Petroleum and Coal

Article **Open Access**

SOME PROBLEMS IN APPROXIMATING PROCESSES AND OBJECTS RELATED TO MAJOR OIL PIPELINES WITH ALGEBRAIC POLYNOMIALS

*A.V. Maystrenko[1], A.A. Svetlakov[1], T.V. Gandzha[1], V.M. Dmitriev[1], Natalia V. Aksenova[2]**

[1] *Tomsk State University of Control Systems and Radio-electronics, Tomsk, Russia*
[2] *Division for foreign languages, School of Core Engineering Education, Tomsk Polytechnic University, Tomsk, Russia*

Received May 22, 2019; Accepted July 31, 2019

Abstract

The article fomulates supposed causes of Hilbert matrices' ill conditioning from analysis of their most important properties. Those causes are used to propose and develop methods to improve the conditioning of the Hilbert matrices on the basis of normalizing their rows and columns.

Keywords: *Matrix, Norm, Element, Scaling, Representation of numbers.*

1. Introduction

Modeling of various processes and objects, as well as the creation of integrated mathematical models of major oil pipelines often involves application of polynomial approximation of some characteristics of the processes [1]. In its turn, it necessitates studies into different algorithms for solution of systems of linear equations (SLE), matrix inversion, etc. Hilbert matrices are widely employed to these ends, are interesting in that they are ill-conditioned even for relatively low matrix dimensions [2]. This very feature determines their wide applicability as some kind of whet stone to check and refine the properties of various algorithms of SLE solving, matrix inversion, etc.

This paper presents an analysis of the most important properties of Hilbert matrices and formulates two suspected causes of their ill-conditionality. Some experimental results are given, that were obtained from research conducted with the aim to confirm or disprove hypothetical causes of such matrices' ill-conditionality.

2. Analysis of features Hilbert matrix

At it is known [2], the Hilbert matrix often appears in solving the problem of approximating a function with algebraic polynomials. At that, approximation error is usually evaluated with Euclidean metric. As this metric is a differentiable function with respect to each unknown coefficient of the polynomial, its minimum is attained when partial derivatives are equal to zero for each of the coefficients. Having completed all the necessary operation, we get a SLE of the following form:

$$H\overset{\downarrow}{c} = \overset{\downarrow}{b}, \quad (1)$$

where H is a square matrix of degree m, whose elements are determined with equations

$$\boldsymbol{h_{ij}} = \frac{\boldsymbol{1}}{\boldsymbol{i+j-1}}; \quad \boldsymbol{i,j} = \overline{\boldsymbol{1,m}}, \quad (2)$$

$\overset{\downarrow}{c}$ is a m -dimensional column vector composed of unknown coefficients c_i; $i = \overline{1,m}$ of the polynomial; $\overset{\downarrow}{b}$ is a m -dimensional column vector, whose components are computed in accordance with the equation:

$$b_i = \int_0^1 y(t)\ t^{(i-1)}dt = 0,\ i = \overline{1,m}, \quad (3)$$

where $y(t)$ is the approximating function.

The matrix H obtained above is a Hilbert matrix, hereinafter designated as H_m, where the index means its degree.

Let us consider those features of the H_m matrix, which is the most interesting. The first three of them are found when considering the equation (2) and are as follows:

1. Matrix H_m is symmetric with respect of its main diagonal and all its elements are rational numbers. At that, its maximum element is the h_{11}, while its minimum element is h_{mm} and they are determined with equations

a) $h_{11} = 1$ and b) $h_{mm} = 1/(2m-1)$,

All the elements of the H matrix, with the exception of $h_{11} = 1$ are strictly less than one.

2. Diagonal elements h_{ii} of the H_m matrix are steadily decreasing with increasing values of the index. With any finite m, the ratio h_{11}/h_{mm} between its maximum element and its minimum element conform to the equality

$h_{11}/h_{mm} = 2m-1.$ (4)

3. The components of the jth column of the matrix H_m with increasing index j decrease steadily and with its unlimited increase tend to zero. For any finite m, the ratio $\delta_j = h_{1j}/h_{mj}$ conforms to the equation

$\delta_j = h_{1j}/h_{mj} = \frac{1}{1+j-1} / \frac{1}{m+j-1} == (m+j-1)/j = 1 + (m-1)/j, j = \overline{1,m}.$ (5)

As it is evident from this equation, for any column of the matrix H_m the ratio δ_j is strictly more than unity. From this equation, it is also evident that:

a) with an increase in index j, the ratios δ_j steadily decrease, while with its increase they tend to unity;

b) for any finite m the ratio of maximum δ_{1j} to minimum δ_{mj} conforms to equation

$\delta_{1j}/\delta_{mj} = \frac{(m+1-1)}{1} : \frac{(m+m-1)}{m} = \quad = \frac{m \cdot m}{2m-1} = \frac{m}{2-1/m},$ (6)

from which it is evident that with increasing m, this ration increases.

The fourth feature of the matrix H_m is vividly illustrated by results taken from [2] and presented in Table 1 below there, the following notational conventions are used: H_m, T_m, $\|H_m\|$ and $\|T_m\|$ are Hilbert matrix, its inverse and their Euclidean norms respectively; t_m^{max} is the maximum element of the matrix T_m; C_m is the conditioning number of the matrix H_m.

Euclidean norms $\|H_m\|$ and $\|T_m\|$ of matrices H_m and T_m respectively, as well as the conditioning number C_m were calculated in accordance with the equations:

a) $\|H_m\| = \left(\sum_{i=1}^{m}\sum_{j=1}^{m} h_{ij}^2\right)^{\frac{1}{2}}$;

b) $\|T_m\| = \left(\sum_{i=1}^{m}\sum_{j=1}^{m} t_{ij}^2\right)^{\frac{1}{2}}$ and

c) $C_m = \|H_m\| \cdot \|T_m\|$, (7)

where t_{ij} is the (i,j) th element of the matrix T_m.

From Table 1 it is evident that the norm $\|H_m\|$ monotonously and gradually increases with increasing m, its increase monotonously slowing down. The norm $\|T_m\|$ is also increasing monotonously with increases in m; however, its increase is of explosive kind.

Table 1. Characteristics of Hilbert matrix obtained from double accuracy calculations (8 bytes).

m	2	3	4	5	6	8	10
$\|H_m\|$	1.27	1.41	1.5	1.57	1.62	1.7	1.75
$\|T_m\|$	15.2	3.72E+2	1.03E+4	3.04E+5	9.24E+6	9E+9	9.15E+12
C_m	1.93E+1	5.24E+2	1.55E+4	4.77E+5	1.5E+7	1.5E+10	1.6E+13
t_m^{max}	1.2E+1	1.92E+2	6.48E+3	1.79E+5	4.41E+6	4.25E+9	3.48E+12

Similarly, with the increase in m the values of the conditioning number C_m and maximum element t_m^{max} change as well.

Taking into account a well-known continuous nature of dependence that each element t_{ij} of the matrix T_m has on the elements h_{ij} of the matrix H_m and the above-noted features of the matrix H_m, it is justified to assume that potential causes of ill-conditionality of this matrix are the following two factors:

1) limited word size of computers used to compute the matrix inverse H_m;
2) difference of scale between the elements h_{ij} of the matrix H_m and Euclidean norms of its rows $\vec{h}_i,\ i=\overline{1,m}$ and columns $\overset{\downarrow}{h}_i,\ i=\overline{1,m}$.

In order to check the validity of these assumptions, experimental research has been conducted.

3. Objectives and conditions of experimental studies of H_m matrix

The objective of the experimental studies was, primarily to:
1) establish, which of the two above mentioned factors and to what extent influences quantitative characteristics of a matrix H_m and accuracy of computing the inverse matrix T_m;
2) compare the quantitative criteria, whose numerical values allow assessing the accuracy of computing the matrix T_m;
3) construct a dependence of the conditionality number C_m and maximum element t_m^{max} of the matrix T_m on the degree m of the matrix.

At that, the values are given in Tables 1–4 were used as quantitative characteristics of the matrix.

As a measure of the accuracy of computing the matrix T_m , the authors used square (Euclidean) norm $||R_m||$ of the matrix R_m, determined with an equation

$$R_m = T_m H_m - E_m, \tag{8}$$

where E_m is an identity matrix of degree m.

Computation of the principal characteristics of the matrices H_m and T_m were performed on an IBM PC machine using various accuracy of an internal representation of real numbers and using the arbitrary precision arithmetic.

4. Research results, characterizing the influence of limited word length of a computer onto accuracy and possibility of computing characteristics of H_m matrix

Three series of experiments have been conducted to study the influence of limited word length. At that, the first of them used single precision in the internal representation, while the second and the third used double precision and maximum available (10 bytes), respectively.

The results of these experiments are given in Table 2 that uses the following notation: Single(4b), Real(8b), Ext (10b) represent the precision of the internal representation – short, double and maximum, respectively; the NAN value means that a word size overflow happened in accumulator or multiplicator and thus the result cannot be obtained.

As it is evident from Table 2, determinant of the matrix T_m and calculation error $||R_m||$ increase very fast, which is determined by the increase in the matrix elements T_m. The determinant of the matrix T_m is accurately computed for orders up to the fourth, while matrix elements are accurately computed up to and including the fifth degree, after that the computation error is steadily increasing. Even with the maximum internal representation precision, the limiting degree of the matrix H_m, for which the matrix T_m is computable with relatively high precision is equal to 14.

Studying Hilbert matrices of higher degrees requires increasing computation accuracy by means of increasing interior representation length. The main issue here is switching to arbitrary precision arithmetic that reneges on the internal representation of real numbers and stores them as arrays of significant digits. In this case, the only limiting factor is the necessity to allocate necessary amount of memory to store and necessary time to process such large arrays of data. In practice it allows studying the principal properties of Hilbert matrix up to $m = 100 - 200$, without using parallel computations or supercomputers. Having 100 significant digits in computations is enough for analysis of Hilbert matrices with m<50.

Comparison between Table 1 and Table 3 reveals the two main differences in the results reflected. The first one is that dimensions of Tables 1 and 3 differ from each other. Table 1 that characterizes the properties of the matrix H_m ends with degree $m = 10$, this is due to the fact that double precision prohibits obtaining the included characteristics for matrices of higher degrees. The second difference is that the values of $||H_m||$ computed with double precision accumulate differences from those computed with arbitrary precision arithmetic with increasing m. It is due to peculiarities of computational algorithms of the modern computers that use 20 significant digits maximum in their representation of real numbers.

Results from the computation of norms of the Hilbert matrices are the same with short and double precision, they are well-known [2] and have been previously described in sufficient detail; thus we are not going to give them here. Let us note, however, that with higher degree of the matrix H_m, the value of its norm grows; this dependence is of monotonously increasing nature. However, application of the arbitrary precision arithmetic revealed some differences in the values of norms of the Hilbert matrices. For example, there are some differences in hundredth starting from the 4th degree (Table 1 and Table 3). It is an indirect evidence that computations that use arbitrary precision arithmetic may significantly expand the scope of tasks that have been unsolvable to regular methods.

Table 2. Characteristics of Hilbert matrix obtained with different internal representations

m	$\Vert R_m \Vert$ Single(4b)	Real(8b)	Ext (10b)	$\det H_m$	$\det T_m$	$t_m^{\max}$
1	0	0	0	1	1	1
2	1.99E-8	3.71E-17	0	0.08333	12	12
3	2.63E-6	8.86E-15	4.34E-19	0.0004629	2160	192
5	0.0020	2.38E-12	3.73E-15	3.7492E-12	**2.667E+11**	179200
7	1.26	3.42E-9	3.65E-12	5.3672E-18	2.07E+24	1.33E+8
8	7.89	1.09E-7	1.01E-10	5.3672E-18	3.65E+32	4.25E+9
9	NAN	3.61E-6	2.32E-9	5.3672E-18	1.03E+42	1.22E+11
11	NAN	0.002	2.21E-6	5.3672E-18	3.31E+64	1.18E+14
13	NAN	2.23	0.0025	5.3672E-18	6.93E+91	1.06E+17
15	NAN	3.23	4.12	5.3672E-18	3.45E+124	2.13E+20

Table 3. Characteristics of Hilbert matrix obtained with arbitrary precision arithmetic

m	2	3	4	5	10	12	15
$\Vert H_m \Vert$	1.27	1.41	1.51	1.58	1.79	1.84	1.9
$\Vert T_m \Vert$	15.2	3.72E+2	1.03E+4	3.04E+5	9.15E+12	9.54E+15	3.31E+20
C_m	19.3	5.26E+2	1.56E+4	4.8E+5	1.63E+13	1.75E+16	6.28E+20
$t_m^{\max}$	12	1.92E+2	6.48E+3	1.79E+5	3.48E+12	3.66E+15	1.15E+20

Indeed, the application of arbitrary precision arithmetic allowed obtaining results that significantly differ from those considered above. Here we are going to go into more detail, as they are of significant interest. Let us provide a number of graphical dependences shown in Fig.1-3 and illustrating correctness of our assumptions. In all the figures, the degree of the studied matrix is plotted along the X axis, while the values of the studied variables are plotted along the Y axis.

From the graphs, it is evident that all the characteristics shown are linear. This development is important enough, as it provides an accurate representation of the behavior of the studied values, which in its turn allows resolving certain problems where Hilbert matrices (or their analogs) may arise with some preliminary knowledge.

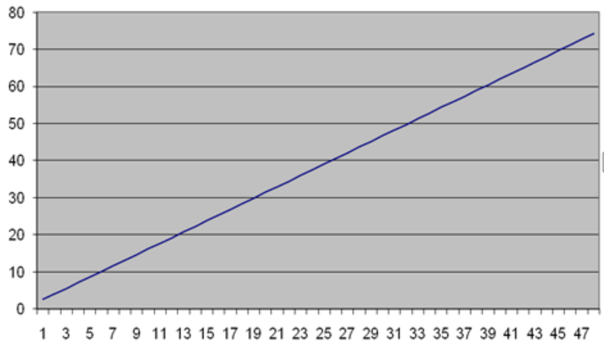

Fig. 1. Common logarithm of Hilbert matrix conditionality number as a function of its degree

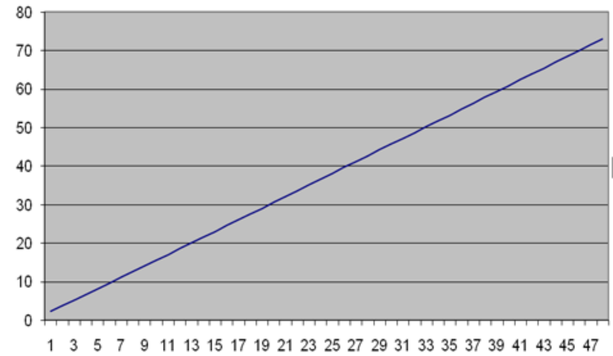

Fig. 2. Common logarithm of Hilbert matrix maximum element as a function of its degree

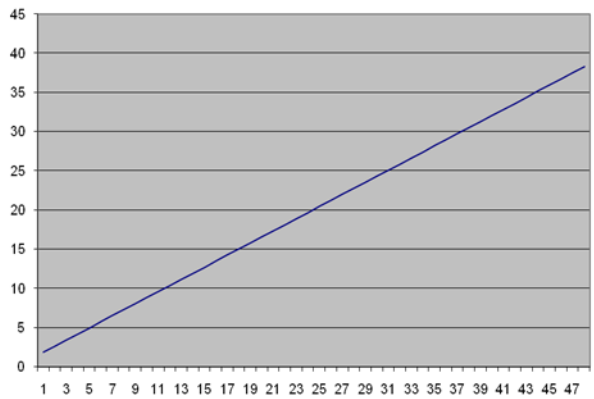

Fig.3. Determinant of Hilbert matrix $\sqrt{-\boldsymbol{log}(\boldsymbol{det}\,\boldsymbol{H_m})}$ as a function of its degree

5. Research results characterizing influence that difference of scale between Hilbert matrix elements has onto its conditionality

To study the influence of difference of scale of matrix elements and Euclidean norms of Hilbert matrix rows and columns onto its conditionality, several experiments were conducted on Hilbert matrices scaled along the rows. A normalized Hilbert matrix H_m^N was produced from the initial matrix with the following sequence of operations:

1) a diagonal matrix D_m of order, m was formed, where along the main diagonal there are elements d_{ii}, computed in accordance with the formula

$d_{ii} = ||\vec{h}_i|| = \left(\sum_{j=1}^{m} h_{ij}^2\right)^{\frac{1}{2}},\ i = \overline{1,m};$

2) normalized Hilbert matrix H_m^N was obtained by postmultiplication of an inverse diagonal matrix D_m^{-1} by the Hilbert matrix

$H_m^N = D_m^{-1} H_m.$

During the research, it has been revealed that the results with single and double precision are practically identical, while application of the arbitrary precision arithmetic allows getting results for matrices of significantly higher orders. At that, it has also been revealed that there are no practical differences between the results for scaling along the rows and along the columns. Thus, only results for row-normalized Hilbert matrices are given below.

Table 4 summarizes comparative characteristics of the values of the norm of matrices, conditionality numbers and maximum element values of the compared matrices. From the table it is evident that with increasing order m of Hilbert matrix H_m, all the above mentioned characteristics of the normalized matrix H_m^N decrease by a factor ranging from 2 to 7 in comparison with the same characteristics of the initial Hilbert matrix. At that, there is a clear dependence on the degree m: the higher the degree, the larger is the difference between the compared values. The values of maximum elements increase with increased order of matrix. The same is characteristic of the determinants of the compared matrices.

Table 4. Comparative characteristics of normalized and non-normalized Hilbert matrices

m	10	20	30	40	50
$\|H_m\|$	1.79E+00	1.97E+00	2.07E+00	2.14E+00	2.19E+00
$\| H_m^N \|$	3.16E+00	4.47E+00	5.48E+00	6.32E+00	7.07E+00
$\|T_m\|$	9.15E+12	1.29E+28	2.13E+43	3.75E+58	6.85E+73
$\|T_m^N\|$	2.72E+12	2.66E+27	3.58E+42	5.45E+57	8.88E+72
C_m	1.63E+13	2.53E+28	4.41E+43	8.03E+58	1.50E+74
C_m^N	8.60E+12	1.19E+28	1.96E+43	3.45E+58	6.28E+73
$t_m^{\max}$	3.48E+12	3.61E+27	5.05E+42	7.84E+57	1.29E+73
$t_m^{\max N}$	1.06E+12	7.56E+26	8.46E+41	1.13E+57	1.66E+72

Thus, it may be concluded that row-scaling a Hilbert matrix may be quite useful in the solution of a certain scope of problems.

6. Conclusion

The conducted research has revealed that the results of studying the principal characteristics of Hilbert matrices differ significantly depending on the interior representation of real numbers employed in computation. At that, the accuracy clearly depends on the internal representation of real numbers: the higher the accuracy of representation, the more accurate the results. The results from single and double precision diverge quite strongly and allow obtaining the main characteristics only for Hilbert matrices of order 10 or lower. Application of maximum available internal representation allows somewhat expanding the possibilities and getting the desired characteristics for matrices up to the order of 15.

Application of arbitrary precision arithmetic allows obtaining the main characteristics of Hilbert matrices of order 100 or higher, which is fundamentally impossible for single, double, or even maximum available floating point representation.

Normalizing rows or columns of a matrix is a method to improve its conditionality. In some cases, the effect is very significant. The comparison of results for Hilbert matrices before and after scaling allows for the following conclusions: scaling allows improving high-order matrix conditionality and transforms the dependence of its main characteristics to a form that is more useful for practical applications and analysis. However, it does not allow radically improving matrix conditionality, nor does it work for arbitrary large orders.

References

[1] Maystrenko AV, Svetlakov AA, Gandzha TV, Dmitriev VM, Aksenova NV. Application of numerical signal differentiation methods to determine stationarity of a process. Petroleum and Coal, 2017; 59(3): 311-318.

[2] Forsythe G, Moler CB. Computer Solution of Linear Algebraic Equations. (Russian Edition) – Moscow: Mir, 1969. – 167 pp.

To whom correspondence should be addressed: Dr.Natalia V. Aksenova, Division for foreign languages of the School of Core Engineering Education, Tomsk Polytechnic University, 30, Lenin Avenue, Tomsk, 634050, Russia; tel.: (+7-923)4577428, E-mail: polozova15@tpu.ru