If the entries of L_n are random variables, they are assumed to be independent. Thus, if we merely say that L_n has entries from a certain distribution, those entries are not only identically distributed but also independent. Of course, only the nonzero entries of L_n are chosen according to that distribution. The condition number always refers to the 2-norm condition number. However, all our results concerning the limits $\lim_{n \rightarrow \infty} \sqrt[3]{\kappa_n}$ apply to all the L_p norms, $1 \leq p \leq \infty$, since

$n^{1/n} \rightarrow 1$ as $n \rightarrow \infty$ and the L_p norms differ by at most a factor of n . The 2-norm condition number of L_n , defined as $\|L_n\|_2 \|L_n^{-1}\|_2$, is denoted by κ_n . The context will make clear the distribution of the entries of L_n .

The analyses and discussions in this paper are phrased for lower, not upper, triangular matrices. However, all the theorems are true for upper triangular matrices as well, as is obvious from the fact that a matrix and its transpose have the same condition number.

We obtain similar results for triangular matrices with entries chosen from the complex normal distribution $\tilde{N}(0, \sigma^2)$. By $\tilde{N}(0, \sigma^2)$ we denote the complex normal distribution of mean 0 and variance σ^2 obtained by taking the real and imaginary parts as independent $N(0, \sigma^2/2)$ variables. Let L_n denote a triangular matrix with $\tilde{N}(0, \sigma^2)$ entries. Then

$$\sqrt[n]{\kappa_n} \rightarrow e^{1/2} \text{ almost surely}$$

as $n \rightarrow \infty$ (Theorem 7.3). Since $e^{1/2} < 2$, triangular matrices with complex normal entries tend to have smaller condition numbers than triangular matrices with real normally distributed entries.

Similarly, let L_n denote a unit lower triangular matrix with $\tilde{N}(0, 1)$ subdiagonal entries. Then

$$\sqrt[n]{\kappa_n} \rightarrow 1.347395784\dots \text{ almost surely}$$

as $n \rightarrow \infty$ (Theorem 7.4). Thus, unit triangular matrices with complex normal entries tend to have slightly bigger condition numbers than unit triangular matrices with real normal entries.

Our results are similar in spirit to results obtained by Silverstein for random dense matrices [16]. Consider a matrix of dimension $n \times (yn)$, where $y \in [0, 1]$, each of whose n^2y entries is an independent $N(0, 1)$ variable. Denote its largest and smallest singular values by σ_{\max} and σ_{\min} , respectively. It is shown in [16] that

$$\frac{\sigma_{\max}}{\sqrt{n}} \rightarrow 1 + \sqrt{y}, \quad \frac{\sigma_{\min}}{\sqrt{n}} \rightarrow 1 - \sqrt{y} \text{ almost surely}$$

as $n \rightarrow \infty$. The complex analogues of these results can be found in [4]. The technique used in [16] is a beautiful combination of what is now known as the Golub–Kahan bidiagonalization step in computing the SVD with the Gerschgorin circle theorem and the Marčenko–Pastur semicircle law. The techniques used in this paper are more direct.

The exponential growth of $\kappa_n = \|L_n\|_2 \|L_n^{-1}\|_2$ is due to the second factor. We outline the approach for determining the rate of exponential growth of κ_n by assuming L_n triangular with $N(0, 1)$ entries. In section 2, we derive the joint probability density function (JPDF) for the entries in any column of L_n^{-1} (Proposition 2.1). If T_k is the 2-norm of column $n - k + 1$ of L_n^{-1} , i.e., the column with k nonzero entries, both positive and negative moments of T_k are explicitly derived in section 3 (Lemma 3.2). These moments allow us to deduce that $\sqrt[n]{\kappa_n}$ converges to 2 almost surely (Theorem 4.3). A similar approach is used to determine the limit of $\sqrt[n]{\kappa_n}$ for L_n unit triangular with $N(0, \sigma^2)$ entries, triangular with $\tilde{N}(0, \sigma^2)$ entries, and unit triangular with $\tilde{N}(0, \sigma^2)$ entries (Theorems 5.3, 7.3, and 7.4, respectively).

The same approach is used more generally to determine the limit of $\sqrt[n]{\kappa_n}$ as $n \rightarrow \infty$ for L_n with entries drawn from any symmetric, strictly stable distribution (Theorems 8.4 and 8.6). These theorems are specialized to the Cauchy distribution, which is symmetric and strictly stable, in Theorems 8.5 and 8.7.

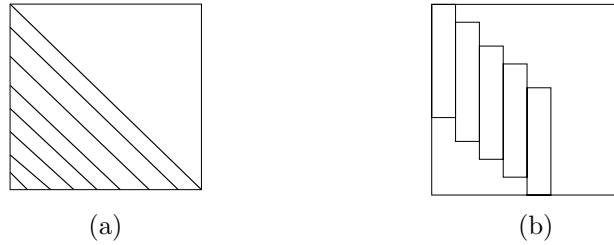


FIG. 2.1. Entries of L_n^{-1} on the same solid line in (a) have the same PDF. Sets of entries of L_n^{-1} in the boxes in (b) have the same JPDP.

2. Inverse of a random triangular matrix. Consider the matrix

$$L_n = \begin{pmatrix} \alpha_{11} & & & \\ -\alpha_{21} & \alpha_{22} & & \\ \vdots & \vdots & \ddots & \\ -\alpha_{n1} & -\alpha_{n2} & \dots & \alpha_{nn} \end{pmatrix},$$

where each α_{ij} is an independent $N(0, 1)$ variable. Then L_n^{-1} is also lower triangular. Denote the first k entries in the first column of L_n^{-1} by t_1, \dots, t_k . The t_i satisfy the following relations:

$$\begin{aligned} t_1 &= 1/\alpha_{11}, \\ t_2 &= (\alpha_{21}t_1)/\alpha_{22}, \\ t_3 &= (\alpha_{31}t_1 + \alpha_{32}t_2)/\alpha_{33}, \\ &\vdots \\ (2.1) \quad t_k &= (\alpha_{k1}t_1 + \dots + \alpha_{k,k-1}t_{k-1})/\alpha_{kk}. \end{aligned}$$

This system of equations can be interpreted as a system of random recurrence relations. The first entry t_1 is the reciprocal of an $N(0, 1)$ variable. The k th entry t_k is obtained by summing the previous entries t_1, \dots, t_{k-1} with independent $N(0, 1)$ variables as coefficients, and dividing that sum by an independent $N(0, 1)$ variable.

Next, consider an arbitrary column of L_n^{-1} and denote the first k entries of that column from the diagonal downwards by t_1, \dots, t_k . The entries t_i satisfy random recurrence relations similar in form to (2.1), but the α_{ij} are a different block of entries in L_n for different columns. For example, any diagonal entry of L_n^{-1} is the reciprocal of an $N(0, 1)$ variable; in particular, the k th diagonal entry is $1/\alpha_{kk}$.

These observations about triangular inversion can be represented pictorially.

Every entry of L_n^{-1} at a fixed distance from the diagonal has the same PDF. We may say that the matrix L_n^{-1} , like L_n , is “statistically Toeplitz.” See Figure 2.1a. Moreover, if we consider the first k entries of a column of L_n^{-1} from the diagonal downwards, those k entries will have the same JPDP irrespective of the column. See Figure 2.1b. The different columns of L_n^{-1} , however, are by no means independent.

The description of triangular inversion above and later arguments are stated in terms of the columns of L_n^{-1} . However, rows and columns are indistinguishable in this problem; we could equally well have framed the analysis in terms of rows.

Denote the JPDP of t_i , $1 \leq i \leq k$, by $f_k = f_k(t_1, \dots, t_k)$. In the next proposition, a recursive formula for f_k is derived. For simplicity, we introduce the further notation

$T_k = \sqrt{t_1^2 + \dots + t_k^2}$. Throughout this section, L_n is the random triangular matrix of dimension n with $N(0, 1)$ entries.

PROPOSITION 2.1. *The JPDPF $f_k = f_k(t_1, \dots, t_k)$ satisfy the following recurrence:*

$$(2.2) \quad f_1 = \frac{\exp(-1/2t_1^2)}{\sqrt{2\pi t_1^2}},$$

$$(2.3) \quad f_k = \frac{1}{\pi} \frac{T_{k-1}}{T_k^2} f_{k-1} \text{ for } k > 1.$$

Proof. The t_k are defined by the random recurrence in (2.1).

The expression for f_1 is easy to get. If x is an $N(0, 1)$ variable, its PDF is

$$\frac{1}{\sqrt{2\pi}} \exp(-x^2/2).$$

The change of variable $x = 1/t_1$ gives (2.2).

To obtain the recursive expression (2.3) for f_k , consider the variable τ_k obtained by summing the variables t_1, \dots, t_{k-1} as $\sum_{i=1}^{k-1} \alpha_{ki} t_i$, where α_{ki} are independent $N(0, 1)$ variables. For fixed values of $t_i, 1 \leq i \leq k-1$, the variable τ_k , being a sum of random normal variables, is itself a random normal variable of mean 0 and variance T_{k-1}^2 . Therefore, the JPDPF of τ_k and t_1, \dots, t_{k-1} is given by

$$\frac{1}{\sqrt{2\pi}} \frac{\exp(-\tau_k^2/2T_{k-1}^2)}{T_{k-1}} f_{k-1}.$$

By (2.1), the variable t_k can be obtained as τ_k/α , where α is an independent $N(0, 1)$ variable. The JPDPF of α, τ_k , and t_1, \dots, t_{k-1} is given by

$$\frac{1}{\sqrt{2\pi}} \exp(-\alpha^2/2) \frac{1}{\sqrt{2\pi}} \frac{\exp(-\tau_k^2/2T_{k-1}^2)}{T_{k-1}} f_{k-1}.$$

Changing the variable τ_k to $t_k = \tau_k/\alpha$ and integrating out α , we obtain

$$f_k = \frac{1}{\pi} \frac{T_{k-1}}{T_{k-1}^2 + t_k^2} f_{k-1} = \frac{1}{\pi} \frac{T_{k-1}}{T_k^2} f_{k-1},$$

i.e., f_k is given by (2.3). \square

Note that the form of the recurrence for f_k in Proposition 2.1 mirrors the random recurrence (2.1) for obtaining t_k from the previous entries t_1, \dots, t_{k-1} . In the following corollary, an explicit expression for f_k in terms of the t_i is stated.

COROLLARY 2.2. *For $k > 1$, the JPDPF $f_k = f_k(t_1, \dots, t_k)$ is given by*

$$f_k = \frac{1}{\pi^{k-1} \sqrt{2\pi}} \frac{1}{(t_1^2 + \dots + t_k^2)} \frac{1}{\sqrt{t_1^2 + \dots + t_{k-1}^2}} \dots \frac{1}{\sqrt{t_1^2 + t_2^2}} \frac{\exp(-1/2t_1^2)}{|t_1|}.$$

3. Moments of T_k . In this section and the next, L_n continues to represent a triangular matrix of dimension n with $N(0, 1)$ entries. As we remarked earlier, the exponential growth of $\kappa_n = \|L_n\|_2 \|L_n^{-1}\|_2$ is due to the second factor $\|L_n^{-1}\|_2$. Since the 2-norm of column $i+1$ of L_n^{-1} has the same distribution as T_{n-i} , we derive formulas for various moments of T_k with the intention of understanding the exponential growth of $\|L_n^{-1}\|_2$ with n .

In the lemma below, we consider the expected value $E(T_k^\xi)$ for both positive and negative values of ξ . By our notation, $T_1 = |t_1|$. The notation $d\Omega_k = dt_k \dots dt_1$ is used to reduce clutter in the proof. As usual, R^k denotes the real Euclidean space of dimension k .

The next lemma is stated as a recurrence to reflect the structure of its proof. Lemma 3.2 contains the same information in a simpler form.

LEMMA 3.1. *For any real $\xi < 1$, $E(T_k^\xi)$ is given by the following recurrence:*

$$(3.1) \quad E(T_1^\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{\exp(-1/2x^2)}{|x|^{2-\xi}} dx,$$

$$(3.2) \quad E(T_k^\xi) = \frac{E(T_{k-1}^\xi)}{\pi} \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{1-\xi/2}} \text{ for } k > 1.$$

For $\xi \geq 1$ and $k \geq 1$, $E(T_k^\xi)$ is infinite.

Proof. To obtain (3.1), use $T_1 = |t_1|$ and the PDF of t_1 given by (2.2). It is easily seen that the integral is convergent if and only if $\xi < 1$.

Next, assume $k > 1$. By definition,

$$E(T_k^\xi) = \int_{R^k} T_k^\xi f_k d\Omega_k.$$

Using the recursive equation (2.3) for f_k , and writing T_k in terms of t_k and T_{k-1} , we get

$$(3.3) \quad \begin{aligned} E(T_k^\xi) &= \frac{1}{\pi} \int_{R^k} \frac{T_{k-1}}{T_k^{2-\xi}} f_{k-1} d\Omega_k \\ &= \frac{1}{\pi} \int_{R^{k-1}} \int_{-\infty}^{\infty} \frac{dt_k}{(t_k^2 + T_{k-1}^2)^{1-\xi/2}} T_{k-1} f_{k-1} d\Omega_{k-1}. \end{aligned}$$

By the substitution $t_k = xT_{k-1}$, the inner integral with respect to dt_k can be reduced to

$$T_{k-1}^{\xi-1} \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{1-\xi/2}}.$$

Inserting this in the multiple integral (3.3) gives the recursive equation (3.2) for $E(T_k^\xi)$. It is easily seen that the integral in (3.2) is convergent if and only if $\xi < 1$. \square

Define γ_ξ by

$$(3.4) \quad \gamma_\xi = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{dx}{(1+x^2)^{1-\xi/2}}.$$

Beginning with the substitution $x = \tan \theta$ in (3.4), it can be shown that $\gamma_\xi = \pi^{-1}B((1-\xi)/2, 1/2)$, where B is the beta function. The relevant expression for the beta function $B(x, y)$ is (6.2.1) in [1]. Also, if x is chosen from the standard Cauchy distribution, then $\gamma_\xi = E((1+x^2)^{\xi/2})$. We do not need γ_ξ in terms of the beta function, however; the integral expression (3.4) is more suitable for our purposes. Lemma 3.1 can be restated in a more convenient form using γ_ξ as follows.

LEMMA 3.2. *For $\xi < 1$, $E(T_k^\xi) = C_\xi \gamma_\xi^k$ for a finite positive constant C_ξ . Also, $\gamma_0 = 1$, $\gamma_\xi < 1$ for $\xi < 0$, and $\gamma_\xi > 1$ for $\xi > 0$.*

Proof. The expression for $E(T_k^\xi)$ is a restatement of Lemma 3.1. By elementary integration, $\gamma_0 = 1$, and by the form of the integral in (3.4), $\gamma_\xi < 1$ for $\xi < 0$ and $\gamma_\xi > 1$ for $\xi > 0$. \square

Lemma 3.2 implies that the positive moments of T_k grow exponentially with k while the negative moments decrease exponentially with k .

Obtaining bounds for $P(T_k > M^k)$ and $P(T_k < m^k)$ is now a simple matter.

LEMMA 3.3. *For $k \geq 1$, $\xi > 0$, and $m > 0$,*

$$P(T_k < m^k) < C_{-\xi}(m/\gamma_{-\xi}^{-1/\xi})^{\xi k}.$$

Proof. Since $\xi > 0$, $P(T_k < m^k) = P(T_k^{-\xi} > m^{-\xi k})$. Use Lemma 3.2 with $\xi = -\xi$ to obtain an expression for $E(T_k^{-\xi})$ and apply Markov's inequality [2]. \square

LEMMA 3.4. *For $k \geq 1$, $0 < \xi < 1$, and $M > 0$,*

$$P(T_k > M^k) < C_\xi(\gamma_\xi^{1/\xi}/M)^{\xi k}.$$

Proof. As in Lemma 3.3, $\xi > 0$ implies that $P(T_k > M^k) = P(T_k^\xi > M^{\xi k})$. Again, the proof can be completed by obtaining an expression for $E(T_k^\xi)$ using Lemma 3.2 followed by an application of Markov's inequality. \square

4. Exponential growth of κ_n . We are now prepared to derive the first main result of the paper, namely, $\sqrt[n]{\kappa_n} \rightarrow 2$ almost surely as $n \rightarrow \infty$ for triangular matrices L_n with $N(0, 1)$ entries. In what follows, *a.s.* means almost surely as $n \rightarrow \infty$. The definition of almost sure convergence for a sequence of random variables can be found in most textbooks on probability; for example, see [2]. Roughly, it means that the convergence holds for a set of sequences of measure 1.

LEMMA 4.1. $\|L_n\|_2^{1/n} \rightarrow 1$ almost surely as $n \rightarrow \infty$.

Proof. The proof is easy. We provide only an outline. The Frobenius norm of L_n , $\|L_n\|_F^2$, is a sum of $n(n+1)/2$ independent χ^2 variables of mean 1. By an argument exactly analogous to the proof of the strong law of large numbers with finite fourth moment assumption [2, p. 80],

$$\frac{\|L_n\|_F^2}{n(n+1)/2} \rightarrow 1 \text{ a.s.}$$

The proof can be completed using the inequalities $n^{-1/2}\|L_n\|_F \leq \|L_n\|_2 \leq \|L_n\|_F$. Note that the suggested proof relies on the existence of the fourth moment of the χ^2 variables. \square

The proof of Lemma 4.2 uses the first Borel–Cantelli lemma in a way that is typical of several proofs in probability. We use $\liminf_{n \rightarrow \infty} x_n$ and $\limsup_{n \rightarrow \infty} x_n$ for $\lim_{n \rightarrow \infty} \inf_{k \geq n} x_k$ and $\lim_{n \rightarrow \infty} \sup_{k \geq n} x_k$ in the following lemma and later.

LEMMA 4.2. *As $n \rightarrow \infty$, for any $0 < \xi < 1$,*

$$\gamma_{-\xi}^{-1/\xi} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{\kappa_n} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{\kappa_n} \leq \gamma_\xi^{1/\xi} \text{ almost surely.}$$

Proof. By Lemma 4.1, it suffices to show that

$$\gamma_{-\xi}^{-1/\xi} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{\|L_n^{-1}\|_2} \leq \limsup_{n \rightarrow \infty} \sqrt[n]{\|L_n^{-1}\|_2} \leq \gamma_\xi^{1/\xi} \text{ a.s.}$$

We consider the lower bound first. The 2-norm of the first column of L_n^{-1} , which has the same distribution as T_n , is less than or equal to $\|L_n^{-1}\|_2$. Therefore, for $0 < \epsilon < 1$,

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} < \gamma_{-\xi}^{-1/\xi} - \epsilon\right) \leq P(T_n < (\gamma_{-\xi}^{-1/\xi} - \epsilon)^n).$$

Using Lemma 3.3 with $k = n$ and $m = \gamma_{-\xi}^{-1/\xi} - \epsilon$, we get

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} < \gamma_{-\xi}^{-1/\xi} - \epsilon\right) < C_{-\xi} \left(\frac{\gamma_{-\xi}^{-1/\xi} - \epsilon}{\gamma_{-\xi}^{-1/\xi}}\right)^{\xi n} = C_{-\xi} p_\epsilon^{\xi n},$$

where $p_\epsilon = \gamma_{-\xi}^{-1/\xi}(\gamma_{-\xi}^{-1/\xi} - \epsilon) < 1$. Since $|p_\epsilon| < 1$, $\sum_{n=1}^\infty p_\epsilon^{\xi n}$ is finite. The first Borel-Cantelli lemma [2] can be applied to obtain

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} < \gamma_{-\xi}^{-1/\xi} - \epsilon \text{ infinitely often as } n \rightarrow \infty\right) = 0.$$

Taking the union of the sets in the above equation over all rational ϵ in $(0, 1)$ and considering the complement of that union, we obtain

$$P\left(\liminf_{n \rightarrow \infty} \sqrt[n]{\|L_n^{-1}\|_2} \geq \gamma_{-\xi}^{-1/\xi} \text{ as } n \rightarrow \infty\right) = 1.$$

In other words, $\gamma_{-\xi}^{-1/\xi} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{\|L_n^{-1}\|_2}$ a.s.

The upper bound can be established similarly. At least one of the columns of L_n^{-1} must have 2-norm greater than or equal to $n^{-1/2}\|L_n^{-1}\|_2$. Since the 2-norm of column $k + 1$ has the same distribution as T_{n-k} ,

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} > \gamma_\xi^{1/\xi} + \epsilon\right) \leq \sum_{k=1}^n P(T_k > n^{-1/2}(\gamma_\xi^{1/\xi} + \epsilon)^n).$$

Bounding each term in the summation using Lemma 3.4 gives

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} > \gamma_\xi^{1/\xi} + \epsilon\right) < C_\xi n^{\xi/2} \sum_{k=1}^n \left(\frac{\gamma_\xi^k}{(\gamma_\xi^{1/\xi} + \epsilon)^{\xi n}}\right).$$

Since $\gamma_\xi > 1$ by Lemma 3.2, the largest term in the summand occurs when $k = n$. Therefore,

$$P\left(\sqrt[n]{\|L_n^{-1}\|_2} > \gamma_\xi^{1/\xi} + \epsilon\right) < C_\xi n^{1+\xi/2} \left(\frac{\gamma_\xi^{1/\xi}}{\gamma_\xi^{1/\xi} + \epsilon}\right)^{\xi n}.$$

From this point, the proof can be completed in the same manner as the proof of the lower bound. \square

THEOREM 4.3. *For random triangular matrices with $N(0, 1)$ entries, as $n \rightarrow \infty$,*

$$\sqrt[n]{\kappa_n} \rightarrow 2 \text{ almost surely.}$$

Proof. By an inequality sometimes called Lyapunov’s [13, p. 144], [2],

$$\gamma_\beta^{1/\beta} < \gamma_\alpha^{1/\alpha}$$

for any real $\beta < \alpha$. Thus, the bounding intervals $[\gamma_{-\xi}^{-1/\xi}, \gamma_\xi^{1/\xi}]$ in Lemma 4.2 shrink as ξ decreases from 1 to 0. A classical theorem [13, p. 139] says that these intervals actually shrink to the following point:

$$\begin{aligned} \lim_{\xi \rightarrow 0} \gamma_\xi^{1/\xi} &= \lim_{\xi \rightarrow 0} \left(\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^{1-\xi/2}} dx \right)^{1/\xi} \\ &= \exp \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\log(1+x^2)}{1+x^2} dx \right). \end{aligned}$$

The exact value of the limit can be evaluated to 2 using the substitution $x = \tan \theta$ followed by complex integration [3, p. 121]. Thus, $\sqrt[n]{\kappa_n} \rightarrow 2$ a.s. \square

Theorem 4.3 holds in exactly the same form if the nonzero entries of L_n are independent $N(0, \sigma^2)$ variables rather than $N(0, 1)$ variables, since the condition number is invariant under scaling.

Our approach to Theorem 4.3 began by showing that $E(T_k^\xi) = C_\xi \gamma_\xi^k$ for both positive and negative ξ . Once these expressions for the moments of T_k were obtained, our arguments did not depend on how the recurrence was computed. The following note summarizes the asymptotic information about a recurrence that can be obtained from a knowledge of its moments.

Note. Let t_1, t_2, \dots be a sequence of random variables. If $E(|t_n|^\xi)$ grows exponentially with n at the rate ν_ξ^n for $\xi > 0$, then $\limsup_{n \rightarrow \infty} \sqrt[n]{|t_n|} \leq \nu_\xi^{1/\xi}$ almost surely. Similarly, if $E(|t_n|^\xi)$ decreases exponentially with n at the rate ν_ξ^n as $n \rightarrow \infty$ for $\xi < 0$, then $\nu_\xi^{1/\xi} \leq \liminf_{n \rightarrow \infty} \sqrt[n]{|t_n|}$ almost surely. Thus, knowledge of any positive moment of t_n yields an upper bound on $\sqrt[n]{|t_n|}$ as $n \rightarrow \infty$, while knowledge of any negative moment yields a lower bound.

5. Unit triangular matrices. So far, we have considered triangular matrices whose nonzero entries are independent, real $N(0, 1)$ variables. In this section and in section 7, we establish the exponential growth of the condition number for other kinds of random triangular matrices with normally distributed entries. The key steps in the sequence of lemmas leading to the analogues of Theorem 4.3 are stated but not proved. The same techniques used in sections 2, 3, and 4 work here, too.

Let L_n be a unit lower triangular matrix of dimension n with $N(0, \sigma^2)$ subdiagonal entries. Let s_1, \dots, s_k be the first k entries from the diagonal downwards of any column of L_n^{-1} . The entries s_i satisfy the recurrence

$$\begin{aligned} s_1 &= 1, \\ s_2 &= \alpha_{21}s_1, \\ s_3 &= \alpha_{31}s_1 + \alpha_{32}s_2, \\ &\vdots \\ (5.1) \quad s_k &= \alpha_{k1}s_1 + \dots + \alpha_{k,k-1}s_{k-1}, \end{aligned}$$

where α_{ij} , $i > j$, are $N(0, \sigma^2)$ variables. The notation $S_k = \sqrt{s_1^2 + s_2^2 + \dots + s_k^2}$ is used below.

PROPOSITION 5.1. *The JPDP of s_1, \dots, s_k , $g_k(s_1, \dots, s_k)$, is given by the recurrence*

$$\begin{aligned} g_2 &= \frac{1}{\sqrt{2\pi\sigma}} \exp(-s_2^2/2\sigma^2), \\ g_k &= \frac{1}{\sqrt{2\pi\sigma}} \frac{\exp(-s_k^2/2\sigma^2 S_{k-1}^2)}{S_{k-1}} g_{k-1} \quad \text{for } k > 2, \end{aligned}$$

and the fact that $s_1 = 1$ identically.

LEMMA 5.2. For any real ξ , $E(S_k^\xi) = \lambda_\xi^{k-1}$, where

$$\lambda_\xi = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} (1+x^2)^{\xi/2} \exp(-x^2/2\sigma^2) dx.$$

The note at the end of section 4 provides part of the link from Lemma 5.2 to the following theorem about κ_n .

THEOREM 5.3. For random unit triangular matrices with $N(0, \sigma^2)$ entries, as $n \rightarrow \infty$,

$$\sqrt[n]{\kappa_n} \rightarrow \exp\left(\frac{1}{2\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} \log(1+x^2)e^{-x^2/2\sigma^2} dx\right) \text{ almost surely.}$$

If this limit is denoted by $C(\sigma)$, then

$$\begin{aligned} C(\sigma) - 1 &\sim \sigma^2/2 \text{ as } \sigma \rightarrow 0, \\ C(\sigma) &\sim K\sigma \text{ as } \sigma \rightarrow \infty, \end{aligned}$$

where $K = \sqrt{\exp(-\gamma)/2} = 0.5298\dots$, with γ being the Euler constant.

Proof. The constant K is given by

$$K = \exp\left(\sqrt{\frac{2}{\pi}} \int_0^\infty \log x \exp(-x^2/2) dx\right).$$

To evaluate K , we used integral 4.333 of [8]. □

In contrast to the situation in Theorem 4.3, the constant that $\sqrt[n]{\kappa_n}$ converges to in Theorem 5.3 depends on σ . This is because changing σ scales only the subdiagonal entries of the unit triangular matrix L_n while leaving the diagonal entries fixed at one. For $\sigma = 1$, the case discussed in the Introduction, numerical integration shows the constant to be 1.305683410\dots

6. A comment on the stability of Gaussian elimination. The conditioning of random unit triangular matrices has a connection with the phenomenon of numerical stability of Gaussian elimination. We pause briefly to explain this connection.

For decades, the standard algorithm for solving general systems of linear equations $Ax = b$ has been Gaussian elimination (with “partial” or row pivoting). This algorithm generates an “LU factorization” $PA = LU$, where P is a permutation matrix, L is unit lower triangular with subdiagonal entries less than or equal to one in absolute value, and U is upper triangular.

In the mid-1940s it was predicted by Hotelling [14] and Goldstine and von Neumann [9] that rounding errors must accumulate exponentially in elimination algorithms of this kind, causing instability for all but small dimensions. In the 1950s, Wilkinson developed a beautiful theory based on backward error analysis that, while it explained a great deal about Gaussian elimination, confirmed that for certain matrices, exponential instability does indeed occur [19]. He showed that amplification of rounding errors by factors on the order of $\|L^{-1}\|$ may take place, and that for certain matrices, $\|L^{-1}\|$ is of order 2^n . Thus, for certain matrices, rounding errors are amplified by $O(2^n)$, causing a catastrophic loss of n bits of precision.

Despite these facts, the experience of 50 years of computing has established that from a practical point of view, Hotelling and von Neumann were wrong: Gaussian elimination is overwhelmingly stable. In fact, it is not clear that a single matrix

problem has ever led to an instability in this algorithm, except for the ones produced by numerical analysts with that end in mind, although Foster [7] and Wright [20] have devised problems leading to instability that plausibly “might have arisen” in applications. The reason appears to be statistical: the matrices A for which $\|L^{-1}\|$ is large occupy an exponentially small proportion of the space of all matrices, so small that such matrices “never” arise in practice. Experimental evidence of this phenomenon is presented in [18].

This raises the question, why are matrices A for which $\|L^{-1}\|$ is large so rare? It is here that the behavior of random unit triangular matrices is relevant. A natural hypothesis would be that the matrices L generated by Gaussian elimination are, to a reasonable approximation, random unit triangular matrices with off-diagonal entries of a size dependent on the dimension n . If such matrices could be shown to be almost always well conditioned, then the stability of Gaussian elimination would be explained.

We have just shown, however, that unit triangular matrices are exponentially ill conditioned. Thus, this attempted explanation of the stability of Gaussian elimination fails, and indeed, the same argument suggests that Gaussian elimination should be unstable in practice as well as in the worst case. The resolution of this apparent paradox is that *the matrices L produced by Gaussian elimination are far from random*. The signs of the entries of these matrices are correlated in special ways that have the effect of keeping $\|L^{-1}\|$ almost always very small. For example, it is reported in [18] that a certain random matrix A with $n = 256$ led to $\|L^{-1}\| = 33.2$, whereas if \tilde{L} was taken to be the same matrix but with the signs of its subdiagonal entries randomized, the result became $\|\tilde{L}^{-1}\| = 2.7 \times 10^8$. In fact, even unpivoted Gaussian elimination does not produce triangular matrices as severely ill conditioned as random triangular matrices [22].

From a comparison of Theorem 5.3 with half a century of the history of Gaussian elimination, then one may conclude that unit triangular factors of random dense matrices are very different from random unit triangular matrices. An explanation of this difference is offered in [17] along the following lines. If A is random, then its successive column spaces are randomly oriented in n -space in the sense that the first column of A is oriented in a random direction, the span of the first two columns is a random two-dimensional space, and so on. Since the span of the first k columns of L is the same as the span of the first k columns of PA , where P is the row permutation matrix produced by partial pivoting, the same holds approximately for the successive column spaces of L . That condition, in turn, implies that large values $\|L^{-1}\|$ can arise only exponentially rarely.

7. Complex matrices. We now consider matrices with complex entries. Let L_n be a lower triangular matrix with $\tilde{N}(0, 1)$ entries. The complex distribution $\tilde{N}(0, 1)$ was defined in the Introduction. Let t_1, \dots, t_k denote the first k entries from the diagonal downwards of any column of L_n^{-1} . The quantities t_k satisfy (2.1), but the α_{ij} are now independent $\tilde{N}(0, 1)$ variables. Let $r_k = |t_k|^2$, and denote $r_1 + \dots + r_k$ by R_k .

PROPOSITION 7.1. *The JPDP of r_1, \dots, r_k , $h_k(r_1, \dots, r_k)$, is given by the recurrence*

$$(7.1) \quad h_1 = \frac{\exp(-1/r_1)}{r_1^2},$$

$$(7.2) \quad h_k = \frac{R_{k-1}}{R_k^2} h_{k-1} \quad \text{for } k > 1,$$

for $r_i \geq 0, 1 \leq i \leq k$.

Proof. We sketch only the details that do not arise in the proof of Proposition 2.1. If x and y are independent $N(0, \sigma^2)$ variables, $x = \sqrt{r} \cos(\theta)$ and $y = \sqrt{r} \sin(\theta)$, then r and θ are independent. Moreover, the distribution of r is exponential with the PDF

$$(7.3) \quad (1/2\sigma^2) \exp(-r/2\sigma^2)$$

for $r > 0$.

Consider the sum $\tau_k = \alpha_{k1}t_1 + \dots + \alpha_{k,k-1}t_{k-1}$ with α_{ki} taken as independent $\tilde{N}(0, 1)$ variables. For fixed t_1, \dots, t_{k-1} , $\text{Re}(\tau_k)$ and $\text{Im}(\tau_k)$ are independent. To see their independence, we write out the equations for $\text{Re}(\tau_k)$ and $\text{Im}(\tau_k)$ as follows:

$$\begin{aligned} \text{Re}(\tau_k) &= \sum_{i=1}^{k-1} \text{Re}(\alpha_{ki})\text{Re}(t_i) - \text{Im}(\alpha_{ki})\text{Im}(t_i), \\ \text{Im}(\tau_k) &= \sum_{i=1}^{k-1} \text{Re}(\alpha_{ki})\text{Im}(t_i) + \text{Im}(\alpha_{ki})\text{Re}(t_i). \end{aligned}$$

The linear combinations of $\text{Re}(\alpha_{ki})$ and $\text{Im}(\alpha_{ki})$ in these two equations can be realized by taking inner products with the two vectors

$$\begin{aligned} v &= [\text{Re}(t_1), \dots, \text{Re}(t_{k-1}), -\text{Im}(t_1), \dots, -\text{Im}(t_{k-1})], \\ w &= [\text{Im}(t_1), \dots, \text{Im}(t_{k-1}), +\text{Re}(t_1), \dots, +\text{Re}(t_{k-1})]. \end{aligned}$$

The independence of $\text{Re}(\tau_k)$ and $\text{Im}(\tau_k)$ is a consequence of the orthogonality of v and w , i.e., $(v, w) = vw' = 0$, and the invariance of the JPDP of independent, identically distributed normal variables under orthogonal transformation [15].

Thus, for fixed t_1, \dots, t_{k-1} , the real and imaginary parts of τ_k are independent normal variables of mean 0 and variance $R_{k-1}/2$. By (7.3), the PDFs of $x = |\tau_k|^2$ and $y = |\alpha_{kk}|^2$ are given by

$$\frac{1}{R_{k-1}} \exp(-x/R_{k-1}), \quad \exp(-y)$$

for positive x, y . The expression (7.2) for h_k can now be obtained using $r_k = |\tau_k|^2/|\alpha_{kk}|^2$. \square

LEMMA 7.2. For any $\xi < 1, E(R_k^\xi) = C\mu_\xi^{k-1}$, where

$$C = \int_0^\infty \frac{\exp(-1/r_1)}{r_1^{2-\xi}} dr_1, \quad \mu_\xi = \int_0^\infty \frac{dx}{(1+x)^{2-\xi}}.$$

The constant μ_ξ in Lemma 7.2 can be reduced to $(1-\xi)^{-1}$ for $\xi < 1$. However, as with γ_ξ in section 3, the integral expression for μ_ξ is more suitable for our purposes. As before, the note at the end of section 4 is an essential part of the link from the previous lemma to the following theorem about κ_n .

THEOREM 7.3. For random triangular matrices with complex $\tilde{N}(0, 1)$ entries, as $n \rightarrow \infty$,

$$\sqrt[n]{\kappa_n} \rightarrow \exp\left(\frac{1}{2} \int_0^\infty \frac{\log(1+x)}{(1+x)^2} dx\right) = e^{1/2} \text{ almost surely.}$$

Theorem 7.3 holds unchanged if the entries are $\tilde{N}(0, \sigma^2)$ variables. As with Theorem 4.3, this is because the condition number is invariant under scaling.

Now, let L_n be a unit lower triangular matrix of dimension n with $\tilde{N}(0, \sigma^2)$ subdiagonal entries. We state only the final theorem about κ_n .

THEOREM 7.4. For random unit triangular matrices with complex $\tilde{N}(0, \sigma^2)$ entries, as $n \rightarrow \infty$,

$$\begin{aligned} \sqrt[n]{\kappa_n} &\rightarrow \exp\left(\frac{1}{4} \int_0^\infty \log(1 + \sigma^2 x/2) e^{-x/2} dx\right) \\ &= \exp(-\exp(\sigma^{-2}) \text{Ei}(-\sigma^{-2})/2) \text{ almost surely,} \end{aligned}$$

where Ei is the exponential integral. If this limit is denoted by $C(\sigma)$, then

$$\begin{aligned} C(\sigma) - 1 &\sim \sigma^2/2 \text{ as } \sigma \rightarrow 0, \\ C(\sigma) &\sim K\sigma \text{ as } \sigma \rightarrow \infty, \end{aligned}$$

where $K = \exp(-\gamma/2) = 0.7493\dots$, with γ being the Euler constant.

Proof. To obtain K , we evaluated

$$K = \exp\left(\frac{1}{4} \int_0^\infty \log(x/2) \exp(-x/2) dx\right)$$

using the Laplace transform of $\log(x)$ given by integral 4.331.1 of [8]. The explicit formula involving $\text{Ei}(\sigma^{-2})$ was obtained using integral 4.337.2 of [8]. \square

For $\sigma^2 = 1$, $\sqrt[n]{\kappa_n}$ converges to $1.347395784\dots$.

8. Matrices with entries from stable distributions. The techniques used to deduce Theorem 4.3 require that we first derive the joint density function of the t_k , defined by recurrence (2.1), as was done in Proposition 2.1. That proposition made use of the fact that when the α_{ki} are independent and normally distributed, and the t_i are fixed, the sum

$$\sum_{i=1}^{k-1} \alpha_{ki} t_i$$

is also normally distributed. This property of the normal distribution holds for any stable distribution.

A distribution is said to be stable if for X_i chosen independently from that distribution,

$$\sum_{i=1}^n X_i$$

has the same distribution as $c_n X + d_n$, where X has the same distribution as X_i and $c_n > 0$ and d_n are constants [6, p. 170]. If $d_n = 0$, the distribution is said to be

strictly stable. As usual, the distribution is symmetric if X has the same distribution as $-X$. A symmetric, strictly stable distribution has exponent a if $c_n = n^{1/a}$. A standard result of probability theory says that any stable distribution has an exponent $0 < a \leq 2$. The normal distribution is stable with exponent $a = 2$ [6].

The techniques used for triangular matrices with normal entries work more generally when the entries are drawn from a symmetric, strictly stable distribution. Let L_n be a unit lower triangular matrix with entries chosen from a symmetric, strictly stable distribution. Denote the PDF of that stable distribution by $\phi(x)$. The recurrence for the entries s_i of the inverse L_n^{-1} is again given by (5.1), but α_{ki} , $k > i$, are now independent random variables with the density function $\phi(x)$.

Our program for deriving the constants that $\sqrt[k]{\kappa_n}$ converge to as $n \rightarrow \infty$ began with Lemma 4.1 in all the previous examples. One of referees pointed out to us that a new proof is needed for that lemma in the present context since a stable distribution of index $a < 2$ does not have the a th or higher moments.

LEMMA 8.1. *For $a < 2$, $\|L_n\|_2^{1/n} \rightarrow 1$ almost surely as $n \rightarrow \infty$.*

Proof. Define $\|L_n\|_\alpha = (\sum_{i,j} |l_{ij}|^\alpha)^{1/\alpha}$ for some $0 < \alpha < a/4$. Then the inequality

$$n^{(1/2-2/\alpha)} \|L_n\|_\alpha \leq \|L_n\|_2 \leq n \|L_n\|_\alpha$$

and the existence of the fourth moment of $|l_{ij}|^\alpha$ make possible a proof analogous to what was outlined for Lemma 4.1. \square

The proposition, the lemma, and the theorem below are analogues of Proposition 5.1, Lemma 5.2, and Theorem 5.3, respectively. If the exponent of the stable distribution is a , denote $(|s_1|^a + \dots + |s_k|^a)^{1/a}$ by S_k .

PROPOSITION 8.2. *If $\phi(x)$ is the density function of a symmetric, strictly stable distribution with exponent a , the JPDF of s_1, \dots, s_k , $g_k(s_1, \dots, s_k)$, is given by the recurrence*

$$g_2 = \phi(s_2),$$

$$g_k = \frac{\phi(s_k/S_{k-1})}{S_{k-1}} g_{k-1} \text{ for } k > 2,$$

and the fact that $s_1 = 1$ identically.

Proof. The proof is very similar to the proof of Proposition 2.1. We note that if α_{ki} , $k > i$, are independent random variables with the PDF $\phi(x)$, and the s_i are fixed, then the sum

$$\alpha_{k1}s_1 + \dots + \alpha_{k,k-1}s_{k-1}$$

has the PDF $\phi(x/S_{k-1})/S_{k-1}$ [6, p. 171]. \square

LEMMA 8.3. *For any real ξ , $E(S_k^\xi) = \lambda_\xi^{k-1}$, where*

$$\lambda_\xi = \int_{-\infty}^{+\infty} (1 + |x|^a)^{\xi/a} \phi(x) dx,$$

with $\lambda_\xi = \infty$ for $\xi \geq a$.

THEOREM 8.4. *For random unit triangular matrices with entries from a symmetric, strictly stable distribution with density function $\phi(x)$ and exponent a , as $n \rightarrow \infty$,*

$$\sqrt[k]{\kappa_n} \rightarrow \exp\left(\frac{1}{a} \int_{-\infty}^{\infty} \log(1 + |x|^a) \phi(x) dx\right) \text{ almost surely.}$$

Theorem 5.3 is a special case of Theorem 8.4 when $\phi(x)$ is the density function for the symmetric, strictly stable distribution $N(0, \sigma^2)$. Another notable symmetric, strictly stable distribution is the Cauchy distribution with the density function

$$\phi(x) = \frac{1}{\pi} \frac{1}{1+x^2}.$$

The exponent a for the Cauchy distribution is 1 [6]. Using Theorem 8.4 we obtain the following.

THEOREM 8.5. *For random unit triangular matrices with entries from the standard Cauchy distribution, as $n \rightarrow \infty$,*

$$\sqrt[n]{\kappa_n} \rightarrow \exp\left(\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\log(1+|x|)}{1+x^2} dx\right) \text{ almost surely.}$$

Numerical integration shows the constant to be 2.533737279...

A similar generalization can be made for triangular matrices without a unit diagonal. However, the analogue of Theorem 8.4 for such matrices involves not $\phi(x)$ but the density function $\psi(x)$ of the quotient $x = y/z$ obtained by taking y, z as independent variables with the PDF ϕ . The distribution ψ can be difficult to compute and work with. We state only the final theorem about κ_n for triangular matrices with entries drawn from a symmetric strictly stable distribution.

THEOREM 8.6. *For random triangular matrices with entries from a symmetric, strictly stable distribution with density function $\phi(x)$ and exponent a , as $n \rightarrow \infty$,*

$$\sqrt[n]{\kappa_n} \rightarrow \exp\left(\frac{1}{a} \int_{-\infty}^{\infty} \log(1+|x|^a) \psi(x) dx\right) \text{ almost surely,}$$

where $\psi(x)$ is the density function of the quotient of two independent variables with the density function $\phi(x)$.

Theorem 4.3 is a special case of Theorem 8.6 when $\phi(x)$ is the density function of the distribution $N(0, \sigma^2)$. The $\psi(x)$ corresponding to $N(0, \sigma^2)$ is the standard Cauchy distribution. To apply Theorem 8.6 for the Cauchy distribution, we note that

$$\psi(x) = \frac{2}{\pi^2} \frac{\log|x|}{x^2-1}$$

is the density function of the quotient if the numerator and the denominator are independent Cauchy variables [11]. Therefore, Theorem 8.6 implies the following.

THEOREM 8.7. *For random triangular matrices with entries from the standard Cauchy distribution, as $n \rightarrow \infty$,*

$$\sqrt[n]{\kappa_n} \rightarrow \exp\left(\frac{2}{\pi^2} \int_{-\infty}^{\infty} \log(1+|x|) \frac{\log|x|}{x^2-1} dx\right) \text{ almost surely.}$$

The constant of convergence in Theorem 8.7 is 3.063094192...

9. Summary. Below is a summary of the exponential growth factors $\lim_{n \rightarrow \infty} \sqrt[n]{\kappa_n}$ that we have established for triangular matrices with normal entries:

Real triangular	2	Theorem 4.3
Real unit triangular, $\sigma^2 = 1$	1.305683410...	Theorem 5.3
Complex triangular	$e^{1/2} = 1.647...$	Theorem 7.3
Complex unit triangular, $\sigma^2 = 1$	1.347395784...	Theorem 7.4

The theorems about unit triangular matrices with normally distributed, real or complex entries apply for any variance σ^2 , not just $\sigma^2 = 1$. Constants of convergence for any symmetric, strictly stable distribution were derived in Theorems 8.4 and 8.6. Those two theorems were specialized to the Cauchy distribution in Theorems 8.5 and 8.7.

Similar results seem to hold more generally, i.e., even when the entries of the random triangular matrix are not from a stable distribution. Moreover, the complete knowledge of moments achieved in Lemma 3.2 and its analogues might be enough to prove stronger limit theorems than Theorem 4.3 and its analogues. We will present limit theorems and results about other kinds of random triangular matrices in a later publication. We will also discuss the connection between random recurrences and products of random matrices, and the pseudospectra of infinite random triangular matrices.

For the random recurrences we have considered here, every new term is a random sum of all the previous terms in the sequence. The exponential increase of successive terms with probability 1 holds even for some random recurrences that generate a new term as a random sum of a fixed number of previous terms. For example, if we define random Fibonacci sequences by $t_1 = t_2 = 1$, and for $n > 2$, $t_n = \pm t_{n-1} \pm t_{n-2}$, where each \pm sign is independent and either $+$ or $-$ with probability $1/2$, then $\sqrt[n]{|t_n|} \rightarrow 1.13198824...$ almost surely [21]. Thus, the condition number increases exponentially even for some random triangular matrices that are banded.

We close with two figures that illustrate the first main result of this paper, namely, for random triangular matrices with $N(0, 1)$ entries, $\sqrt[n]{\kappa_n} \rightarrow 2$ almost surely as $n \rightarrow \infty$ (Theorem 4.3). Figure 9.1 plots the results of a single run of the random recurrence (2.1) to 100,000 steps, confirming the constant 2 to about two digits. The expense involved in implementing the full recurrence (2.1) for so many steps would be prohibitive. However, since t_k grows at the rate 2^k , we need include only a fixed number of terms in (2.1) to compute t_k to machine precision. For the figure, we used 200 terms, although half as many would have been sufficient. Careful scaling was necessary to avoid overflow while computing this figure.

Figure 9.2 plots the condition number of a single random triangular matrix for each dimension from 1 to 200. The exponential trend at the rate 2^n is clear, but as in Figure 1.1, the convergence as $n \rightarrow \infty$ is slow.

Acknowledgments. We thank D. Coppersmith, P. Diaconis, H. Kesten, A. Odlyzko, J. Sethna, H. Wilf, and the referees for helpful discussions and comments. We are especially grateful to Prof. Diaconis for introducing us to stable distributions and to one of the referees for a very careful reading of this paper.

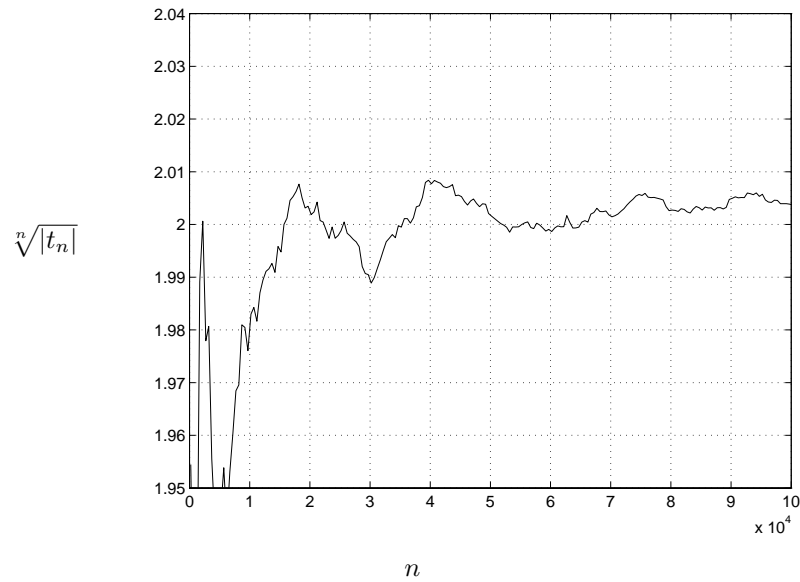


FIG. 9.1. Illustration of Theorem 4.3. After 100,000 steps of the random recurrence (2.1), $\sqrt[n]{|t_n|}$ has settled to within 1% of its limiting value 2. The implementation is explained in the text.

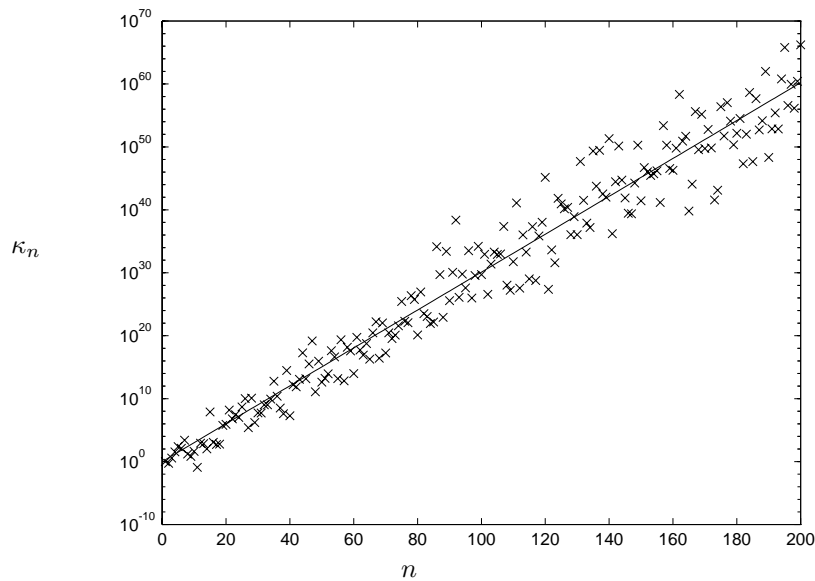


FIG. 9.2. Another illustration of Theorem 4.3. Each cross is obtained by computing the condition number κ_n for one random triangular matrix of dimension n with $N(0, 1)$ entries. The solid line represents 2^n .

REFERENCES

- [1] M. ABRAMOWITZ AND I.A. STEGUN, EDs., *Handbook of Mathematical Functions*, Dover, New York, 1970.
- [2] P. BILLINGSLEY, *Probability and Measure*, 2nd ed., John Wiley, New York, 1986.

- [3] J. B. CONWAY, *Functions of One Complex Variable*, Springer, New York, 1995.
- [4] A. EDELMAN, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. dissertation and Numerical Analysis Report 89-7, Massachusetts Institute of Technology, Cambridge, MA, 1989.
- [5] A. EDELMAN, *Eigenvalues and condition numbers of random matrices*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 543–560.
- [6] W. FELLER, *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed., John Wiley, New York, 1971.
- [7] L. V. FOSTER, *Gaussian elimination with partial pivoting can fail in practice*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1354–1362.
- [8] I. S. GRADSHTEYN AND I. W. RYZHIK, *Table of Integrals, Series, and Products*, 4th ed., Academic Press, New York, 1965.
- [9] H. H. GOLDSTINE AND J. VON NEUMANN, *Numerical inverting of matrices of high order*, Amer. Math. Soc. Bull., 53 (1947), pp. 1021–1099.
- [10] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD, 1996.
- [11] A. K. GUPTA AND Z. GOVINDARAJULU, *Distribution of the quotient of two independent Hotelling T^2 -variates*, Comm. Statist., 4 (1975), pp. 449–453.
- [12] N. J. HIGHAM, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, PA, 1996.
- [13] G. HARDY, J. E. LITTLEWOOD, AND G. POLYA, *Inequalities*, 2nd ed., Cambridge University Press, Cambridge, UK, 1988.
- [14] H. HOTELLING, *Some new methods in matrix calculation*, Ann. Math. Statist., 14 (1943), pp. 1–34.
- [15] M. L. MEHTA, *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York, 1967.
- [16] J. W. SILVERSTEIN, *The smallest eigenvalue of a large-dimensional Wishart matrix*, Ann. Probab., 13 (1985), pp. 1364–1368.
- [17] L. N. TREFETHEN AND D. BAU, III, *Numerical Linear Algebra*, SIAM, Philadelphia, PA, 1997.
- [18] L. N. TREFETHEN AND R. S. SCHREIBER, *Average-case stability of Gaussian elimination*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 335–360.
- [19] J. H. WILKINSON, *Error Analysis of direct methods of matrix inversion*, J. Assoc. Comput. Mach., 8 (1961), pp. 281–330.
- [20] S. J. WRIGHT, *A collection of problems for which Gaussian elimination with partial pivoting is unstable*, SIAM J. Sci. Comput., 14 (1993), pp. 231–238.
- [21] D. VISWANATH, *Random Fibonacci sequences and the number 1.13198824...*, Math. Comp., submitted.
- [22] M. C. YEUNG AND T. F. CHAN, *Probabilistic analysis of Gaussian elimination without pivoting*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 499–517.