

A Method for the Generation of a Class of Ill-conditioned Matrices

Tetsuo Nishi[†], Takeshi Ogita[†], Shin'ichi Oishi[†], and S. Rump[‡]

[†]Faculty of Science and Engineering, Waseda University
 Okubo, Shinjuku, Tokyo, Japan

[‡]Hamburg University of Technology
 Hamburg, Germany

Email: nishi-t@waseda.jp, ogita@waseda.jp, oishi@waseda.jp, rump@tu-harburg.de,

Abstract—This paper shows a generalization of Rump's method, which generates a class of matrices with extremely large condition number.

1. Introduction

S. Rump proposed a method to generate a class of *extremely ill-conditioned matrices* [1]. Here the *ill-conditioned* matrix implies those having its condition number such as $10^{50} \sim 10^{100}$ or larger in double precision arithmetic. These matrices are very useful to examine the quality of *accuracy-guaranteed algorithms* for solving linear simultaneous equations [2]–[4]. For this purpose we desire more variety of extremely ill-conditioned matrices.

Rump used a *tricky* technique (shown in Sec. 2.1) to generate an ill-conditioned matrix A . One of key points of his method is to generate a 2×2 matrix V whose determinant is one and whose elements are extremely large. To generate V Rump utilized a very special equation, the *PELL EQUATION*. Since his method includes some freedom, we have a class of ill-conditioned matrices of an arbitrary size. *However the class of matrices is not so wide and therefore it is still desirable to generate more variety of ill-conditioned matrices.* This is the motivation of this research and this paper studies on the following three items.

Item 1. We show that almost the same discussion as Rump's is possible when we use the matrix V generated by the Euclid algorithm instead of the Pell Equation. This extends the Rump's matrices considerably. *Instead of his tricky and rather complicated calculation, we use a simple matrix manipulation, by which the following Item 2 can easily be derived.*

Item 2. We show an extension to the case where the matrix V is a 3×3 matrix.

Item 3. Related to Items 1 and 2 above, we show that we can find a 3×3 integer matrix $A' = [a'_{ij}]$ with $|a'_{ij}| < \mu$ and with $|A'| = 1$ by bordering a 2×2 integer matrix.

2. A generalization of Rump's method by means of simple matrix manipulation (Item 1)

2.1. Three key steps in Rump's method

Rump's method is composed of three key steps as:

1. Generation of a 2×2 integer matrix V such that

$$V = \begin{bmatrix} P & F \\ Q & G \end{bmatrix}, \quad |V| = \begin{vmatrix} P & F \\ Q & G \end{vmatrix} = 1 \quad (1)$$

where P, Q, F and G are *extremely large positive integers* such as 10^{50} .

2. P, Q, F and G are particularly chosen as

$$F = kQ, \quad G = P \quad (2)$$

and P and $Q > 0$ satisfy the *PELL Equation*:

$$P^2 - kQ^2 = 1 \Rightarrow V = \begin{bmatrix} P & kQ \\ Q & P \end{bmatrix}, \quad |V| = 1 \quad (3)$$

3. P, Q, F and G are realized by decomposing it into relatively small integers and by using the matrix similar to the *companion matrix* as shown later.

As the result he could give an explicit form of the inverse matrix A^{-1} as well as the condition number of A .

2.2. Formulation of the problem

In a similar way as Rump's method we consider:

Problem 1: Generate a class of *integer matrices*, say $B = [b_{ij}]$ satisfying

$$|B| = \pm 1 \quad (4)$$

$$|b_{ij}| < \sigma' \quad (5)$$

where σ' is a number such as 10^8 (for single precision), $10^{16}, 2^{53}$ (for double precision), but may possibly be 2, 10 or 1000 depending on the applications.

2.3. Generalization of Rump's method

Instead of the Pell equation we use the *Euclid algorithm* to determine F and G in Eq.(1) for the prescribed P and Q . As is well-known, this is possible if P and Q have no common factor[5]. For example, we choose P and Q as:

$$P = 2^k, \quad Q = 3^m \quad \text{or} \quad (6)$$

$$P = 2^{k1}5^{k2}11^{k3}, \quad Q = 3^{m1}7^{m2} \quad (7)$$

Let σ be a large number such as 10^8 , 10^{16} and 2^{53} . The extremely large numbers P , Q , F and G can be expanded as:

$$\begin{aligned} P &= p_n\sigma^n + p_{n-1}\sigma^{n-1} + p_{n-2}\sigma^{n-2} + \cdots + p_1\sigma + p_0 \\ Q &= q_n\sigma^n + q_{n-1}\sigma^{n-1} + q_{n-2}\sigma^{n-2} + \cdots + q_1\sigma + q_0 \\ F &= f_n\sigma^n + f_{n-1}\sigma^{n-1} + f_{n-2}\sigma^{n-2} + \cdots + f_1\sigma + f_0 \\ G &= g_n\sigma^n + g_{n-1}\sigma^{n-1} + g_{n-2}\sigma^{n-2} + \cdots + g_1\sigma + g_0 \\ &0 < P, Q, F, G < \sigma^{n+1}, \quad 0 \leq p_i, q_i, f_i, g_i < \sigma \end{aligned}$$

Using the above, we define a $(2n+2) \times (2n+2)$ matrix, A as

$$A = \begin{array}{c} \begin{array}{cccccccc} p_n & p_{n-1} & p_{n-2} & p_{n-3} & \cdots & p_1 & p_0 & \\ q_n & q_{n-1} & q_{n-2} & q_{n-3} & \cdots & q_1 & q_0 & \\ \hline 1 & -\sigma & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 1 & -\sigma & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 1 & -\sigma & \cdots & 0 & 0 & \\ \hline 0 & 0 & 0 & 1 & \ddots & 0 & 0 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -\sigma & \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ \hline f_n & f_{n-1} & f_{n-2} & f_{n-3} & \cdots & f_1 & f_0 & \\ g_n & g_{n-1} & g_{n-2} & g_{n-3} & \cdots & g_1 & g_0 & \\ \hline 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \\ \hline 1 & -\sigma & 0 & 0 & \cdots & 0 & 0 & \\ 0 & 1 & -\sigma & 0 & \cdots & 0 & 0 & \\ 0 & 0 & 1 & -\sigma & \cdots & 0 & 0 & \\ \hline 0 & 0 & 0 & 1 & \ddots & 0 & 0 & \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & 0 & \cdots & 1 & -\sigma & \end{array} \end{array}$$

2.4. Calculation of the inverse matrix A^{-1}

We will calculate A^{-1} for (10). Let

$$\Gamma \equiv H \oplus H, \quad H \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & \sigma^n \\ 0 & 1 & 0 & 0 & \cdots & 0 & \sigma^{n-1} \\ 0 & 0 & 1 & 0 & \cdots & 0 & \sigma^{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & \sigma \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Then we have:

$$A' \equiv A\Gamma \Rightarrow A = A'\Gamma^{-1} \Rightarrow A^{-1} = \Gamma(A')^{-1}$$

Let p, q, f, g be row vectors of order n and Σ_n^* be $n \times n$ matrix as follows:

$$\begin{aligned} p &= [p_n, p_{n-1}, \cdots, p_1], \quad q = [q_n, q_{n-1}, \cdots, q_1] \\ f &= [f_n, f_{n-1}, \cdots, f_1], \quad g = [g_n, g_{n-1}, \cdots, g_1] \end{aligned}$$

$$\Sigma_n^* \equiv \begin{bmatrix} 1 & -\sigma & 0 & \cdots & 0 \\ 0 & 1 & -\sigma & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -\sigma \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

and let O_{ml} be $m \times l$ matrix with all entries 0.

Then A' can be rewritten as:

$$A' = \left[\begin{array}{cc|cc} p & P & f & F \\ q & Q & g & G \\ \Sigma_n^* & O_{n1} & O_{n,n} & O_{n1} \\ O_{n,n} & O_{n1} & \Sigma_n^* & O_{n1} \end{array} \right] \quad (9)$$

Let the $(2n+2) \times (2n+2)$ permutation matrix P_{erm} changing the order of the columns of A' such that

$$A'' \equiv A'P_{erm} = \left[\begin{array}{cc|cc} p & f & P & F \\ q & g & Q & G \\ \Sigma_n^* & O_{n,n} & O_{n1} & O_{n1} \\ O_{n,n} & \Sigma_n^* & O_{n1} & O_{n1} \end{array} \right] \quad (10)$$

Then we have:

$$A'' \equiv A'P_{erm} \Rightarrow A' = A''P_{erm}^{-1} \Rightarrow (A')^{-1} = P_{erm}(A'')^{-1} \quad (11)$$

Let A'' be rewritten as:

$$A'' = \begin{bmatrix} U & V \\ W & 0 \end{bmatrix} \quad U : 2 \times 2n, \quad V : 2 \times 2, \quad W : 2n \times 2n \quad (12)$$

Then we have

$$(A'')^{-1} = \begin{bmatrix} 0 & W^{-1} \\ V^{-1} & -V^{-1}UW^{-1} \end{bmatrix} \quad (13)$$

(8) Since $|V| = 1$, we have:

$$V^{-1} = \begin{bmatrix} G & -F \\ -Q & P \end{bmatrix} \quad (14)$$

Since W can be written as:

$$W = \Sigma_* \oplus \Sigma_* \quad (15)$$

$$W^{-1} = \Sigma_*^{-1} \oplus \Sigma_*^{-1} \quad (16)$$

$$\Sigma_*^{-1} = \begin{bmatrix} 1 & \sigma & \sigma^2 & \sigma^3 & \cdots & \sigma^{n-1} \\ 0 & 1 & \sigma & \sigma^2 & \cdots & \sigma^{n-2} \\ 0 & 0 & 1 & \sigma & \cdots & \sigma^{n-3} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & \sigma \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

We can easily calculate $-V^{-1}UW^{-1}$. So we have

$$A^{-1} = \Gamma P_{erm}(A'')^{-1} = \Gamma P_{erm} \begin{bmatrix} 0 & W^{-1} \\ V^{-1} & -V^{-1}UW^{-1} \end{bmatrix} \quad (17)$$

of which first three columns are as follows:

$$\left[\begin{array}{cc|cc} \sigma^n G & -\sigma^n F & 1 + \sigma^n(-G\tilde{p}_n + F\tilde{q}_n) & \cdots \\ \sigma^{n-1} G & -\sigma^{n-1} F & 0 + \sigma^{n-1}(-G\tilde{p}_n + F\tilde{q}_n) & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \hline -\sigma^n Q & \sigma^n P & \sigma^n(-G\tilde{f}_n + F\tilde{g}_n) & \cdots \\ -\sigma^{n-1} Q & \sigma^{n-1} P & \sigma^{n-1}(-G\tilde{f}_n + F\tilde{g}_n) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array} \right] \quad (18)$$

2.5. Infinity-norm condition number of A

From Eqs.(18) and (10) we have

$$\begin{aligned} \|A^{-1}\|_{\infty} &> \max[\sigma^n(G+F), \sigma^n(Q+P)] \\ \|A\|_{\infty} &> \max\left[\sum_0^n \{p_i + f_i\}, \sum_0^n \{q_i + g_i\}\right] \\ \sigma^n \sum_0^n p_i &\geq P, \sigma^n \sum_0^n q_i \geq Q, \sigma^n \sum_0^n f_i \geq F, \sigma^n \sum_0^n g_i \geq G, \end{aligned}$$

$$\|A\|_{\infty} \cdot \|A^{-1}\|_{\infty} > \max[(P+F), (Q+G)] \cdot \max[(G+F), (Q+P)]$$

Assume without loss of generality that

$$P > Q, P > F \Rightarrow F > G$$

Then we have a final result (Proof is omitted):

$$\|A\|_{\infty} \cdot \|A^{-1}\|_{\infty} > (P+F)(Q+P) \approx O(\sigma^{2(n+1)})$$

This is a generalization of Rump's result.

3. Extension to Case where V is a 3×3 matrix (Item 2)

3.1. Outline of extension

Let $P, Q, R, F, G, H, L, M, N$ be extremely large integers with the magnitude less than σ^{n+1} and let

$$V = \begin{bmatrix} P & F & L \\ Q & G & M \\ R & H & N \end{bmatrix}, \quad |V| = 1$$

$$P = p_n \sigma^n + p_{n-1} \sigma^{n-1} + \dots + p_0, \quad |p_i| < \sigma \text{ etc.},$$

Let A be a $3(n+1) \times 3(n+1)$ matrix such that

$$A = \begin{bmatrix} \tilde{p} & \tilde{f} & \tilde{l} \\ \tilde{q} & \tilde{g} & \tilde{m} \\ \tilde{r} & \tilde{h} & \tilde{n} \\ \Sigma & 0 & 0 \\ 0 & \Sigma & 0 \\ 0 & 0 & \Sigma \end{bmatrix}$$

$\tilde{p}, \tilde{q}, \tilde{r}, \tilde{f}, \dots, \tilde{n}$: row vectors of order $n+1$

i.e., $\tilde{p} = [p_n, p_{n-1}, \dots, p_0]$, etc.

Let Σ be an $n \times (n+1)$ matrix such that

$$\Sigma_n = \begin{bmatrix} 1 & -\sigma & 0 & 0 & \dots & 0 \\ 0 & 1 & -\sigma & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 & -\sigma \end{bmatrix} \quad (19)$$

A^{-1} can similarly be calculated as previously. Most different point is the calculation of V^{-1} :

$$V^{-1} = \begin{bmatrix} w_{11} & w_{21} & w_{31} \\ w_{12} & w_{22} & w_{32} \\ w_{13} & w_{23} & w_{33} \end{bmatrix}$$

where w_{ij} is a cofactor of the (i, j) element of V .

3.2. Infinity-norm condition number of A

Infinity-norm condition number of A can be calculated as follows:

$$\begin{aligned} \|A^{-1}\|_{\infty} &> \max[\sigma^n(|w_{11}| + |w_{21}| + |w_{31}|), \\ &\sigma^n(|w_{12}| + |w_{22}| + |w_{32}|), \\ &\sigma^n(|w_{11}| + |w_{21}| + |w_{31}|)] \end{aligned}$$

$$\begin{aligned} \|A\|_{\infty} &> \max\left[\sum_0^n \{\tilde{p}_i + \tilde{f}_i + \tilde{l}_i\}, \sum_0^n \{\tilde{q}_i + \tilde{g}_i + \tilde{m}_i\}, \right. \\ &\left. \sum_0^n \{\tilde{r}_i + \tilde{h}_i + \tilde{n}_i\}\right] \end{aligned}$$

Using the formula:

$$|V|V \begin{pmatrix} i & j \\ k & l \end{pmatrix} = w_{ik}w_{jl} - w_{il}w_{jk}$$

we can derive

$$P + F + L = \frac{N}{L}(P + Q + R), \quad \text{etc.},$$

Without loss of generality we assume

$$P, Q, F, G, \text{ and } w_{33} \text{ are very large.}$$

Then we have the final result:

$$\|A\|_{\infty} \cdot \|A^{-1}\|_{\infty} > (P+F+L)(|w_{31}| + |w_{32}| + |w_{33}|) \approx O(\sigma^{3(n+1)})$$

4. Generation of a third order integer matrix with the determinant one

Problem 2: Find a 3×3 integer matrix $A' = [a'_{ij}]$ by bordering the prescribed 2×2 matrix $A = [a_{ij}]$. Here $|a'_{ij}| < \mu$, where μ corresponds to σ^{n+1} in the previous sections.

$$A' = \begin{bmatrix} a_{11} & a_{12} & y_1 \\ a_{21} & a_{22} & y_2 \\ x_1 & x_2 & z \end{bmatrix}, \quad |A'| = 1 \quad (20)$$

$$|a_{ij}| < \mu, \quad |x_i| < \mu, \quad |y_i| < \mu, \quad |z| < \mu \quad (21)$$

Of course we have to impose some restrictions on a_{ij} . **Theorem 1:** An $n \times n$ integer matrix A can be bordered so that its determinant is equal to one, *only if* the $(n-1)$ th determinant (common) divisor is one.

In the case of $n=2$ this means that there is no common factor among all a_{ij} ($i, j = 1, 2$). As a special case we impose:

Assumption 1: a_{21} and a_{22} have no common factor.

Theorem 2: On Assumption 1 we can find x_1, x_2, y_1, y_2 and z satisfying the conditions in Problem 2.

Proof of Theorem 2)

Lemma 1: By the *Euclid algorithm* we can choose x_1 and x_2 such that

$$\begin{vmatrix} a_{21} & a_{22} \\ x_1 & x_2 \end{vmatrix} = 1, \quad |x_i| < \mu$$

Then $|A'|$ can be written as:

$$\begin{aligned} |A'| &= y_1 \begin{vmatrix} a_{21} & a_{22} \\ x_1 & x_2 \end{vmatrix} - y_2 \begin{vmatrix} a_{11} & a_{12} \\ x_1 & x_2 \end{vmatrix} + z \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \\ &\equiv y_1 + ay_2 + bz \end{aligned}$$

Here in general

$$|a| < \mu^2, \quad |b| < \mu^2$$

Lemma 2: Let a and b are integers satisfying

$$|a|, |b| < \mu^2.$$

Then there exist y_2 and z satisfying $|y_2|, |z| < \mu$ and

$$|ay_2 + bz| < \mu.$$

We can prove Lemma 2 by using Euclid algorithm. Using Lemma 2, we can choose y_1 such that

$$|y_1 + ay_2 + bz| = 1, |y_1| < \mu$$

Proof of Lemma 2) Assume that $a > b > 0$ and let $x_0 \equiv a$ and $x_1 \equiv b$. The Euclid algorithm derives the following:

$$\begin{aligned} x_0 &= k_0x_1 + x_2 \quad (0 < x_2 < x_1) \\ x_1 &= k_1x_2 + x_3 \quad (0 < x_3 < x_2) \\ x_2 &= k_2x_3 + x_4 \quad (0 < x_4 < x_3) \\ &\vdots \\ x_{n-1} &= k_{n-1}x_n + x_{n+1} \quad (0 < x_{n+1} < x_n) \\ x_n &= k_nx_{n+1} + x_{n+2} \quad (0 < x_{n+2} < x_{n+1}) \end{aligned}$$

From this we have the following expression:

$$\begin{aligned} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} &= \begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix} \begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix} \\ p_0 &= 1, \quad p_1 = k_0, \quad p_n = p_{n-1}k_{n-1} + p_{n-2} \\ q_0 &= 0, \quad q_1 = 1, \quad q_n = q_{n-1}k_{n-1} + q_{n-2} \end{aligned}$$

The following important formula holds:

$$\begin{vmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{vmatrix} = (-1)^n$$

We therefore have:

$$\begin{bmatrix} x_n \\ x_{n+1} \end{bmatrix} = (-1)^n \begin{bmatrix} q_{n-1} & -p_{n-1} \\ -q_n & p_n \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \end{bmatrix}$$

or

$$\begin{aligned} x_n &= (-1)^n [q_{n-1}x_0 - p_{n-1}x_1] \\ x_{n+1} &= (-1)^n [-q_nx_0 + p_nx_1] \end{aligned}$$

Lemma 2 can be proved by showing $0 < p_n < \mu$ from the following theorem:

Theorem 3: Assume that

$$\mu^2 > x_0 > x_1 > \mu > 0 \quad (22)$$

and let n be an integer such that

$$x_n > \mu, \quad x_{n+1} < \mu \quad (23)$$

Then

$$\mu > p_n > q_n (> 0) \quad (24)$$

Proof of Theorem 3) Assume first that n is even. Then the above conditions means

$$\begin{aligned} q_{n-1}x_0 - p_{n-1}x_1 &> \mu \\ 0 < -q_nx_0 + p_nx_1 &< \mu \end{aligned}$$

Since $p_nq_{n-1} - p_{n-1}q_n = 1$ holds, we have:

$$q_{n-1} = \frac{1 + p_{n-1}q_n}{p_n}$$

which is substituted into the above, we have:

$$\frac{1 + p_{n-1}q_n}{p_n}x_0 - p_{n-1}x_1 > \mu$$

from which we have

$$x_0 + p_{n-1}(q_nx_0 - p_nx_1) > p_n\mu$$

Since $\mu^2 > x_0$ and $0 > q_nx_0 - p_nx_1 > -\mu$ hold, we have:

$$\mu^2 > p_n\mu, \quad \text{i.e., } p_n < \mu$$

This completes the proof for n even. The case of n odd can be treated similarly.

5. Conclusion

We show some generalization of Rump's method to generate extremely ill-condition matrices.

Acknowledgment

This research was supported in part by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Specially Promoted Research, no. 17002012, 2005-2010, on "Establishment of Verified Numerical Computation" and by the Grant-in-Aid for Scientific Research (C) (No. 20560374) of the Ministry of Education, Culture, and Science in Japan.

References

- [1] S.M. Rump, "A class of arbitrarily ill conditioned floating-point matrices", SIAM Matrix Anal. Appl. vol.12, no.4. pp.645-653, Oct. 1991.
- [2] S. Oishi, K. Tanabe, T. Ogita, S. M. Rump: Convergence of Rump's Method for Inverting Arbitrarily Ill-conditioned Matrices, Journal of Computational and Applied Mathematics, 205:1 (2007), 533-544.
- [3] T. Ohta, T. Ogita, S. M. Rump, S. Oishi: Numerical Verification Method for Arbitrarily Ill-conditioned Linear Systems, Trans. JSIAM, 15:3 (2005), 269-287. [in Japanese]
- [4] S. Rump, "Inversion of extremely ill-conditioned matrices in floating-point," to be submitted, 2008.
- [5] Teiji Takagi, "Lectures on Algebra (2nd ed.)," vol.20, pp.130-141, 1965.[in Japanese]
- [6] Teiji Takagi, "Lectures on Elements of Integer Theory," vol.20, pp.130-141, 1971.[in Japanese]