92-247

сообщения
объединенного
института
ядерных
исследований
дубна

E10-92-247

V.I.Ilyushchenko

INFORMATION LOSSES IN SOLVING UNSTABLE INVERSE PROBLEMS

1992

1. Introduction

The vast majority of integral and differential equations of mathematical physics are known to be discretized in the form of systems of linear algebraic equations (SLAE). Moreover, analogous systems result from the analysis of the general information processing structure:

(1)

Where T stands for a true (complete) input information, B^2 denotes an information processing Black Box and F is an output (experimental) data set specified by some error (noise) component N.

The term T can be exemplified by an ideal δ-function-like gamma-spectrum, with B^2 incorporating a gamma-ray detector and/or processing unit, while F is an experimental gamma-spectrum with residual background, statistical and systematic errors included into N.

Another example deals with structure functions $F_2(x,Q^2)$ studied in high energy physics by means of giant supercolliders and superdetectors. Here T is a lepton-lepton, lepton-hadron, or hadron-hadron interaction product, B^2 - a detector+computer unit and $F = F_2(x)$ - a published one-dimensional structure function averaged over some Q^2-range.

2. Unstable inverse problems (UIP)

All the above-mentioned mathematical structures can be discretized as the SLAE

$$At = f + n \qquad (2)$$

where A - a matrix apparatus function, t - a true solution vector, and f - a vector set of experimental data measured with an additive error n.

The simplest inverse filter (IF) solution of the UIP for the SLAE (2) looks like

$$t = A^{-}(f + n) \quad (3)$$

which is valid only for square nonsingular (nondegenerate) A-matrices with inverses A^{-}.

In case of A-matrix being rectangular or/and singular (degenerate) the relevant solution is

$$t = A^{+}(f + n) \quad (4)$$

where A^{+} stands for a Moore-Penrose pseudomatrix which affords a solution in the least squares (LSQ) sense.

For a square nonsingular A-matrix both inverses coincide, $A^{-}= A^{+}$.

The seeming mathematical simplicity of the general solutions (3) and (4) leads to a generally accepted false conclusion about their stability and reliability. This misleading conclusion is especially popular among the members of the vast community of push-button users participating in international collaborations with standard computer libraries and professional packages.

However, at the beginning of this century the French mathematician J.Hadamard (1865-1963) discovered that inverse problems are generally ill-posed and,as a consequence, their solutions are ill-conditioned, i.e. unstable [1]. In addition, it is useful to recall here that the first analytical solution of one of the UIP's, i.e. that of reduction to an ideal spectrometer, was obtained as early as 1871 by the English physicist Rayleigh (1842-1919) [2].

The subsequent development of events has shown the solution of the UIP to be one of the most difficult items of the mathematical processing of experimental data.

3. Main sources of information losses

3.1 A simplified frequency spectrum model of information

Let us assume that the true solution

$$T = \sum_{i} A_i W_i \quad (5)$$

where A_i are amplitudes and W_i frequencies. Then each constituent member of B^2 (1) will act as a low-pass filter (LPF). In most cases, after interaction, detection, tracking, identification, processing etc one observes irretrievable losses of useful physical information contained in T, i.e. there takes place a set of successive truncations of the high-frequency part of information and a, generally nonlinear, distortion of its low-frequency part with an accompanying increase in the noise component, N. In short, one observes here an order (information)-disorder (chaos) transition within the informative system under study.

It will be appropriate to note here that such highly popular (user's folklore) processing means as smoothing of experimental data, LSQ curve-fitting etc intended to "improve" information content of the available input information sets act as pure LPF information cutters. The subsequent attempts to retrieve the lost useful information are vain except of some rare exclusions.

For example, let us consider the case of a muon resonance spectrum composed of several broad peaks. By applying a standard Fast Fourier Transform (FFT) code one obtains an image spectrum with spectral peaks placed in the low-frequency domain and a noise component, in the high-frequency one. By means of some direct (analytical) or indirect (iterative) algorithm it is possible to optimize the separation and a subsequent elimination of the noise component from the input image spectrum. This is equivalent to the application of a stabilizing rectangular window LPF. This done, one can use an analytical continuation of the "pure" spectrum contained within the stabilizing window to obtain the high-frequency part lost during the filtering. Unfortunately, this is possible only for highly analytical "pure" spectra, i.e. monotone, differentiable etc.

Here it will be useful to remind about an information trap in the form of the direct and inverse FFT's which are intrinsically unstable in the UIP sense.

Another example of using compensating physical hypotheses is an assumption of positiveness of the true spectrum, T. Here again one can miss an information trap due to an incorrect background subtraction from experimental raw data. If this background is underestimated, the resulting spectrum will contain a smooth underlying remnant part. In case of an overestimated background the resulting spectrum will acquire negative elements protruding below the abscissa axis. In addition, the problem of statistical errors is complicated here due to the correlated nature of the spectrum+background compound produced by the same projectile beam.

Notwithstanding these and similar facts, the introduction of a complete set of additional physical and mathematical hypotheses is the only effective strategy to remedy the information losses mentioned above.

3.2 Structure of A-matrix

Other aspects of information losses are due to an involved structure of the apparatus A-matrix and multiple error sources in both A and F.

First, even a simple A-matrix can be factorized as

$$A = T*R*X*D \quad (6)$$

where T - an interaction transform factor, R - a resolution factor, X - an acceptance factor and D - a last term to account for residual - and usually unknown - information about A. In case of D = I, where I is the diagonal identity matrix one assumes A =T*R*X to be a rigorous factorization of the A.

Second, any A-matrix is prone to the following errors [3]:

$$A = A_0 + E_P + E_M + E_C \quad (7)$$

where A_0 - an error-free core, E_P - the physical errors due to detector inhomogeneities, false tracking etc, E_M - the mathematical errors caused by, e.g. an inadequate or erraneous mathematical model of acceptance computation in the form of the so-called factor method [4], and E_C - the computational errors arising from rounding-off effects, e.g. in a single-precision floating point mode of operation, fl_1[3].

For example, our test studies indicate the threshold for an E_C -induced instability to arise at values << M^3*MACHEPS, where M is the dimensionality of the A-matrix and MACHEPS is the machine epsilon. Our IBM-compatible computer EC-1055M is specified by

$$MACHEPS = 0.4E-6 \quad (8)$$

and the E_C -induced instability threshold corresponds to

$$M \le 20 \quad (9)$$

There are some evident means to overcome this problem, i.e. an A-matrix dimension windowing or tearing, a transfer to an interval processing [5] or error-free computations [6].

The most important conclusion from our computational experience is that all three instability thresholds induced by E_P, E_M or E_C are comparable.

However, our studies of the dependence of spectral patterns (envelopes) demonstrate a very persistent nature of the spectrum pattern conservation. For example, in processing the test Rayleigh doublet in Fourier original and image domains with E_P = 100% one observes solution modes with upside-down negative spectrum patterns, i.e. with very large accompanying CHI2-estimates but quite acceptable peak feature reconstructions.

These observations contradict many previous results obtained for low-level Ep-errors.

4. An illustrative example

The thorough inspection of the standard matrix-vector SLAE form (2) demonstrates the irreversible character of its equality sign. From the point of view of reversible mathematics [7] the "equation" (2) must be considered as a reduced equation (requation) like

$$At \longrightarrow f + n \qquad (10)$$

because it is impossible to retrieve its left-hand-side (l.h.s.), At, from the right-hand-side (r.h.s.), f + n. Moreover, given t,f and n, it is impossible to reconstruct the A-matrix.

The most effective way of repairing this main information loss is to transform the matrix-vector SLAE (2) into an all-matrix SLAE [8] which affords a few independent vector versions of the true solution, T, instead of a single standard one, t. Our computational experience shows this transform to be one of the most effective novel tools in analyzing the potential sources of different instabilities and information losses.

5. The main strategy to eliminate noise by A-matrix spectral corrections

Let us use a simplified all-matrix SLAE form like:

$$AT = F \qquad (11)$$

and recall that the well-known eigenproblem decomposition looks like

$$AV = VL \qquad (12)$$

which is valid for a nonsingular eigenvector matrix V and where L stands for a diagonal eigenvalue matrix of A. By substituting $A^{-} = V^{-}L^{-}V$ from (12) into the IF-solution of (11) one obtains

$$T = (V^{-}L^{-}V)*F \qquad (13)$$

the subsequent rather trivial transformations lead to

$$T \approx 1/l_i \qquad (14)$$

thus underlying the dominant effect of small eigenvalues, l_i, in computing the true spectrum, T.

Our computer code was based on a modified version of this algorithm to enable the correction of final T through a truncation or correction of the lowest eigenvalues. With relative r.h.s. errors of 1-100% the acceptable reconstruction of a test Rayleigh doublet was obtained for the UIP's with a number of discretization bins, NBIN = 4-128.

The same strategy of suppressing the destabilizing noise effect by a correction of A-matrix spectrum was implemented with the pseudomatrix A^+ through singular value decomposition [9, 10, 11]. Here again it is possible to use a singular value spectrum truncation, correction or both.

The processing results obtained by means of these two techniques served to crosscheck the outputs provided by different other techniques including those based on FFT, complex pseudoinverse etc.

6. Orthogonal erosion of T

One of the mathematical sources of instabilities in solving the UIP's is due to the well-known Riemann-Lebesgue lemma which claims that

$$\int_{\alpha \to \text{OXO}} A(x,s)\sin(\alpha s)ds \longrightarrow 0 \qquad (15)$$

by translating this lemma into a slightly different semantic field, it is important to note the orthogonal nature of the second integrand factor, i.e. $\sin(\alpha s)$. A trivial conclusion is that any T-orthogonal factor in the l.h.s. of requation (15) is matched by a corresponding zero-matrix term, Z, in its r.h.s.

Thus, in performing a multiplication of matrices A and T one can meet a case when

$$A*T_O = Z \qquad (16)$$

where T_O is an A-orthogonal part of T-matrix [10]. Here the pure mathematical loss of information is most evident since the l.h.s. of req.(16) contains the two non-zero factors, A and T_O , while its r.h.s. is the primitive zero matrix, Z.

However, although the authors of [12] claim the presence of a nonorthogonal "visible" component, T_{vis}, and an A-orthogonal "invisible" component, T_O, with $T = T_O + T_{vis}$, the principal nonexistence of such a unique matrix additive

decomposition and the total independence observed between T_O and F press for a negative final conclusion about the real existence of this two-component solution. On the other hand, such partial additive decompositions can be easily realized via a standard QR- or QL-algorithm based on a modified Gram-Schmidt orthogonalization procedure [13]:

$$A = QL = LQ \qquad (17)$$

$$LQ(T_O + T_{vis}) = F \qquad (18)$$

$$Q(T_O + T_{vis}) = L^{-} F \qquad (19)$$

$$QT_{vis} = L^{-} F + Z \qquad (20)$$

In principle, T_{vis} can be further expressed in an analogous additive factorization form etc.

This orthogonality erosion problem needs to be investigated in more detail.

7. Conclusions

The ill-posed nature of the UIP's leads to their ill-conditioned solutions. This ill-conditioning is due to multiple sources of information losses and/or errors which can be classified as physical, mathematical, computational etc. The different input r.h.s. errors are enhanced by the condition number of the apparatus A-matrix, COND(A), thus giving rise to large distortions in final solutions, T. A natural remedy to correct this disastrous effect is the control of eigenspectrum and/or singular value spectrum of the relevant A-matrix.

To sum up, the author expresses his deep gratitude to the late Profs. C.Lanczos, A.S.Householder and G.E.Forsythe for their profound insight into the fundamental problems under study.

References

1. J.Hadamard, Bull.Univ. Princeton, 13 (1902).
2. Rayleigh, Sci. Papers, 1, 135 (1871).
3. L.Fox and D.F.Mayers, Computing Methods for Scientists and Engineers, Clarendon Press, Oxford (1968).
4. V.Blobel, 1984 CERN School of Computing, CERN 85-09 (1985).
5. G.Alefeld and G.Herzberger, Introduction to Interval Computations, Academic Press, N.Y.(1983).
6. R.T.Gregory and E.V.Krishnamurthy, Methods and Applications of Error-free Computations, Springer, N.Y. (1984).
7. V.I.Ilyushchenko, Reversible Mathematics - An Artificial Science or A Scientific Art ?, JINR Preprint (to be published) (1992).
8. V.I.Ilyushchenko, From Matrix-vector SLAE'S to All-matrix ones, JINR Preprint (to be published) (1992).
9. C.Lanczos, Linear Differnetial Operators, Van Nostrand, N.Y.,(1961).
10. A.S.Householder, The Theory of Matrices in Numerical Analysis, Blaisdell, Johnson (1964).
11. G.E.Forsythe, M.A.Malcolm, and C.B.Moler, Computer Methods for Mathematical Computations, Prentice-Hall,Englewood Cliffs (1977).
12. V.B.Anykeyev, A.A.Spiridonov and V.P.Zhigunov, NIM, A303, 350 (1991).
13. L.Strang, Linear Algebra and its Applications, Academic Press, N.Y. (1976).

Received by Publishing Department
on June 9, 1992.