

УДК 512.54

Numerical Evaluation of the Truncated Singular Value Decomposition Within the Seismic Traveltimes Tomography Framework

Alexandr S. Serdyukov*

Novosibirsk State University,
Pirogova, 2, Novosibirsk, 630090
Russia

Andrey V. Patutin[†]

Tatiana V. Shilova[‡]

Institute of Mining,
Krasny, 54, Novosibirsk, 630091
Russia

Received 20.08.2013, received in revised form 25.12.2013, accepted 20.02.2014

The method of truncated singular value decomposition is a powerful tool of regularization and solution of inverse problems. Application of this method is limited by the memory requirements for the calculation of singular value decomposition of matrices. The solution is a usage of iterative procedures, that provide the possibility to evaluate only the largest singular values and corresponding vectors. For the practical application of this methodology one should answer the questions: what part of the spectrum can be determined numerically and whether the number of singular vectors is enough for the solution. These problems are considered in the present paper within the seismic traveltimes inversion framework.

Keywords: inverse problems, numerical methods, truncated singular value decomposition, geophysics, seismic traveltimes.

Introduction

The seismic traveltimes tomography is the conventional method of the elastic medium parameters estimation, using the times that are expended for the seismic waves propagation from the sources to receivers [1]. After the linearisation the problem is reduced to the solution of integral equation with the compact operator (tomography operator) in the right hand side. In the case of the least squares solution the most natural regularization approach is the truncated singular value decomposition (SVD). In practice the infinite dimensional tomography operator is replaced by the matrix approximation (tomography matrix), constructed by the discretization of the data and model spaces. Thus one comes to a question: how strong is the identity between the tomography matrix SVD, used for the numerical data inversion, and the SVD of the initial tomography operator. In the present paper we observe the behaviour of the tomography matrix SVD during the discretization step splitting via the model numerical experiment.

*AleksanderSerdyukov@yandex.ru

[†]ss3032@yandex.ru

[‡]ShilovaTanya@yandex.ru

© Siberian Federal University. All rights reserved

1. Theory

Firstly we'll give the brief introduction to the seismic traveltimes tomography. Than, we'll introduce the r-solution concept, following the paper [2], which is the theoretical basement of the truncated SVD approach.

1.1. Seismic traveltimes tomography

Seismic tomography is the group of methods for the Earth's subsurface structure recovering, using the measurements of seismic waves. One can use the whole seismograms, recorded in the number of receivers, as the data, that leads to the inverse problem for the elastic wave equation. This approach is known as *full wave form inversion* or *waveform tomography*. As an alternative, one can pick up the traveltimes from the seismograms and use them for inversion, that leads (in the high-frequency limit approximation) to the inverse problem for eikonal equation. This kinematic approach, known as *traveltimes tomography*, is considered in the present paper.

It turns out that the traveltimes are more linearly related to large-scale velocity perturbations than the waveforms, so one should resort to traveltime tomography to resolve these structures. Also the traveltimes inversion is preferable when the dynamical characteristics of the seismic waves (waveforms and amplitudes) can't be recorded accurately, for instance in the case of unmeasurable conditions of wave initializing in the sources. Actually, the kinematic and dynamic inversion methods are complementing each other during the seismic data inversion [3].

The seismic traveltime tomography is conventionally considered in the high-frequency approximation framework, called *geometrical seismics*. The seismic waves are considered as the propagation of the weak discontinuities of the wave equation solution. Let's consider the isotropic elastic equations:

$$\frac{\partial}{\partial x_k} \left(\lambda \frac{\partial u_i}{\partial x_i} \right) + \frac{\partial}{\partial x_i} \left(\mu \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right) \right) = \rho \frac{\partial^2 u_k}{\partial t^2}, \quad (1)$$

where u_i , $i = 1 \dots 3$ are the components of displacement vector, λ , μ and ρ are the Lamé parameters and density of the elastic medium. We use the notation of the summation over the repeated indexes in the left hand side of the equation (1).

The weak discontinuities of the solution of the elastic equation (1) can exist only on the characteristics. The characteristic relation leads to the eiqonal equation [4]:

$$|\nabla \tau|^2 = \frac{1}{V_P^2}, \quad (2)$$

where $\tau(\mathbf{x})$ is the traveltime for the space point \mathbf{x} , $V_P^2 = (\lambda + 2\mu)/\rho$. V_P is the pressure wave velocity (we are going to consider only this type of elastic waves). The inverse quantity for the velocity, $s = 1/V_P$ standing in the right hand side of the eikonal equation is the slowness.

The seismic traveltimes tomography is the inverse problem for the eikonal equation (2). The limitation of the acquisition system and strong curvilinearity of the seismic rays (rays are the characteristics of eiqonal equation) in the complex velocity models make the procedures of back-projection and convolution, that are used in X-ray tomography and are based on the Fourier transform, not really suitable for seismic tomography. The most general and strong for seismic traveltimes inversion is the matrix approach [5], based on the linearization and discretization that reduce the problem to a solution of the large linear algebraic system.

Let's introduce the small disturbance δs of the known slowness s . Consider the elastic medium with the unknown slowness that is the sum: $s + \delta s$. The traveltimes in this medium would be the sums: $\tau + \delta \tau$. The linearisation of the eikonal equation (2) leads to the following relation for every source-receiver pair:

$$\delta \tau = \int_{\Gamma_0} \delta s \, dl, \quad (3)$$

where the integral in the right hand side is taken over the ray trajectory (i.e the characteristic of the eikonal equation) connecting source with the receiver. The problem is to find the small disturbance of the slowness δs , using the traveltimes residuals $d\tau$.

Let's consider the rectangular grid discretization of the target area in the model space. The disturbances of the slowness are supposed to be constant in each cell. The integral from the relation (3) for every source-receiver pair would be the corresponding sum. Thus the problem is reduced to the determination of the unknown vector of the disturbances values $\vec{\delta s}$ in the rectangular grid cells, using the traveltime data vector $\vec{\delta\tau}$, which is composed by the traveltimes residuals for all source-receiver pairs. That leads to the solution of the tomographic linear algebraic system:

$$B\vec{\delta s} = \vec{\delta\tau},$$

where B is the matrix approximation of the tomographic operator, called tomography matrix. As it is seen from the relation (3), one needs only to know the length of the ray path through each cell of the rectangular grid to construct the tomography matrix.

In the majority of the researches the effect of the discretization parameters during the matrix approximation on the solution of the tomography problem is explored just by synthetic data inversion. It is clear that such approach is not rigorous and not general enough. Actually, there are some theoretical results, connecting the discretization parameters with the resolution of the tomography method. These results are based on the usage of the Fourier transform. They don't take into account the compactness of the tomography operator and the present of the data and approximation errors [5]. That's why the matrix approximation SVD convergence during the discretization step splitting is need to be observed.

1.2. R-solution

In this section we'll introduce the r-solution concept and some theoretical facts, following the paper [2] Let's consider the operator equation:

$$Ax = f, \tag{4}$$

where $A: X \rightarrow Y$ is linear compact operator, acting from separable Hilbert space X into the Hilbert space Y .

By virtue of compact operator singular value decomposition theorem [6], the operator $A: X \rightarrow Y$ could be decomposed. It means that there exist orthonormal basis in the model space: $\{x_n \in X, n \in N\}$ — *right singular vectors*, orthonormal basis in the data space: $\{y_n \in Y, n \in N\}$ — *left singular vectors* and set of nonnegative scalar values: $\{s_n \geq 0, n \in N\}$ — *singular values* such that:

$$\begin{aligned} Ax_n &= s_n y_n, \\ A^* y_n &= s_n x_n, \\ \overline{span}\{x_n\} &= \overline{R}(A^*) = N(A)^\perp, \\ \overline{span}\{y_n\} &= \overline{R}(A) = N(A^*)^\perp, \end{aligned} \tag{5}$$

where $R(A)$, $(R(A^*))$ is an image and $N(A)$, $(N(A^*))$ is a kernel of operator A (A^*), $span(x_n)$ and $span(y_n)$ are the linear subspaces, spanned by the n singular vectors. The set $\{s_n, n \in N\}$ doesn't have any nonzero limit points.

Note that compact operators have the following important property [7]: *the multiplicities of the nonzero singular values are equal to one*. The singular values and corresponding vectors are ordered in the singular value descending order: $s_1 \geq s_2 \geq s_3 \geq \dots \geq s_n \dots$.

For any right hand side of the equation (4), satisfying the condition:

$$f \in R(A) \oplus N(A^*),$$

the minimal norm extremal solution is defined as follows [8] :

$$x^{**} = A^\dagger f, \quad (6)$$

where A^\dagger is the pseudoinverse operator, that is determined by using the SVD (5):

$$A^\dagger = \sum_{s_n \neq 0} \frac{(f, y_n)}{s_n} x_n.$$

Definition. *r-solution is the element of the model space that is defined by the following formula:*

$$x_{[r]} = A_r^\dagger f = \sum_{n=1}^r \frac{(f, y_n)}{s_n} x_n.$$

If the minimal norm extremal solution (6) exist, r-solution is represented by its projection on the linear subspace, spanned by the r largest right singular vectors. The operator A_r^\dagger is called the r-pseudoinverse. This operator is defined and limited on the whole space Y . Moreover, for every $f \in D(A^\dagger)$ and $\forall \delta > 0$ always exist $r(\delta)$, such that for all $r > r(\delta)$ the following inequality holds: $\| A^\dagger f - A_r^\dagger f \| < \delta$. Therefore, the set of operators A_r^\dagger is approximating the pseudoinverse operator A^\dagger on its definition domain. Thus it can be considered as a regularization of the inversion [9].

Within the numerical solution of equation (4) neither the operator A , nor the right-hand-side f are known exactly. One has \tilde{A} and \tilde{f} instead. The approximation and data errors are given as follows:

$$\| A - \tilde{A} \| \leq \delta_1 \quad \| f - \tilde{f} \| \leq \delta_2.$$

The disturbed right-hand side \tilde{f} may not belong to the definition domain of the pseudoinverse operator \tilde{A}^\dagger even if $f \in D(A^\dagger)$. That's why the observation of the difference $\| A^\dagger f - \tilde{A}^\dagger \tilde{f} \|$ during $\delta_{1,2} \rightarrow 0$ has no sense. However, since the operators A_r^\dagger and \tilde{A}_r^\dagger are defined and limited on the whole space Y , the difference $\| A_r^\dagger f - \tilde{A}_r^\dagger \tilde{f} \|$ specifies the evaluation accuracy of the exact equation (4) solution projection on the first r right singular vectors for the given levels of operator approximation and data errors.

The solution of some inverse problems consists in finding a function, depending on several variables, in some limited target area. As usual, the number of variables is not more than three. In particular, the seismic tomography problems are consist in finding the unknown elastic velocity distribution, that is the function of spatial variables. In practice the data turns out to be a vector, so the data space is finite dimensional. Almost all numerical methods of solving such inverse problems are based on finding decomposition coefficients, corresponding to the finite set of some basis functions. The common choice is the 'step' functions, that equal to one in the grid cells and zeros outside. Thus the passage to the finite model and data dimensions X and Y is made. The equation (4) becomes the large linear algebraic system in this case. The r-solution and other related notions, can be introduced similar way in the case of algebraic system. The singular value decomposition of the matrix corresponds to the SVD (5) of initial operator. It should be noted that, the matrix, approximating the initial compact operator, derives it's properties. The linear system is ill-conditioned. In the case of exact operator the number r for which the r-solution can be stably determined for the known data error level is given by the relation:

$$\frac{\| \delta x_{[r]} \|}{\| x_{[r]} \|} \leq \frac{s_1}{s_r} \frac{\| \delta f \|}{\| f \|}, \quad (7)$$

where $\delta x_{[r]}$ is the residual, coming out from finding the r -solution $x_{[r]}$, using the data f , containing the error δf .

The matrix approximation of the initial operator leads to the errors. So, the estimate (7) can't be used. In the paper [2] there is the estimate of the r -solution accuracy in the case of inexact operator. From this estimate follows that the r -solutions of the linear algebraic systems, that approximate initial operator, are converging to the r -solution of the initial operator by the splitting of the discretization step. However, the estimate doesn't provide the number of singular vectors that ensure the stability of solution for the given discretization parameters. The goal of our research is to produce the numerical exploration of the SVD of the seismic traveltime tomography operator matrix approximation.

2. The numerical study of the tomography matrix SVD convergence during the discretization step splitting

As was previously pointed out, the stable recovering subspaces are spanned by the 'largest' right singular vectors. Let's observe the effect of the discretization step splitting on the tomography matrix SVD by considering the model example. The target area is the square, on the opposing sides of which N sources and N receivers are placed, see the Fig. 1. Such acquisition geometry is common for cross-well tomography. For the reason of simplicity and numerical experiment purity, we suppose that the initial undisturbed (containing) medium is homogeneous. In this medium the rays are just straight lines, so the tracing for every source-receiver pair and the ray path length computation for each grid cell are evaluated analytically and don't bring the additional bias. The initial infinite dimensional operator appears when $N \rightarrow \infty$: the sources and receivers are placed in every point of the corresponding side of the square and velocity is the two dimensional function, defined in the considered target area.

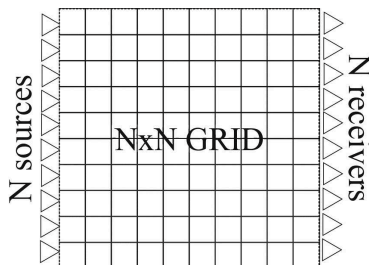


Fig. 1. Model exaple: aquisition system and discretization

We explore the behavior of the singular spectrum via increasing the number N . Let's compare the first largest 500 singular values and right singular values for the sequence: $N = 100, 200, 400, 800$.

We have performed numerical SVD by using the ARPACK package. Although the tomography matrix is sparse (and stored in the special format) the memory space of usual PC was not enough in the case $N = 800$, so the numerical computations were made on four-processor cluster with common-pool with 256 GB memory space. The codes were automatically threaded and optimised.

On the Fig. 2 on can see the plots of the relative residuals between the singular values:

$$s_{N,2N}(i) = \frac{s_N(i) - s_{2N}(i)}{s_{2N}(i)}, \quad (8)$$

where $s_N(i)$ is the i -th singular value of tomography matrix for the corresponding N .

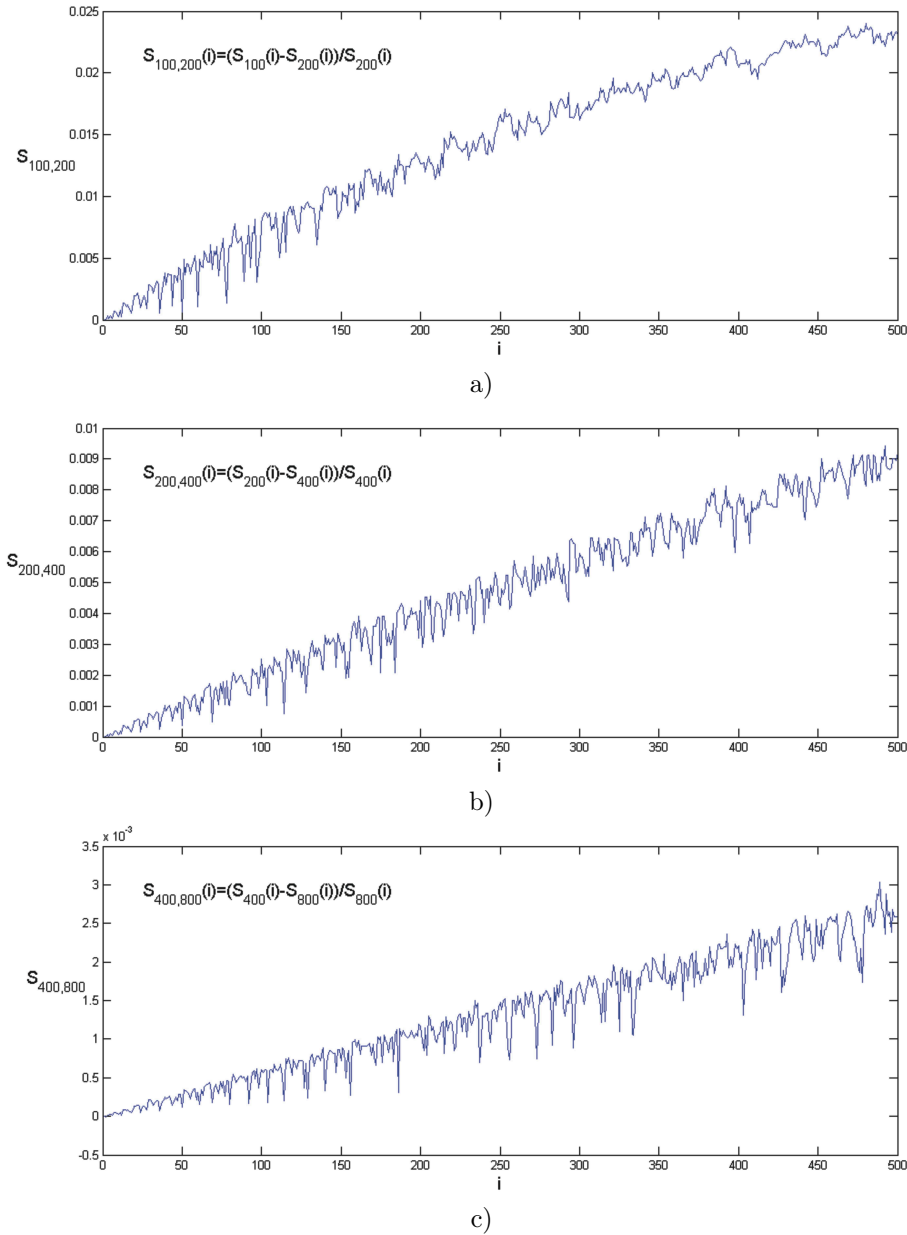


Fig. 2. The convergence of singular spectrum for the cross-well tomography. The relative difference between singular values for a) $N = 100$ and $N = 200$, b) $N = 200$ and $N = 400$, c) $N = 400$ and $N = 800$

As it seen form the Fig. 2 the spectrum begins to converge from the largest singular values during the mesh splitting. As follows form the change of the singular spectrum by the duplication of the number N , the tomography matrix approximate the initial infinite dimensional operator at least with first order accuracy. Using the given comparison one can estimate the precision of singular values numerical evaluation.

The right singular vectors produce the basis in the model space. Thus the accuracy of their

computation is also important. For the convergence measurement lets consider the following relation:

$$\delta_{N,2N}(i) = \frac{\| \overline{V_N(i)} - V_{2N}(i) \|}{\| V_{2N}(i) \|} \quad (9)$$

where $V_{2N}(i)$ is the i -th right singular vector for the $2N \times 2N$ grid. $\overline{V_N(i)}$ is the vector defined on the same grid $2N \times 2N$, produced from the vector $V_N(i)$ computed on the grid $N \times N$ by linear interpolation.

On the Fig. 3 one can see $\delta_{N,2N}$ for the sequence $N = 100, 200, 400$ for the cross-well tomography. From these plots one can estimate what part of the singular vectors are computed reliably for the given grid. Note, that the first ‘largest’ (corresponding to the largest singular values) right singular vectors are computed more accurately. For instance, from the Fig. 3 it is seen, that about 70 right singular vectors are the same for the whole sequence of N .

By observing the singular vector images, one can define some properties. The sequence of singular vectors $V_{800}(i)$, $i = 10, 68, 376$ is shown on the Fig. 4 a), b), c). Note, that the ‘pattern’ structure becomes more fine during the rank number increasing. This means, that low spatial frequency velocity structures are restored better during the inversion. One can also see, that the ‘pattern’ is more fine in the middle part of the target area. Thus, finer velocity structures can be resolved there.

Note that resolution of the spacial grid is not enough for the imaging of singular vectors starting from some large number. As it seen from the Fig. 4 c), the grid 100×100 turns out to be too coarse for the approximation of 376-th right singular vector. It is different from the vector, shown Fig. 4 b). From the point of resolution concept, it is more important to explore the convergence of the right singular values subspaces instead of single vectors. The measure of the two subspaces equivalence is the angle between them (see [10]). Let’s introduce the definition of this angle. Consider the Hilbert separable space H . Let X_1, X_2 be the two its N -dimensional subspaces. Introduce the matrix:

$$G = \left[(x_i^{(1)}, x_j^{(2)}) \right]_{i,j=1}^N,$$

composed of the scalar products of the orthonormal bases of X_1 and X_2 . The orthogonal substitutions of the X_1 and X_2 bases lead to the matrix transformation by the orthogonal matrices V, U . Using the SVD theorem for the matrix G , one can find V and U such that there would be the diagonal matrix in the result:

$$G = \Sigma = \text{diag}(\sigma_j)$$

with the elements $\sigma_j \geq 0$. This means that biorthogonal bases are chosen in the subspaces X_1 and X_2 .

$$\begin{aligned} (x_j^{(1)}, x_j^{(2)}) &= \sigma_j, \\ (x_i^{(1)}, x_j^{(2)}) &= 0, \quad i \neq j \end{aligned} \quad (10)$$

where $i, j = 1, \dots, n$. By the definition $\sigma_j = \cos \phi_j$, where $\pi/2 \geq \phi_j \geq 0$ are the angles between vectors $x_j^{(1)}$ и $x_j^{(2)}$. The increase-order sorted set of these angles $\{\phi_1, \phi_2, \dots, \phi_N\}$ is called the *series of angles* between the subspaces X_1 and X_2 , the maximum of them is the angle between subspaces.

On the Fig. 5 the series of angles between subspaces, spanned by the i largest singular vectors for $N = 100$ and $N = 200$, are given. The vectors of the coarse grid are transported to the finer grid by the linear interpolation. From the Fig. 5 it is seen that the angles are so small that the accuracy of their computation is doubtful. From the similar plots for the finer grids one can see

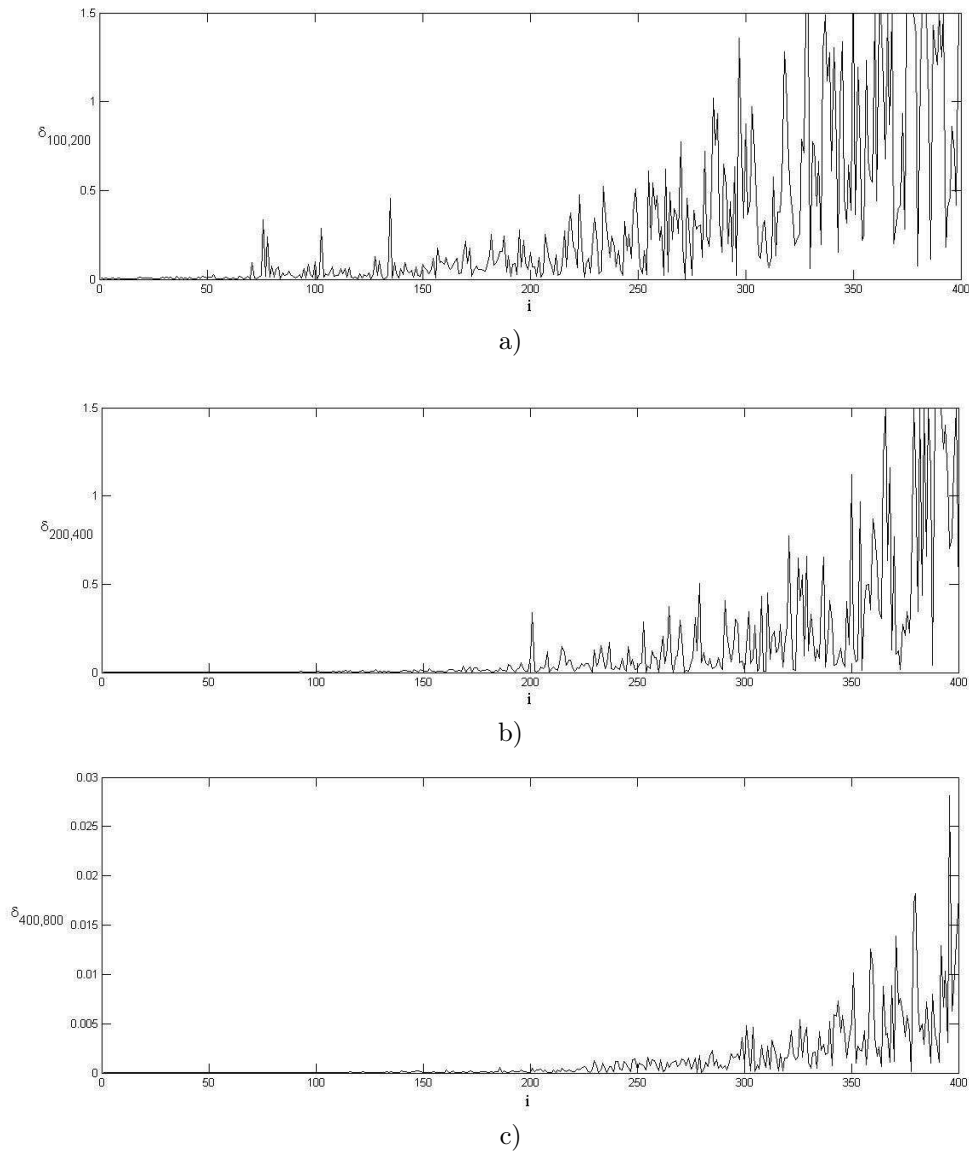


Fig. 3. The residuals between singular vectors for different grids; a) $N=100$ и $N=200$, б) $N=200$ и $N=400$, в) $N=400$ и $N=800$

that the order of angles is decreasing about two times by the dividing into two the mesh step. We will not present here the other plots, since they are spurious. Although the angles between subspaces seem to be small, it is hard to deduce the subspaces equivalence from that, because the dimensions of the considered spaces are huge.

Let's consider the example of the singular value decomposition usage for the solution of cross-well tomography problem. The target region is the square 400×400 meters, split by the rectangular grid with the step 1 meter. The 400 sources and 400 receivers were used. The acquisition system was similar to the previous (see Fig. 1). The synthetic velocity model and the result of the inversion are shown on the Fig. 6. The homogeneous medium with pressure wave velocity equal to 2700 m/s was used as initial model. The 400 hundred 'first' singular

value vectors was used for synthetic data inversion. The usage of the large part of the singular spectrum has no sense, cause the rest part of the spectrum is computed unreliably for the given spatial discretization step.

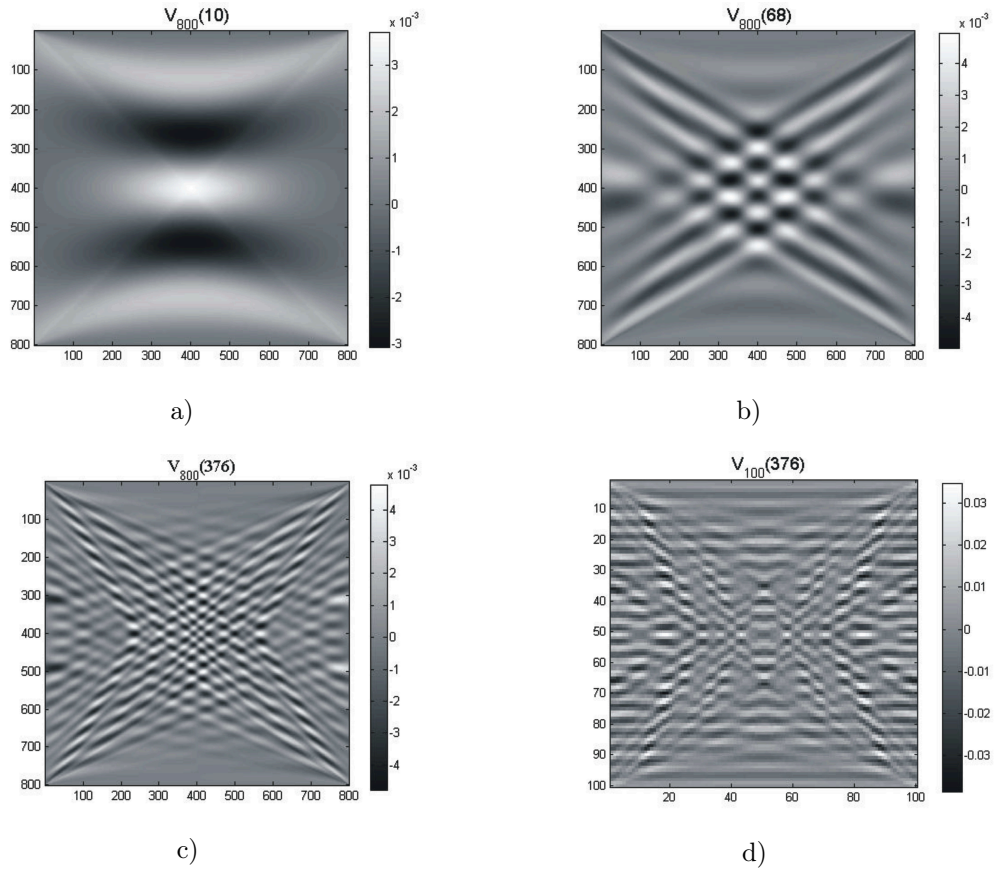


Fig. 4. Right singular vectors $V_N(i)$: a) $N = 800, i = 10$, b) $N = 800, i = 68$, c) $N = 800, i = 376$, d) $N = 100, i = 376$

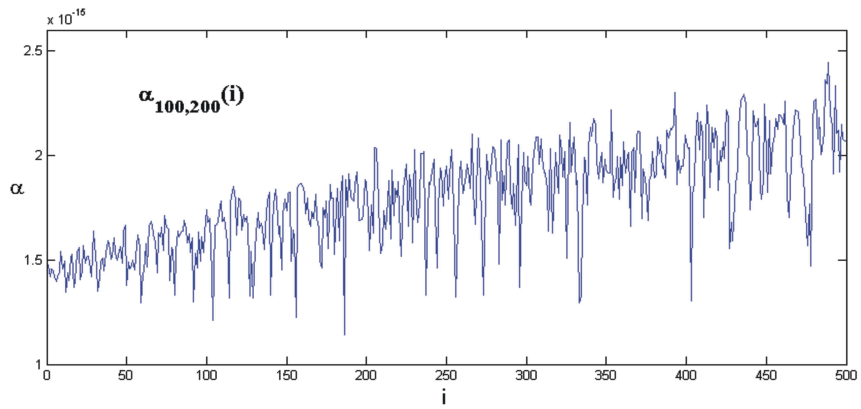


Fig. 5. The right singular subspaces angles for $N=100, N=200$

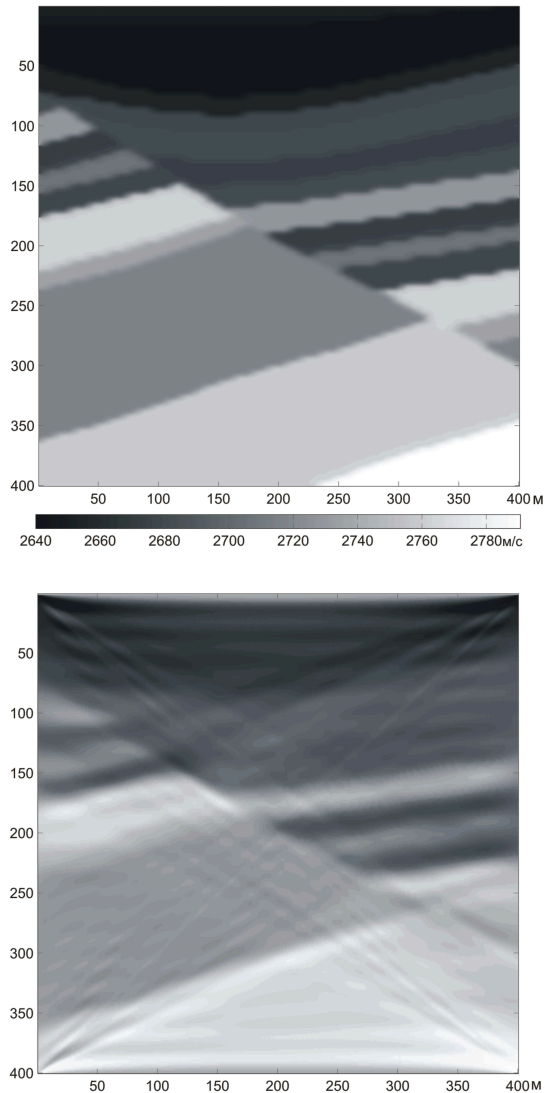


Fig. 6. The synthetic velocity model (on the top) and its tomography recovery result

3. Conclusion

The analysis of the matrix approximation SVD convergence by the spatial grid splitting for the tomography cross-well operator has been performed in the present paper.

The reliably computed part of the singular spectrum has been picked out for the given discretization and model example.

From the presented research follows that the spatial frequencies of the right singular vector are increasing with their rank. There is a number, such that the vector pattern can't be approximated using the given rectangular grid. The comparison of the singular spectrum for the sequence of splitting meshes helps to find the regularization parameter - the number of largest singular values and vector, used for the inversion.

From our point of view the analysis of the Fourier transform of the right singular vectors is

the perspective direction of the study. The left singular vectors subspaces are also need to be observed, because they form the basis in the data space. We believe, that their study would allow to develop some new algorithm of data filtration and preprocessing.

The research was supported by the Russian Federation President grant for young scientists (number MK-2598.2014.5).

References

- [1] L.Hatton, M.H.Worthington, J.Makin, Seismic data processing: theory and practice, *Merlin Profiles Ltd.*, 1986.
- [2] V.A.Tcheverda, V.I.Kostin, R-pseudoinverse for a compact operator, *Sib. Elektron. Mat. Izv.*, **7**(2010), 258–282 (in Russian).
- [3] P.Mora, Inversion=migration+tomography, *Geophysics*, **54**(1989), 1575–1586.
- [4] S.V.Goldin, Introduction to geometric seismic, Novosibirsky Gos. Univ., 2005 (in Russian).
- [5] F.Natterer, The Mathematics of Computerized Tomography, Vieweg+Teubner Verlag, 1986.
- [6] T.Kato, Perturbation theory for linear operators, Springer-Verlag, 1995.
- [7] L.V.Kantorovich, G.P.Akilov, Functional analysis, Oxford etc., Pergamon Press, 1982.
- [8] M.Z.Nashed, G.F.Fortuba, A Unified Theory of Generalized Inverses, Generalized Inverses and Applications, Academic Press, New York–San Francisco–London, 1976.
- [9] A.N.Tikhonov, V.Y.Arsenin, Methods for solving ill-posed problems, Nauka, Moscow, 1979 (in Russian).
- [10] S.K.Godunov, A.G.Antonov, O.P.Kiriluk, V.I.Kostin, Guaranteed Accuracy in Numerical Linear Algebra, Kluwer Academic Publishers, 1993.

Вычислительные аспекты применения метода усеченного сингулярного разложения при решении задачи сейсмической томографии

Александр С. Сердюков
Андрей В. Патутин
Татьяна В. Шилова

Метод усеченного сингулярного разложения является мощным методом регуляризации и решения обратных задач. Применение данного метода на практике сильно ограничено объемами машинной памяти, требуемой на вычисление сингулярного разложения матриц большой размерности. В этой связи актуально использование итерационных процедур поиска старшей части сингулярного спектра. При этом возникают вопросы, какая часть спектра может быть найдена достоверно и достаточно ли найденное количество сингулярных векторов для построения решения. Данные вопросы исследуются в настоящей работе численно на примере обратной кинематической задачи геофизики.

Ключевые слова: обратные задачи, численные методы, усеченное сингулярное разложение, геофизика, времена пробега сейсмических волн.