ENTROPY ESTIMATE OF THE REGULARIZATION PARAMETER IN SOLVING THE INVERSE PROBLEMS OF MICROWAVE OPTICS

V.P. Yakubov

V. D. Kuznetsov Siberian Physicotechnical Scientific–Research Institute at the Tomsk State University, Tomsk

Received December 6, 1990

A technique of estimating the regularization parameter in solving the inverse problems of microwave optics based on maximization of the ratio of entropies of the power spectra of the residual noise and information-bearing process extracted from the input signal is considered. An example illustrating the applicability of the proposed technique is given.

One of the central questions of regularization of inverse problems of microwave optics consists in determining the principle for assigning the regularization parameter.[1–3] The Tikhonov principle of the discrepancy has been now widely used according to which the regularization parameter α is determined from the condition that the residual discrepancy is consistent with the *a priori* known error in the initial data.[1] In the present paper the value of α is proposed to determine from the condition of optimizing a certain entropy functional when the noise level is unknown *a priori*.

Many problems of microwave optics may be reduced to a linear equation

$$\mathrm{A}\mathbf{x} = \mathbf{y}\ , \tag{1}$$

where $\mathbf{y} = \{y_i\}$ is the M–dimensional vector of the measured variables and $\mathbf{x} = \{x_i\}$ is the N–dimensional vector of the unknown parameters, A is the $M{\times}N$ matrix, and $M \times N$. To solve equation (1), the Tikhonov regularizing functional is introduced,[1] whose minimization leads to a standard equation:

$$(\mathrm{A}^{\mathrm{T}}\mathrm{A} + \alpha)\mathbf{x}_\alpha = \mathrm{A}^{\mathrm{T}}\mathbf{y}. \tag{2}$$

Here $x = \mathbf{x}_\alpha$ is the regularized solution, α is the regularization parameter, and "T" denotes the transposition operation. It can be shown[4] that the solution of the standard equation (2) coincides with the pseudosolution $\mathbf{x}_\alpha = B^{+}\mathbf{y}$ of the equation $B\mathbf{x} = \mathbf{y}$, where $B = A + \alpha(A^{+})^{\mathrm{T}}$ and B^{+} denotes a pseudoinverse matrix.[5,6] Then the approximate solution $\mathbf{x}_\alpha$ differs from the expected ideal solution $\mathbf{x}_0$ by the value[6]

$$\Delta\mathbf{x} = \mathbf{x}_\alpha - \mathbf{x}_0 = B^{+}[\Delta\mathbf{y} - (B\mathbf{x}_0 - \mathbf{y}_0)]\ ,$$

where $\Delta\mathbf{y}$ is the deviation of the initially measured data from the ideal and exact $\mathbf{y}_0 = A\mathbf{x}_0$. It follows then for the relative deviation norm $d_x = ||\Delta\mathbf{x}|| / ||\mathbf{x}_0||$ that $d_x \le \varepsilon(\alpha)$ cond A, where $\mathrm{cond}\,A = ||A^{+}||\cdot||A||$ is the conditional number of the matrix A and

$$\varepsilon(a) = \frac{||B^{+}[\Delta\mathbf{y} - (B\mathbf{x}_0 - \mathbf{y}_0)]||}{||A^{+}||\cdot||\mathbf{y}_0||} \tag{3}$$

is the standard deviation of the initial data which might decrease in comparison with the initial standard deviation for $\alpha = 0$, namely, $\varepsilon(\alpha) < \varepsilon(0)$. Regularization is the more successful, the more accurately the discrepancy of the change of the operator $(B\mathbf{x}_0 - \mathbf{y}_0)$ in Eq. (3) compensates for the deviation of the initial data $\Delta\mathbf{y}$. The regularization parameter α plays the role of the adjustable parameter. The value $\alpha = \alpha_0$ is optimal if $\varepsilon(\alpha)$ is minimal. In general, the value α_0 will obviously depend on the deviation level $\Delta\mathbf{y}$, on the form of the solution $\mathbf{x}_0$, and on the matrix A. When α is determined based on the discrepancy, the level of the discrepancy $\delta = ||\mathrm{A}\mathbf{x}_0 - \mathbf{y}||$ is made consistent with the level of deviation of the initial data $||\Delta\mathbf{y}||$ which is assumed to be well known. The value $\alpha = \alpha_\delta$ determined from this condition disregards the form of the exact solution x_0 and the form of the operator A and is, therefore, not optimal.

When the deviation level $||\Delta\mathbf{y}||$ is unknown, the determination based on the discrepancy becomes inapplicable, and any supplemental information, for example, about the shape of the noise spectrum is required. Determination of the regularization parameter α and the corresponding solution obtained with it are equivalent to the separation of the information–bearing process $\mathbf{y}_\alpha = \mathrm{A}\mathbf{x}_\alpha$ and noise $\Delta\mathbf{y}_\alpha = \mathbf{y} - \mathbf{y}$ from the initial signal $\mathbf{y}$. The criterion of selection of the value of α proposed here is based on the assumption that the power spectrum of noise $\Delta\mathbf{y}$ is much more uniform than the energy spectrum of the information–bearing process $\mathbf{y}_0$. The entropy functional $\gamma(\alpha) = H[\Delta\mathbf{y}_\alpha]/H[\mathbf{y}_\alpha]$ can be used for the measure of contrast between these spectra. Here $H(z) = -\sum\limits_j p_j \ln p_j$ is the average entropy of the normalized power spectrum $p_j\left(\sum\limits_j p_j = 1\right)$ where $z = \Delta\mathbf{y}_\alpha$ for noise and $z = \mathbf{y}_\alpha$ for the signal. The requirement that the functional $\gamma(\alpha)$ reaches its maximum provides the value $\alpha = \alpha_\gamma$ which is optimal in the sense of the principle of the maximum entropy (PME). We underscore here that although the PME is somewhat similar to the well-known method of maximum entropy, which is widely used to estimate the power spectra of signals based on the autoregression models,[7] it still differs from it.

To find the pseudoinverse matrix and the pseudosolution, the singular expansion of the matrix

A = U^TCV is convenient, where U and V are the orthogonal matrices, C is a quasidiagonal matrix of coefficients c_i ($i = 1, \ldots, N$) called singular numbers.[4,5] The pseudoinverse matrix B^+ is calculated in the form of a product $B^+ = V^TSU$, in which the coefficients of the quasidiagonal matrix S are the numbers $s_i = [c_i + \alpha/c_i]^{-1}$. For $\alpha = 0$, these numbers are equal to $s_i = c_i$ and $B^+ = A^+$ while selection of $\alpha \neq 0$ means perturbation of the singular numbers of the matrix A. To calculate the power spectra, certain standard techniques may be used, including the algorithms of the fast Fourier transform, the Walsh transform, etc.

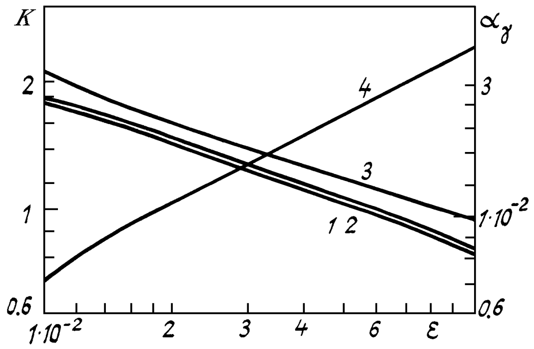

FIG. 1. Actual conditionality of the model inverse problem after regularization (curves 1–3), and the regularization parameter (curve 4) vs the error in the initial data.

We have chosen a discrete analog of the Fredholm integral equation of the first kind of the convolution type (see Ref. 2, p. 109) for our test example which usually serves to test various regularizing procedures. For $M = 41$ and $N = 13$ the conditional number was taken to be cond $A = \max c_i / \min c_i = 285$. Without regularization, ($\alpha = 0$) the relative standard deviation of the solution exceeded 100% given that the noise level was $\varepsilon = ||\Delta \mathbf{y}||/||\mathbf{y}_0|| \geq 10^{-2}$ while regularization ($\alpha \neq 0$) resulted in the relative standard deviation less than 9%. Figure 1 shows the dependence of the actual conditionality of the problem $K = d_x/\varepsilon$ on the noise level ε when α was optimal: $\alpha = \alpha_0$ (curve *1*), when α was determined based on the discrepancy: $\alpha = \alpha_\delta$ (curve *3*), and when α was determined based on the technique proposed in this paper: $\alpha = \alpha_\gamma$ (curve *2*).

It can be seen from Fig. 1 that despite the lack of an *a priori* data on the noise level, the selection $\alpha = \alpha_\gamma$ provides the accuracy not worse than $\alpha = \alpha_\delta$ based on the discrepancy; moreover, it is even closer to the optimal value $\alpha = \alpha_0$. Curve *4* shows the obtained dependence $\alpha = \alpha_\gamma$ on the noise level and may be approximated by the formula $\alpha_\gamma = [6{\cdot}10^{-3}\varepsilon]^{1/2}$.

The presented test example testifies to the efficiency of the technique for estimating the regularization parameter. The applicability of this technique to inverse problems has been tested by solving an improperly posed inverse problem of atmospheric refraction.[6] The entropy functional proposed here was successfully used to separate the Doppler and noise components of the actual spectrum of the signal obtained in urban surroundings and to determine the velocity of the relative motion of the transmitter. Optimization of such a functional may be useful in the adaptive methods of image processing.

REFERENCES

1. A.N. Tikhonov and V.Ya. Arsenin, *Methods for Solving Improperly Posed Problems* (Nauka, Moscow, 1986), 288 p.
2. A.N. Tikhonov, A.V. Goncharskii, V.V. Stepanov, and A.G. Yagola, *Regularization Algorithms and an* A *Priori Information* (Nauka, Moscow, 1983), 200 pp.
3. A.B. Bakushinskii and A.V. Goncharskii, *Improperly Posed Problems, Numerical Techniques and Applications* (Moscow State University Publishing House, Moscow, 1989), 199 pp.
4. C.L. Lawson and R.J. Hanson, *Solving Least–Squares Problems* (Prenlice–Holl, Englewoord Clipps, New Iork, 1974), 232 pp.
5. G.E. Forsythe, M.A. Malcolm, and C.B. Moler, *Computer Techniques for Mathematical Calculations* [Russian translation] (Mir, Moscow, 1980), 279 pp.
6. N.A. Simakova and V.P. Yakubov, Radiotekh. Elektron. **34**, No. 7, 1329-1336 (1989).
7. E.A. Robison, Proc. IEEE (transl. into Russ) **70**, No. 9, 6-33 (1982).