Resolution Estimates for Selected Coordinate Descent: Identification of Seismic Structure in the Area of Geothermal Plants

Tatyana A. Smaglichenko, Alexander Smaglichenko

Resolution Estimates for Selected Coordinate Descent: Identification of Seismic Structure in the Area of Geothermal Plants*

Tatyana A. Smaglichenko[1] [0000−0001−9786−6959] and Alexander Smaglichenko[2] [0000−0003−1221−2281]

[1] Research Oil and Gas Institute, Russian Academy of Sciences, 3 Gubkina street, Moscow 119333, Russia
t.a.smaglichenko@gmail.com
[2] V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, 65 Profsoyuznaya street, Moscow 117997, Russia

Abstract. The coordinate descent method is a traditional inverse solver to optimization problems. In modern sectors of production research: in computer graphics, computer tomography, a theory of pattern recognition various algorithms of the coordinate descent have been applied. In this paper, we investigate the novel algorithm of selected coordinate descent and outline the difference between this algorithm and the classical coordinate descent. The solution selection is performed owing to the search of the maximum among values of the specific parameter. The maximum indicates a single direction, which is responsible for the minimum of the function in the least square sense. We develop the technique for defining the explicit expression for the resolution measure of the linear systems, which are solved using the proposed algorithm. The algorithm and its resolution tool are applied to seismic observations collected in the area of the Krafla and Theistareykir geothermal power plants, northern Iceland. The result confirms that the distinctive feature of the algorithm is its effectiveness when large-size structures are retrieved. The analysis of the resolution parameter values shows that the calculation of this parameter might be helpful to recognize the true structure.

Keywords: Coordinate descent · Linear equations · Tomography · Structure recognition.

Introduction

The simplicity and capabilities of the iterative solution construction are features of the method coordinate descent (CD). Therefore this method is preferred over others for the computer graphics industry. The iterative process calculates the unknown parameters of the animation object moving. The known vector corresponds to the needful position of the part (end effector).of an object [1].

* The research was carried out within the framework of the state projects №AAAA-A19-119013190038-2, №10.331-17

In tomography, the unknown vector has the physical meaning of the object properties in limited volumes on the grid. A set of observations are related to the object characteristics leads us to the fundamental problem of solving the system of linear equations. The most popular method is the LSQR algorithm introduced by Paige & Saunders [2]. The algorithm is similar to the conjugate gradients method. It applies the Lanczos process and processes a symmetric system [3]. The symmetric matrix might be obtained owing to the multiplication of both parts of the system to the transposed system matrix. We suppose that such numerical transformations may distort the original matrix, which reflects the initial data of the physical experiment. The other problem might be the use of the complicated recurrence formulas that is a reason for the presence of rounding-off errors. Therefore, the interest is still not lost to traditional iterative methods as the Kaczmarz algorithm [4] and the studied CD method. A review of the CD application has been made in [5] analyzing the application sectors: optical diffusion tomography and cryo-electron tomography.

In the theory of pattern recognition, the input vector might be presented by various data: information about the image pixels (components of intensity), speech (acoustic signals), writing language (logograph), etc. For recognition tasks, the CD method applied to the binary classification problem [6, 7] by updating of construction of multiclass predictors, where misclassification error has a linear bound [8]. In [7], experiments have been made to recognize documents. On the other hand, the direct solver is commonly used in digital image processing. The singular value decomposition (SVD) method is often applied in the field of image compression [9]. It is known that SVD is equivalent to the Lanczos process, which is a base of LSQR.

This paper proposes the Selected CD (SCD) algorithm, which might be a contribution to computing image and image structure recognition. The SCD method originated in seismic tomography and was examined comparing with LSQR under equal conditions of conducted numerical experiments [10]. The testing on various arbitrary models showed that SCD may be more effective than LSQR when a simple large-size structure is retrieved. The SCD convergence properties were designed in [10].

There is a difference between the traditional CD and SCD. When solving the system of equations we often apply the least-square approach to the function that is the difference between the system parts. CD calculates the current approximation on the base of components of the gradient, which gives the possible direction of the function minimization. In the traditional CD, indexes of the gradient vector components might be chosen in a cycle fashion or randomly with the probability criterion [5]. SCD selects the index, which provides the function minimum satisfying the special condition. In [11], SCD was established by analyzing the solution errors and updating the condition for big and sparse matrices.

This paper aims to reveal the difference between the SCD algorithm and the classical CD. Another aim of this paper is to present the technique for determining the SCD resolution measure. We describe the SCD application to the real data of local seismic events that were observed at the end of the Krafla rifting episode, during three years 1986-1989. Nowadays two power plants are located in the being investigated area. The production process leads to the cooling of a deep underground medium [12]. Therefore the robust evaluation of seismic structure is essential for the environmental issues.

1 SCD with its relationship to CD

The solution of the system of equations might be considered as the least square minimization of the function $f(x)$ in x. In the case of the system of linear equations $Ax = b$ the functional can be written in the following form:

$$f(x) = (Ax - b, Ax - b), \tag{1}$$

where $(Ax - b, Ax - b)$ denotes a scalar product.

The conventional CD [5, 13] builds the iterative solution as:

$$x^i = x^{i-1} - \left|\nabla f(x^i)\right|_k e_k, \tag{2}$$

where $\left|\nabla f(x^i)\right|_k$ is a component of the gradient, e_k is the vector in the direction of coordinate, $i \in \{1, 2, 3 \ldots n\}$.

Components of the gradient are found by taking the first derivative $f(x^i)$ and setting it to zero. In [11], we have found $\left|\nabla f(x^i)\right|_k$ as:

$$\left|\nabla f(x^i)\right|_k = \frac{(Ax^{i-1} - b, Ae_k)}{(Ae_k, Ae_k)}, \tag{3}$$

where Ae_k is the k-th column of the matrix A. Thus, we set:

$$x^i = x^{i-1} - \frac{(Ax^{i-1} - b, Ae_k)}{(Ae_k, Ae_k)}. \tag{4}$$

If the CD and SCD starting points are equal to zero, then we can see that the iterative approximation (4) is similar to CD for the linear regression [13]. The explanation of several CD algorithms can be found in works [14, 15] and in the presentation of the teaching course of the University of Wisconsin-Madison that was made by authors [14, 15].

Thus, both CD and the developed SCD can calculate the gradient components in the same manner. Note, setting the first derivative $f(x)$ to zero gives the function extremum. The question arises. What is the index provides the direction to the function minimum? CD ordinary uses the cyclic coordinate descent and thus, index is cyclically selected. Applying the SCD approach continues the search of the descent direction (the index k). Namely, if we substitute the equation (2) for x^i in the formula (1), then after a few transformations we obtain $f(x^i)$ in the following form [11]:

$$f(x^i) = (Ax^{i-1} - b, Ax^{i-1} - b) - \frac{(Ax^{i-1} - b, Ae_k)^2}{(Ae_k, Ae_k)}. \tag{5}$$

It is clear that the index with the maximal absolute value of the fractional expression gives the direction of the function minimum.

A proof of the SCD convergence and the convergence analysis were described in [10]. The convergence rate was determined through the span of the angle between the directing vectors to the hyper-planes, to which the vectors $Ax^{i-1}-b_*$ and $Ax^i - b_*$ belong. Here the vector b_* corresponds to the accurate solution x_* in the least square sense $Ax_* = b_*$. Thus, in SCD the degree of convergence was made in terms of the space behavior of multidimensional vectors. Note that the CD convergence rate is determined by the values of the characteristic vector [5, 15].

2 The SCD resolution parameter

In this section, we develop the technique to determine the SCD resolution parameter.

Definition. Let x be a point in the n-dimensional space R^n. For any $b \in R^m$ there are exist a $x \in R^n$ such that $Ax = b$.

Consider the equation (5) as the following:

$$f(x^i) = \|Ax^{i-1} - b\|^2 - \frac{(Ax^{i-1} - b, Ae_k)^2}{\|Ae_k\|^2}. \tag{6}$$

By multiplying the numerator and denominator of the fractional expression by the same value of the scalar product $(Ax^{i-1} - b, Ax^{i-1} - b)$ we get:

$$f(x^i) = \|Ax^{i-1} - b\|^2 (1.0 - \frac{(Ax^{i-1} - b, Ae_k)^2}{\|Ae_k\|^2 \|Ax^{i-1} - b\|^2}). \tag{7}$$

Let x_* be the minimum norm solution of the system in the least square sense. Suppose, $Ax_* = b_*$. Denote by R_k the value of one of the factors in the right part of the equation (7):

$$R_k = 1.0 - \frac{(Ax^{i-1} - b, Ae_k)^2}{\|Ae_k\|^2 \|Ax^{i-1} - b\|^2}. \tag{8}$$

Then, we have the following expression:

$$\|Ax^i - b_*\|^2 = \|Ax^{i-1} - b_*\|^2 R_k. \tag{9}$$

The more values of the parameter R_k are close to 1.0 the more the closeness of the solution x^i to the minimum norm solution x_*. The solution accuracy of the system is estimated by the standard deviation value of the vector $Ax^i - b$. This value mainly depends on the modeling error and the errors of the observed data set. We assume that the iterations are repeated until we got the difference between neighboring approximations x^{i-1} and x^i is not bigger than the observation error.

3 The SCD inversion of seismic data gathered after the Krafla rifting episode

The Krafla rifting episode occurred during 1975–1984 in northern Iceland. Cyclic inflation and deflation of the magma chamber within the Krafla caldera finally led to kilometer-scale volcanic deformation [16]. Research groups from Iceland and other countries participated in the collection of various data related to the Krafla rifting episode. In this paper, we analyze the SCD inversion result, when input data sets are P-wave arrivals of seismic waves from local 11 events recorded by 12 temporary stations during the period 1986–1989. The installation of stations and monitoring of seismicity were performed by the researchers from Mainz University, Germany.

Figure 1(upper part) shows the location of volcanoes (white squares), power stations (white triangles), seismic stations (black triangles), epicenters of earthquakes (black circles). The figure is built in the GMT program by using data from ASTER GDEM v.2 (METI and NASA product). Figure 1(bottom part) displays the obtained P-wave velocity structure for the depth range 0–5 *km*.Volcanoes, hydrothermal stations, hypocenters are denoted by open squares, triangles, circles, respectively. To the right of fig. 1, the scale determines the correspondence of the numerical values of the calculated P-wave velocity to different shades of gray.

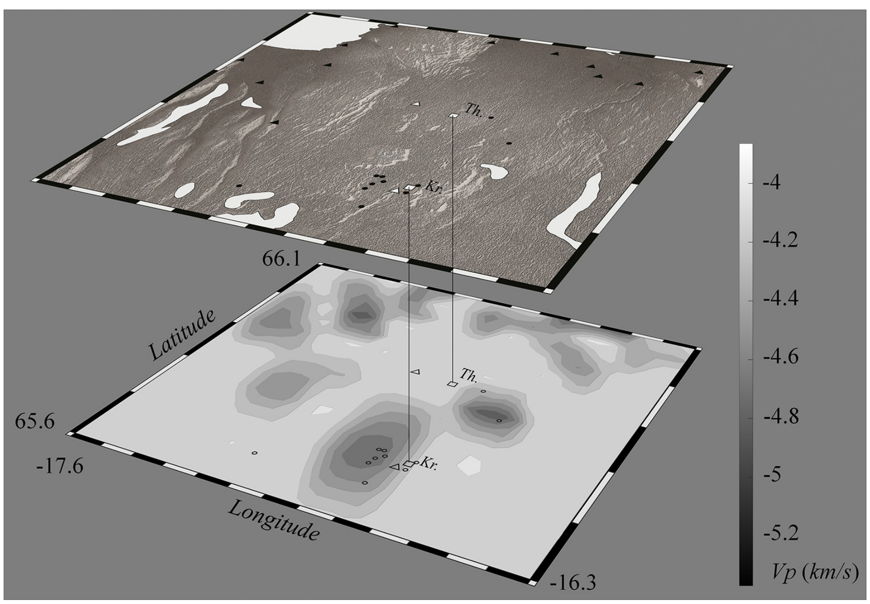


Fig. 1. Upper part. Relief map showing the location of volcanoes Krafla (Kr) and Theistareykir (Th), hydrothermal stations, seismic stations, and epicenters. Bottom part. Seismic velocity distribution. Solid vertical lines connect the locations of volcanoes

The Krafla and Theistareykir geothermal fields are located close to the volcanoes Krafla and Theistareykir in the limits of the studied area. In the vicinity of volcanic calderas, geothermal reservoirs are exploited by the Krafla and Theistareykir power plants for electricity production [12, 17]. The Krafla geothermal power station began operations in 1978. The Theistareykir station turbines have been operating since 2018.

One can see that SCD revealed a large-scale underground structure (high seismic velocities) that mainly correlates with the uplifts in the northern landscape zone. In the vicinity of the Krafla volcano, the anomalous high velocity (dark gray) characterizes the deep formations located to the west, the northwest, and the southwest of the volcano. The Theistareykir volcano is in few kilometers from the high-velocity anomaly to the southeast of the volcano.

4 The SCD resolution measure as practical instrument to identify a structure

The application of the SCD solver to the initial data set, the computation of the SCD resolution parameter, preparation of all data for visualizations, the data processing were carried out using the FORTRAN and MATLAB environments and the corresponding programs designed by the authors of this paper.

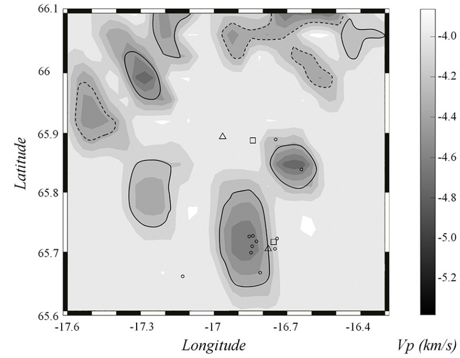
Figure 2 illustrates (a) the P-wave velocity image, (b) the calculated values of the resolution parameter, (c) the relief map. Figures 2-a and 2-c respond to the bottom and upper parts of fig. 1. Domains of poor resolution ($R_k \approx 0.1$) and acceptable resolution ($R_k \in [0.6; 1.0]$) are delineated and shown by dotted and solid lines, respectively. Note that most of the R_k values are in the range of good resolution (from 0.86 till 0.99).

Comparison of the different resolution domains with geological structure (fig. 2-c) demonstrates that the acceptable values of the resolution parameter distinguish the relief uplifts from lowlands. For lowlands, high velocities were determined with poor resolution.Confirmation that the underground structure in the Krafla volcanic caldera is characterized by the increase of P-velocity can be found in the recent work on drilling [18].

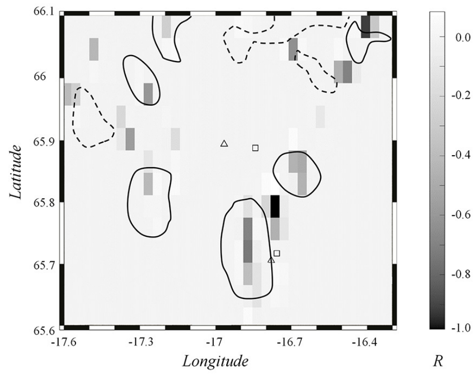
Conclusion

In this paper, we described the SCD algorithm and compared it with the traditional CD method. The main difference between SCD and CD is the following. In CD, the direction of coordinate descent is formed applying the extremum condition for the least square function. SCD requires the developed condition to provide the function minimum.

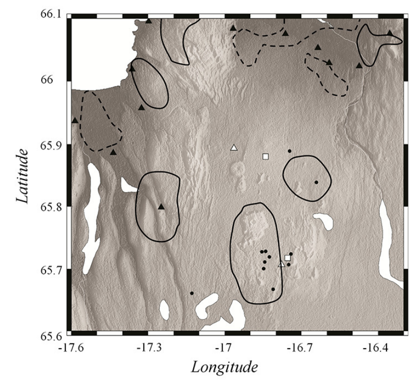
Our main results are the following. The explicit expression for the SCD resolution parameter was obtained. The SCD application to real observations supports the previous statement for synthetic models [10] and confirms that SCD is robust to reconstruct the simple large-size structure. The application of the SCD resolution tool revealed its capability to identify geological structures.



a)



b)



c)

Fig. 2. The domains of poor and acceptable resolution (dotted and solid lines) of computing seismic images in comparison with real geological structures in the area of the geothermal energy production

Acknowledgments

We highly appreciate to Prof. Wolfgang Jacoby, who organized the fieldwork at the northern coast of Iceland in 1986–1989. We also thank him for fruitful help in promotion of this research.

References

1. Wang, L., Chen, C.: A Combined Optimization Method for Solving the Inverse Kinematics Problem of Mechanical Manipulators. IEEE Transactions On Robotics and Automation **7**, 489–498 (1991)
2. Paige, C., Saunders, M.: LSQR: an algorithm for sparse linear equations and sparse least squares. ACM Trans. Math. Soft. **8**(1), 43–71 (1982)
3. Lanczos, K.: An iteration method for the solution of the eigenvalue problem of linear differential and integral operators. Journal Research N.B.S. **45**, 255–282 (1950)

4. Kaczmarz, S.: Angen aherte au osung von systemen linearer gleichungen. Bulletin International de l'Academie Polonaise des Sciences et des Lettres **35**, 355–357 (1937)
5. Wright, S.: Coordinate descent algorithms. Mathematical Programming **151**(1), 3–34 (2015)
6. Hsieh, C., Chang, K., Lin, C., Keerthi, S., Sundararajan, S.: A dual coordinate descent method for large-scale linear SVM. In: Proceedings of the Twenty-fifth International Conference on Machine Learning, pp. 1369–1398. Helsinki, Finland (2008)
7. Keerthi, S., Sundararajan, S., Chang, K., Hsieh, C., Lin, C.: A sequential dual method for large scale multiclass linear SVMs. In: Proceeding of the 14th ACM SIGKDD international conference on Knowedge discovery and data mining KDD, pp. 408–413 (2008). https://doi.org/10.1145/1401890.1401942
8. Crammer, K., Singer, Y.: On the algorithmic implementation of multiclass kernel-based vector machines. JMLR **2**, 265–292 (2001)
9. Pandey, J., Singh Umrao, L.: Digital Image Processing using Singular Value Decomposition. In: Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE), (2019). https://doi.org/10.2139/ssrn.3350278
10. Smaglichenko, T., Nikolaev, A., Horiuchi, S., Hasegawa, A.: The method for consecutive subtraction of selected anomalies: the estimated crustal velocity structure in the 1996 Onikobe (M = 5.9) earthquake area, northeastern Japan. Geophys. J. Int. **153**, 627–644 (2003)
11. Smaglichenko, T., Jacoby, W., Smaglichenko, A.: Alternative 3D Tomography Methods and Their Applications to Identify Seismic Structure Around the Hydrothermal Gas Field In: Tsvirkun, A. (ed.) Proceedings of 2020 Eleventh International Conference "Management of large-scale system development" (MLSD). Institute of Electrical and Electronics Engineers (IEEE), Moscow (2020). https://doi.org/10.1109/MLSD49919.2020.9247697
12. Ali, S., Feigl, K., Carr, B., Masterlark, T., Sigmundsson, F.: Geodetic measurements and numerical models of rifting in Northern Iceland for 1993–2008. Geophys. J. Int. **196**(3), 1267–1280 (2014)
13. Ali, A., Kolter, Z., Tibshirani, R.: A Continuous-Time View of Early Stopping for Least Squares Regression. In: International Conference on Artificial Intelligence and Statistics. (2019)
14. Gordon, G.: Generalized Linear Models. In: Becker, S., Thrun, S., Obermayer, K. (eds.) Advances in Neural Information Processing Systems 15 (NIPS 2002) NeurIPS Proceedings, the MIT Press (2002).
15. Tibshirani, R.: Dykstra's Algorithm, ADMM, and Coordinate Descent: Connections, Insights, and Extensions. In: 31st Conference on Neural Information Processing Systems (NIPS 2017). Long Beach, CA, USA (2017).
16. Heimisson, E., Einarsson, P., Sigmundsson, F., Brandsdóttir, B.: Kilometer-scale Kaiser effect identified in Krafla volcano, Iceland. Geophysical Research Letters **42**(19), 7958–7965 (2015)
17. Khodayar, M., Kristinsson, S., Karlsdóttir, R.: Structural Drilling Targets from Platforms A, B, and F at Theistareykir. Northern Rift Zone and Tjörnes Fracture Zone LV report no: LV-2016-060. Prepared by Iceland GeoSurvey (ÍSOR) for Landsvirkjun 24 p. (2016)
18. Millett, J.M. et al.: Sub-surface geology and velocity structure of the Krafla high temperature geothermal field, Iceland: Integrated ditch cuttings, wireline and zero

offset vertical seismic profile analysis. Journal of Volcanology and Geothermal Research **391**, 106342 (2020)