IMPROVED NUMERICAL METHOD FOR SOLVING INVERSE PROBLEMS OF ATMOSPHERIC OPTICS

Dubovik O.V., Oshchepkov S.L., Lapyonok T.V.
Institute of Physics, Belarus Academy of Sciences,
Scorina pr. 68, 220072, Minsk, Belarus

A *new numerical procedure is proposed for solving a set of algebraic equations being involved in typical problems to monitor optically atmospheric objects. The method is based on the statistical estimation theory, can use arbitrary a priori information, and provides* an *increased stability of the solution.*

The numerical solution is a universal way to solve a mathematical equation. For example, the Fredholm integral equation of the first kind $\sigma(\theta)=\int_0^\infty K(\theta,r)\varphi(r)dr$ is reduced very often to a set of linear algebraic equations

$$\vec{\sigma}^* = K\,\vec{\varphi} + \vec{\Delta}\,, \qquad (1)$$

where $\vec{\sigma}^*$ is the vector-column of measurements (σ^*_1 ...σ^*_m), $\vec{\varphi}$ the vector-column of searched parameters (φ_1 ...φ_n), K the matrix of first derivatives $K_{ij}=\partial\sigma_j/\partial\varphi_i$. Eq. (1) is common for many problems, such as the estimation of aerosol size distribution, auroral and atmospheric tomography, temperature and humidity atmosphere profiling, etc.

The main difference of the inverse ($\vec{\sigma}^* \rightarrow \hat{\vec{\varphi}}$) and of the direct ($\vec{\varphi} \rightarrow \vec{\sigma}$) problems is the capability of noise $\vec{\Delta}$ in initial data $\vec{\sigma}^*$ to influence strongly on the quality of a solution $\hat{\vec{\varphi}}$. That is why a procedure of solution optimization concerning the noise influence is necessary to interpret real measurements $\vec{\sigma}^*$. Because of a random character of the noise, the statistical estimation approach is quite suitable for this [1-3]. The optimal transformation of the probability density function (PDF, likelihood function) of initial data $P(\vec{\sigma}(\vec{\varphi})|\vec{\sigma}^*)$ to PDF of the searched parameters $P(\vec{\varphi}|\vec{\sigma}^*)$ is the purpose of the statistical estimation. The method of maximum likelihood produces the optimal solution characterized by the minimal variance of any parameter $\hat{\varphi}_i$ estimation. For normal noise (substantial noise for many situations [1]), the maximum likelihood solution of (1) is the LSM solution:

$$(K^T C^{-1}_{\sigma^*} K)\,\vec{\varphi} = K^T C^{-1}_{\sigma} \vec{\sigma}^* , \qquad (2)$$

where C_{σ^*} - covariance matrix of $\vec{\sigma}^*$. Consequently the LSM solution is optimal and unsatisfactory estimations of LSM can be explained by the following reasons

1) the insufficient information content of $\vec{\sigma}^*$ concerning searched parameters φ_i;
2) an incorrect preliminary assumption (e.g. noise distribution assumption);
3) disadvantages of standard operations to solve Eq.(2)

For a situation of low information content of basic experiment, the only way to improve solution is to add new data. To perform this operation, there are many methods formulated on principals different from above ones. For example,

$$(K^T K + \gamma I)\,\vec{\varphi} = K^T \vec{\sigma}^* + \gamma\,\vec{\varphi}^0 \quad \text{- Twomey [4],} \qquad (3)$$

$$(K^T K + \gamma \Omega)\,\vec{\varphi} = K^T \vec{\sigma}^* \quad \text{- Phillips [5],} \qquad (4)$$

$$(K^T K + C^{-1}) \vec{\varphi} = K^T \vec{\sigma}^* \qquad \text{- statistical regularization method [6],} \tag{5}$$

where I is the n x n - unit matrix, $\vec{\varphi}^0$ the sample solution, Ω the smoothing matrix, C the *a priori* covariance matrix, γ the regularization parameter.

In statistical approach the use of additional data can be implemented by solving the joint system

$$\begin{aligned} \vec{\sigma}^* &= K \vec{\varphi} + \vec{\Delta} \\ \vec{\beta}^* &= B \vec{\varphi} + \vec{\Delta}_1 \end{aligned} \tag{6}$$

where $\vec{\beta}^* = B \vec{\varphi}$ - a set of equations for the known additional data $\vec{\beta}^*$. Assumption on normally distributed noise in all the data of (6) and on the statistical independence of $\vec{\sigma}^*$ and $\vec{\beta}^*$ gives the LSM solution of (6) to be a solution of the equation system

$$(K^T C^{-1}_{\sigma^*} K + B^T C^{-1}_{\beta^*} B) \vec{\varphi} = K^T C^{-1}_{\sigma^*} \vec{\sigma}^* + B^T C^{-1}_{\beta^*} \vec{\beta}^*, \tag{7}$$

where C_{β^*} - covariance matrix of β^*.

In particular cases, Eq. (7) can be transformed to (3)-(5) and all parameters receive suitable interpretation. For example, if $\vec{\beta}^* = \vec{\varphi}^*$, $C_{\varphi^*} = I_{m' \times m'} \varepsilon_1^2$, $C_{\sigma^*} = I_{m \times m} \varepsilon_0^2$, Eq. (3) is obtained with parameters $\gamma = (\varepsilon_0 / \varepsilon_1)^2$, $\vec{\varphi}^0 = \vec{\varphi}^*$. Correspondingly, the solution is expected in the region $\hat{\varphi}_i \in [\varphi_i^* - 2\, \varepsilon_0 \gamma^{1/2};\ \varphi_i^* + 2\, \varepsilon_0 \gamma^{1/2}]$ with probability 96% .

The reasons 2) and 3) should also be considered because of the disagreement between the LSM theory and real experience. Firstly, *a priori* information about non-negativity can not be correctly included in (7). Secondly, the LSM solution can be less stable than the solution of Eq. (1) by other methods, such as

$$\varphi_i^{p+1} = \varphi_i^p \, (\sigma_i^* / \sigma_i^p) \tag{8}$$

Chahine method [7] (for square matrix K with diagonal elements prevailed);

$$\varphi_i^{p+1} = \varphi_i^p \prod_{j=1}^{m} (1 + (\sigma^*_j / \sigma^p_j - 1) K_{ji}) \qquad (9)$$

Twomey-Chahine method [8]. These methods have not been formulated on the statistical basis and can not be reduced to a particular case of (2) or (7). They provide the non-negative and quite a stable solution. Often, it is better than the LSM solution [8].

A priori information about non-negativity of measured or searched values can be, in the scope of a statistical approach, included in the solution by changing the strategy of normal noise to the log-normal noise distribution [9], because of:

- a log-normally distributed random value σ_j is obviously to be positively defined;
- there is a number of the theoretical and experimental reasons [1] showing that log-normal noise is quite close to real one;
- using the log-normal PDF for σ_j is simple to be performed by transforming the problem to the space of normally distributed logarithms $f^*_j=\ln\sigma^*_j$, $a_i=\ln\varphi_i$.

Thus, solving Eq. (6) for the logarithmic space $a_i=\ln\varphi_i$, $f_j=\ln\sigma_j$, $d_i=\ln\beta_i$ transforms this problem into a non-linear problem. Accordingly, the solution of (6) is obtained by the iterative scheme of Newton-Gauss

$$\vec{a}^{p+1}=\vec{a}^p-(U_p^T C_{f^*}^{-1} U_p + D_p^T C_{d^*}^{-1} D_p)^{-1} (U_p^T C_{f^*}^{-1} (\vec{f}^p-\vec{f}^*)+D_p^T C_{d^*}^{-1} (\vec{d}^p-\vec{d}^*)) \qquad (10)$$

where U_p and D_p are matrices of first derivatives $\partial\sigma_j/\partial a_i|_{\vec{a}^p}$, $\partial d_j/\partial a_i|_{\vec{a}^p}$; C_{f^*} and C_{d^*} the covariance matrices of $\vec{f}^*$ and $\vec{d}^*$.

Besides, let us use the simple linear iterations [10] to stabilize solution

$$\vec{\varphi}_i^{p+1} = \vec{\varphi}_i^p - H (K \vec{\varphi}^p - \vec{\sigma}^*) , \qquad (11)$$

where H - diagonal matrix instead of the standard operator of matrix inversion $(K^T K)^{-1}$. The main question here is the formulation of H - matrix that provides the convergence of iterations. Statistical approach applies the iterative inversion to solve (6)

$$\vec{\varphi}^{p+1} = \vec{\varphi}^{p} - H\,(\Phi\,\vec{\varphi}^{p} - K^T C^{-1}_{\sigma^*}\,\vec{\sigma}^* - B^T C^{-1}_{\beta^*}\,\vec{\beta}^*)\,. \qquad (12)$$

In this process, the following formulation of a diagonal matrix H is very fruitful [9]

$$\{H\}_{ii} = (\sum_{k=1}^{n} |\{\Phi\}_{ik}|)^{-1}\,, \qquad (13)$$

where $|\{\Phi\}_{ik}|$ - the absolute values of a ik-th element of the Fisher matrix Φ. The great advantage of H-matrix choice (13) is the fact that even in a case of $\det\Phi=0$, the process (12) provides the optimization of the solution $\vec{\varphi}^{p}$ with the peculiarity that this optimization is implemented in the space of the solution with reduced dimension n' (n>n' is the Φ - matrix rank).

So, the method (12) is attractive to be used to solve linear inverse problems (1),(6) due to it is:

-quite a stable (even in searching large number of parameters);

-based on the statistical estimation approach;

-simple for implementation with the universal H-matrix.

The non-linear scheme (10) can be caused to include the process (12) by joining the two iterative processes: general non-linear one (10) and process of linear solution at each step of (10).

$$\vec{a}^{q+1} = \vec{a}^{q} - H_q\,(U^T_q C^{-1}_{f^*}\,(\vec{f}^{q} - \vec{f}^*) + D^T_q C^{-1}_{d^*}(\vec{d}^{q} - \vec{d}^*)), \qquad (14)$$

where $\{H_q\}_{ii} = [\sum_{i=1}^{n}(|\{U^T_q C^{-1}_{f^*} U_q + D^T_q C^{-1}_{d^*} D_q\}_{ik}|)]^{-1} = \sum_{i=1}^{n}(|\{\Phi\}_{ik}|)$.

This joining is consonant to the Levenberg-Marquardt idea [11] improving non-linear inversion by "cutting" the length of the p-th step $\Delta\vec{a}^{p}$. The length $\Delta\vec{a}^{q}$ in

(14) is obvious to be shorter than $\Delta\vec{a}^p$ of (10).

Thus, this paper is aimed to consider, in the scope of a single approach, the two kinds of numerical inversion methods: methods including the matrix inversion operator and iterative ones without that operator. They are very often considered as opposite to each other. As a result, the procedure (14) uniting the advantages of the methods of both groups is proposed.

References

1. Tarantola A. *Inverse problem theory: methods for data fitting and model parameter estimation.* (ELSEVIER., Amsterdam-Oxford-New-York-Tokyo, 1987).
2. Rogers C.D. In *Inversion methods in atmospheric remote sounding.* (ed. A. Deepak Academic Press, 1977), 117-134.
3. Oshchepkov S.L. and Dubovik O.V. J.Phys.D: Appl. Phys. 26, 728-732 (1993).
4. Twomey S. J.Assoc.Mach. 10, 97-101 (1963).
5. Phillips B.L. J.Assoc.Comp.Mach. 9, 84-97(1962).
6. Turchin V.F., Kozlov V.P., Malkevich M.S. Usp.Fiz. Nauk. 102, 345-386 (1970).
7. Chahine M.T. JOSA. 12, 1634-1637 (1968).
8. Twomey S. J.Comp.Phys. 18, 188-200 (1975).
9. Oshchepkov S.L., Dubovik O.V. Izv.RAN, Fiz.Atm. Okeana. 30, 165-172 (1994).
10. Ortega J.M. *Introduction to parallel and vector solution of linear systems.* (Plenum Press, New York, 1988).
11. Ortega J.M. and Rheinboldt W.C. *Iterative solution of nonlinear equations in several variables.* (Academic Press, New York and London, 1970).