DOI: https://doi.org/10.15407/kvt210.04.026

REVUNOVA E.G.[1], DSc (Engineering),
Leading Researcher, Department of Neural Information
Processing Technologies
https://orcid.org/0000-0002-3053-7090, e-mail: egrevunova@gmail.com
TYSHCHUK O.V.[2],
Senior Software Engineer
e-mail: avtyshcuk@gmail.com
DESIATERYK O.O.[3], PhD (Phys&Math)
Assistant Professor of the Department
of Algebra and Computer Mathematics
e-mail: sasha.desyaterik@gmail.com
[1] International Research and Training Center for Information
Technologies and Systems of the National Academy of Sciences
of Ukraine and Ministry of Education and Science of Ukraine,
40, Acad. Glushkova av., Kyiv, 03187, Ukraine
[2] Roku Inc., Kyiv, Ukraine
[3] Taras Shevchenko Kyiv National University,
Faculty of Mechanics and Mathematics,
4e, Acad. Glushkova av., Kyiv, 03127, Ukraine

THE TECHNOLOGY OF THE STABLE SOLUTION FOR DISCRETE ILL-POSED PROBLEMS BY MODIFIED RANDOM PROJECTION METHOD

Introduction. *Ill-posed problems solution is actual for many areas of science and technology. For example, discrete ill-posed problems (DIP) appears after discretization of the integral equations in the spectrometry, gravimetry, magnitometry, electrical prospecting and others.*

In the case of linear DIP the matrix, which model some measuring system, makes a linear transformation of input vector to the output vector. Usually DIP output vector contains noise and singular values series of the matrix smoothly decrease to zero. In this case, the solution (input vector estimation) using the inversion of the transformation matrix is unstable and inaccurate. To overcome instability and increase accuracy we use regularization methods.

We develop an approach which uses regularizing properties of random projection to obtain a stable solution of DIP. However, the development of effective sustainable methods for solving DIP continues to be a problem of current interest.

The purpose *of the paper is to increase the accuracy of DIP solution by the random projection method.*

Results. *In this paper we developed the method of stable solution of DIP by the modified method of random projection. For this modification the regularization by random projection is complemented by the regularization in the ridge regression style.*

For the our method we obtained expressions which connect in the direct way the solution error components with the matrix specter and the regularization parameter. For the developed method the experimental research of the accuracy is conducted on the test problems.

Conclusions. *The modified method of random projecting is characterized by stability and increased accuracy of the solution. This achieved by simultaneous ridge regression style regularization and random projecting. The representation of the solution error in the form where error components are related to the matrix specter and regularization parameter is important for further study of the error.*

Keywords: *discrete ill-posed problem, random projection.*

INTRODUCTION

The need to solve inverse problems arises in many areas of science and technology in connection with the recovery of the object signal based on the results of indirect remote measurements. The transformation of the object signal when interacting with the environment and the measuring system is modeled by a linear input-output transformation matrix. The transformation matrix and the vector of the results of indirect measurements (the output vector) are known, it is required to determine the vector of the input signal (the input vector, i.e., the solution vector).

Usually DIP output vector contains noise and singular values series of the matrix smoothly decrease to zero. In this case, the solution (input vector estimation) using the inversion of the transformation matrix is unstable and inaccurate. To overcome instability and increase accuracy we use regularization methods.

Furthermore, in problems of statistics and machine learning, a situation often arises when the solution by existing methods is unstable, i.e. small changes in the input data (conditions of the problem) lead to a large change in the solution. Such unstable solutions are inaccurate and cannot be used in practice. To remove the instability of the solution, the regularization approach is use.

Regularization imposes stability constraints on the sought solution. For example, a compromise of accuracy and stability is provided by choosing a regularization parameter that weights the ratio of the magnitude of the norm of the difference between the vectors of the reconstructed and the observed output, as well as the magnitudes of the norm of the solution vector (that is, the reconstructed input).

Our studies of the regularizing properties of random projection began in 2009 [1]. Later other researchers began to explore the regularizing properties of random projection, for example, for classification problems [2] and machine learning [3], and, more recently, for solving inverse problems [4, 5, 6]. Since the approach of random projection, along with improving the accuracy of the solution by regularization, reduces the computational complexity of the solution, we have managed to develop algorithms that provide an accurate and fast solution for discrete inverse problems [7, 8, 9].

METHOD OF THE STABLE SOLUTION FOR DISCRETE ILL-POSED PROBLEMS BY MODIFIED RANDOM PROJECTION APPROACH

We study the recovery problem of an input vector $\mathbf{x} \in \Re^N$ for a model $\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{e}$. We consider the case when the matrix $\mathbf{A} \in \Re^{N \times N}$ and the output vector $\mathbf{b} \in \Re^N$ are known. In particular $\mathbf{b} = \mathbf{b}_0 + \mathbf{e}$ where $\mathbf{b}_0 = \mathbf{A}\mathbf{x}$ and series of singular values of the matrix $\mathbf{A}$ smoothly descend to the zero (but do not reach it). The input vector recovery problem with considered properties belongs to the class of ill-posed problems. In this case the matrix $\mathbf{A}$ is a full rank matrix. To solve this problem by random projection approach [10–15] each sides of the basic equation $\mathbf{b} = \mathbf{Ax} + \mathbf{e}$ we multiply by the matrix $\mathbf{R}_k \in \Re^{k \times N}$. Elements of the matrix $\mathbf{R}_k$ are realizations of random variable with normal distribution, zero mean and unit variance. After the randomization we obtain the next equation: $\mathbf{R}_k \mathbf{Ax} = \mathbf{R}_k \mathbf{b}$. The problem of input vector estimation (recovery) is written in the next form: $\mathbf{x}_R = \arg\min_x \| \mathbf{R}_k \mathbf{Ax} - \mathbf{R}_k \mathbf{b} \|^2$. The input vector estimation $\mathbf{x}_R^* = (\mathbf{R}_k \mathbf{A})^+ \mathbf{R}_k \mathbf{b}$, where $(\mathbf{R}_k \mathbf{A})^+$ is a pseudoinverse matrix. To the statement of random projection problem we introduce additional regularization:

$$\mathbf{x} = \arg\min_x \| \mathbf{R}_k \mathbf{Ax} - \mathbf{R}_k \mathbf{b} \|^2 + \lambda \| \mathbf{x} \|^2 . \tag{1}$$

This problem formulation we call the modified method of random projection. This modification means that additional increase of stability is reached by determining of the regularization parameter which gives a balance between the Euclidean norm of solution vector discrepancy and the solution vector norm.

The input vector estimation for modified method of random projection we obtain as

$$\mathbf{x}_{MR}^* = ((\mathbf{R}_k \mathbf{A})^{\mathrm{T}} \mathbf{R}_k \mathbf{A} + \lambda \mathbf{I})^{-1} (\mathbf{R}_k \mathbf{A})^{\mathrm{T}} \mathbf{R}_k \mathbf{b} .$$

The corresponding output vector estimation is the next

$$\mathbf{b}_{MR}^* = \mathbf{Ax}_{MR}^* = \mathbf{A}((\mathbf{R}_k \mathbf{A})^{\mathrm{T}} \mathbf{R}_k \mathbf{A} + \lambda \mathbf{I})^{-1} (\mathbf{R}_k \mathbf{A})^{\mathrm{T}} \mathbf{R}_k \mathbf{b} .$$

The problem formulation (1) is similar to the problem formulations for ridge regression and Tikhonov regularization.

The difference of the formulation (1) is the fact that it includes two regularization parameters. First parameter λ is balancing the Euclidean norm of discrepancy vector and the norm of the solution vector. Second parameter k is a number of rows in the random matrix.

To compare methods we consider a dependency of input and output recovery error components from the parameter of the Tikhonov regularization.

REGULARIZATION IN THE RIDGE REGRESSION STYLE AND BY RANDOM PROJECTION

The classical method which used to solve ill-posed problems is the Tikhonov regularization. The standard Tikhonov regularization problem has the next form $\mathbf{x}_{reg} = \arg\min_{\mathbf{x}} \left(\|\mathbf{Ax} - \mathbf{b}\|^2 + \lambda \|\mathbf{x}\|^2 \right)$ where λ is a regularization parameter.

For the error of the *recovery of the true signal* (input) by the Tikhonov regularization method $e_x^{tikh} = \left\| (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^{\mathrm{T}} \mathbf{b} - \mathbf{x} \right\|^2$ we can obtain deterministic e_d^{tikh} and stochastic e_s^{tikh} components. Such that $e_x^{tikh} = e_d^{tikh} + e_s^{tikh}$ and analytical formulas for these components are obtained

$$e_d^{tikh} = \left\| (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^{\mathrm{T}} \mathbf{b}_0 - \mathbf{x} \right\|^2 = \left\| ((\mathbf{A}^{\mathrm{T}}\mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^{\mathrm{T}} \mathbf{A} - \mathbf{I})\mathbf{x} \right\|^2 ,$$

$$e_s^{tikh}=\left\|(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}\mathbf{e}\right\|^2.$$

Consider a behavior of the deterministic component depend on regularization parameter.

Since $(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}=\mathbf{V}diag(\frac{s_i^2}{s_i^2+\lambda^2}\frac{1}{s_i})\mathbf{U}^{\mathrm T}$ for the deterministic error component e_d^{tikh} we get

$$e_d^{tikx}(\lambda)=\mathbf{x}^{\mathrm T}\mathbf{V}diag\left(\frac{1}{(s_i^2/\lambda^2+1)^2}\right)\mathbf{V}^{\mathrm T}\mathbf{x}=\sum_{i=1}^{N}\mathbf{x}^{\mathrm T}\mathbf{v}_i\left(\frac{1}{(s_i^2/\lambda^2+1)^2}\right)\mathbf{v}_i^{\mathrm T}\mathbf{x}.$$

Since $\frac{1}{(s_i^2/\lambda^2+1)^2}<\frac{1}{(s_i^2/(\lambda+\delta)^2+1)^2}$ where $0<\lambda<1$ and $0<(\lambda+\delta)<1$ the deterministic error component increases with the increase of λ: $\mathbf{x}^{\mathrm T}\mathbf{V}diag\left(\frac{1}{(s_i^2/\lambda^2+1)^2}\right)\mathbf{V}^{\mathrm T}\mathbf{x}<\mathbf{x}^{\mathrm T}\mathbf{V}diag\left(\frac{1}{(s_i^2/(\lambda+\delta)^2+1)^2}\right)\mathbf{V}^{\mathrm T}\mathbf{x}$.

Further we consider a behavior of the stochastic component dependence on regularization parameter. Thus we note that

$$e_s^{tikh}=\left\|(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}\mathbf{e}\right\|^2=\mathbf{e}^{\mathrm T}\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\frac{1}{s_i}\right)^2\mathbf{U}^{\mathrm T}\mathbf{e}$$

and average over the noise realizations

$$\mathrm{E}\{e_s^{tikh}\}=\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\frac{1}{s_i}\right)^2\mathbf{U}^{\mathrm T}\right).$$

Since $\left(\frac{s_i}{s_i^2+\lambda^2}\right)^2>\left(\frac{s_i}{s_i^2+(\lambda+\delta)^2}\right)^2$ where $0<\lambda<1$ and $0<(\lambda+\delta)<1$, then averaged by noise realizations stochastic error component decreases by increase of λ:

$$\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm T}\right)>\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i}{s_i^2+(\lambda+\delta)^2}\right)^2\mathbf{U}^{\mathrm T}\right).$$

The stochastic error component without the averaging over the noise also decrease with increase of λ:

$$\mathbf{e}^{\mathrm T}\mathbf{U}diag\left(\frac{s_i}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm T}\mathbf{e}>\mathbf{e}^{\mathrm T}\mathbf{U}diag\left(\frac{s_i}{s_i^2+(\lambda+\delta)^2}\right)^2\mathbf{U}^{\mathrm T}\mathbf{e}.$$

For the error of *the output recovery* by the Tikhonov regularization method $e_y^{tikh}=\left\|\mathbf{A}(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}\mathbf{b}-\mathbf{b}_0\right\|^2$ we obtain deterministic e_{yd}^{tikh} and stochastic e_{ys}^{tikh} components. Such that $e_y^{tikh}=e_{yd}^{tikh}+e_{ys}^{tikh}$, $e_{yd}^{tikh}=\left\|(\mathbf{A}(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}-\mathbf{I})\mathbf{b}_0\right\|^2$, $e_{ys}^{tikh}=\left\|\mathbf{A}(\mathbf{A}^{\mathrm T}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm T}\mathbf{e}\right\|^2$.

We consider the behavior of deterministic component dependence on the regularization parameter.

Since $\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}=\mathbf{U}diag(\frac{s_i^2}{s_i^2+\lambda^2})\mathbf{U}^{\mathrm{T}}$ and after rewriting

$$\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}-\mathbf{I}=\mathbf{U}diag(\frac{s_i^2}{s_i^2+\lambda^2}-1)\mathbf{U}^{\mathrm{T}}$$

we obtain $e_{yd}^{tikh}(\lambda)=\mathbf{b}_0^{\mathrm{T}}\mathbf{U}diag(\frac{s_i^2}{s_i^2+\lambda^2}-1)^2\,\mathbf{U}^{\mathrm{T}}\mathbf{b}_0$,

$$e_1^{tikh}(\lambda)=\mathbf{b}_0^{\mathrm{T}}\mathbf{U}diag\left(\frac{1}{(s_i^2/\lambda^2+1)^2}\right)\mathbf{U}^{\mathrm{T}}\mathbf{b}_0.$$

Since $\frac{1}{(s_i^2/\lambda^2+1)^2}<\frac{1}{(s_i^2/(\lambda+\delta)^2+1)^2}$, where $0<\lambda<1$ and $0<(\lambda+\delta)<1$ the deterministic error component increases with the increase of λ:

$$\mathbf{b}_0^{\mathrm{T}}\mathbf{U}diag\left(\frac{1}{(s_i^2/\lambda^2+1)^2}\right)\mathbf{U}^{\mathrm{T}}\mathbf{b}_0<\mathbf{b}_0^{\mathrm{T}}\mathbf{U}diag\left(\frac{1}{(s_i^2/(\lambda+\delta)^2+1)^2}\right)\mathbf{U}^{\mathrm{T}}\mathbf{b}_0.$$

Further we consider the behavior of stochastic component dependence on regularization parameter.

Thus, we note that $e_{ys}^{tikh}=\left\|\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A}+\lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{e}\right\|^2=\mathbf{e}^{\mathrm{T}}\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm{T}}\mathbf{e}$

and average by the noise realizations

$$\mathrm{E}\{e_{ys}^{tikh}\}=\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm{T}}\right).$$

Since $\left(\frac{s_i^2}{s_i^2+\lambda^2}\right)^2>\left(\frac{s_i^2}{s_i^2+(\lambda+\delta)^2}\right)^2$, where $0<\lambda<1$ and $0<(\lambda+\delta)<1$, then averaged by noise realizations stochastic error component decreases with increase of λ:

$$\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm{T}}\right)>\sigma^2 trace\left(\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+(\lambda+\delta)^2}\right)^2\mathbf{U}^{\mathrm{T}}\right).$$

The error stochastic component with no averaging by noise realizations decreases with increase of λ:

$$\mathbf{e}^{\mathrm{T}}\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+\lambda^2}\right)^2\mathbf{U}^{\mathrm{T}}\mathbf{e}>\mathbf{e}^{\mathrm{T}}\mathbf{U}diag\left(\frac{s_i^2}{s_i^2+(\lambda+\delta)^2}\right)^2\mathbf{U}^{\mathrm{T}}\mathbf{e}.$$

An example of increase of deterministic and decrease of stochastic component of the error by the increase of the regularization parameter is presented in the Fig. 1. Since the increasing (stochastic) error component is scaling by the noise level σ^2, then the position of error minimum with the increase of the noise level shifts to the area of greater values of regularization parameter (Fig. 2).

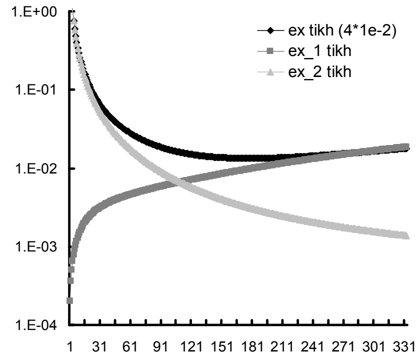

Fig. 1. An example of the dependency of deterministic and stochastic components of the error e_x by λ.

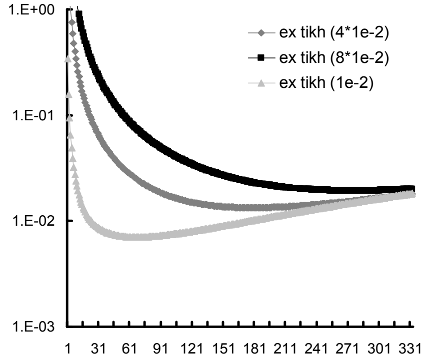

Fig. 2. Dependency $e_x(\lambda)$ for different noise levels.

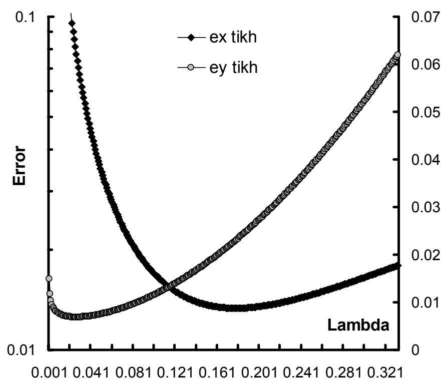

Fig. 3. Dependency of the true signal recovery error and output recovery error by λ.

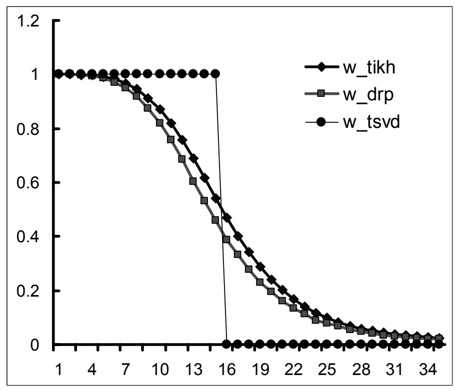

Fig. 4. The vector of weights for the Tikhonov regularization, DRP and TSVD methods.

We considered the dependence of the true signal recovery error and the output recovery error by the regularization parameter for the Phillips problem. There exists the optimal regularization parameter value for which true signal recovery error is minimal. Similar behavior is observed for output vector recovery error. True signal recovery error by Tikhonov regularization (as a truncated singular decomposition error) has two components. One of components decreases with increase of regularization parameter and another component increases.Solve DIP mean as accurate as possible recover the true signal. Thus, find regularization parameter value such as true signal recovery error is close to minimal. The choice of regularization parameter is based on output data which are corrupted by noise. There exists another class of problems with transformation input-output matrix has a DIP properties. These are true output signal recovery problems by noise corrupted data (measurements), for example, for the forecast of next output values. In problems from this class the value of regularization parameter is chosen such as the output vector recovery error is close to the minimal.

The value of regularization parameter for which true signal recovery error λ_x^{opt} is minimal and the value of regularization parameter with the minimal output vector recovery error λ_y^{opt} are different. The functional of Tikhonov method is build such that by «regulation» of parameter λ it is possible to reach both the minimization of true signal recovery error and the minimization of output vector recovery error.

When compare the Tikhonov regularization method and the truncated singular value decomposition method [9] by the approach to the solution of DIP based on choice of optimal number of components of linear model it appears that TSVD chooses the optimal number k_{opt} of the linear model components to minimize true signal recovery error. The rest of the $(N-k_{opt})$ components of singular value decomposition is not included to the model. The weight factor takes values 1 or 0.

The Tikhonov regularization method includes to the model all components of singular value decomposition and weighted each component (weights are real numbers). When compare experimentally obtained weights for TSVD and for Tikhonov regularization (Fig. 4), we see that weights of Tikhonov method form so called «soft threshold». Strongly sign-changing components (with big sequence number) have very small weights this prevents from approximation of noise by model. However, since these components are not excluded from the model they can work for approximation of signal of the complex form and thus provide greater (then TSVD) accuracy of Tikhonov regularization method.

We improved the accuracy of the basic random projection method using analytical averaging by random matrices [16]. Averaging over the realizations of matrices (in the experimental investigation) leads to a smoothing of the dependence of the error on k and a decrease in the number of local minima. This makes it easier to find the optimal value of k and increases the accuracy of the solution. Therefore, we did an analytical averaging over random matrices. The following expression was averaged

$$\mathrm{E}_R\{\mathrm{E}_\varepsilon\{e\}\} = \mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathrm{E}_R\{\mathbf{R}_k^{\mathrm{T}}(\mathbf{R}_k\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{R}_k^{\mathrm{T}})^{-1}\mathbf{R}_k\}\mathbf{A}\mathbf{x} + \\ + \sigma^2 trace\ \mathrm{E}_R\{\mathbf{R}_k^{\mathrm{T}}(\mathbf{R}_k\mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{R}_k^{\mathrm{T}})^{-1}\mathbf{R}_k\}.$$

The expressions for the error after averaging over random matrices are the next

$$\mathrm{E}_R\{\mathrm{E}_\varepsilon\{e\}\} = \mathbf{x}^{\mathrm{T}}\mathbf{x} - \mathbf{x}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{U}\mathbf{D}_k\mathbf{U}^{\mathrm{T}}\mathbf{A}\mathbf{x} + \sigma^2 trace\ (\mathbf{U}\mathbf{D}_k\mathbf{U}^{\mathrm{T}}).$$

$$\mathbf{D}_k \equiv \mathrm{E}_R\{\mathbf{R}_k^{\mathrm{T}}(\mathbf{R}_k\mathbf{S}^2\mathbf{R}_k^{\mathrm{T}})^{-1}\mathbf{R}_k\} = diag(\lambda_1,...,\lambda_m,\mu\mathbf{I}_{n-m}).$$

Averaging over random matrices leads to the diagonalization of the matrix included in both error components (deterministic and stochastic). An experimental study of $\mathbf{D}_k$ showed that the sequence $\lambda_1,\ldots,\lambda_m$ is bounded by the sequence $s_1^{-2},\ldots,s_m^{-2}$ from above and that several initial values of the diagonal of the matrix $\mathbf{D}_k$ approach s_i^{-2} with great accuracy. The values of the diagonal elements vary monotonically with k. This leads to a smoothing of the $e_x(k)$ and $e_y(k)$ dependences and to a decrease in the number of local minima [17–20].

Solution methods of DIP such as Tikhonov regularization, truncated singular value decomposition, "deterministic" random projection DRP to obtain the solution vector in one or another way weight singular values. The expression for input vector estimation in general case has a form: $\mathbf{x}^* = \mathbf{V}diag(s_i^{-1}w_i)\mathbf{U}^{\mathrm{T}}\mathbf{b}$. As known for the Tikhonov regularization weights are smoothly decrease (with the increase of indices of singular values):

$$\mathbf{x}_{tikh} = (\mathbf{A}^{\mathrm{T}}\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} = \mathbf{V}diag\left(\frac{s_i^2}{(s_i^2+\lambda)}\frac{1}{s_i}\right)\mathbf{U}^{\mathrm{T}}\mathbf{b},\ \ w_{tikh} = \frac{s_i^2}{(s_i^2+\lambda)}.$$

For the truncated singular value decomposition weights for $s_1 \ldots s_k$ are equal 1, and for $s_{k+1} \ldots s_n$ weights are equal 0. For the "deterministic" random projection DRP:

$$\mathbf{x}_{DR} = \mathbf{A}^{\mathrm{T}}\mathbf{U}\mathbf{D}_k\mathbf{U}^{\mathrm{T}}\mathbf{b} = \mathbf{V}\mathbf{S}\mathbf{D}_k\mathbf{U}^{\mathrm{T}}\mathbf{b} = \mathbf{V}diag\left(s_i^2 d_{ki}\frac{1}{s_i}\right)\mathbf{U}^{\mathrm{T}}\mathbf{b},\ w_{DR} = s_i^2 d_{ki},$$

values of weights smoothly decrease with the increase of the number of singular value. This is similar to the behavior of weights for the Tikhonov regularization method.

THE ERROR ANALISIS FOR MODIFIED RANDOM PROJECTION METHOD

To estimate the output vector of modified random projection method $\mathbf{x}^*_{MR}$ we study the input vector recovery error $e_{MR} = \left\|\mathbf{x} - \mathbf{x}^*_{MR}\right\|^2$. Taking into the account that output vector $\mathbf{b}$ can be represented as $\mathbf{Ax} + \mathbf{e}$ we obtain an expression for the input vector recovery error

$$e_{MR} = \left\|((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{Ax} - \mathbf{x} + ((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{e}\right\|^2 .$$

Average e_{MR} by noise realizations in output vector

$$e' = \mathrm{E}_\varepsilon\{e_{MR}\} = \left\|((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{Ax} - \mathbf{x}\right\|^2 +$$
$$+ \mathrm{E}_\varepsilon\{\left\|((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{e}\right\|^2\} = e'_1 + e'_2 .$$

Here e'_1 is deterministic and e'_2 is stochastic component of the error. Consider dependence of error components by the number of rows of random matrix and by the parameter λ. We represent error components with the use of singular value decomposition $svd(\mathbf{R}_k\mathbf{A}) = \mathbf{U}_r\mathbf{S}_r\mathbf{V}_r^{\mathrm{T}}$. We transform $((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A}$ by presenting $(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A}$ as a singular value decomposition taking into the account that $\mathbf{U}_r^{\mathrm{T}}\mathbf{U}_r = \mathbf{I}_k$, $\mathbf{V}_r^{\mathrm{T}}\mathbf{V}_r = \mathbf{I}_n$. Then we obtain

$$((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} = \mathbf{V}_r(\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{S}_r^{\mathrm{T}}\mathbf{S}\mathbf{V}_r^{\mathrm{T}} . \qquad (2)$$

For the deterministic component of the error by (2) we obtain:

$$e'_1 = \mathbf{x}^{\mathrm{T}}(\mathbf{V}_r((\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{S}_r^{\mathrm{T}}\mathbf{S}(\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r - 2(\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{S}_r^{\mathrm{T}}\mathbf{S})\mathbf{V}_r^{\mathrm{T}} + \mathbf{I})\mathbf{x},$$

$$e'_1(\lambda) = \mathbf{x}^{\mathrm{T}}\mathbf{V}_r diag\left(\frac{s_i^4}{(s_i^2 + \lambda)^2}\right)\mathbf{V}_r^{\mathrm{T}}\mathbf{x} - 2\mathbf{x}^{\mathrm{T}}\mathbf{V}_r diag\left(\frac{s_i^2}{(s_i^2 + \lambda)}\right)\mathbf{V}_r^{\mathrm{T}\,\mathrm{T}}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{x},$$

$$e'_1(\lambda) = \sum_{i=1}^{N}\mathbf{x}^{\mathrm{T}}\mathbf{v}_i\left(\frac{s_i^4 - 2(s_i^2 + \lambda)^2}{(s_i^2 + \lambda)^2}\right)\mathbf{v}_i^{\mathrm{T}}\mathbf{x} + \mathbf{x}^{\mathrm{T}}\mathbf{x} . \qquad (3)$$

The stochastic error component:

$$e'_2 = \mathrm{E}_\varepsilon\{\left\|((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}(\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{e}\right\|^2\} =$$
$$= \sigma^2 trace(\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k\mathbf{A}((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1\mathrm{T}}((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k) .$$

To the ortonormal matrix $\mathbf{R}_k$, $\mathbf{R}_k\mathbf{R}_k^{\mathrm{T}} = \mathbf{I}$, which is obtained in singular value decomposition of the matrix $\mathbf{G}$ with elements $\mathbf{G}_{ij} \sim N(0, \text{y}^2)$, the expression for e'$_2$ got a form

$$e'_2 = \sigma^2 trace(\mathbf{R}_k\mathbf{A}((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1\mathrm{T}}((\mathbf{R}_k\mathbf{A})^{\mathrm{T}}\mathbf{R}_k\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{R}_k^{\mathrm{T}}).$$

Using the singular value decomposition we obtain and since $\mathbf{U}_r^{\mathrm{T}}\mathbf{U}_r = \mathbf{I}_k$ and using the cyclic property of $trace(\cdot)$ we get that

$$e'_2 = \sigma^2 trace(\mathbf{S}_r\mathbf{V}_r^{\mathrm{T}}\mathbf{V}_r(\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{V}_r^{\mathrm{T}}\mathbf{V}_r(\mathbf{S}_r^{\mathrm{T}}\mathbf{S}_r + \lambda\mathbf{I})^{-1}\mathbf{V}_r^{\mathrm{T}}\mathbf{V}_r\mathbf{S}_r^{\mathrm{T}}).$$

Since $\mathbf{V}_r^{\mathrm{T}}\mathbf{V}_r = \mathbf{I}_n$,

$$e'_2(\lambda) = trace(diag\left(\frac{s_i^2}{(s_i^2+\lambda)^2}\right)) = \sum_{i=1}^{N}\left(\frac{s_i^2}{(s_i^2+\lambda)^2}\right). \quad (4)$$

Expressions (3), (4) allows to study components of error for the modified method of random projection depending on regularization parameter.

Table 1. The mean value of the solution error and its standard deviation, mean value of *k* and its standard deviation for problems of Phillips (N1), Carasso (N2), Delves (N3) [22].

N	σ^2		e	CR_Q	Cp	AIC	MDL
1	10^{-2}	E(e)	6.44E-3	8.15E-3	7.60E-3	8.15E-3	7.30E-3
		std	1.56E-3	5.85E-3	2.48E-3	5.85E-3	2.17E-3
		E(k)	7.70	7.60	7.20	7.60	7.26
		std	0.68	0.57	0.64	0.57	0.63
	10^{-4}	E(e)	3.05E-4	3.37E-4	5.46E-4	3.38E-4	6.28E-4
		std	3.84E-5	6.02E-5	1.65E-4	5.02E-5	2.20E-4
		E(k)	16.76	15.48	12.34	15.06	11.92
		std	1.64	1.43	1.19	1.25	1.05
2	10^{-2}	E(e)	7.69E-1	1.04	1.06	1.07	1.00
		std	1.81E-1	3.80E-1	1.41E-1	3.21E-1	1.34E-1
		E(k)	6.86	7.16	3.12	5.98	4.04
		std	2.00	2.32	0.94	2.15	1.23
	10^{-4}	E(e)	1.50E-2	2.79E-2	3.35E-2	1.99E-2	5.09E-2
		std	4.91E-3	2.05E-2	1.86E-2	5.89E-3	3.68E-2
		E(k)	21.7	23.34	16.82	19.34	15.46
		std	1.93	3.15	1.99	1.92	2.58
3	10^{-3}	E(e)	6.95E-2	1.91E-1	9.21E-2	1.91E-1	8.05E-2
		std	1.36E-2	8.02E-1	2.50E-2	8.02E-1	1.86E-2
		E(k)	3.30	2.48	1.64	2.48	2.14
		std	0.953	0.91	0.49	0.91	0.45
	10^{-5}	E(e)	1.24E-2	1.67E-2	2.02E-2	1.46E-2	2.09E-2
		std	1.84E-3	1.21E-2	3.89E-3	1.76E-3	4.39E-3
		E(k)	16.14	15.34	8.76	12.72	8.4
		std	2.14	2.58	1.72	2.03	1.84

In the Table 1 we presented results of the solution of ill-posed problems by modified random projection method, the matrix **A** is 40×40. Optimal value of k was defined by model selection criteria: Cp Mallows, AIC Akaike, MDL minimal description length and CR_Q [23, 15]. Results analysis showed that usage of the modified random projection method with criteria CR_Q and AIC Akaike assures the error value close to the optimal.

CONCLUSIONS

Modified random projection method (MRP) is marked with solution stability and increased accuracy. It is achieved by simultaneous regularization by ridge regression and random projection. Previously we proved that random projection has a regularizing properties and increases solution stability. The modification is represented by that the additional increase of stability is reached by choosing the regularization parameter. This parameter assures the balance between the Euclidean norm of the solution vector discrepancy and the norm of the vector itself. Thus for MRP we obtained expressions for deterministic and stochastic components of the solution error. Dependencies of error components on regularization parameters (as on the regularization parameter of ridge regression and on the number of rows of random projector) were also obtained. These results are approved by experimental study. For solving the discrete ill-posed problems by MRP the optimal size of random projection matrix is defined by the developed by us special criteria and the regularization parameter for ridge regression is defined by the generalized discrepancy criteria.

Hence the modified method of random projecting is characterized by stability and increased accuracy of the solution. This achieved by simultaneous ridge regression style regularization and random projecting. The representation of the solution error in the form where error components are related to the matrix specter and regularization parameter is important for further study of the error.

REFERENCES

1. Revunova E.G., Rachkovskij D.A. Using randomized algorithms for solving discrete ill-posed problems. *Intern. Journal Information Theories and Applications*. 2009, Vol. 2, N. 16, pp.176–192.
2. Durrant R.J., Kaban A. Random projections as regularizers: learning a linear discriminant from fewer observations than dimensions. *Machine Learning*. 2015, vol. 99, N 2, pp. 257–286.
3. R.J. Durrant , A. Kaban. Compressed Fisher Linear Discriminant Analysis: Classification of Randomly Projected Data. *Proceedings of 16th ACM SIGKDD* Conference on Knowledge Discovery and Data Mining (KDD 2010), 2010.
4. Xiang H., Zou J. Regularization with randomized SVD for large-scale discrete inverse problems. *Inverse Problems*. 29(8):085008, 2013.
5. Xiang H., Zou J. Randomized algorithms for large-scale inverse problems with general Tikhonov regularizations. *Inverse Problems*. 2015, Vol. 31, N 8:085008, pp. 1–24.
6. Wei Y., Xie P., Zhang L. Tikhonov regularization and randomized GSVD. *SIAM J. Matrix Anal. Appl.* 2016, Vol. 37, N 2, pp. 649–675.
7. Hansen P. Rank-deficient and discrete ill-posed problems. Numerical aspects of linear inversion. Philadelphia: SIAM. 1998. 247 p.
8. Tikhonov A., Arsenin, V. Solution of ill-posed problems. Washington: V.H. Winston. 1977. 231 p.
9. Hansen P.C. The truncated SVD as a method for regularization. BIT 27, (1987), pp. 534–553.

10. Rachkovskij D.A., Revunova E.G. Randomized method for solving discrete ill-posed problems. *Cybernetics and Systems Analysis*. 2012, Vol. 48, N. 4. pp. 621–635.
11. Revunova EG, Rachkovskij DA, Stable transformation of a linear system output to the output of system with a given basis by random projections. The 5th Int. Workshop on Inductive Modelling (IWIM'2012), Kyiv, 2012, pp. 37–41.
12. Revunova E.G. Randomization approach to the reconstruction of signals resulted from indirect measurements. *Proc. 4th International Conference on Inductive Modelling* (ICIM'2013), Kyiv, 2013, pp. 203–208.
13. Revunova E.G., Tyshchuk A.V. A model selection criterion for solution of discrete ill-posed problems based on the singular value decomposition. *The 7th International Workshop on Inductive Modelling* (IWIM'2015), Kyiv-Zhukyn, 2015, pp.43–47.
14. Revunova E.G. Analytical study of the error components for the solution of discrete ill-posed problems using random projections. *Cybernetics and Systems Analysis*. 2015, Vol. 51, N. 6, pp. 978–991.
15. Revunova E.G. Model selection criteria for a linear model to solve discrete ill-posed problems on the basis of singular decomposition and random projection. *Cybernetics and Systems Analysis.* 2016, Vol. 52, N.4, pp.647–664.
16. Revunova E.G. Averaging over matrices in solving discrete ill-posed problems on the basis of random projection. Proc. CSIT'17. 2017. Vol. 1. P. 473–478.
17. Revunova E.G. Solution of the Discrete ill-posed problem on the basis of singular value decomposition and random projection. Advances in Intelligent Systems and Computing II. Cham: Springer, 2017, pp. 434–449.
18. Revunova E.G. Improving the accuracy of the solution of discrete ill-posed problem by random projection. *Cybernetics and Systems Analysis.* 2018. Vol. 54, N 5, pp. 842–852.
19. Revunova E.G., Tyshcuk O.V., Desiateryk O.O. On the generalization of the random projection method for problems of the recovery of object signal described by models of convolution type. *Control Systems and Computers.* 2021, N 5–6, pp. 25–34.
20. Tyshchuk O.V., Desiateryk O.O., Volkov O.E., Revunova E.G., Rachkovskij D.A., A linear system output transformation for sparse approximation. *Cybernetics and Systems Analysis.* 2022, Vol. 58, N. 5, pp. 840–850.
21. Marzetta T., Tucci G., Simon S. A random matrix-theoretic approach to handling singular covariance estimates. *IEEE Trans. Information Theory*. 2011. Vol. 57, N 9, pp. 6256–6271.
22. Hansen P. C. Regularization Tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Algorithms*. 1994, Vol. 6, N 1, pp. 1–35.
23. Rachkovskij D.A, Revunova E.G. Intelligent gamma-ray data processing for environmental monitoring. In: Intelligent data analysis in global monitoring for environment and security. Kiev-Sofia: ITHEA. 2009, pp. 124–145.

Received 04.10.2022

Ревунова О.Г.[1], докт.техн. наук, старш. наук. співроб.,
провід. наук. співроб. відд. нейромережевих технологій оброблення інформації,
https://orcid.org/0000-0002-3053-7090, e-mail: egrevunova@gmail.com
Тищук О.В.[2],
старший інженер-програміст,
e-mail: avtyshcuk@gmail.com
Десятерик О.О.[3], канд.фіз.-мат. наук,
асистент кафедри алгебри і комп'ютерної математики,
механіко-математичного факультету
e-mail: sasha.desyaterik@gmail.com

[1] Міжнародний науково-навчальний центр інформаційних технологій
та систем НАН України та МОН України,
40, пр. Акад. Глушкова, м. Київ, 03187, Україна,
[2] Roku Inc., м. Київ, Україна,
[3] Київський національний університет імені Тараса Шевченка,
4е, пр. Акад. Глушкова, м. Київ, 03127, Україна

ТЕХНОЛОГІЯ ДЛЯ СТІЙКОГО РОЗВ'ЯЗАННЯ ДИСКРЕТНИХ НЕКОРЕКТНИХ ЗАВДАНЬ МОДИФІКОВАНИМ МЕТОДОМ ВИПАДКОВОГО ПРОЕКТУВАННЯ

Вступ. *Розв'язання некоректних завдань є актуальним для багатьох галузей науки та техніки. Дискретні некоректні завдання (ДНЗ) виникають, наприклад, у разі дискретизації інтегральних рівнянь у таких галузях як спектрометрія, гравіметрія, магнітометрія, електророзвідка тощо.*

У разі лінійної ДНЗ матриця, яка моделює деяку вимірювальну систему, виконує лінійне перетворення вхідного вектора у вектор виходу. Для ДНЗ характерно, що вектор виходу містить шум і ряд сингулярних чисел матриці спадає до нуля плавно. Водночас розв'язок (оцінка вхідного вектора) з використанням обернення матриці перетворення є нестійким та неточним. Для подолання нестійкості та підвищення точності розв'язку використовують методи регуляризації.

Нами розробляється підхід, в якому для отримання стійкого розв'язку ДНЗ використовуються регуляризувальні властивості випадкового проектування. Розроблення ефективних стійких методів розв'язання ДНЗ продовжує залишатися актуальною проблемою.

Мета. *Підвищення точності методу розв'язання ДНЗ на основі випадкової проекції.*

Результати. *Розроблено метод стійкого розв'язання ДНЗ на основі модифікованого методу випадкової проекції. Модифікація полягає в тому, що регуляризація на основі випадкової проекції доповнюється регуляризацією у стилі гребеневої регресії.*

Для розробленого методу отримано вирази, що пов'язують у явному вигляді складові помилки розв'язку зі спектром матриці та параметром регуляризації. Також проведено експериментальне дослідження точності розробленого метода на тестових задачах.

Висновки. *Модифікований метод випадкового проектування відрізняється стійкістю та підвищеною точністю розв'язку, які досягаються за рахунок одночасної регуляризації у стилі гребеневої регресії та випадковим проектуванням. Подання помилки розв'язку, де складові помилки розв'язку пов'язані зі спектром матриці та параметром регуляризації, є важливим для подальшого вивчення меж зміни помилки.*

Ключові слова: *дискретна некоректна задача, випадкова проекція.*