Method for search of quasi-solutions in the photoacoustic laser gas analysis

M.L. Belov, V.A. Gorodnichev, V.I. Kozintsev, and Yu.V. Fedotov

Scientific Research Institute of Radio Electronics and Laser Equipment, Moscow
N.E. Bauman Moscow State Technical University

Received May 21, 2001

The problem of reconstructing gas concentrations in laser photoacoustic gas analysis is considered. Signal processing procedures based on the search for quasi-solutions of the ill-posed problems are described for the differential absorption method. The results of mathematical modeling and processing of experimental data are presented. It is shown that using the search for quasi-solutions based on neural network algorithms it is possible to solve efficiently the problem of gas concentration reconstruction for both few- and many-component gas mixtures without involving additional information about the noise level in the measurement channels.

The on-line monitoring of atmospheric pollution and the industrial level of hazardous gaseous pollutants, as well as real-time control of technological processes demand fast-speed cheap automated devices allowing high-accuracy analysis of gas mixtures. Laser photoacoustic gas analyzers (LPGA) represent one class of such devices.[1–5]

If the problem of identification of gas mixture composition is solved (it is considered, for instance, in Refs. 6 and 7) or the composition of the analyzed gas mixture is known, then the main problem in processing measurement data is to reconstruct concentrations of gas mixture components from LPGA data.

To solve the problem on reconstructing gas concentrations in multicomponent mixtures, Tikhonov regularization method[8] can be successfully used with application of various methods (both deterministic and statistical) for selection of the regularization parameter (see, for instance, Refs. 9–13). However, this method has two disadvantages:

1. For a few-component mixture, use of regularization method for solving the system of linear algebraic equations of laser gas analysis usually gives larger errors in gas concentrations than standard methods.

2. Selection of the regularization parameter usually involves the use of either *a priori* information (for example, about the level of measurement noise) or a set of independent measurements (what is impossible for monitoring in real time).

This paper is devoted to the method of search for quasi-solutions, which is more promising for gas analysis and free of the above disadvantages.[8]

When reconstructing the gas concentrations in an N-component mixture from LPGA data (in the case if the width of laser lines is much smaller than the spectral width of absorption lines of the component gases), one faces the problem of solution of a system of M ($M > N$) linear algebraic equations having the following form[14]:

$$\begin{cases} k_1^1 c_1 + k_2^1 c_2 + \ldots + k_N^1 c_N + \alpha_1 = y_1, \\ k_1^2 c_1 + k_2^2 c_2 + \ldots + k_N^2 c_N + \alpha_2 = y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ k_1^M c_1 + k_2^M c_2 + \ldots + k_N^M c_N + \alpha_M = y_M, \end{cases} \quad (1)$$

where k_i^j is the absorption coefficient of the ith component gas for the jth spectral measurement channel; c_i is the sought concentration of the ith component gas; y_j is the measured reduced signal for the jth spectral measurement channel; α_j is the coefficient of nonselective absorption for the jth spectral measurement channel.

To eliminate the effect of nonselective absorption, the differential absorption method is used. In this case, the number of spectral measurement channels (and, correspondingly, the number of equations in the system (1)) is taken $M = 2N$, and the spectral measurement channels themselves are selected close to each other by pairs, so that in every pair the coefficients of nonselective absorption can be thought equal. Subtracting the pair equations one from another, we can have the following system of difference equations:

$$\begin{cases} \Delta k_1^1 c_1 + \Delta k_2^1 c_2 + \ldots + \Delta k_N^1 c_N = \Delta y_1, \\ \Delta k_1^2 c_1 + \Delta k_2^2 c_2 + \ldots + \Delta k_N^2 c_N = \Delta y_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \Delta k_1^N c_1 + \Delta k_2^N c_2 + \ldots + \Delta k_N^N c_N = \Delta y_N, \end{cases} \quad (2)$$

where

$$\Delta k_i^j = k_i^{2j-1} - k_i^{2j}; \;\; \Delta y_j = y_{2j-1} - y_{2j}.$$

In the matrix form Eq. (2) can be written as

$$\Delta K \, \mathbf{C} = \Delta \mathbf{Y}, \quad (3)$$

where $\Delta \mathbf{Y}$ is the vector of differences of reduced signals Δy_j; ΔK is the matrix of differences of the absorption coefficients Δk_i^j; $\mathbf{C}$ is the vector of sought concentrations c_i.

0235-6880/02/04 349-05 $02.00

The fitting method is widely used in the computational practice for approximate solution of the ill-posed problems.[8] This method can be efficiently used for solution of the system of equations of laser photoacoustic gas analysis in the case that only the approximate solution, or the right-hand side $\Delta\mathbf{Y}$ of Eq. (3), rather than the exact one, is known and (or) the matrix of the absorption coefficients is inaccurate.

For such problems, the concept of quasi-solution is introduced and the fitting method allows one to find some approximation of a quasi-solution. The fitting method in this case consists in the following. The direct problem is solved for $\mathbf{C}$ elements in some bounded region D satisfying the physical sense (the operator $\Delta K\,\mathbf{C}$ is calculated) and the element of $\tilde{\mathbf{C}} \in D$ minimizing the discrepancy between $\Delta K\,\tilde{\mathbf{C}}$ and $\Delta\mathbf{Y}$ is called the quasi-solution of Eq. (3), that is, the quasi-solution is determined by the following equation:

$$\rho(\Delta K\tilde{\mathbf{C}}, \Delta\mathbf{Y}) = \inf_{\mathbf{C}\in M} \rho(\Delta K\,\mathbf{C}, \Delta\mathbf{Y}) \qquad (4)$$

under the condition that $0 \le \tilde{c}_i \le 1$; $\sum_{i=1}^{N} \tilde{c}_i \le 1$.

The evolution-genetic method (EGM)[15,16] is a promising computational fitting method that allows determination of a quasi-solution to Eq. (3). Its advantage is the possibility of efficiently solving complex optimization problems. Although the discrepancy functional has one extreme (minimum), it acquires the ravine shape (strongly elongated along one direction) in the case of ill-posed problems, and application of traditional minimum search algorithms becomes inefficient.[17]

In doing this work, we have realized a modification of the algorithm of evolution-genetic strategy described in Ref. 6. Let us introduce the needed terms: chromosome is some realization of the vector of unknown concentrations, the conditions of the domain of definition and the physical sense of the sought solution are imposed on a chromosome; population is a set of chromosomes.

Let the discrepancy functional

$$F(\tilde{\mathbf{C}}) = \rho(\Delta K\tilde{\mathbf{C}}, \Delta\mathbf{Y})$$

be a target (optimized) function calculated for every chromosome.

The algorithm of search for the quasi-solution involves the following stages:

1. At the initial stage, the starting parent population of P chromosomes is formed in a random manner, that is, P versions of the vectors of initial gas concentrations are generated: $\mathbf{C}_0^p(c_{01}^p, c_{02}^p, \ldots c_{0i}^p, \ldots c_{0N}^p)$, $p = 1\ldots P$. The values of Δy_i are calculated from the initial concentrations c_{0i}^p by the system of equations (2). The starting population is sorted by the target function $F(\mathbf{C}_0^p)$ in the decreasing order (the smallest superscript p corresponds to the smallest value of the target function).

2. The next stage involves reproduction. At the stage of reproduction, L new vectors are generated instead of the P initial concentration vectors: every new concentration vector $\mathbf{C}_{1\mathrm{r}}^{l}$ ($l = 1\ldots L$) is formed of two randomly selected old vectors $\mathbf{C}_0^{p1}$ and $\mathbf{C}_0^{p2}$ ($p1$ and $p2$ are values from the range $1\ldots P$), and the vectors with the smaller value of the target function $F(\mathbf{C}_0^p)$ are taken with higher probability. The components of the vectors are formed in the following way:

$$c_{1\mathrm{r}i}^{l} = \begin{cases} c_{0i}^{p1}, & \text{if } s = 0, \\ c_{0i}^{p2}, & \text{if } s = 1, \end{cases}$$

where $c_{1\mathrm{r}i}^{l}$ are the concentration values after the first reproduction; s is a random value equal to 0 or 1.

3. Then the vectors obtained as a result of reproduction mutate. At this stage, the concentration vectors $\mathbf{C}_{1\mathrm{r}}^{l}$ are changed. The number of versions of the concentration vectors keeps equal to L, while the concentrations change by the following law:

$$c_{1\mathrm{m}i}^{l} = \begin{cases} c_{1\mathrm{r}i}^{l}, & \text{if } s_1 = 0, \\ c_{1\mathrm{r}i}^{l} + w, & \text{if } s_1 = 1, \end{cases}$$

where $c_{1\mathrm{m}i}^{l}$ are the concentration values after the first mutation; s_1 is a random value equal to 0 or 1; w is a random parameter distributed normally with the mathematical expectation equal to 0 and the variance σ^2. The values of Δy_i and new values of the target function $F(\mathbf{C}_{1\mathrm{m}}^{l})$ are calculated from the obtained concentrations c_{1i}^{m} by the system of equations (2).

4. Of all L versions of the concentrations $\mathbf{C}_{1\mathrm{m}}^{l}$, P the best ones (from the viewpoint of smallness of the target function) are selected and, at the next stage, they are used as the initial ones for mutation and reproduction.

5. The reproduction and mutation cycle are reiterated a fixed number of times (generations) G. During a cycle, the variance σ^2 of a random parameter w varies in the process of evolution depending on the frequency of decrease of the target function calculated for several iterations.

6. As the cycle is terminated, only one, the best (from the viewpoint of the smallest value of the target function), vector of concentrations $\mathbf{C}_{G\mathrm{m}}^{l}$ is taken as a quasi-solution.

For comparison of EGM operation with the Tikhonov regularization method and the standard method of solution of a system of linear algebraic equations, we conducted mathematical simulation and processing of experimental LPGA data. As a method for solution of a system of linear algebraic equations, we used the Gauss method with column-by-column selection of the main element[17] (hereinafter in this paper, it is referred to as the method of direct solution). When constructing the regularized solution, the discrepancy method was used for determination of the regularization parameter.[13]

Mathematical simulation was conducted in a closed cycle: the composition of a gas mixture was specified, the vector of reduced signals of the system (1) was calculated for this mixture composition, the components of the vector were disturbed by additive noise distributed by the normal law with the given relative standard deviation. Then the concentrations of the mixture components were reconstructed from these model data. The measured data were obtained with LPGA based on a tunable CO_2 low-pressure laser and a nonresonance measurement cell described in Ref. 10.

Figure 1 depicts the results of algorithm operation with a few number of mixture components $N = 3$ [method of direct solution (*1*), search for quasi-solution over 1000 generations with the use of EGM (*2*), and the regularization method (*3*)].

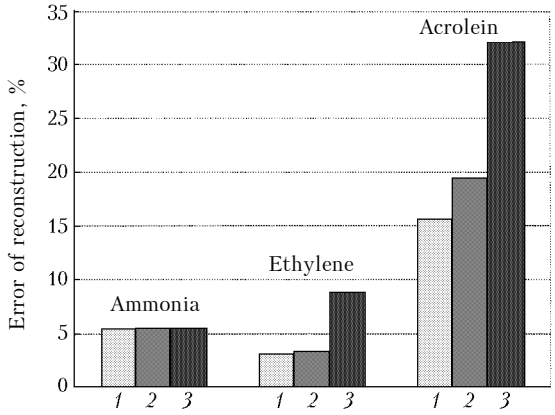

Fig. 1. Errors of reconstruction of gas concentration in a three-component mixture.

The following mixture was taken for analysis: ammonia–ethylene–acrolein ($10^{-3} : 5 \cdot 10^{-3} : 9 \cdot 10^{-4}$). The pairs of spectral measurement channels were 10*R*24–10*R*18, 10*P*34–10*P*30, and 10*P*6–10*P*4. The "measurement" data are the results of numerical simulation. The condition number for the matrix of difference of absorption coefficients was 5.3. The relative standard noise in all measurement channels was 10%. For reconstruction we used the data averaged over a series of 10 measurements. Simulation was repeated 10 times, and the mean error for 10 series of experiments was calculated.

It is seen from Fig. 1 that at a small number of mixture components, application of the processing algorithms based on the regularization method is inefficient (this was also noticed in Ref. 13). The mean error of reconstruction with the use of the search for quasi-solutions is roughly at the same level with the mean error of the method of direct solution, and the error of reconstruction by the regularization method is much larger.

Figures 2 and 3 illustrate how the error of reconstruction of gas mixture components depends on the composition of the analyzed mixture (with unchanged number of mixture components). Figures 2 and 3 depict the results of mathematical simulation of reconstruction of gas concentrations in two six-component mixtures [method of direct solution (*1*), search for quasi-solution over 1000 generations with the use of EGM (*2*), search for quasi-solution over 200 generations with the use of EGM (*3*), and the regularization method (*4*)]. The relative standard noise was 5%, and no averaging was conducted. The reconstruction error calculated from a series of ten single measurements is presented.

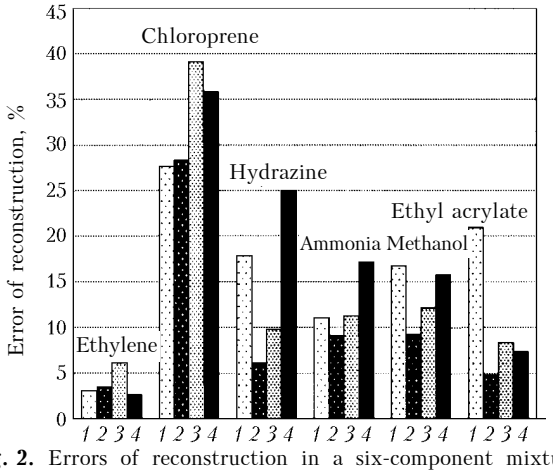

Fig. 2. Errors of reconstruction in a six-component mixture ethylene – chloroprene – hydrazine – ammonia – methanol – – ethyl acrylate.

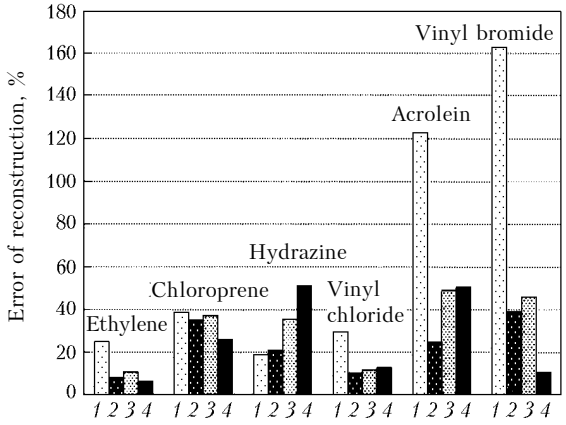

Fig. 3. Errors of reconstruction in a six-component mixture ethylene – chloroprene – hydrazine – vinyl chloride – acrolein – – vinyl bromide.

Figure 2 shows the results of reconstruction for the mixture ethylene–chloroprene–hydrazine–ammonia–methanol–ethyl acrylate ($5 \cdot 10^{-4} : 9 \cdot 10^{-5} : 4 \times 10^{-4} : : 10^{-4} : 2 \cdot 10^{-4} : 3 \cdot 10^{-4}$). The pairs of the measurement channels were: 10*R*6–10*R*12, 10*P*14–10*P*10, 10*R*18–10*R*22, 9*P*14–9*P*22, 10*P*22–10*P*18, and 9*R*14–9*R*22.

Figure 3 shows the results of reconstruction for the mixture ethylene–chloroprene–hydrazine–vinyl chloride–acrolein–vinyl bromide ($10^{-4} : 9 \cdot 10^{-5} : 2 \cdot 10^{-4} : 5 \times \times 10^{-4} : 4 \cdot 10^{-4} : 3 \cdot 10^{-4}$). The pairs of measurement channels were: 10*P*14–10*P*10, 10*P*22–10*P*18, 10*R*18–10*R*22, 10*P*32–10*P*28, 9*P*14–9*P*12, and 9*R*20–9*R*12.

For the mixture shown in Fig. 2, the errors of reconstruction are not very large and application of specialized reconstruction algorithms does not provide a

significant advantage as compared with the method of direct solution. However, even here the search for quasi-solutions over 1000 generations with the use of EGM is preferable as compared with the regularization method and the method of direct solution.

For the mixture shown in Fig. 3, two components (acrolein and vinyl bromide) cannot be reconstructed by the method of direct solution (errors are higher than 100%) and here the mean error of reconstruction for most gases decreases markedly with the use of the regularization method and the search for quasi-solutions (the errors for these two methods are on the same order). The increase in the error of reconstruction for EGM with the decrease of the number of generations from 1000 to 200 (the computational time in this case decreased five times) is not very significant due to irregular convergence rate of this algorithm. With the further decrease of the number of generations, the reconstruction error increased drastically and it is not shown in Figs. 2 and 3 because of too large difference in the values.

It should be noted that the condition number of the matrix of difference of the absorption coefficients for the gas mixture in Fig. 2 equals 17.8, and for Fig. 3 it is 1176. The condition number characterizes the maximum possible coefficient of transfer of the error in the right-hand side or the error of the matrix in the left-hand side to the error of solution of the system of linear equations.[18] Thus (as is seen from Figs. 2 and 3), the condition number can serve a qualitative criterion of the need to apply specialized algorithms for solution of equations of laser gas analysis.

Figure 4 depicts the results of processing the LPGA experimental data. In this case, we analyzed a six-component mixture of ethylene – carbon dioxide – – ammonia – methanol – ethanol – isopropanol ($1.6 \times \times 10^{-4} : 9.6 \times 10^{-2} : 7.3 \cdot 10^{-3} : 8.5 \cdot 10^{-3} : 6.6 \cdot 10^{-3} : : 9.2 \cdot 10^{-3}$). (The results on reconstruction of carbon dioxide are not shown, since we failed to reconstruct it in all cases.) The pairs of the spectral measurement channels were $10P14$–$10P20$, $9R18$–$9R10$, $9R16$–$9R12$, $9P34$–$9P32$, $9P40$–$9P42$, and $10R12$–$10R30$. The relative standard noise value in the spectral measurement channels at single measurements was from 1 to 9%. Reconstruction was conducted from averaged (over ten measurements) data. The condition number of the matrix of difference of the absorption coefficients was $4.7 \cdot 10^4$. For simulation of the effect of errors in absorption coefficients, the system matrix was distorted (imperfect values of the absorption coefficients for isopropanol and carbon dioxide at $10R30$ wavelength were input). The input distortions of the matrix of absorption coefficients were small:

$$\frac{\|\Delta K_{ERR}\|_M}{\|\Delta K\|_M} = 0.34\%,$$

where $\|\bullet\|_M$ is the matrix norm; ΔK is the matrix of difference of absorption coefficients; ΔK_{ERR} is the matrix of errors in the absorption coefficients.

The condition number of the distorted matrix was $2.4 \cdot 10^5$.

Figure 4 shows the errors of reconstruction performed with the use of the distorted [method of direct solution (*1*), search for quasi-solution over 1000 generations with the use of EGM (*2*), and the regularization method (*3*)] and actual matrix of absorption coefficients [method of direct solution (*4*), search for quasi-solution over 1000 generations with the use of EGM (*5*), and the regularization method (*6*)].

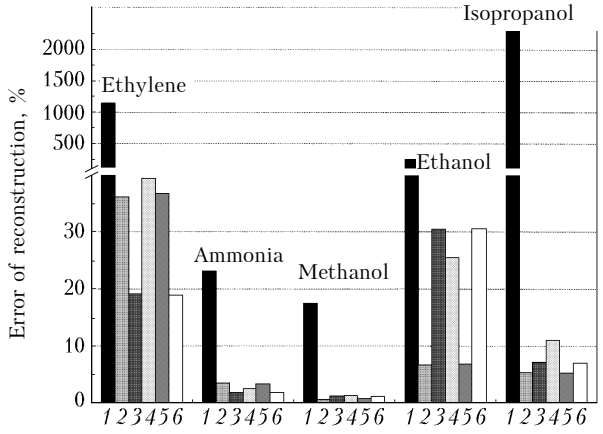

Fig. 4. Errors of gas reconstruction in a six-component mixture ethylene – carbon dioxide – ammonia – methanol – ethanol – – isopropanol.

It can be seen from Fig. 4 that for the distorted matrix the method of direct solution gives very large errors (hundreds and thousands percent) for most gases. (The reconstructed concentrations not only differ from the actual ones by several orders of magnitude, but they may even have physically meaningless values.) At the same time, these errors are not too large for the actual matrix of absorption coefficients (although they can achieve 40%).

It is interesting to note that the level of reconstruction errors for the method of search for quasi-solutions and the regularization method is almost independent of errors in the matrix of gas absorption coefficients – both of these methods allow reconstruction of the concentrations of component gases with acceptable accuracy (errors of both methods are on the same order of magnitude).

Thus, the results of mathematical simulation and processing of measurements show that the method of search for quasi-solutions (using the evolution-genetic fitting algorithm) allows reconstruction of gas concentrations in laser photoacoustic gas analysis for both few-component and many-component gas mixtures. This method does not demand extra information about the level of noise in measurement channels.

A disadvantage of the method of search of quasi-solutions is relatively long (as compared with the regularization method) computational time. However, this disadvantage becomes insignificant when using up-to-date computers. Besides, with the use of EGM, it

becomes possible to regulate, to a certain extent, the computational time by varying the number of generations. The EGM speed can be increased due to paralleling of this algorithm and use of multiprocessor systems.[15,16]

In conclusion, we would like to note one more significant advantage of the method of search for quasi-solutions – it allows solution of not only the system of linear algebraic equations (1), but also more general (without the condition of the small spectral width of a laser line as compared with the width of a gas absorption line) system of integral equations of the laser gas analysis.

References

1. A.B. Antipov, V.A. Kapitanov, Yu.N. Ponomarev, and V.A. Sapozhnikova, *Photoacoustic Method in Laser Spectroscopy of Molecular Gases* (Nauka, Novosibirsk, 1984), 128 pp.
2. V.P. Zharov and V.S. Letokhov, *Laser Photoacoustic Spectroscopy* (Nauka, Moscow, 1984), 320 pp.
3. V.S. Antonov, G.I. Bekov, M.A. Bol'shov, V.P. Zharov, V.S. Letokhov, Yu.A. Kuritsyn, R.I. Personov, and A.N. Shibanov, *Laser Analytical Spectroscopy* (Nauka, Moscow, 1986), 320 pp.
4. Yu.N. Ponomarev, Atmos. Oceanic Opt. **8**, Nos. 1–2, 116–124 (1995).
5. V.E. Zuev, Yu.S. Makushkin, and Yu.N. Ponomarev, *Atmospheric Spectroscopy* (Gidrometeoizdat, Leningrad, 1987), 247 pp.
6. M.Yu. Kataev and A.A. Mitsel', Avtometriya, No. 4, 15–20 (1985).
7. M.Yu. Kataev, A.A. Mitsel', and S.R. Tarasova, Atm. Opt. **3**, No. 8, 763–771 (1990).
8. A.N. Tikhonov and V.Ya. Arsenin, *Methods for Solution of Ill-Posed Problems* (Nauka, Moscow, 1979), 288 pp.
9. Yu.S. Makushkin, A.A. Mitsel', and G.S. Khmel'nitskii, Zh. Prikl. Spektrosk. **35**, No. 5, 785–790 (1981).
10. V.I. Kozintsev, Atmos. Oceanic Opt. **9**, No. 8, 689–691 (1996).
11. V.I. Kozintsev, Atmos. Oceanic Opt. **9**, No. 10, 871–874 (1996).
12. V.V. Avdienko, M.L. Belov, V.A. Gorodnichev, V.I. Kozintsev, and D.B. Dobritsa, Zh. Prikl. Spektrosk. **63**, No. 5, 755–759 (1996).
13. M.L. Belov, V.A. Gorodnichev, V.I. Kozintsev, and D.B. Dobritsa, Atmos. Oceanic Opt. **13**, No. 2, 131–135 (2000).
14. M. Sigrist, M.Yu. Kataev, A.A. Mitsel', Yu.N. Ponomarev, and A. Thony, Atmos. Oceanic Opt. **7**, Nos. 11–12, 795–799 (1994).
15. E. Schoeneburg, F. Heinmann, and S. Feddersen, *Genetische Algorithmen und Evolutionsstrategien: Eine Einfuerung in Theorie und Praxis der Simulierten Evolution* (Addison-Wesley, Bonn–Paris–Reading, Mass., 1994), 321 pp.
16. V.M. Kureichik, *Genetic Algorithms* (TRTU, Taganrog, 1998), 117 pp.
17. A.A. Amosov, Yu.A. Dubinskii, N.V. Kopchenova, *Computational Methods for Engineers* (Vysshaya Shkola, Moscow, 1994), 544 pp.
18. V.M. Verzhibitskii, *Numerical Methods. Linear Algebra and Nonlinear Equations* (Vysshaya Shkola, Moscow, 2000), 266 pp.