

NUMERICAL INVERTING OF MATRICES OF HIGH ORDER. II

HERMAN H. GOLDSTINE AND JOHN VON NEUMANN

TABLE OF CONTENTS

PREFACE.....	188
CHAPTER VIII. Probabilistic estimates for bounds of matrices	
8.1 A result of Bargmann, Montgomery and von Neumann.....	188
8.2 An estimate for the length of a vector.....	191
8.3 The fundamental lemma.....	192
8.4 Some discrete distributions.....	194
8.5 Continuation.....	196
8.6 Two applications of (8.16).....	198
CHAPTER IX. The error estimates	
9.1 Reconsideration of the estimates (6.42)–(6.44) and their consequences..	199
9.2 The general A_I	200
9.3 Concluding evaluation.....	200

PREFACE

In an earlier paper,¹ to which the present one is a sequel, we derived rigorous error estimates in connection with inverting matrices of high order. In this paper we reconsider the problem from a probabilistic point of view and reassess our critical estimates in this light. Our conclusions are given in Chapter IX and are summarized in §9.3.

As in the earlier paper we have here made no effort to obtain optimal numerical estimates. However, we do feel that within the framework of our method our estimates give the correct orders of magnitude.

This work has been made possible by the generous support of the Office of Naval Research under Contract N-7-onr-388. We wish also to acknowledge the use by us of some, as yet unpublished, results of V. Bargmann, D. Montgomery, and G. Hunt, which were privately communicated to us.

CHAPTER VIII. PROBABILISTIC ESTIMATES FOR BOUNDS OF MATRICES

8.1 A result of Bargmann, Montgomery and von Neumann. We seek a probabilistic result on the size of the upper bound of an n th order matrix which was first established by Bargmann, Montgomery, and von Neumann. Since the result is contained in an as yet unpublished paper, we give below an outline of the proof and also

Received by the editors February 27, 1950.

¹ *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc. vol. 53 (1947) pp. 1021–1099. The numbering of chapters and sections in the present paper follows directly upon those of the one just cited.

we modify slightly the final theorem. Let

$$(8.1) \quad A = (a_{ij}) \quad (i, j = 1, 2, \dots, n)$$

be a matrix whose elements are independent random variables, each of which is normally distributed with mean zero and dispersion σ . Under these assumptions it is well known² that

(a) The joint distribution function for the elements a_{ij} ($i, j = 1, 2, \dots, n$) is given by

$$\frac{1}{\sigma^{n^2} (2\pi)^{n^2/2}} \exp \left[- \frac{\sum_{i,j=1}^n a_{ij}^2}{2\sigma^2} \right] \prod_{i,j=1}^n da_{ij};$$

(b) The joint distribution function for the elements of the definite matrix $A^*A = B = (b_{ij})$ ($i, j = 1, 2, \dots, n$) is given by

$$W_{n,n} \left(b_{ij}; \frac{\delta_{ij}}{\sigma^2} \right) = \frac{(\text{Det } B)^{-1/2}}{(2^{1/2}\sigma)^{n^2} \pi^{n(n-1)/4} \prod_{i=1}^n \Gamma \left(\frac{n+1-i}{2} \right)} \exp \left[- \frac{\sum_{i=1}^n b_{ii}}{2\sigma^2} \right];$$

(c) The joint distribution function for the n proper values, $\lambda_1, \lambda_2, \dots, \lambda_n$ of the matrix B is given by

$$K_n \exp \left[- \frac{\sum_{i=1}^n \lambda_i}{2\sigma^2} \right] \prod_{i < j} (\lambda_i - \lambda_j) \prod_{i=1}^n \lambda_i^{-1/2} d\lambda_i,$$

where³

$$(8.2) \quad K_m = \frac{\pi^{m/2}}{(2\sigma^2)^{m^2/2} \prod_{i=1}^m \left(\Gamma \left(\frac{m+1-i}{2} \right) \right)^2} = \frac{\pi^{m/2}}{(2\sigma^2)^{m^2/2} \prod_{i=1}^m \left(\Gamma \left(\frac{i}{2} \right) \right)^2} \quad (m = 1, 2, \dots).$$

² Cf. for example, S. S. Wilks, *Mathematical statistics*, Princeton, 1943, pp. 226 ff.

³ Cf. S. S. Wilks, op. cit., pp. 262-263. In his results $n-1$ should be replaced by n , λ by σ^2 , and l_i by λ_i ($i=1, 2, \dots, n$) to make them consistent with our notational practices.

Next

(8.3) Assume that $\lambda = \lambda_1$ is the largest proper value of B and that $\phi(\lambda)$ is its probability density. Then, if $r \geq 1$,

$$\text{Prob}(\lambda > 2\sigma^2 rn) < \left(\frac{2r}{e^{r-1}}\right)^n \times \frac{1}{4(r-1)(r\pi n)^{1/2}}.$$

To prove (8.3) let R be the region of $(n-1)$ -dimensional space $\lambda_i \geq 0$ ($i=2, 3, \dots, n$) and for each $\lambda > 0$ let R_λ be the sub-region of R consisting of all $(\lambda_2, \lambda_3, \dots, \lambda_n)$ such that $\lambda \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. We then see that

$$\begin{aligned} \phi(\lambda) &= K_n \lambda^{-1/2} \exp\left(-\frac{\lambda}{2\sigma^2}\right) \int_{R_\lambda} \exp\left[-\frac{\sum_{i=2}^n \lambda_i}{2\sigma^2}\right] \\ &\quad \cdot \prod_{2 \leq i < j} (\lambda_i - \lambda_j) \prod_{j=2}^n (\lambda - \lambda_j) \lambda_j^{-1/2} d\lambda_j \\ &\cong K_n \lambda^{n-3/2} \exp\left(-\frac{\lambda}{2\sigma^2}\right) \int_{R_\lambda} \exp\left[-\frac{\sum_{i=2}^n \lambda_i}{2\sigma^2}\right] \\ (8.4.a) \quad &\quad \cdot \prod_{2 \leq i < j} (\lambda_i - \lambda_j) \prod_{j=2}^n \lambda_j^{-1/2} d\lambda_j \\ &\leq K_n \lambda^{n-3/2} \exp\left(-\frac{\lambda}{2\sigma^2}\right) \int_R \exp\left[-\frac{\sum_{i=2}^n \lambda_i}{2\sigma^2}\right] \\ &\quad \cdot \prod_{2 \leq i < j} (\lambda_i - \lambda_j) \prod_{j=2}^n \lambda_j^{-1/2} d\lambda_j \\ &= \lambda^{n-3/2} \exp\left(-\frac{\lambda}{2\sigma^2}\right) K_n K_{n-1}^{-1}; \end{aligned}$$

the first inequality follows at once from the fact that $\lambda - \lambda_j \leq \lambda$ ($j=2, 3, \dots, n$), the second from the fact that $R_\lambda \subset R$ and the last equality is a known result.⁴ It now follows easily from (8.2) that

$$(8.4.b) \quad \frac{K_n}{K_{n-1}} = \frac{\pi^{1/2}}{(2\sigma^2)^{n-1/2} (\Gamma(n/2))^2};$$

and hence summing up, we have

⁴ Cf. S. S. Wilks, *op. cit.*, pp. 263-264.

$$(8.5) \quad \phi(\lambda) \leq \frac{\pi^{1/2}}{(2\sigma^2)^{n-1/2}(\Gamma(n/2))^2} \lambda^{n-3/2} e^{-\lambda/2\sigma^2}.$$

With the help of (8.5) and the substitution $2\sigma^2\mu = \lambda - 2\sigma^2rn$ we find that

$$\begin{aligned} & \text{Prob}(\lambda > 2\sigma^2rn) \\ &= \int_{2\sigma^2rn}^{\infty} \phi(\lambda) d\lambda \leq \frac{\pi^{1/2}}{(2\sigma^2)^{n-1/2}(\Gamma(n/2))^2} \int_{2\sigma^2rn}^{\infty} \lambda^{n-3/2} e^{-\lambda/2\sigma^2} d\lambda \\ &= \frac{\pi^{1/2} e^{-rn}}{(\Gamma(n/2))^2} \int_0^{\infty} e^{-\mu} (\mu + rn)^{n-3/2} d\mu \\ (8.6) \quad &= \frac{(rn)^{n-3/2} e^{-rn} \pi^{1/2}}{(\Gamma(n/2))^2} \int_0^{\infty} e^{-\mu} \left(1 + \frac{\mu}{rn}\right)^{n-3/2} d\mu \\ &< \frac{(rn)^{n-3/2} e^{-rn} \pi^{1/2}}{(\Gamma(n/2))^2} \int_0^{\infty} e^{-\mu[1-(n-3/2)rn]} d\mu \\ &= \frac{(rn)^{n-3/2} e^{-rn} \pi^{1/2}}{(\Gamma(n/2))^2 (1 - ((n-3/2)/rn))} < \frac{(rn)^{n-1/2} e^{-rn} \pi^{1/2}}{(\Gamma(n/2))^2 (r-1)n}. \end{aligned}$$

Finally we recall with the help of Stirling's formula that

$$(8.7) \quad \left(\Gamma\left(\frac{n}{2}\right)\right)^2 > \frac{\pi n^{n-1}}{e^n \cdot 2^{n-2}} \quad (n = 1, 2, \dots);$$

now combining (8.6) and (8.7) we obtain our desired result:

$$\begin{aligned} (8.8) \quad \text{Prob}(\lambda > 2\sigma^2rn) &< \frac{(rn)^{n-1/2} e^{-rn} \pi^{1/2} e^n \cdot 2^{n-2}}{\pi n^{n-1} (r-1)n} \\ &= \left(\frac{2r}{e^{r-1}}\right)^n \times \frac{1}{4(r-1)(r\pi n)^{1/2}}. \end{aligned}$$

We sum up in the following theorem:

(8.9) The probability that the upper bound $|A|$ of the matrix A of (8.1) exceeds $2.72\sigma n^{1/2}$ is less than $.027 \times 2^{-n} n^{-1/2}$, that is, with probability greater than 99% the upper bound of A is less than $2.72\sigma n^{1/2}$ for $n = 2, 3, \dots$.

This follows at once by taking $r = 3.70$.

8.2 An estimate for the length of a vector. It is well known that

(8.10) If a_1, a_2, \dots, a_n are independent random variables each of which is normally distributed with mean 0 and dispersion σ^2 and if $|a|$ is the length of the vector $a = (a_1, a_2, \dots, a_n)$, then

$$(8.11) \quad P(|a| > (\sigma^2 rn)^{1/2}) = \frac{2}{2^{n/2} \sigma^n \Gamma(n/2)} \int_{(\sigma^2 rn)^{1/2}}^{\infty} x^{n-1} e^{-x^2/2\sigma^2} dx.^5$$

Let $n = 2N$ and $y = x^2$; then (8.11) becomes

$$P(|a| < (2\sigma^2 rN)^{1/2}) = \frac{1}{2^N \sigma^{2N} \Gamma(N)} \int_{2\sigma^2 rN}^{\infty} y^{N-1} e^{-y/2\sigma^2} dy.$$

Since the integral on the right is of the same form as the one in (8.6), we see by the same sort of argument which led to (8.8) that

$$\begin{aligned} P(|a| > (2\sigma^2 rN)^{1/2}) &< \frac{(rN)^N}{e^{rN} N(r-1)\Gamma(N)} \\ &< \left(\frac{r}{e^{r-1}}\right)^N \times \frac{1}{\pi^{1/2}(r-1)(2N)^{1/2}}, \end{aligned}$$

that is, we have

(8.12) For random vectors a , as defined in (8.10), we have

$$(8.13.a) \quad P(|a| > \sigma(rn)^{1/2}) < (re^{1-r})^{n/2} \times \frac{1}{(\pi n)^{1/2}(r-1)},$$

$$(8.13.b) \quad P(|a| > 1.92\sigma n^{1/2}) < 0.21 \times 2^{-n}.$$

The last inequality follows at once from (8.13.a) by setting $r = 3.70$.

8.3 The fundamental lemma. To describe the result upon which the body of the paper depends we need two definitions.

(8.14) Designate by $U_l(\alpha)$ the l th convolution of the uniform distribution on the interval $-\alpha, +\alpha$.

(8.15) Let ϕ be a non-negative valued function defined on a real M -dimensional euclidean space and let it have the properties:

(a) ϕ is continuous and convex, that is, for $a \neq b$, $\phi((a+b)/2) \leq (\phi(a) + \phi(b))/2$.

(b) ϕ is positively homogeneous, that is, $\phi(sa) = |s| \cdot \phi(a)$ for every real number s .

(c) ϕ has bound 1, that is if $|a| \leq 1$, $\phi(a) \leq 1$.

If $M = n^2$ and if the point $a = (a_1, a_2, \dots, a_M)$ is interpreted as an n th order matrix, then the bound satisfies the properties (a)–(c) of (8.15). Similarly, if $M = n$ and if a is interpreted as a vector, then the

⁵ Cf. for example, Cramér, *Mathematical methods of statistics*, Princeton, 1946, p. 236.

length of a vector satisfies the properties (a)–(c) of (8.15). We introduce the function ϕ to unify the discussion of §§8.4–8.6 below, which are devoted to probabilistic estimates for the bounds of matrices and lengths of vectors.

Our principal result is

(8.16) Assume the following:

(a) $h = (h_1, h_2, \dots, h_M)$ is a vector whose elements h_m are independent random variables such that each h_m has the distribution $U_{l_m}(\beta^{-s}/2)$ and let l be a number greater than or equal to l_m ($m = 1, 2, \dots, M$).

(b) $a = (a_1, a_2, \dots, a_M)$ is a vector whose elements are independent random variables such that each a_m is normally distributed with mean 0 and dispersion $\sigma^2 = (l+1)\beta^{-2s}/4$. Then we have for $\kappa > 0$

$$P \left[\phi(h) > \frac{(l+1)^{1/2}}{2} (\kappa + 2M^{1/2}l^{-1/2})\beta^{-s} \right] \leq 2e^{3M/4l} P \left[\phi(a) > \frac{(l+1)^{1/2}}{2} \kappa\beta^{-s} \right] \times (\log_2 2l).$$

As a first step in establishing this result we state and prove an unpublished theorem of G. A. Hunt.

(8.17) Assume the following:

(a) h is a vector having the property (a) of (8.16).

(b) $g = (g_1, g_2, \dots, g_M)$ is a vector whose elements are independent and which have the distribution $U_i(\beta^{-s}/2)$.

Then for $a > 0$

$$P(\phi(g) > a) \geq P(\phi(h) > a)/2.$$

Hunt's proof, which is very simple and elegant, can in this context be given as follows: Let f be a vector whose elements are independent and which have the distribution $U_{l-l_m}(\beta^{-s}/2)$. Then clearly g and $f+h$ have the same distribution. Since ϕ is convex we always have $\text{Max}(\phi(h+f), \phi(h-f)) \geq \phi(h)$; and since the distribution $U_i(\alpha)$ is symmetric about the mean 0, with probability at least 1/2, we have $\phi(h+f) \geq \phi(h)$ for each vector h . Thus, by integrating over the set $\phi(h) > a$ we find $P(\phi(g) > a) = P(\phi(h+f) > a) \geq P(\phi(h) > a)/2$.

With the help of (8.17) we see that in order to establish (8.16) it suffices to show

(8.18) Let g be a vector with the properties described in (b) of (8.17) and let a be a vector with the properties described in (b) of (8.16). Then

$$\begin{aligned}
 P \left[\phi(g) > \frac{(l+1)^{1/2}}{2} (\kappa + 2M^{1/2}l^{-1/2})\beta^{-s} \right] \\
 \leq e^{3M/l} P \left[\phi(a) > \frac{(l+1)^{1/2}}{2} \kappa\beta^{-s} \right] \times (\log_2 2l).
 \end{aligned}$$

8.4 Some discrete distributions. To establish (8.18) we consider a number of related distributions.

(8.19) Assume that

(a) $b = (b_1, b_2, \dots, b_M)$ is a vector whose elements are independent and have the discrete distribution

$$P \left(\beta = \frac{2h\sigma}{(l+1)^{1/2}} \right) = \frac{1}{\sigma(2\pi)^{1/2}} \int_{((2h-1)/(l+1)^{1/2})\sigma}^{((2h+1)/(l+1)^{1/2})\sigma} e^{-x^2/2\sigma^2} dx$$

($h = 0, \pm 1, \pm 2, \dots$).

Then we have

$$\begin{aligned}
 P(\phi(b) > \sigma(\kappa + M^{1/2}l^{-1/2})) \\
 \leq P(\phi(a) > \sigma\kappa) \\
 = P_0,
 \end{aligned}$$

(8.20)

where $a = (a_1, a_2, \dots, a_M)$ is the vector described in (b) of (8.16).

To show this we note first that each b_m is expressible in the form

$$\begin{aligned}
 b_m &= \frac{2\sigma}{(l+1)^{1/2}} \left[a_m \frac{(l+1)^{1/2}}{2\sigma} + \frac{1}{2} \right] \\
 &= a_m + \frac{2\sigma}{(l+1)^{1/2}} \left\{ \left[a_m \frac{(l+1)^{1/2}}{2\sigma} + \frac{1}{2} \right] - a_m \frac{(l+1)^{1/2}}{2\sigma} \right\},
 \end{aligned}$$

where $a = (a_1, a_2, \dots, a_M)$ has the property (b) of (8.16) and where $[x]$ means the integer part of x . Now, if we define a vector $c = (c_1, c_2, \dots, c_M)$ by

$$c_m = \frac{2\sigma}{(l+1)^{1/2}} \left[a_m \frac{(l+1)^{1/2}}{2\sigma} + \frac{1}{2} \right] - a_m,$$

then $b = a + c$ and by (a), (b) of (8.15) we have $\phi(b) \leq \phi(a) + \phi(c)$; moreover, we have at once $|c_m| \leq \sigma/(l+1)^{1/2}$ and thus by (c) of (8.15) $\phi(c) \leq \sigma(M/l+1)^{1/2}$. If $\phi(b) > \sigma(\kappa + (M/l)^{1/2})$, then $\phi(a) > \kappa\sigma + \sigma(M/l)^{1/2} - \sigma(M/(l+1))^{1/2} > \sigma\kappa$; our conclusion (8.20) now follows at once.

We interrupt momentarily our chain of argument to show

$$\begin{aligned}
 (8.21) \quad \frac{1}{2^l} C_{l, l/2+h} &\leq \frac{e^{3/4l}}{\sigma(2\pi)^{1/2}} \int_{((2h-1)/(l+1))^{1/2}\sigma}^{((2h+1)/(l+1))^{1/2}\sigma} e^{-x^2/2\sigma^2} dx \\
 &= e^{3/4l} \left(\frac{2}{\pi(l+1)} \right)^{1/2} \int_{h-1/2}^{h+1/2} e^{-2y^2/l+1} dy \\
 &\quad (h = 0, \pm 1, \pm 2, \dots, \pm l/2),
 \end{aligned}$$

where $C_{l,m}$ is the binomial coefficient.

Let

$$f(h) = e^{2h^2/(l+1)} C_{l, l/2+h}, \quad g(h) = e^{2h^2/(l+1)} \int_{h-1/2}^{h+1/2} e^{-2y^2/(l+1)} dy.$$

Then clearly we have for $h \geq 0$

$$\begin{aligned}
 \frac{f(h+1)}{f(h)} &= \frac{l/2 - h}{l/2 + h + 1} e^{2((h+1)^2 - h^2)/(l+1)} \\
 &= \frac{(l+1) - (2h+1)}{(l+1) + (2h+1)} e^{2(2h+1)/(l+1)} \\
 &= \frac{1 - (2h+1)/(l+1)}{1 + (2h+1)/(l+1)} e^{2(2h+1)/(l+1)} \\
 &= e^{2(2h+1)/(l+1) + \ln(1 - (2h+1)/(l+1)) / (1 + (2h+1)/(l+1))} \leq 1.
 \end{aligned}$$

Thus for $h \geq 0$, $f(h+1) \leq f(h) \leq \dots \leq f(0)$; but by Stirling's formula

$$\begin{aligned}
 f(0) = C_{l, l/2} &= \frac{l!}{((l/2)!)^2} \leq \frac{(2\pi)^{1/2} e^{-l} l^{l+1/2} e^{1/12l}}{2\pi e^{-l} (l/2)^{l+1} e^{1/8(l+1)}} \\
 &\leq 2^l e^{(1-3l)/12l(l+1)} \left(\frac{2}{\pi l} \right)^{1/2},
 \end{aligned}$$

and thus for $h \geq 0$

$$\begin{aligned}
 (8.22) \quad f(h) &\leq 2^l e^{(1-3l)/12l(l+1)} \left(\frac{2}{\pi l} \right)^{1/2} \\
 &= 2^l e^{(1-3l)/12l(l+1)} \left(\frac{2}{\pi(l+1)} \right)^{1/2} \left(\frac{l+1}{l} \right)^{1/2} \\
 &\leq 2^l e^{(3l+7)/12l(l+1)} \left(\frac{2}{\pi(l+1)} \right)^{1/2}.
 \end{aligned}$$

But $f(h)$ is even for all h and thus (8.22) is valid for all h .

Next we underestimate $g(h)$. Notice that

$$\begin{aligned}
 g(h) &= e^{2h^2/(l+1)} \int_{-1/2}^{1/2} e^{-2(h+x)^2/(l+1)} dx = \int_{-1/2}^{1/2} e^{-(4h/(l+1))x - (2/(l+1))x^2} dx \\
 &\geq e^{-1/2(l+1)} \int_{-1/2}^{1/2} e^{-(4h/(l+1))x} dx \\
 &= e^{-1/2(l+1)} \frac{\sinh(2h/(l+1))}{2h/(l+1)} \geq e^{-1/2(l+1)};
 \end{aligned}$$

combining this estimate for $g(h)$ with (8.22) we obtain

$$\frac{f(h)}{g(h)} = C_{l, l/2+h} / \int_{h-1/2}^{h+1/2} e^{-2v^2/(l+1)} dy \leq 2^l \left(\frac{2}{\pi(l+1)} \right)^{1/2} e^{3/4l},$$

which implies (8.21) directly.

We pick up again the thread of our previous argument in

(8.23) Assume that

(a) $c = (c_1, c_2, \dots, c_M)$ is a vector whose elements are independent and have the distribution

$$(8.24) \quad P\left(\gamma = \frac{2h\sigma}{(l+1)^{1/2}}\right) = \frac{1}{2^l} C_{l, l/2+h} \quad (h = 0, \pm 1, \pm 2, \dots, \pm l/2).$$

Then we have

$$P(\phi(c) > \sigma(\kappa + M^{1/2}l^{-1/2})) \leq e^{3M/4l} P_0,$$

where P_0 is defined in (8.20).

We see with the help of (8.21) that

$$P\left(\gamma = \frac{2h\sigma}{(l+1)^{1/2}}\right) \leq e^{3/4l} P\left(\beta = \frac{2h\sigma}{(l+1)^{1/2}}\right)$$

and thus (8.23) follows directly from (8.19).

It is of some interest to notice that the distribution defined in (8.24) is the l th convolution of the distribution $\pm\sigma/(l+1)^{1/2}$ with probability 1/2.

8.5 Continuation. The result (8.23) is true of course for σ arbitrary and, hence, is also true for σ replaced by $\sigma/2^p$.

(8.25) Let $d^{(p)} = (d_1^{(p)}, d_2^{(p)}, \dots, d_M^{(p)})$ with $p=1, 2, \dots, q$ be a vector whose elements are independent and have as their distributions

$$P\left(\delta = \frac{2h\sigma}{2^p(l+1)^{1/2}}\right) = \frac{1}{2^l} C_{l, l/2+h} \quad (h = 0, \pm 1, \pm 2, \dots, \pm l/2).$$

Since ϕ is homogeneous (cf. (b) of 8.15), $P_0 = P(\phi(a) > \sigma\kappa)$ is seen to be independent of σ . It then follows directly from (8.23) that $P(\phi(d^{(p)}) > \sigma(\kappa + M^{1/2}l^{-1/2})2^{-p}) \leq e^{3M/4l}P_0$. Let

$$(8.26) \quad e^{(q)} = \sum_{p=1}^q d^{(p)}.$$

Now if $\phi(d^{(p)}) \leq 2^{-p}\sigma(\kappa + M^{1/2}l^{-1/2})$ for all $p=1, 2, \dots, q$, then by (8.15), $\phi(e^{(q)}) < \sigma(\kappa + M^{1/2}l^{-1/2})$. Hence if $\phi(e^{(q)}) > \sigma(\kappa + M^{1/2}l^{-1/2})$, then for some p we must have $\phi(d^{(p)}) > 2^{-p}\sigma(\kappa + M^{1/2}l^{-1/2})$ and thus we have

$$(8.27) \quad P(\phi(e^{(q)}) > \sigma(\kappa + M^{1/2}l^{-1/2})) \leq qe^{3M/4l}P_0.$$

(8.28) Let q_0 be the smallest integer q' for which $2^{-q'} \leq 1/l$ and let the vector $e = e^{(q_0)} = \sum_{p=1}^{q_0} d^{(p)}$.

Then if the q in (8.25), (8.26) is greater than q_0 , $e^{(q)} = e + \sum_{p=q_0+1}^q d^{(p)}$. In the event that $\phi(e^{(q)}) > \sigma(\kappa + 2M^{1/2}l^{-1/2})$, then by (8.15)

$$(8.29) \quad \sigma(\kappa + 2M^{1/2}l^{-1/2}) < \phi(e) + \sum_{q_0+1}^q \phi(d^{(p)}).$$

To estimate $\phi(d^{(p)})$ we recall from the definition of $d^{(p)}$ that each of its elements is in absolute value less than or equal to $2^{-p}l\sigma/(l+1)^{1/2} \leq 2^{-p}\sigma l^{1/2}$ and thus that $|d^{(p)}| \leq 2^{-p}\sigma M^{1/2}l^{1/2}$. Hence $\phi(d^{(p)}) \leq 2^{-p}\sigma M^{1/2}l^{1/2}$ and

$$\sum_{q_0+1}^q \phi(d^{(p)}) \leq \sigma M^{1/2}l^{1/2} \sum_{q_0+1}^q 2^{-p} < 2^{-q_0}\sigma M^{1/2}l^{1/2} \leq \sigma M^{1/2}l^{-1/2}$$

by (8.28). From this and (8.29) it now follows that

$$\phi(e) > \sigma(\kappa + 2M^{1/2}l^{-1/2}) - \sigma M^{1/2}l^{-1/2} = \sigma(\kappa + M^{1/2}l^{-1/2});$$

that is, $\phi(e^{(q)}) > \sigma(\kappa + 2M^{1/2}l^{-1/2})$ implies that $\phi(e) > \sigma(\kappa + M^{1/2}l^{-1/2})$. Hence by (8.27) with $q = q_0$ we have for $q \geq q_0$

$$(8.27') \quad P(\phi(e^{(q)}) > \sigma(\kappa + 2M^{1/2}l^{-1/2})) \leq q_0 e^{3M/4l} P_0.$$

Next if $q < q_0$

$$\begin{aligned} P(\phi(e^{(q)}) > \sigma(\kappa + 2M^{1/2}l^{-1/2})) &\leq P(\phi(e^{(q)}) > \sigma(\kappa + M^{1/2}l^{-1/2})) \\ &\leq q e^{3M/4l} P_0 < q_0 e^{3M/4l} P_0. \end{aligned}$$

Thus (8.27') is valid for all q .

(8.30) If $e^{(q)}$ ($q=1, 2, \dots$) is the vector defined in (8.26), (8.25) then

$$P(\phi(e^{(q)}) > \sigma(\kappa + 2M^{1/2}l^{-1/2})) < e^{3M/4l}P_0 \cdot (\log_2 2l).$$

This follows at once from (8.28) since

$$q_0 \leq [\log_2 l] + 1 \leq \log_2 l + 1 = \log_2 2l.$$

Also since ϕ is a continuous function we see that

(8.31) If f is a vector whose elements are independent and have the distribution $U_l(\sigma/(l+1)^{1/2})$, then

$$P(\phi(f) > \sigma(\kappa + 2M^{1/2}l^{-1/2})) \leq e^{3M/4l}P_0 \cdot (\log_2 2l).$$

To complete the proof of (8.18) we have only to adjust the size of σ . Let $\sigma/(l+1)^{1/2} = \beta^{-s}/2$ and take the vector f of (8.31) to be the g of (8.18). Then by the definition of P_0 in (8.20) we have (8.18) and also (8.16).

8.6 Two applications of (8.16). We are, of course, interested in (8.16) primarily for estimating probabilistically the bounds of matrices and norms of vectors whose elements are sums of one or more rounding errors. Stated more precisely:

(8.32) Assume the following:

(a) $H = (h_{ij})$ is a matrix whose elements are independent random variables having the distributions $U_{l_{ij}}(\beta^{-s}/2)$ with $l_{ij} \leq un$ ($u \geq 1$).

(b) $x = (x_1, x_2, \dots, x_m)$ is a vector whose elements are independent random variables having the distributions $U_{m_i}(\beta^{-s}/2)$, $m_i \leq un$.

Then we have for $4 \geq u \geq 1$, $n \geq 10$

$$(c) P(|H| > (u + .1)^{1/2}(1.53 + 1/u^{1/2})n\beta^{-s}) < .07 \times 2^{-n},$$

$$(d) P(|x| > (u + .1)^{1/2}(1.34n\beta^{-s})) < .03 \times 2^{-n/2}.$$

We now apply (8.8) and (8.16) to the derivation of (c) above. In (8.16) we take $M = n^2$ and interpret the n^2 -dimensional vectors h and a there as n th order matrices H and A , respectively, with $\phi = | \cdot |$, the bound of a matrix. Then (8.8) is applicable to A with $\sigma^2 = (l+1)\beta^{-2s}/4$ and we choose $l = un$ ($4 \geq u \geq 1$), $\kappa = (2rn)^{1/2}$ and recall from §7.2 that $n \geq 10$. Thus we have by (8.8) and (8.16)

$$\begin{aligned} P(|H| > (u + .1)^{1/2}((2r)^{1/2} + 2/u^{1/2})n\beta^{-s}/2) \\ &\leq P(|H| > (un + 1)^{1/2}((2rn)^{1/2} + 2(n/u)^{1/2})\beta^{-s}/2) \\ (8.33) \quad &\leq 2e^{3n/4u}P(|A| > ((un + 1)(2rn))^{1/2}\beta^{-s}/2) \times (\log_2 2un) \\ &< (\ln 8n)(2re^{1+3/4u-r})^n \times ((r - 1)(r\pi n)^{1/2} \ln 2)^{-1}. \end{aligned}$$

To complete the argument we note that $u \geq 1$ and take $r = 4.7$. Then $2re^{1+3/4-r} \leq 1/2$ and for $n \geq 10$, $(\ln 8n)/n^{1/2} \leq (\ln 80)/10^{1/2}$.

We next apply (8.13.a) and (8.16) to the derivation of (d) above. In (8.16) we, this time, take $M = n$ and ϕ to be the length of a vector. Then (8.13.a) is applicable to the vector a with $\sigma^2 = (l+1)\beta^{-2s}/4$ and we choose $l = un, \kappa = (rn)^{1/2}$. Thus, we have by (8.13.a.) and (8.16)

$$\begin{aligned} P(|x| > (u+.1)^{1/2}((rn)^{1/2} + 2/u^{1/2})n^{1/2}\beta^{-s}/2) \\ \leq P(|x| > (un + 1)^{1/2}((rn)^{1/2} + 2/u^{1/2})\beta^{-s}/2) \\ \leq 2e^{3/4u}P(|a| > ((un + 1)rn)^{1/2} \beta^{-s}/2 \times (\log_2 2un)) \\ < (2 \ln 8n)e^{3/4}(re^{1-r})^{n/2} \cdot ((r - 1)(\pi n)^{1/2} \ln 2)^{-1} \\ = 2e^{3/4} \frac{(\ln 8n)}{n^{1/2}} (re^{1-r})^{n/2} ((r - 1)\pi^{1/2} \ln 2)^{-1}. \end{aligned}$$

To complete the argument we take $r = 4$ and note that $(u+.1)^{1/2} \times (1.34n) \geq (u+.1)^{1/2}(n^{1/2} + 1/u^{1/2})n^{1/2}$ for $n \geq 10$.

CHAPTER IX. THE ERROR ESTIMATES

9.1 Reconsideration of the estimates (6.42)–(6.44) and their consequences. The decisive estimates of §§6.3–6.10 are (6.42), (6.43), and (6.44) and they yield (6.96) and (6.102) almost directly. We proceed now to replace the strict estimates (6.42)–(6.44) by probabilistic ones. These replacements are all based upon (8.32); and since we assume $n \geq 10$ they have probabilities of at least 99.9%.

First, we apply (8.32) to the matrix of (6.33) and see by (6.29) that the u of (8.32) can be taken to be 2. Thus with high probability we can replace (6.34) and the equivalent (6.42) by

$$(9.1) \quad |\bar{A} - \bar{B}^* \bar{D} \bar{B}| \leq 3.25n\beta^{-s}.$$

Second, we apply (8.32) to the matrix U of (6.39) and see by (6.39) that u can be taken to be 4. Thus we replace (6.43) by

$$(9.2) \quad |U| \leq 4.12n\beta^{-s}.$$

Third, we apply (8.32) to the vectors $U^{(j)}$ and see by (6.39) that u can be taken to be 4. Thus we replace (6.44) by

$$(9.3) \quad |U^{(j)}| \leq 2.72n\beta^{-s}.$$

Next we define

$$(9.4) \quad \alpha' = 8n\beta^{-s}/\mu,$$

and notice that (9.2)–(9.4) imply

$$(9.5) \quad |\bar{A} - \bar{B}^* \bar{D} \bar{B}| < .42\mu\alpha',$$

$$(9.6) \quad |U| < .58\mu\alpha',$$

$$(9.7) \quad |U^{(i)}| \leq .34\mu\alpha'.$$

Thus (9.5)–(9.7) are in form the same as (6.42)–(6.44) except that α has been replaced by α' .

The estimate (6.65'') will now be valid with α replaced by α' , provided only that (6.62) and (6.67) are valid under this same replacement. Using (8.32) with $u=4$, we find that (6.56) becomes

$$(9.9) \quad |\bar{W}^{(q)} - 2^{-q} \bar{Z} \Delta^{-2} \bar{D}^{-1} \bar{Z}^*| \leq 4.12n\beta^{-s} < .91\mu\alpha'.$$

If now we assume that

$$(9.10) \quad \alpha' \leq .1,$$

then we have

$$(9.11) \quad |2^{q_0} \bar{A} \bar{W}_0 - I| \leq (1.98 + 10.28\lambda)\alpha',$$

where q_0 is defined in (6.59).

9.2 The general A_I . By analogy with (6.73) and (9.4) we define

$$(9.12) \quad \alpha'_I = 8n\beta^{-s} / \mu_I^2.$$

Then with the help of (8.32) with $u=1$ we can replace (6.75) by

$$(9.13) \quad |\bar{A} - \bar{A}_I \bar{A}_I^*| < 2.66n\beta^{-s} < \mu_I^2 \alpha'_I / 2,$$

and thus (6.78') and (9.10) can be replaced by

$$\alpha'_I \leq .095.$$

All the results of §§6.9–6.11 are now valid with α_I replaced by α'_I and in particular (6.102) is now

$$(9.14) \quad |2^{q_1} \bar{A}_I \bar{S} - I| \leq (5.29 + 14.55\lambda_I^2) \alpha'_I,$$

where q_1 is defined by (6.92).

9.3 Concluding evaluation. In closing it is desirable for us to evaluate once again our conclusions in a final capitulation.

First, the requirements of §§7.2 and 7.3 are unchanged except for (7.2), (7.2_I) which are now

$$(9.15) \quad \alpha' \leq .1, \quad \text{that is } \mu \geq 80n\beta^{-s},$$

$$(9.15_I) \quad \alpha'_I \leq .095, \quad \text{that is } \mu_I^2 \geq 84.2n\beta^{-s}.$$

Second in the light of (9.11), (9.14), (7.5'), (7.5'_I) now become

$$(9.16) \quad |2^{20}\overline{A}\overline{W}_0 - I| \leq 114(\lambda/\mu)n\beta^{-s},$$

$$(9.16_I) \quad |2^{21}\overline{A}_I\overline{S} - I| \leq 293(\lambda_I^2/\mu_I^2)n\beta^{-s},$$

and the conditions which are the negations of (9.15), (9.15_I) are

$$(9.17) \quad \mu < 80n\beta^{-s},$$

$$(9.17_I) \quad \mu_I^2 < 84.2n\beta^{-s}.$$

Thus we have either (9.16) or (9.17) for \overline{A} and either (9.16_I) or (9.17_I) for \overline{A}_I . The conditions (9.16), (9.16_I) express that $2^{20}\overline{W}_0$ and $2^{21}\overline{S}$ are approximate inverses for \overline{A} and \overline{A}_I , respectively, while the conditions (9.17), (9.17_I) express that \overline{A} and \overline{A}_I , respectively, are approximately singular.

It is perhaps worth remarking that the coefficients 114 and 293 appearing in (9.16), (9.16_I) and the coefficients 80, 84.2 in (9.15), (9.15_I) are probably not optimal, but result from the crudeness of our probabilistic techniques. However, our principal aim was not to produce "best possible" estimates, but was merely to obtain a feeling for the sizes of the errors likely to be encountered.

By analogy with §7.8 we conclude with some numerical evaluations of our estimates (9.16), (9.17) for \overline{A} and (9.16_I), (9.17_I) for \overline{A}_I . Let us accept the probabilistic estimates (7.14'_I.a)–(7.14'_I.c) for λ_I, μ_I and (7.14'.a)–(7.14'.c) for λ, μ . Then (9.17), (9.17_I) imply approximately

$$(9.18) \quad n \geq .01\beta^{s/2};$$

that is, an approximate inverse will usually be found if

$$(9.18') \quad n < .01\beta^{s/2}.$$

Furthermore, the right-hand sides of the estimates (9.16), (9.16_I) become (again approximately)

$$(9.19) \quad \leq 20,000n^3\beta^{-s}.$$

(In the case of (9.16) the factor 20,000 should have been ~11,000 and in the case of (9.16_I) ~29,000. Cf. the remarks immediately succeeding (7.16); the remarks made there are substantially applicable here.)

For this to be less than one, we have

$$(9.18'') \quad n < (\beta^s/20)^{1/3}/10 < .04\beta^{s/3};$$

that is, an approximate inverse will usually be found if n satisfies the condition (9.18''). If we assume that $\beta^s \geq 10^4$, which is certainly

reasonable, then condition (9.18'') is more stringent than (9.18') and we therefore regard it as the critical condition regarding the relative precision of the approximate inverse.

Considering the precisions (7.17.a)–(17.17c.), we find that (9.18'') becomes

$$(9.19.a) \quad n < 19,$$

$$(9.19.b) \quad n < 86,$$

$$(9.19.c) \quad n < 400.$$

INSTITUTE FOR ADVANCED STUDY

ON SERIES OF WALSH EIGENFUNCTIONS

W. H. SPRAGENS

Walsh [4]¹ considered expansions in terms of eigenfunctions satisfying the equation

$$u''(x) + [\rho^2 - g(x)]u(x) = 0, \quad 0 \leq x \leq 1,$$

and the boundary conditions

$$u(0) = 0, \quad u(1) = 0,$$

the function $g(x)$ being assumed continuous on $0 \leq x \leq 1$. He used the asymptotic formula for the k th eigenfunction

$$(1) \quad \bar{u}_k(x) = (2)^{1/2} [\sin k\pi x + (1/k)\phi_k(x)], \quad |\phi_k(x)| \leq C.$$

Comparing series of these functions with corresponding series of the functions

$$(2) \quad u_k(x) = (2)^{1/2} \sin k\pi x,$$

he proved that if a function in L^2 is expanded in terms of both sets of functions, then the series of term-by-term differences converges uniformly and absolutely to zero on $0 \leq x \leq 1$.

This result, as Walsh points out, is closely related to that of Haar [1], who considered the Sturm-Liouville problem with boundary conditions $u' + Hu = 0$, and compared the series of the resulting functions with the Fourier cosine series. He obtained equiconvergence in

Received by the editors April 3, 1950.

¹ Numbers in brackets refer to the references cited at the end of the paper.