Journal of Computational and Applied Mathematics 22 (1988) 219–230
North-Holland

Complexity theory of numerical linear algebra

Eric KOSTLAN

Received 1 April 1987

Abstract: In this paper the statistical properties of problems that occur in numerical linear algebra are studied. Bounds are calculated for the average preformance of the power method for the calculation of eigenvectors of symmetric and Hermitian matrices, thus an upper bound is found for the average complexity of eigenvector calculation. The condition number of a matrix is studied and sharp bounds are calculated for the average (and the variance) loss of precision encountered when one solves a system of linear equations.

Keywords: Complexity, numerical analysis, linear algebra, large systems, random matrices, Gaussian ensembles, probability densities, power method, condition number.

Introduction

In 1937 and 1940, Goldstine and Von Neumann published a paper in two parts on the numerical problems involved in solving large systems of equations. Only the second part discussed statistical issues, and this analysis is rather incomplete (Part I of their work is 78 pages, whereas part II is only 14 pages.) Here we explore in depth the statistical properties of random matrices and random systems of linear equations. We will rely heavily on information about the joint probability distribution of the eigenvalues and singular values of random matrices and will develope tools (really one tool) to estimate integrals against these densities. Once we develope these tools we use them to calculate the average preformance of the 'power method' of numerical linear algebra. Although the power method is not used in unmodified form for the calculation of eigenvectors, it is in some sense the prototype of RQI and of the QR algorithms that are in use. Also the power method serves as a simple example of a generally convergent iterative algorithm and at least gives us some upper bound on the complexity of eigenvector calculation. Although the (population) average number of iterations required by the power method (to give an ϵ-eigenvector) is infinite, we show that we can expect sample averages to be small. Our methods of estimation work as well for singular values as they do for eigenvalues, and thus we can estimate the distribution function of the condition number of a random matrix. The condition number has well known connections with error estimates and we use our estimates of the probability density for the condition number to give statistical error estimates for linear problems.

1. Random matrices

We consider four cases: $S(n)$, the set of all $n \times n$ real symmetric matrices, $H(n)$, the set of all $n \times n$ complex Hermitian matrices, $A(n)$, the set of all $n \times n$ real matrices.

We will concentrate on the following two-parameter family of probability densities:

$$\rho(M) = C \exp\left\{ -\frac{1}{2\sigma^2} \operatorname{tr}\left[(M - \mu I)^{\dagger} (M - \mu I) \right] \right\} \tag{1.1}$$

where $\dagger$ denotes the Hermitian conjugation (i.e. is the identity for $S(n)$ and $H(n)$), and the normalization constant C is given by

$$C^{-1} = \begin{cases} (\sqrt{2\pi}\,\sigma)^{n(n+1)/2} 2^{n(1-n)/2} & \text{for } S(n), \\ (\sqrt{2\pi}\,\sigma)^{n^2} 2^{n(1-n)} & \text{for } H(n), \\ (\sqrt{2\pi}\,\sigma)^{-2n^2} & \text{for } A(n). \end{cases}$$

σ is assumed to be positive; μ may be any real number except for the case $A(n)$, in which μ may be any complex number.

We adopt the terminology of Mehta and call the resulting random variables on $S(n)$, $H(n)$, and $A(n)$ the 'Gaussian ensembles'.

The primary reason for restricting our attention to these densities is that important information concerning the statistical behavior of the eigenvalues, the spacings of the eigenvalues, the singular values, etc., is well understood for these densities [4]. Furthermore, these densities seem like natural ones to be singled out, as is suggested by the following two alternate characterizations of the Gaussian ensembles.

Note. By "the Gaussian variable $N(\mu, \sigma^2)$" we will mean the real-valued random variable with density

$$\rho(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2\sigma^2} (x - \mu)^2 \right]$$

where σ is the standard deviation, μ is the mean, or we will mean the complex-valued random variable whose real and complex parts are independent Gaussian variables with the same (real, positive) standard deviation σ, and whose means are respectively the real and imaginary parts of μ (which is now allowed to be any complex number), depending on the context.

I. The Gaussian ensembles on $S(n)$, $H(n)$, and $A(n)$ are characterized by the following properties:

(A) The entries of the matrices are independent Gaussian variables.

(B) The means of the off-diagonal elements are all equal to zero.

(C) The means of the diagonal elements are all equal.

(D) For $A(n)$ the standard deviation of all the elements are equal; for $S(n)$ and $H(n)$ the variance of any diagonal entry is twice the variance of any off-diagonal entry. (Using the notation of equation (1.1), the density of any diagonal element is $N(\mu, \sigma^2)$).

Proof. This is a trivial calculation using (1.1). □

II. The Gaussian ensembles on $S(n)$, $H(n)$, and $A(n)$ are characterized by
(A) statistical independence of the entries of the matrices,
(B) statistical independence of the real and imaginary parts of each entry, and
(C) invariance of the density under conjugation by elements of the orthogonal group $O(n)$, in the case of $S(n)$, or by elements of the unitary group $U(n)$, in the cases of $H(n)$ and $A(n)$.

Proof. See [4]. □

2. The probability density of the eigenvalues

We now summarize the essential statistical information for the eigenvalues of the Gaussian ensembles [4]. We parameterize the Gaussian ensembles as in equation (1.1) of the last section. The main result for the cases of real symmetric or Hermitian matrices is the following theorem.

Theorem 2.1. *The joint probability density for the eigenvalues,* $\lambda_1, \ldots, \lambda_n$, *of the Gaussian ensembles* $S(n)$ *and* $H(n)$ *is given by*

$$B\left\{\prod_{i<j} |\lambda_i - \lambda_j|^r\right\} \exp\left[-\frac{r}{2\sigma^2}\sum_{i=1}^{n}(\lambda_i - \mu)^2\right] \tag{2.1}$$

where $r = 1$ *for* $S(n)$ *and* $r = 2$ *for* $H(n)$. *The normalization constant* B *is given by*

$$B^{-1} = \sigma^{(rn(n+1)/2)+n}(2\pi)^{n/2} r^{-n^2/2}\left[\Gamma\left(1+\tfrac{1}{2}r\right)\right]^{-n}\prod_{j=1}^{n}\Gamma\left(1+\tfrac{1}{2}rj\right).$$

Note. Mehta [4] writes $r^{-n/2-(rn(n-1)/4)}$ instead of $r^{-n^2/2}$ which makes his formula correct for his "symplectic ensembles" when $r = 4$.

For Gaussian ensembles of arbitrary complex-valued matrices we have the following theorem.

Theorem 2.2. *The joint probability density for the eigenvalues,* $\lambda_1, \ldots, \lambda_n$, *of the Gaussian ensembles* $A(n)$ *and* $H(n)$ *is given by*

$$B\left\{\prod_{i<j} |\lambda_i - \lambda_j|^2\right\} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}|\lambda_i - \mu|^2\right] \tag{2.2}$$

where the normalization constant B *is given by*

$$B^{-1} = \sigma^{n(n+3)}\pi^n \sum_{j=1}^{n}\Gamma(1+j).$$

3. Main computational lemma

We see that in our analysis we are faced with the problem of making calculations involving complicated integrals. The following results are elementary, but they are quite useful.

Lemma 3.1. *Let* $\psi(x)$ *denote the Euler psi function* $\mathrm{d}(\ln \Gamma(x))/\mathrm{d}x$. *Then we have the following*:

$$[x > 0] \qquad \ln(x) - 1/x < \psi(x) < \ln(x) - 1/2x$$

$$[x \neq 0, -1, \ldots] \qquad \psi(x+1) = \psi(x) + 1/x; \quad \psi(1) = -\gamma; \quad \psi(\tfrac{1}{2}) = -\gamma - 2\ln 2$$

$$[x \neq 0, -1, \ldots] \qquad \psi'(x+1) = \psi(x) - 1/x^2; \quad \psi'(1) = \tfrac{1}{6}\pi^2; \quad \psi'(\tfrac{1}{2}) = \tfrac{1}{2}\pi^2$$

where $\gamma = 0.577215\ldots$ *is Euler's constant.*

Proof. The inequality follows immediately from the exact formula

$$[\mathrm{Re}(x) > 0] \quad \psi(x) = \ln(x) - \frac{1}{2x} - 2\int_0^\infty \frac{t\,\mathrm{d}t}{(t^2 + x^2)(\mathrm{e}^{2\pi t} - 1)}.$$

See, for example, [3]. □

Lemma 3.2. *The gamma function satisfies the following inequalities whenever a and r are positive*

$$\left(\frac{a}{\mathrm{e}^{1/a}}\right)^r < \frac{\Gamma(a+r)}{\Gamma(a)} < \left(\frac{a+r}{\mathrm{e}^{1/2(a+r)}}\right)^r$$

Proof. This follows by applying the mean value theorem to the logarithm of the gamma function, and by using the preceding lemma. □

Lemma 3.3. *If P is a polynomial with all real roots and with no roots in the interval* (0, 1), *if f is a monotone increasing function and if* $c \in (0, 1)$ *then*

$$L_P(f, c) = \frac{\int_0^c f(t)P(t)\,\mathrm{d}t}{\int_0^c P(t)\,\mathrm{d}t} \leqslant \frac{\int_0^c f(t)t^n\,\mathrm{d}t}{\int_0^c t^n\,\mathrm{d}t}.$$

Proof. We show that for all such P, L_P is maximized for $P = t^{\deg P}$. First, the maximizing P cannot have roots greater than or equal to one, for omitting the factors corresponding to the roots will increase L_P and decrease the degree of P, which is absurd. Thus the maximizing P has all negative roots and thus all positive coefficients. Thus L_P is a convex combination of $L_{t^1}, L_{t^2}, \ldots, L_{t^{\deg P}}$, and since L_{t^i} an increasing function in i, we are done. □

Lemma 3.4. $\mathrm{Vol}\ S^{n-1} = n\ \mathrm{Vol}\ B^n = 2\pi^{n/2}/\Gamma(\frac{1}{2}n)$.

See, for example, [6].

Lemma 3.5. *If* π: $R^n \to R^k$ *is an orthogonal projection, then we have the following equalities*

$$\operatorname*{Av}_{x \in S^{n-1}} (-\ln \| \pi x \|) = \operatorname*{Av}_{x \in B^n} (\ln \| x \| - \ln \| \pi x \|) = \tfrac{1}{2}\left[\psi\left(\tfrac{1}{2}n\right) - \psi\left(\tfrac{1}{2}k\right)\right],$$

$$\operatorname*{Av}_{x \in S^{n-1}} (-\ln \| \pi x \|)^2 = \tfrac{1}{4}\left[\psi\left(\tfrac{1}{2}n\right) - \psi\left(\tfrac{1}{2}k\right)\right]^2 + \psi\left(\tfrac{1}{2}k\right) - \psi\left(\tfrac{1}{2}n\right).$$

Proof. The first equality follows from the homogeneity of the middle term; the second and third equalities reduce to integrals of one variable that can be found in [3, formulas 4.245, 4.261.17]. □

Lemma 3.6. *If* π: $R^n \to R^k$ *is an orthogonal projection, then we have the following equality*

$$\operatorname*{Av}_{x \in S^{n-1}} (\| \pi x \|^r) = \frac{\Gamma\left[\frac{1}{2}(k+r)\right]\Gamma\left[\frac{1}{2}n\right]}{\Gamma\left[\frac{1}{2}(n+r)\right]\Gamma\left[\frac{1}{2}k\right]}.$$

Proof. This follows immediately from the definition of the beta function and its well-known relationship to the gamma function, and from Lemma 3.4. □

Lemma 3.7. *Let* D *be a convex, compact region in* R^n (*eigenvalue space*), *and give* R^n *a density given in* (2.1) *with the parameter* μ *equal to* 0. *Lt* g: $D \to R^1$ *be measurable and homogeneous. Assume that for all* $x \in D$ *and for all* i, j, $x_i \neq x_j$. *Let* $\Lambda = \{ x \in R^n \colon L(x) = 0 \}$ *be the hyperplane defined by a positive linear functional* L, *and assume* $\Lambda \cap D$ *is empty; let* $\delta = \max_{x \in D} L(x)$, *and choose* $y \in D$ *such that* $L(y) = \delta$. *Let* f: $R^1 \to R^1$ *be a monotonically decreasing function such that for all* $x \in D$ $fL(x) \geqslant g(x)$. *We then have the following inequality*:

$$\operatorname*{Av}_{x \in D} g(x) \leqslant \left[n(n-1)/\beta + n\right] \int_0^1 f(\delta(1-t)) t^{n(n-1)/\beta + n - 1} \, \mathrm{d}t.$$

Here $\beta = 2$ *for the real symmetric case and* $\beta = 1$ *for the complex Hermitian case.*

Proof. Since $g(x)$ is homogeneous in x, we can simplify the measures given in (2.1) by omitting the exponential term and renormalizing. Now D is convex so it may be decomposed into segments starting at the point y. It suffices to establish the lemma for an arbitrarily small 'cone' of such segments surrounding each segment l. Use the linear parameterization of l, T: $(0, c) \to R^1$ defined by the properties $T(0) = y$ and $L(T(t)) = \delta(1-t)$. Note that $c \in [0, 1]$. We let Π: $R^n \to R^1$ be the function $\Pi_{i>j} |x_i - x_j|$. The proof of lemma reduces to analyzing the quotient

$$\frac{\int_0^c fLT(t)\Pi(T(t)) t^{n-1} \, \mathrm{d}t}{\int_0^c \Pi(T(t)) t^{n-1} \, \mathrm{d}t}.$$

Observe that ΠT is a polynomial of degree $\frac{1}{2}n(n-1)$ such that all of its roots are real and none of its root lies in (0, 1). Its roots occur exactly where $x_i = x_j$. So we can apply Lemma 3.3 and we are done.

4. Some results on the complexity theory of the power method

We use the statistical information about the eigenvalues of matrices to derive information about the complexity of the power method of numerical linear algebra. By the power method we mean the induced action of a linear automorphism of some vector space on the associated projective space. Choosing an arbitrary point p in projective space (the 'starting point') we apply this induced map to p repeatedly, thus producing a sequence of points in projective space. For 'almost all' matrices, this sequence will converge to an eigenvalue of the matrix. This justifies the term 'power method' and also suggests the sort of problems to consider:

(I) What is the average performance of the power method, that is, how quickly the sequences of points described above converge on the average?

(II) How often will we run into 'problem matrices' for which the power method performs poorly?

(III) How important is the choice of starting points; can the average performance of the power method be improved substantially by using more than one starting point?

We start with the simplest case: the real symmetric matrices. It will be seen that the available statistical information concerning the eigenvalues of real symmetric matrices will suffice to reduce many questions of the sort posed above (once restated precisely) to problems of integral calculus. The resulting calculus problems are by no means trivial, but crude bounds on complexity can be obtained without too much difficulty. The properties of the Hopf fibration allow us to 'lift' the analysis of the Hermitian case from $CP(n-1)$ to $RP(2n-1)$ and thus the analysis of Hermitian matrix is reduced to the analysis of a degenerate real symmetric matrix. Therefore, we have no difficulty modifying the results given for the real symmetric case to give results for the Hermitian case, which are stated without proof. The case of non-Hermitian matrices is subtler and even the questions asked in the Hermitian cases cannot be directly generalized. We therefore will treat the question of non-Hermitian matrices separately.

To start we choose a positive integer $n > 1$ and we consider an arbitrary nondegenerate $n \times n$ real symmetric matrix M. We use the standard metric on $RP(n-1)$ (the metric induced from the unit sphere in Euclidean n-space by the quotient map), and as a probability measure on $RP(n-1)$ we use the renormalized Riemannian volume. We use the convention of Shub, and define an ϵ-eigenvector of M as a point in projective space whose Riemannian distance from an eigenvector of M is $\leqslant \epsilon$. By a dominant ϵ-eigenvector we mean an ϵ-eigenvector corresponding to a dominant eigenvector, i.e. an eigenvector whose eigenvalue has magnitude greater than or equal to the magnitudes of all the other eigenvalues.

We start with the case where the eigenvalues of M have only two distinct non-zero absolute values. R^n decomposes into the orthogonal direct sum of three M-invariant subspaces R_1 and R_2. R_1 is spanned by eigenvalues of absolute value $|\lambda_1|$, R_2 is spanned by eigenvalues of absolute value $|\lambda_2|$, and R_0, the kernel of M. Assume, without loss of generality, that $|\lambda_1| > |\lambda_s|$. Let π_1, π_2 and π_0 be the projections from R^n onto R_1, R_2 and R_0 respectively. We let J be the canonical projection $J\colon R^n \to RP(n-1)$. Let N_ϵ be the ϵ-tubular neighborhood of $J(R_1)$ in $RP(n-1)$ and let W_ϵ be the inverse image of N_ϵ under the canonical projection $J\colon R^n \to RP(n-1)$; equivalently,

$$W_\epsilon = \left\{ x \in R^n \colon \frac{\| \pi_1 x \|}{\sqrt{\| \pi_2 x \|^2 + \| \pi_0 x \|^2}} \geqslant \cot \epsilon \right\}.$$

We let $N_\epsilon(s)$ denote the set of points in $RP(n-1)$ that after s steps of the power method (i.e. after s iterated applications of M to $RP(n-1)$) yield points in N_ϵ. Equivalently for $s>0$ we have,

$$J^{-1}N_\epsilon(s) = \left\{ x \in R^n : \frac{|\lambda_1|^s \|\pi_1 x\|}{|\lambda_2|^s \|\pi_2 x\|} \geqslant \cot \epsilon \right\}$$

$$= \left\{ x \in R^n : s \geqslant \frac{\ln \|\pi_2 x\| - \ln \|\pi_1 x\| + \ln \cot \epsilon}{\ln |\lambda_1| - \ln |\lambda_2|} \right\}.$$

This last formula has the nice feature that the numerator depends only the starting point and ϵ, whereas the denominator depends only on the eigenvalues of M. Thus the numerator and denominator can be treated separately. Here we state a reasonable bound for the integral of the numerator over $RP(n-1) - N_\epsilon$ (k and l are the dimensions of R_1 and R_2 respectively.).

Lemma 4.1.

$$\tfrac{1}{2}\left[\psi\left(\tfrac{1}{2}l\right) - \psi\left(\tfrac{1}{2}k\right)\right] + \ln \cot \epsilon < \int_{x \in RP(n-1)-N_\epsilon} \left[\ln \|\pi_2 x\| - \ln \|\pi_1 x\| + \ln \cot \epsilon\right]$$

$$< \tfrac{1}{2}\left[\psi\left(\tfrac{1}{2}n\right) - \psi\left(\tfrac{1}{2}k\right)\right] + \max(\ln \cot \epsilon, 0).$$

Proof. Lemma (3.5) gives us integrals over $RP(n-1)$ instead of $RP(n-1) - N_\epsilon$. Since the integrand is negative on N_ϵ we immediately get the first inequality, but for an upper bound we must replace the integrand with a larger function that is positive on N_ϵ before integrating over all of $RP(n-1)$. Here we have replaced the integrand with $[-\ln \|\pi_1\| + \max(\ln \cot \epsilon, 0)]$ and have applied Lemma 3.5. □

We now concentrate on the case $k=1$. (The case $k=2$ is used for the analysis of Hermitian matrices.). For any such M we let $\rho_\epsilon(M)$ denote the number of steps required by the power method to give a dominant ϵ-eigenvector, averaged over all possible starting points. In the notation above we have:

$$\rho_\epsilon(M) = \frac{1}{\text{Vol } RP(n-1)} \sum_{s=1}^{\infty} s \text{ Vol}\left[N_t(s) - N_\epsilon(s-1)\right].$$

The above lemma with $k=1$ gives us sharp bounds for $(\ln |\lambda_1| - \ln |\lambda_2|)\rho_\epsilon(M)$ if M has the simple form assumed in the theorem.

To extend the above result to an arbitrary real symmetric matrix M we will compare M to two matrices M_{bad} and M_{good}. These are defined as follows: first diagonalize M so that the eigenvalues appear in decreasing magnitude; that is,

$$M = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}, \quad \text{where } |\lambda_1| > |\lambda_2| > \cdots > |\lambda_0|;$$

in this frame of reference we define

$$M_{\text{bad}} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_2 \end{pmatrix}; \qquad M_{\text{good}} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_n & & \\ & & \ddots & \\ & & & 0 \end{pmatrix}.$$

We have at once $\rho_\epsilon(M_{\text{good}}) \leqslant \rho_\epsilon(M) \leqslant \rho_\epsilon(M_{\text{bad}})$ and thus by the estimates above we have for example the following theorem.

Theorem 4.2.

$$\frac{\ln \cot \epsilon}{\ln |\lambda_1| - \ln |\lambda_2|} < \rho_\epsilon(M) < \frac{\frac{1}{2}[\psi(n/\beta) - \psi(1/\beta)] + \max(\ln \cot \epsilon, 0)}{\ln |\lambda_1| - \ln |\lambda_2|} + 1.$$

Where $\beta = 2$ for M real symmetric and $\beta = 1$ for M complex Hermitian.

As rough as these estimates are, they immediately give us several results, including:

Theorem 4.3. *The number of iterations required by the power methods to give a dominant ϵ-eigenvector, when averaged over all starting points and over all real symmetric matrices (using a Gaussian ensemble) is infinite.*

Proof. If $\epsilon < \frac{1}{4}\pi$ we simply inspect the integral of the denominator of the left side of (4.1) with respect to the measure given by (2.1). A singularity occurs at the hyperplane (in eigenvalue space) given by $\lambda_1 = -\lambda_2$ (but *not* at the hyperplane given by $\lambda_1 = \lambda_2$). If $\epsilon \geqslant \frac{1}{4}\pi$ the argument must be modified slightly, but this is only because the lemma above has been adapted for small ϵ. □

In particular, increasing the number of starting points cannot make the average number of iterations required by the power method finite (unless the number of starting points depends on ϵ). These infinite average results, when modified in the obvious way, are true for Hermitian Gaussian ensembles as well.

These results are interesting because except for a set of measure zero the number of iterations required by the power method is always finite and usually quite small. To make this precise we introduce the notion of a 'generalized average' that we will denote as $\rho_{\epsilon,\eta}(\mu, \sigma)$ and can be thought of as the average number of iterations required to give a dominant ϵ-eigenvector if we are allowed to ignore a set whose normalized measure in $RP(n-1) \times S(n)$ (resp. $CP(n-1) \times H(n)$) is equal to η. This notion has been used extensively by Smale.

We now concentrate on $\mu = 0$ Gaussian Ensembles.

Theorem 4.4. *For the Gaussian ensemble with parameters $(0, \sigma^2)$ the generalized average satisfies*

$$\rho_{\epsilon,\eta}(0, \sigma) < \frac{3}{2\eta\sqrt{2}}\{n(n-1)/\beta + n\}[\psi(n/\beta) - \psi(1/\beta) + 2\max(\ln \cot \epsilon, 0)] + 1.$$

Where $\beta = 2$ for the real symmetric case and $\beta = 1$ for the complex Hermitian case.

Proof. We will use Lemma 3.7. We approximate $Z = \{x \in B^n: |x_1| = |x_{II}|\}$ (where $|x_I| \geqslant |x_{II}| \geqslant \cdots$) by $\{x_1 = x_2$ or $x_1 = -x_n$ or $x_{n-1} = x_n$ where $x_1 > \cdots > x_n\}$. The homogeneity of the problem allows us to restrict our attention to the unit ball in eigenvalue space minus some hyperplane. Now apply the preceding lemma with the function g of the form $\chi(X)$ where $X = \{x: 1/(1-\omega) \geqslant |\pm x_i/x_j| \geqslant 1-\omega\}$ (Note that outside X $|\ln|\lambda_i| - \ln|\lambda_j|| > \omega$), and $L(x)$ of the form $d(x,$ hyperplane). For the function f we use $\chi\{t \leqslant \alpha\}$, where we let $\alpha = \sqrt{\frac{1}{2}}\,\omega/\sqrt{(1-\omega)^2+1}$. The lemma then gives us the following estimate:

$$\overline{\mathrm{Vol}(\omega)} < \left[\tfrac{1}{2}n(n-1)+n\right]\omega/\sqrt{2}\,.$$

where $\overline{\mathrm{Vol}}$ is the normalized volume. Combining this result with Theorem 4.2 yields the desired result. □

The case of positive definite matrices yields a finite result:

Theorem 4.5. *If we restrict the density given by a $\mu = 0$ Gaussian ensemble to the set of positive (or negative) definite matrices (and renormalize), average number of iterations required by the power method to give a dominant ϵ-eigenvector is less than*

$$\tfrac{1}{2}\{n(n-1)/\beta + n\}[\psi(n/\beta) - \psi(1/\beta) + 2\max(\ln\cot\epsilon, 0)] + 1$$

where $\beta = 2$ for the real symmetric case and $\beta = 1$ for the complex Hermitian case.

Proof. We apply an argument similar to the one used in the previous theorem, but first we must replace Π with $\Pi^0 = \Pi/(\lambda_1 - \lambda_2)$ where $\lambda_1 > \lambda_2 > \cdots$. Observe that:

$$\frac{1}{\ln(\lambda_1) - \ln(\lambda_2)} < \frac{\lambda_1}{\lambda_1 - \lambda_2}\,.$$

Thus we can use the reasoning used in the last problem to analyze each integral that arises here. □

We see that although the average complexity of the power method is infinite, the power method is in some sense fast on the average.

5. Condition numbers

The condition number of a matrix is more closely related to the questions of solving linear equations than are the eigenvalues. We state results analogous to the results in Section 2 for the condition number of an arbitrary real or complex $n \times n$ matrices. Since the condition numbers are less well-known than the eigenvalues, we give here five definitions, which are equivalent if the matrix is nonsingular. M is an arbitrary $n \times n$ complex valued matrix and the condition number is denoted by C. We let Σ denote the set of singular matrices and let $\|\ \|$ denote the operator norm.

(I) $C = \|M\| \times \|M^{-1}\|$.

(II) $C = \|M\|/\mathrm{d}(M, \Sigma)$ where d is the operator distance.

(III) $C = \| M \| / \mathrm{d}(M, \Sigma)$ where d is the Euclidean distance in $C^{n \times n}$.
(IV) Let $M = UH$ where U is unitary and H is positive definite Hermitian; then C is equal to the ratio of the largest and smallest eigenvalue of H.
(V) Let $Y = M^{\dagger}M$, where $\dagger$ denotes Hermitian conjugation, then C is equal to the square root of the ratio of the largest and smallest eigenvalue of Y.

The condition number is of interest in numerical linear algebra because it bounds the ratio of relative error of a vector x to the relative error of b where $b = Mx$ (see Section 6). Another interesting quantity to study is $\ln(C)$ which bounds the loss in precision involved in solving the system [1,9]. We let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of H (or equivalently the square root s of the eigenvalues of Y). We now concentrate on the case of real matrices.

Theorem 5.1. *The space $A_R(n)$ with the probability density given* (1.1) *induces the joint probability density on $\lambda_1, \ldots, \lambda_n$ given by*

$$B\left\{\prod_{i<j}\left|\lambda_i^2 - \lambda_j^2\right|\right\} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\lambda_i^2\right]$$

when the parameter μ is 0 (see, for example, [10]; *see also* [2]).

As in the case of the above study of the power method, the knowledge of this density immediately gives us several results; the proofs of these results are similar to proofs of theorems given in Section 4. and are therefore omitted.

Theorem 5.2. *The average condition number for real $n \times n$ matrices with the density given by* (1.1) *with parameters $(0, \sigma^2)$ is infinite.*

Theorem 5.3. *For all η there exists a region V of $A_R(n)$ of volume less than η in the Gaussian ensemble with parameters $(0, \sigma^2)$ such that the condition number for all $M \in A_R(n) - V$ is less than $n^2\sqrt{n}/\eta$.*

What we have here is an estimate of the distribution function of the condition number on $A_R(n)$ (with the given density). We let $L(n)$ denote the average, over a Gaussian ensemble on $A_R(n)$ with parameters $(0, \sigma^2)$, of the natural logarithm of the condition number (this is, of course, independent of σ^2).

Theorem 5.4. *The average log-condition number $L(n)$ for $A_R(n)$ with the Gaussian ensemble with parameters $(0, \sigma^2)$ is less than*

$$\tfrac{5}{2}\ln(n) + 1.$$

Proof. We simply integrate the log of the bound given in the previous theorem against η.

Adrian Ocneanu had previously calculated a logarithmic bound for $L(n)$ for $\mu = 0$ real Gaussian ensembles:

Theorem 5.5. (Ocneanu). *The average log-condition number* $L(n)$ *for* $A_R(n)$ *with the Gaussian ensemble with parameters* $(0, \sigma^2)$ *satisfies*

$$\left(\tfrac{2}{3} - \epsilon\right) \ln(n) < L(n) < (3 + \epsilon) \ln(n)$$

where ϵ *can be made as small as desired by setting a lower bound on* n.

6. Random linear systems

We now look more closely at the average loss of accuracy or precision in solving a linear system of equations. For an $n \times n$ invertible complex [resp. real] matrix M consider the equation

$$M(x + \Delta x) = b + \Delta b$$

where b and Δb are independent random vectors in C^n. We will assume that these random vectors have densities invariant under the usual action of the unitary group on C^n (for example we could use standard Gaussian variables as the components of b and Δb). Here x represents the solution to the equation

$$M(x) = b.$$

Δb can be thought of as a random error and Δx is the resulting error in the answer. For the triplet $(M, b, \Delta b)$ we define the *error ratio*, $\alpha(M, b, \Delta b)$ by

$$\alpha(M, b, \Delta b) = \left(\frac{\| \Delta x \|}{\| x \|}\right) \cdot \left(\frac{\| \Delta b \|}{\| b \|}\right)^{-1}$$

We see that α represents the ratio of the relative error of the solution x to the relative error of the vector b. In particular, we see that

$$C(M) = \operatorname*{Max}_{b,\, \Delta b \neq 0} \alpha(M, b, \Delta b)$$

is exactly the condition number of M as defined and studied in the previous section.

Note. There are other ways of defining the condition number using error ratios. These other definitions are of interest and can be found in several texts on numerical analysis, or can be worked out as an exercise.

We define the *average error ratio* for a matrix M, $\bar{\alpha}(M)$, by

$$\bar{\alpha}(M) = \operatorname*{Av}_{b,\, \Delta b} \alpha(M, b, \Delta b).$$

We could also define the *average loss in precision* for the matrix M as

$$\bar{L}(M) = \operatorname*{Av}_{b,\, \Delta b} \ln \alpha(M, b, \Delta b)$$

but a simple symmetry argument shows that for all M, $\bar{L}(M) = 0$. We therefore study instead the *variance of the loss of precision* for the matrix M, which is defined as

$$\operatorname{Var}(M) = \operatorname*{Av}_{b,\, \Delta b} [\ln \alpha(M, b, \Delta b)]^2.$$

For each n it is easy to find numbers $f(n)$ and $g(n)$ such that for every $n \times n$ complex matrix M, $\bar{\alpha}(M) < f(n)$ and, for every real or complex M, $\mathrm{Var}(M) < g(n)$; indeed we can show that f and g may grow quite slowly. However, much more is probably true; the following two conjectures seem very likely to be true and if they are they are sharp.

Conjecture 1. *For $n \geqslant 2$ and for all invertible complex $n \times n$ matrices M, we have*

$$1 \leqslant \bar{\alpha}(M) < \tfrac{1}{2}\pi.$$

Conjecture 2. *For $n \geqslant 2$ and for all invertible complex (resp. real) $n \times n$ matrices M, we have*

$$0 \leqslant \mathrm{Var}(M) < \tfrac{1}{12}\pi^2 \quad (\textit{resp.}\ \tfrac{1}{4}\pi^2).$$

Note that conjecture 1 has no possible analog for the real case.

References

[1] L. Blum and M. Shub, Evaluating rational functions: infinite precision is finite cost and tractable on average, to appear.
[2] H. Goldstine and J. Neumann, *Numerical Inverting of Matrices of High Order, Collected Works Volume V* (Pergamon Press, Oxford, 1963).
[3] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integrals, Series and Products* (Academic Press, New York, 1965).
[4] M.L. Mehta, *Random Matrices and the Statistical Theory of Energy Levels* (Academic Press, New York, 1967).
[5] A. Ocneanu, On the stability of large linear systems, to appear.
[6] L. Santalo, *Integral Geometry and Geometric Probability* (Addison-Wesley, Reading, MA, 1976).
[7] M. Shub, The geometry and topology of dynamical systems and algorithms for numerical problems, prepared for lectures given at D.D.4, Peking University, Beijing, China, Aug. Sept., 1983.
[8] S. Smale, The fundamental theorem of algebra and complexity theory, *Bull. AMS* **4** (1) (1981).
[9] S. Smale, *Complexity Theory*, to appear.
[10] S. Wilks, *Mathematical Statistics* (Wiley, New York, 1987).