2013

TR-2013015: Estimating the Norms of Random Circulant and Toeplitz Matrices and Their Inverses

Victor Y. Pan

Guoliang Qian

Estimating the Norms of Random Circulant and Toeplitz Matrices and Their Inverses *

Victor Y. Pan[1,2],[a] and Guoliang Qian[2],[b]

[1] Department of Mathematics and Computer Science
Lehman College of the City University of New York
Bronx, NY 10468 USA
[2] Ph.D. Programs in Mathematics and Computer Science
The Graduate Center of the City University of New York
New York, NY 10036 USA
[a] victor.pan@lehman.cuny.edu
http://comet.lehman.cuny.edu/vpan/
[b] gqian@gc.cuny.edu

Abstract

We estimate the norms of standard Gaussian random Toeplitz and circulant matrices and their inverses, mostly by means of combining some basic techniques of linear algebra. In the case of circulant matrices we obtain sharp probabilistic estimates, which show that these matrices are expected to be very well conditioned. Our probabilistic estimates for the norms of standard Gaussian random Toeplitz matrices are within a factor of $\sqrt{2}$ from those in the circulant case. We also achieve partial progress in estimating the norms of Toeplitz inverses. Namely we yield reasonable probabilistic upper estimates for these norms assuming certain bounds on the absolute values of two corner entries of the inverse. Empirically we observe that the condition numbers of random Toeplitz and general matrices tend to be of the same order. As the matrix size grows, these numbers grow equally slowly, although faster than for random circulant matrices.

2000 Math. Subject Classification: 15A52, 15A12, 65F22, 65F35

Key Words: Matrix norms, Condition numbers, Toeplitz matrices, Circulant matrices, Random matrices

1 Introduction

Estimating the condition numbers $\kappa(A) = ||A||\ ||A^{-1}||$ of random structured matrices A is a well known challenge [SST06], linked to the design of efficient randomized matrix algorithms, e.g., in the papers [HMT11], [XXG12], [PQZ13]. We seek such estimates for standard Gaussian random Toeplitz and circulant matrices, mostly by employing and combining some basic techniques of linear algebra. In the case of circulant matrices A our sharp probabilistic estimates for the norms $||A||$ and $||A^{-1}||$ show that the matrices are expected to be very well conditioned. In the case of Toeplitz

*Some results of this paper have been presented at the ACM-SIGSAM International Symposium on Symbolic and Algebraic Computation (ISSAC '2011), San Jose, CA, 2011, the 7th International Congress on Industrial and Applied Mathematics (ICIAM 2011), in Vancouver, British Columbia, Canada, July 18-22, 2011, the SIAM International Conference on Linear Algebra, in Valencia, Spain, June 18-22, 2012, and the 18th Conference of the International Linear Algebra Society (ILAS'2013), Providence, RI, 2013

matrices A our estimates for the norm $||A||$ are within a factor of $\sqrt{2}$ from the bounds in the circulant case. Estimating the norms $||A^{-1}||$ turns out to be harder, however. Namely we obtain reasonable probabilistic upper bounds on that norm, and consequently on the condition number $\kappa(A) = ||A||\ ||A^{-1}||$, assuming that the norms of the first row and the first column of the inverse do not exceed dramatically the absolute values of their two corner entries, $(A^{-1})_{1,n}$ and $(A^{-1})_{n,1}$, respectively.

For some large and important special classes of $n \times n$ Toeplitz matrices the condition numbers grow exponentially in n as $n \to \infty$ [BG05], but in our tests with both random general and random Toeplitz matrices their condition numbers consistently grew with the same reasonably slow rate as n grew large, although the growth was still significantly faster than in the case of circulant matrices.

Our study can be immediately extended to the cases of factor circulant and Hankel matrices and partly to the case of other than Gaussian probability distributions.

We organize our paper as follows. We recall some definitions and basic results on general matrix computations in the next section and on Toeplitz and circulant matrices in Section 3. In Section 4 we deduce sharp estimates on the norms of circulant and Toeplitz matrices in terms of their generating vectors. In Section 5 we similarly estimate the norms of the inverses of these matrices. We recall the definition of Gaussian random matrices and some basic facts about them in Section 6. In Section 7 we extend our estimates of Sections 4 and 5 to standard Gaussian random Toeplitz and circulant matrices and their inverses. In Section 8 we cover numerical tests, which are the contribution of the second author. In the Appendix we discuss nondegeneration of random matrices.

2 Some definitions and basic results

In ths section we recall some customary definitions and basic properties on matrix computations [GL96], [S98]. A^T is the transpose of a matrix A, and A^H is its Hermitian transpose. $A^H = A^T$ for a real matrix A. $A^T = A$ if A is a real symmetric matrix, $A^H = A$ if A is a Hermitian matrix.

A square matrix A is unitary if $A^H A = AA^H = I$, and is orthogonal if $A^T A = AA^T = I$. For a vector $\mathbf{v} = (v_i)_{i=1}^n$ define the norms $||\mathbf{v}||_1 = \sum_i |v_i|$, $||\mathbf{v}||_2 = (\sum_i |v_i|^2)^{1/2}$, and $||\mathbf{v}||_\infty = \max_i |v_i|$. For a matrix $A = (a_{ij})_{i,j=1}^n$ define its h-norms $||A||_h = \inf_{||\mathbf{v}||_h=1} ||A\mathbf{v}||_h$ for $h = 1, 2, \infty$ and its Frobenius norm $||A||_F = (\sum_{i,j=1}^n |a_{ij}|^2)^{1/2}$. We write $||\mathbf{v}|| = ||\mathbf{v}||_2$ and $||A|| = ||A||_2$. $||A||$ is called the spectral norm of a matrix A. For all matrices A it holds that

$$||A||_1 = ||A^T||_\infty = ||A^H||_\infty = \max_j \sum i, j|a_{ij}|, \tag{2.1}$$

$$\frac{1}{\sqrt{n}}||A||_h \leq ||A|| \leq \sqrt{n}||A||_h \text{ for } h = 1 \text{ and } h = \infty,\ ||A||^2 \leq ||A||_1||A||_\infty, \tag{2.2}$$

$$||D|| = ||D||_1 = ||D||_\infty,\ ||D||_F^2 = \sum_{i=1}^n |d_i|^2 \text{ for a diagonal matrix } D = \text{diag}(d_i)_{i=1}^n, \tag{2.3}$$

$$||UAV|| = ||A|| \text{ and} ||UAV||_F = ||A||_F \text{ for unitary matrices } U \text{ and } V, \tag{2.4}$$

$$||A|| \leq ||A||_F \leq \sqrt{\rho}||A|| \text{ where } \rho = \text{rank}(A), \tag{2.5}$$

$$||A + B||_h \leq ||A||_h + ||B||_h,\ ||FG||_h \leq ||F||_h||G||_h \text{ for } h = 1, 2, \infty. \tag{2.6}$$

For the latter bounds the matrix sizes must match, to define the matrices $A + B$ and FG.

$\kappa(A) = ||A||\ ||A^{-1}||$ is the *condition number* of a nonsingular matrix A. Such matrix is *ill conditioned* if its condition number is large and is *well conditioned* if it is reasonably bounded.

3 Toeplitz and f-circulant matrices and their inverses

$\mathbf{e}_i$ is the ith coordinate vector of a dimension n for $i = 1, \ldots, n$. Define two vectors

$$\mathbf{t}_+ = (t_i)_{i=1-n}^{n-1} \text{ and } \mathbf{t} = (t_i)_{i=0}^{n-1} \tag{3.1}$$

of dimensions n and $2n-1$, respectively. They in turn define a *Toeplitz* $n \times n$ matrix $T_n = T(\mathbf{t}_+) = (t_{i-j})_{i,j=1}^n$, a lower *triangular Toeplitz* $n \times n$ matrix $Z(\mathbf{t}) = (t_{i-j})_{i,j=1}^n$ (where $t_k = 0$ for $k < 0$), and its transpose $Z(\mathbf{t})^T = (Z(\mathbf{t}))^T$. Define the $n \times n$ matrices $Z = Z_0 = Z(\mathbf{e}_2)$ of downshift and $Z_f = Z + f\mathbf{e}_1^T\mathbf{e}_n$, for $f \neq 0$, of *f-circular shift* (see equation (3.2) below). It holds that $Z\mathbf{t} = (t_{i-1})_{i=0}^{n-1}$ and $Z(\mathbf{t}) = Z_0(\mathbf{t}) = \sum_{i=1}^n t_h Z^{i-1}$. An *f-circulant* (or factor circulant) matrix $Z_f(\mathbf{t}) = \sum_{i=1}^n t_i Z_f^{i-1}$ is a special Toeplitz $n \times n$ matrix defined by its first column vector $\mathbf{t} = (t_i)_{i=1}^n$ and a scalar f. f-circulant matrix is called *circulant* if $f = 1$.

$$T_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{1-n} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & & \cdots & & 0 \\ 1 & \ddots & & & \\ \vdots & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & \\ 0 & & \cdots & 1 & 0 \end{pmatrix}, \quad Z_f = \begin{pmatrix} 0 & & \cdots & & f \\ 1 & \ddots & & & \\ \vdots & \ddots & \ddots & & \vdots \\ & & \ddots & 0 & \\ 0 & & \cdots & 1 & 0 \end{pmatrix}. \tag{3.2}$$

$J = J_n = (\mathbf{e}_n \mid \ldots \mid \mathbf{e}_1)$ is the reflection matrix of size $n \times n$, $J = J^T = J^{-1}$. A Toeplitz matrix T_n and its inverse (if defined) are *persymmetric*, that is,

$$JT_nJ = T_n \text{ and } JT_n^{-1}J = T_n^{-1}. \tag{3.3}$$

Hereafter $\omega = \exp(2\pi\sqrt{-1}/n)$ is a primitive nth root of unity. $\Omega = (\omega^{(i-1)(j-1)})_{i,j=0}^{n-1}$ is the matrix of the discrete Fourier transform at n points.

Theorem 3.1. *$\Omega^H\Omega = nI$, that is $\frac{1}{\sqrt{n}}\Omega$ is a unitary matrix.*

The following theorem implies that the inverses (wherever they are defined) and pairwise products of f-circulant matrices are f-circulant.

Theorem 3.2. *(See [CPW74].) We have $Z_1(\mathbf{t}) = \Omega^{-1}D(\Omega\mathbf{t})\Omega$. More generally, for any $f \neq 0$, we have $Z_{f^n}(\mathbf{t}) = U_f^{-1}D(U_f\mathbf{t})U_f$ where $U_f = \Omega D(\mathbf{f})$, $\mathbf{f} = (f^i)_{i=0}^{n-1}$, $D(\mathbf{u}) = \operatorname{diag}(u_i)_{i=0}^{n-1}$ for a vector $\mathbf{u} = (u_i)_{i=0}^{n-1}$.*

Theorem 3.3. *Write $T_k = (t_{i-j})_{i,j=0}^{k-1}$ for $k = n, n+1$.*

(a) Let the matrix T_n be nonsingular and write $\mathbf{p} = T_n^{-1}\mathbf{e}_1$ and $\mathbf{q} = T_n^{-1}\mathbf{e}_n$. If $p_1 = \mathbf{e}_1^T\mathbf{p} \neq 0$, then $p_1T_n^{-1} = Z(\mathbf{p})Z(J\mathbf{q})^T - Z(Z\mathbf{q})Z(ZJ\mathbf{p})^T$.

In parts (b) and (c) below let a Toeplitz $(n+1) \times (n+1)$ matrix T_{n+1} be nonsingular and write $\widehat{\mathbf{v}} = (v_i)_{i=0}^n = T_{n+1}^{-1}\mathbf{e}_1$, $\mathbf{v} = (v_i)_{i=0}^{n-1}$, $\mathbf{v}' = (v_i)_{i=1}^n$, $\widehat{\mathbf{w}} = (w_i)_{i=0}^n = T_{n+1}^{-1}\mathbf{e}_{n+1}$, $\mathbf{w} = (w_i)_{i=0}^{n-1}$, and $\mathbf{w}' = (w_i)_{i=1}^n$.

(b) If $v_0 \neq 0$, then the matrix T_n is nonsingular and $v_0T_n^{-1} = Z(\mathbf{v})Z(J\mathbf{w}')^T - Z(\mathbf{w})Z(J\mathbf{v}')^T$.

(c) If $v_n \neq 0$, then the matrix $T_{1,0} = (t_{i-j})_{i=1,j=0}^{n,n-1}$ is nonsingular and $v_nT_{1,0}^{-1} = Z(\mathbf{w})Z(J\mathbf{v}')^T - Z(\mathbf{v})Z(J\mathbf{w}')^T$.

Proof. See [GS72] on parts (a) and (b); see [GK72] on part (c). □

4 The norm bounds for circulant and Toeplitz matrices

Theorem 4.1. *Assume a pair of vectors $\mathbf{t}$ and $\mathbf{t}_+$ of (3.1), defining the circulant and Toeplitz $n \times n$ matrices $Z_1(\mathbf{t})$ and $T_n = T(\mathbf{t}_+)$. Then it holds that*

$$||Z(\mathbf{t})||_h \leq ||Z_1(\mathbf{t})||_h \text{ for } h = F, 1, 2, \infty, \tag{4.1}$$

$$||Z_1(\mathbf{t})|| \leq ||Z_1(\mathbf{t})||_1 = ||Z_1(\mathbf{t})||_\infty = ||\mathbf{t}||_1 \leq \sqrt{n}\ ||\mathbf{t}||, \tag{4.2}$$

$$||Z_1(\mathbf{t})||_F = \sqrt{n}\ ||\mathbf{t}||, \tag{4.3}$$

$$||T_n|| \leq ||T_n||_1 = ||T_n||_\infty \leq ||\mathbf{t}_+||_1 \leq \sqrt{2n-1}\ ||\mathbf{t}_+|| \text{ and } ||T_n||_F \leq \sqrt{2n-1}\ ||\mathbf{t}_+||. \tag{4.4}$$

Proof. Readily verify equation (4.1). Combine equations (2.1) and (2.2) and obtain relationships (4.2). Combine Theorems 3.1 and 3.2 with relationships (2.3) and (2.4) and obtain that $||Z_1(\mathbf{t})||_F^2 = ||D(\Omega\mathbf{t})||_F^2 = ||(\Omega\mathbf{t})||^2 = n\ ||\mathbf{t}||^2$, yielding equation (4.3). Embed the $n \times n$ matrix T_n into the circulant $(2n-1)\times(2n-1)$ matrix $Z_1(\mathbf{t}_+)$ and obtain that $||T_n||_h \leq ||Z_1(\mathbf{t}_+)||_h$ where h can stand for F,1,2, and ∞. Together with relationships (4.2) and (4.3), this implies relationships (4.4). □

Remark 4.1. *(Extension to the case of f-circulant matrices.) Theorem 3.2 implies that*

$$\frac{1}{g(f)}||Z_1(\mathbf{v})|| \leq ||Z_f(\mathbf{v})|| \leq g(f)||Z_1(\mathbf{v})||$$

and if the matrices $Z_1(\mathbf{v})$ and $Z_f(\mathbf{v})$ are nonsingular, then also

$$\frac{1}{g(f)}||Z_1(\mathbf{v})^{-1}|| \leq ||Z_f(\mathbf{v})^{-1}|| \leq g(f)||Z_1(\mathbf{v})^{-1}||$$

for all vectors $\mathbf{v}$, scalars $f \neq 0$, $g(f) = \max\{|f|, 1/|f|\}$, and $j = 1,\dots,n$. Therefore we can readily extend our norm estimates from circulant to f-circulant matrices for $f \neq 0$. In particular the former estimates do not change where $|f| = 1$.

Remark 4.2. *(Extension to the case of Hankel matrices.) All our estimates for Toeplitz matrices are immediately extended to the case of Hankel matrices $H_n = (h_{i+j})_{i,j=1}^n$ because the products $H_nJ = T_n$ and $JH_n = T_n$ are $n \times n$ Toeplitz matrices.*

5 The norm bounds for the inverses of Toeplitz and circulant matrices

Hereafter we write $((v_i)_{i=1}^n)^{-1} = (1/v_i)_{i=1}^n)$. Similarly to equation (4.3) deduce that

$$||Z_1(\mathbf{t})^{-1}||_F = ||(D(\Omega\mathbf{t}))^{-1}||_F = ||(\Omega\mathbf{t})^{-1}||,\ ||Z_1(\mathbf{t})^{-1}|| = ||(D(\Omega\mathbf{t}))^{-1}|| = ||(\Omega\mathbf{t})^{-1}||_\infty. \qquad (5.1)$$

Combine this bound with (2.2) and obtain that

$$||Z_1(\mathbf{t})^{-1}||_h \leq \sqrt{n}\ ||(\Omega\mathbf{t})^{-1}||_\infty \text{ for } h = 1 \text{ and } h = \infty.$$

Theorem 5.1. *Under the assumptions of part (a) of Theorem 3.2 it holds that $||p_1T_n^{-1}||_h \leq 2||\mathbf{p}||_1\ ||\mathbf{q}||_1 \leq 2n||\mathbf{p}||\ ||\mathbf{q}||$ for $h = 1, 2, \infty$.*

Proof. Recall from part (a) of Theorem 3.3 that $p_1T_n^{-1} = Z(\mathbf{p})Z(J\mathbf{q})^T - Z(Z\mathbf{q})Z(ZJ\mathbf{p})^T$. Therefore (cf. (2.6)) $||p_1T_n^{-1}||_h \leq ||Z(\mathbf{p})||_h\ ||Z(J\mathbf{q})^T||_h + ||Z(Z\mathbf{q})||_h\ ||Z(ZJ\mathbf{p})^T||_h$ for $h = 1, 2, \infty$. Combine this bound with relationships (4.1) and (4.2) and deduce that $||p_1T_n^{-1}||_h \leq ||\mathbf{p}||_1\ ||J\mathbf{q}||_1 + ||Z\mathbf{q}||_1\ ||ZJ\mathbf{p}||_1$ due to (4.2). Note that $||J\mathbf{w}||_1 = ||\mathbf{w}||_1$ and $||Z\mathbf{w}||_1 \leq ||\mathbf{w}||_1$ for every vector $\mathbf{w}$ and obtain Theorem 5.1. □

6 Gaussian random matrices

By extending the norm bounds of the two previous sections we can estimate the norms of circulant and Toeplitz matrices with random entries selected under various probability distributions. We are going to do this just under the Gaussian distribution, which enables simpler and stronger estimates.

Definition 6.1. $F_\gamma(y) =$ *Probability*$\{\gamma \leq y\}$ *is the* cumulative distribution function (cdf) *of a real random variable γ evaluated at a real point y.* $F_{g(\mu,\sigma)}(y) = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{y}\exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$ *for a Gaussian random variable $g(\mu,\sigma)$ with a mean μ, a positive variance σ^2, and a positive standard deviation σ.*

It follows that

$$|y-\mu|/\sigma \geq s \text{ with the probability } \sqrt{\frac{2}{\pi}}\int_s^{+\infty}\exp(-\frac{x^2}{2})dx, \tag{6.1}$$

that is the probability decays very fast as s grows large.

Definition 6.2. *A matrix (or a vector) is a* Gaussian random matrix (or vector) *with a mean μ and a positive variance σ^2 if it is filled with independent identically distributed Gaussian random variables, all having the mean μ and variance σ^2. $\mathcal{G}_{\mu,\sigma}^{m\times n}$ is the set of such Gaussian random $m\times n$ matrices, which are* standard *where $\mu=0$ and $\sigma^2=1$. By restricting this set to Toeplitz or f-circulant matrices we obtain the sets $\mathcal{T}_{\mu,\sigma}^{m\times n}$ and $\mathcal{Z}_{f,\mu,\sigma}^{n\times n}$ of* Gaussian random Toeplitz *and* Gaussian random f-circulant matrices*, respectively, which are standard where $\mu=0$ and $\sigma^2=1$.*

Definition 6.3. *$\chi_{\mu,\sigma,n}(y)$ is the cdf of the norm $||\mathbf{v}||=(\sum_{i=1}^n v_i^2)^{1/2}$ of a Gaussian random vector $\mathbf{v}=(v_i)_{i=1}^n\in\mathcal{G}_{\mu,\sigma}^{n\times 1}$. For $y\geq 0$ we have $\chi_{0,1,n}(y)=\frac{2}{2^{n/2}\Gamma(n/2)}\int_0^y x^{n-1}\exp(-x^2/2)dx$ where $\Gamma(h)=\int_0^\infty x^{h-1}\exp(-x)dx$ is the Gamma function, $\Gamma(n+1)=n!$ for nonnegative integers n.*

We recall the following basic results.

Lemma 6.1. *Suppose G is Gaussian matrix, S and T are square orthogonal matrices, and the products SG and GT are well defined. Then SG and GT are Gaussian matrices.*

Lemma 6.2. *[SST06, Lemma A.2]. For a nonnegative scalar y, a unit vector $\mathbf{t}\in\mathbb{R}^{n\times 1}$, and a vector $\mathbf{b}\in\mathcal{G}_{\mu,\sigma}^{n\times 1}$, we have $F_{|\mathbf{t}^T\mathbf{b}|}(y)\leq\sqrt{\frac{2}{\pi}}\frac{y}{\sigma}$.*

Remark 6.1. *The latter bound is independent of μ and n. It holds for any μ even if all coordinates of the vector $\mathbf{b}$ are fixed except for a single coordinate in $\mathcal{G}_{\mu,\sigma}$.*

Hereafter we assume by default that *Gaussian random general, Toeplitz and circulant matrices have full rank* (see Appendix A).

7 Norm bounds for Gaussian random Toeplitz and circulant matrices and their inverses

7.1 The norms of Gaussian random Toeplitz and circulant matrices

Combine our estimates of Theorem 4.1 with Definition 6.2 and obtain the following upper bounds on the norms of Gaussian random Toeplitz and circulant matrices.

Corollary 7.1. *For Gaussian random Toeplitz $n\times n$ matrix T_n and Gaussian random circulant $n\times n$ matrix $Z_1(\mathbf{t})$ it holds that $F_{||T_n||_h}(y)\geq\chi_{\mu,\sigma,2n-1}(y/\sqrt{2n-1})$ and $F_{||Z_1(\mathbf{t}||_h}(y)\geq\chi_{\mu,\sigma,n}(y/\sqrt{n})$ where h can stand for F,1,2, or ∞.*

7.2 The expected norms of the inverses of Gaussian random matrices have order at least $1/\sigma$ where the mean μ greatly exceeds the standard deviation σ

Bounds (6.1) imply that the matrix $M\in\mathcal{G}_{\mu,\sigma}^{n\times n}$ is expected to be approximated within the norm bound of order σ by the matrix $\mu\mathbf{e}\mathbf{e}^T$ of rank 1, for the vector $\mathbf{e}=(1,\dots,1)^T$ provided that $mu|\gg\sigma$, and if indeed so, then the norm $||M^{-1}||$ has order of at least $1/\sigma$. Similar argument applies to Gaussian random Toeplitz and circulant matrices. In the nexct section we derive our estimates for any pair of (μ,σ), but one can avoid the latter undesired growth of the norm by restricting the study to standard Gaussian random Toeplitz and circulant matrices.

7.3 The norm of the inverse of Gaussian random circulant matrix

Theorem 7.1. *(Cf. Table 8.4.) Assume $y \geq 0$ and vector $\mathbf{t} \in \mathcal{G}_{\mu,\sigma}^{n\times 1}$. Then $F_{||Z_1(\mathbf{t})^{-1}||}(z) = 1-(1-q)^n$, for $q = \frac{1}{\sigma\sqrt{2\pi}}\int_{-1/z}^{1/z}\exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$, and so $F_{||Z_1(\mathbf{t})^{-1}||}(z) \approx nq$ where the value q is small, that is where the value $|z|$ is large or the distance between μ and the range $[-1/z, 1/z]$ is large.*

Remark 7.1. *By applying bounds (2.2) and (2.5) we can extend the theorem to estimate the cdfs $F_{||Z_1(\mathbf{t})^{-1}||_h}(z)$ where h can stand for F, 1, or ∞.*

Proof. Equation (5.1) implies that $||Z_1(\mathbf{t})^{-1}|| = 1/\min_{i=1}^n |u_i|$ where $\mathbf{u} = (u_i)_{i=1}^n = \Omega\mathbf{t}$. Apply Lemma 6.1 and obtain that $\mathbf{u}$ is a Gaussian random vector $\mathbf{q}(\mu,\sigma)$. For any i, $i = 1,\dots,n$, it holds that $|u_i| \leq y$ with probability $1-p$ where $p = \frac{1}{\sigma\sqrt{2\pi}}\int_{-y}^{y}\exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$ (cf. Definition 6.1), and so $\min_{i=1}^n |u_i| \leq y$ with probability $(1-p)^n$ because u_i are independent random variables. Equivalently $1/\min_{i=1}^n |u_i| \geq 1/y$ with probability $(1-p)^n$. Therefore $||Z_1(\mathbf{t})^{-1}|| = 1/\min_{i=1}^n |u_i| \leq 1/y$ with probability $1-(1-p)^n \geq np$. Substitute $z = 1/y$ and obtain the theorem. □

7.4 Norm bounds for the inverse of Gaussian random Toeplitz matrix

Our next subject is the estimates for the norm $||T_n^{-1}||_h$ for $h = 1, 2, \infty$ and Gaussian random Toeplitz matrix $T_n \in \mathcal{T}_{\mu,\sigma}^{n\times n}$, which is known to be nonsingular with probability 1. We can extend these estimates to the norm $||T_n^{-1}||$ by using (2.5).

Theorem 7.2. *Given a matrix $T_n = (t_{i-j})_{i,j=1}^n \in \mathcal{T}_{0,1}^{n\times n}$, assumed to be nonsingular, write $\mathbf{p} = (p_i)_{i=1}^n = T_n^{-1}\mathbf{e}_1$, $\mathbf{q} = (q_i)_{i=1}^n = T_n^{-1}\mathbf{e}_1$, $u_n = p_n/||\mathbf{p}||$, $v_1 = q_1/||\mathbf{q}||$, and $p_1 = \mathbf{e}_n^T\mathbf{p} = \mathbf{e}_1^T\mathbf{q} = q_n$ (cf. (3.3)). Then $||p_1 T_n^{-1}||_h \leq 2n/(\alpha\beta)$ for $h = 1, 2, \infty$ and two random variables α and β such that*

$$F_\alpha(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma|u_n|} \text{ and } F_\beta(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma|v_1|} \text{ for } y \geq 0. \tag{7.1}$$

Proof. By virtue of Theorem 5.1 we just need to estimate the two random variables $||\mathbf{p}||$ and $||\mathbf{q}||$. By virtue of its definition the vector $\mathbf{p}$ is orthogonal to the vectors $T_n\mathbf{e}_2,\dots,T_n\mathbf{e}_n$, whereas $\mathbf{p}^T T_n\mathbf{e}_1 = 1$ (cf. [SST06]). Consequently the vectors $T_n\mathbf{e}_2,\dots,T_n\mathbf{e}_n$ uniquely define the vector $\mathbf{u} = (u_i)_{i=1}^n = \mathbf{p}/||\mathbf{p}||$, whereas $|\mathbf{u}^T T_n\mathbf{e}_1| = 1/||\mathbf{p}||$. The last coordinate t_{n-1} of the vector $T_n\mathbf{e}_1$ is independent of the vectors $T_n\mathbf{e}_2,\dots,T_n\mathbf{e}_n$ and consequently of the vector $\mathbf{u}$. Apply Remark 6.1 to estimate the cdf of the random variable $\alpha/|u_n| = 1/(||\mathbf{p}||\ |u_n|) = |\mathbf{u}^T T_n\mathbf{e}_1|/|u_n|$ and obtain that $F_\alpha(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma|u_n|}$ for $y \geq 0$. Likewise we deduce that $F_\beta(y) \leq \sqrt{\frac{2n}{\pi}}\frac{y}{\sigma|v_1|}$ for $y \geq 0$. Finally combine both bounds (7.1) on the cdfs $F_{\alpha/|u_n|}(y)$ and $F_{\beta/|v_1|}(y)$ with Theorem 5.1. □

7.5 Bounding the leading entry of the inverse

Theorem 7.2 bounds the norm $||T_n^{-1}||$, in terms of the random variables $|u_n|$, $|v_1|$, and $|p_1| = |q_n|$. By applying parts (b) and (c) of Theorem 3.3 instead of its part (a), we similarly deduce the bounds $||v_0 T_{n+1}^{-1}|| \leq 2/(\alpha\beta)$ and $||v_n T_{n+1}^{-1}|| \leq 2/(\alpha\beta)$ for two pairs of random variables α and β that satisfy (7.1) for $n+1$ replacing n. Note that $p_1 = \frac{\det T_{n-1}}{\det T_n}$, $v_0 = \frac{\det T_n}{\det T_{n+1}}$, and $v_n = \frac{\det T_{0,1}}{\det T_{n+1}}$ for $T_{0,1} = (t_{i-j})_{i=0,j=1}^{n-1,n}$. Next we bound the geometric means of the ratios $|\frac{\det T_{h+1}}{\det T_h}|$ for $h = 1,\dots,k-1$. $1/|p_1|$ and $1/|v_0|$ are such ratios for $k = n-1$ and $k = n$, respectively, whereas the ratio $1/|v_n|$ is similar to $1/|v_0|$, under slightly distinct notation.

Theorem 7.3. *Let $T_h \neq O$ denote $h\times h$ matrices for $h = 1,\dots,k$ whose entries have absolute values at most t for a random variable t, e.g., for $t = ||T||$. Furthermore let $T_1 = (t)$. Then the geometric mean $(\prod_{h=1}^{k-1}|\frac{\det T_{h+1}}{\det T_h}|)^{1/(k-1)} = \frac{1}{t}|\det T_k|^{1/(k-1)}$ is at most $k^{\frac{1}{2}(1+\frac{1}{k-1})}t$.*

Proof. The theorem follows from Hadamard's upper bound $|\det M| \leq k^{k/2}t^k$, which holds for any $k\times k$ matrix $M = (m_{i,j})_{i,j=1}^k$ with $\max_{i,j=1}^k |m_{i,j}| \leq t$. □

The theorem shows that the geometric mean of the ratios $|\frac{\det T_{h+1}}{\det T_h}|$ for $h = 1, \ldots, k-1$ is not greater than $k^{0.5+\epsilon(k)}t$ where $\epsilon(k) \to 0$ as $k \to \infty$. Furthermore if $T_n \in \mathcal{T}_{\mu,\sigma}^{n\times n}$ we can write t and recall Definition 6.1 to define the cdf of the Gaussian random variable t. This implies a reasonable lower bound on the expected value $|p_1| = |q_n|$.

7.6 The generic corner property, empirical results, and a link of Toeplitz inversion to polynomial computations

Our study in the two previous subsections implies that the norms of the inverses of Gaussian random Toeplitz matrices and consequently their condition numbers are expected to be reasonably bounded provided that the values $|u_n|$ and $|v_1|$ are not small. We call the latter provision the *generic property of two corners of the inverse*. Proving it may be hard. No proof could work if the mean value μ greatly exceeds the standard deviation σ because in this case the norm $||T_n^{-1}||$ is likely to be large (see Section 7.2), which would contradict to the generic property by virtue of Theorems 7.2 and 7.3. Empirically, however, standard Gaussian random Toeplitz matrices of reasonable sizes do tend to be reasonably well conditioned. Our tests in the next section show this directly, whereas the numerical tests in [XXG12] and [PQZ13], provide additional indirect support. Namely these tests succeed by employing random Toeplitz multipliers, whereas they would have been expected to fail numerically if the multipliers were ill conditioned.

We conclude this section by recalling a link of Toeplitz matrix inversion to some polynomial computations. Namely it is well known (see [P01, Section 2.11]) that the equation $T_n\mathbf{p} = \mathbf{e}_1$ is equivalent to the polynomial equation $t(x)p(x) \mod x^{2n+1} = r(x)$ where $r(x)$ is a monic polynomial of degree n, whereas $p(x)$ and $t(x)$ are two polynomials of degrees n and $2n-1$, respectively, with the coefficient vectors $\mathbf{p}$ and $\mathbf{t}_+$, respectively. For a given vector $\mathbf{t}_+$, the Euclidean algorithm computes the coefficients of the polynomials $p(x)$ and $r(x)$ (apart from the cases of degeneracy, occurring with probability 0 for Gaussian random input) but provides no explicit expressions for these coefficients.

8 Numerical Experiments

Our numerical experiments with random general, Hankel, Toeplitz and circulant matrices have been performed in the Graduate Center of the City University of New York on a Dell server with a dual core 1.86 GHz Xeon processor and 2G memory running Windows Server 2003 R2. The test Fortran code was compiled with the GNU gfortran compiler within the Cygwin environment. Random numbers were generated with the random _ number intrinsic Fortran function, assuming the uniform probability distribution over the range $\{x : -1 \leq x < 1\}$. The tests have been designed by the first author and performed by his coauthor.

We have computed the condition numbers of Gaussian random general $n \times n$ matrices for $n = 2^k$, $k = 5, 6, \ldots$, with the entries sampled in the range $[-1, 1)$ as well as complex general, Toeplitz, and circulant matrices whose entries had real and imaginary parts sampled at random in the same range $[-1, 1)$. We performed 100 tests for each class of inputs, each dimension n, and each nullity r. Tables 8.2–8.4 display the test results. The last four columns of each table display the average (mean), minimum, maximum, and standard deviation of the computed condition numbers of the input matrices, respectively. Namely we computed the values $\kappa(A) = ||A||\ ||A^{-1}||$ for general, Toeplitz, and circulant matrices A and the values $\kappa_1(A) = ||A||_1\ ||A^{-1}||_1$ for Toeplitz matrices A. We computed and displayed in Table 8.3 the condition numbers $\kappa_1(A) = ||A||_1||A^{-1}||_1$ by using 1-norms of Toeplitz matrices and their inverses rather than their spectral norms. This shift from 2-norms to 1-norms facilitated the computations in the case of the inputs of large sizes and made no significant impact on the output condition numbers. Indeed already relationships (2.2) link the 1-norms and 2-norms to one another, but the empirical data in Table 8.1 consistently show even much closer links, in all cases of general, Toeplitz, and circulant $n \times n$ matrices A where $n = 32, 64, \ldots, 1024$. Our estimates in Section 7.3 are sharp, and so we were most interested in testing the condition numbers of random Toeplitz matrices. The displayed data show that these condition numbers grow substantially faster with the growth of the matrix size than in the case of random circulant matrices but not faster than in the case of random general matrices, expected to be well conditioned according to the results of the intensive formal and experimental study in [D88], [E88], [ES05], [CD05], [SST06].

Table 8.1: The norms of random general, Toeplitz and circulant $n \times n$ matrices and of their inverses

matrix A	n	$\|\|A\|\|_1$	$\|\|A\|\|_2$	$\frac{\|\|A\|\|_1}{\|\|A\|\|_2}$	$\|\|A^{-1}\|\|_1$	$\|\|A^{-1}\|\|_2$	$\frac{\|\|A^{-1}\|\|_1}{\|\|A^{-1}\|\|_2}$
General	32	1.9×10^1	1.8×10^1	1.0×10^0	4.0×10^2	2.1×10^2	1.9×10^0
General	64	3.7×10^1	3.7×10^1	1.0×10^0	1.2×10^2	6.2×10^1	2.0×10^0
General	128	7.2×10^1	7.4×10^1	9.8×10^{-1}	3.7×10^2	1.8×10^2	2.1×10^0
General	256	1.4×10^2	1.5×10^2	9.5×10^{-1}	5.4×10^2	2.5×10^2	2.2×10^0
General	512	2.8×10^2	3.0×10^2	9.3×10^{-1}	1.0×10^3	4.1×10^2	2.5×10^0
General	1024	5.4×10^2	5.9×10^2	9.2×10^{-1}	1.1×10^3	4.0×10^2	2.7×10^0
Toeplitz	32	1.8×10^1	1.9×10^1	9.5×10^{-1}	2.2×10^1	1.3×10^1	1.7×10^0
Toeplitz	64	3.4×10^1	3.7×10^1	9.3×10^{-1}	4.6×10^1	2.4×10^1	2.0×10^0
Toeplitz	128	6.8×10^1	7.4×10^1	9.1×10^{-1}	1.0×10^2	4.6×10^1	2.2×10^0
Toeplitz	256	1.3×10^2	1.5×10^2	9.0×10^{-1}	5.7×10^2	2.5×10^2	2.3×10^0
Toeplitz	512	2.6×10^2	3.0×10^2	8.9×10^{-1}	6.9×10^2	2.6×10^2	2.6×10^0
Toeplitz	1024	5.2×10^2	5.9×10^2	8.8×10^{-1}	3.4×10^2	1.4×10^2	2.4×10^0
Circulant	32	1.6×10^1	1.8×10^1	8.7×10^{-1}	9.3×10^0	1.0×10^1	9.2×10^{-1}
Circulant	64	3.2×10^1	3.7×10^1	8.7×10^{-1}	5.8×10^0	6.8×10^0	8.6×10^{-1}
Circulant	128	6.4×10^1	7.4×10^1	8.6×10^{-1}	4.9×10^0	5.7×10^0	8.5×10^{-1}
Circulant	256	1.3×10^2	1.5×10^2	8.7×10^{-1}	4.7×10^0	5.6×10^0	8.4×10^{-1}
Circulant	512	2.6×10^2	3.0×10^2	8.7×10^{-1}	4.5×10^0	5.4×10^0	8.3×10^{-1}
Circulant	1024	5.1×10^2	5.9×10^2	8.7×10^{-1}	5.5×10^0	6.6×10^0	8.3×10^{-1}

Table 8.2: The condition numbers $\kappa(A)$ of random $n \times n$ matrices A

n	**input**	**min**	**max**	**mean**	**std**
32	real	2.4×10^1	1.8×10^3	2.4×10^2	3.3×10^2
64	real	4.6×10^1	1.1×10^4	5.0×10^2	1.1×10^3
128	real	1.0×10^2	2.7×10^4	1.1×10^3	3.0×10^3
256	real	2.4×10^2	8.4×10^4	3.7×10^3	9.7×10^3
512	real	3.9×10^2	7.4×10^5	1.8×10^4	8.5×10^4
1024	real	8.8×10^2	2.3×10^5	8.8×10^3	2.4×10^4
2048	real	2.1×10^3	2.0×10^5	1.8×10^4	3.2×10^4

Table 8.3: The condition numbers $\kappa_1(A) = ||A||_1 \; ||A^{-1}||_1$ of random Toeplitz $n \times n$ matrices A

n	**min**	**mean**	**max**	**std**
256	9.1×10^2	9.2×10^3	1.3×10^5	1.8×10^4
512	2.3×10^3	3.0×10^4	2.4×10^5	4.9×10^4
1024	5.6×10^3	7.0×10^4	1.8×10^6	2.0×10^5
2048	1.7×10^4	1.8×10^5	4.2×10^6	5.4×10^5
4096	4.3×10^4	2.7×10^5	1.9×10^6	3.4×10^5
8192	8.8×10^4	1.2×10^6	1.3×10^7	2.2×10^6

Table 8.4: The condition numbers $\kappa(A)$ of random circulant $n \times n$ matrices A

n	**min**	**mean**	**max**	**std**
256	9.6×10^0	1.1×10^2	3.5×10^3	4.0×10^2
512	1.4×10^1	8.5×10^1	1.1×10^3	1.3×10^2
1024	1.9×10^1	1.0×10^2	5.9×10^2	8.6×10^1
2048	4.2×10^1	1.4×10^2	5.7×10^2	1.0×10^2
4096	6.0×10^1	2.6×10^2	3.5×10^3	4.2×10^2
8192	9.5×10^1	3.0×10^2	1.5×10^3	2.5×10^2
16384	1.2×10^2	4.2×10^2	3.6×10^3	4.5×10^2
32768	2.3×10^2	7.5×10^2	5.6×10^3	7.1×10^2
65536	2.4×10^2	1.0×10^3	1.2×10^4	1.3×10^3
131072	3.9×10^2	1.4×10^3	5.5×10^3	9.0×10^2
262144	6.3×10^2	3.7×10^3	1.1×10^5	1.1×10^4
524288	8.0×10^2	3.2×10^3	3.1×10^4	3.7×10^3
1048576	1.2×10^3	4.8×10^3	3.1×10^4	5.1×10^3

Appendix

A Randomness and nonsingularity

The total degree of a multivariate monomial is the sum of its degrees in all its variables. The total degree of a polynomial is the maximal total degree of its monomials.

Lemma A.1. *[DL78], [S80], [Z79]. For a set Δ of a cardinality $|\Delta|$ in any fixed ring let a polynomial in m variables have a total degree d and let it not vanish identically on this set. Then the polynomial vanishes in at most $d|\Delta|^{m-1}$ points.*

We assume that Gaussian random variables range over infinite sets Δ, usually over the real line or its interval. Then the lemma implies that a nonzero polynomial vanishes with probability 0. Consequently a square Gaussian random general, Toeplitz or circulant matrix is nonsingular with probability 1 because its determinant is a polynomials in the entries. Likewise rectangular Gaussian random general, Toeplitz and circulant matrices have full rank with probability 1, and similarly under the other probability distributions whose measures are absolutely continuous relatively to Lebesgue's measure. These results can be also adapted to the case of probability distribution over finite sets [DL78], [S80], [Z79].

Acknowledgements: Our research has been supported by NSF Grant CCF–1116736 and PSC CUNY Awards 64512–0042 and 65792–0043.

References

[BG05] A. Böttcher, S. M. Grudsky, *Spectral Properties of Banded Toeplitz Matrices*, SIAM Publications, Philadelphia, 2005.

[CD05] Z. Chen, J. J. Dongarra, Condition Numbers of Gaussian Random Matrices, *SIAM. J. on Matrix Analysis and Applications*, **27**, 603–620, 2005.

[CPW74] R. E. Cline, R. J. Plemmons, and G. Worm, Generalized Inverses of Certain Toeplitz Matrices, *Linear Algebra and Its Applications,* **8**, 25–33, 1974.

[D88] J. Demmel, The Probability That a Numerical Analysis Problem Is Difficult, *Math. of Computation*, **50**, 449–480, 1988.

[DL78] R. A. Demillo, R. J. Lipton, A Probabilistic Remark on Algebraic Program Testing, *Information Processing Letters*, **7**, **4**, 193–195, 1978.

[E88] A. Edelman, Eigenvalues and Condition Numbers of Random Matrices, *SIAM J. on Matrix Analysis and Applications*, **9**, **4**, 543–560, 1988.

[ES05] A. Edelman, B. D. Sutton, Tails of Condition Number Distributions, *SIAM J. on Matrix Analysis and Applications*, **27**, **2**, 547–560, 2005.

[GK72] I. Gohberg, N. Y. Krupnick, A Formula for the Inversion of Finite Toeplitz Matrices, *Matematicheskiie Issledovaniia* (in Russian), **7**, **2**, 272–283, 1972.

[GL96] G. H. Golub, C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, Maryland, 1996 (third addition).

[GS72] I. Gohberg, A. Sementsul, On the Inversion of Finite Toeplitz Matrices and Their Continuous Analogs, *Matematicheskiie Issledovaniia* (in Russian), **7**, **2**, 187–224, 1972.

[HMT11] N. Halko, P. G. Martinsson, J. A. Tropp, Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions, *SIAM Review*, **53, 2**, 217–288, 2011.

[P01] V. Y. Pan, *Structured Matrices and Polynomials: Unified Superfast Algorithms*, Birkhäuser/Springer, Boston/New York, 2001.

[PQZ13] V. Y. Pan, G. Qian, A. Zheng, Randomized Preprocessing versus Pivoting, *Linear Algebra and Its Applications*, **438, 4**, 1883–1899, 2013.

[S80] J. T. Schwartz, Fast Probabilistic Algorithms for Verification of Polynomial Identities, *Journal of ACM*, **27**, **4**, 701–717, 1980.

[S98] G. W. Stewart, *Matrix Algorithms, Vol I: Basic Decompositions*, SIAM, 1998.

[SST06] A. Sankar, D. Spielman, S.-H. Teng, Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices, *SIAM J. on Matrix Analysis*, **28**, **2**, 446–476, 2006.

[XXG12] J. Xia, Y. Xi, M. Gu, A Superfast Structured Solver for Toeplitz Linear Systems via Randomized Sampling, *SIAM J. Matrix Anal. Appl.*, **33**, 837–858, 2012.

[Z79] R. E. Zippel, Probabilistic Algorithms for Sparse Polynomials, *Proceedings of EUROSAM'79, Lecture Notes in Computer Science*, **72**, 216–226, Springer, Berlin, 1979.