Analyzing of the advantages of the MM algorithm in increasing the temporal resolution of GPR data

Sadegh Moghaddam[1], Behrooz Oskooi[1], Alireza Goudarzi[2] and Asghar Azadi[*3]

1Institute of Geophysics, University of Tehran, Tehran, Iran.sadegh136789@yahoo.com
2Graduate University of advanced technology, Kerman, Iran.
3Payam Noor University of Parand, Tehran, Iran.asghr-azadi-2007@yahoo.com

ABSTRACT

In this paper, the advantages of sparse signal processing using the Majorization-Minimization (MM) method in GPR signal compression is investigated. In this method, minimizing the cost function is determined with L1 and L2 norms; also, the banded structures of matrices resulting from the sparse deconvolution problem is regarded. In order to increasing the temporal resolution of the GPR signals, the MM algorithm has been implemented with least-square deconvolution (LSQR) on the synthetic data. Analysis of the outputs reported that the reflection coefficient improved significantly by application of the MM algorithm to the synthetic data compared with the least square deconvolution which only filters the data. The power spectrum after using the MM algorithm shows acceptable compression. Moreover, this algorithm leads to a considerable improvement on the amplitudes so that the hidden anomalies are better restored.

Keywords: Ground-penetrating radar (GPR), resolution, Majorization-Minimization (MM) method, Least Square Deconvolution (LSQR).

INTRODUCTION

Ground-Penetrating Radar (GPR) is a non-invasive geophysical technique that has been a very powerful tool for the high-resolution imaging of the subsurface structures. This technique uses the electromagnetic waves (EM) to visualize the subsurface in geological, archaeological and other environmental applications (Bednarczyk and szynkiewicz, 2015). Detailed interpretation of the GPR features basically depends on the temporal resolution. Absorption is a function of frequency, and the Earth acts as a low-pass filter which decreases the amplitudes of the higher frequencies more than the lower ones. Hence, in the case of absorption, the dominant frequency in the spectral range is in the low frequencies, and an increase in the wavelength is expected. Indeed, the time width of the wavelet that propagates through the earth increases after it is received again on the surface. It can lead to a reduction in the temporal resolution. If the compressed wavelet in the time domain is desired, the width of the frequency spectrum should be increased. Wavelet compression could lead to an increase in the temporal resolution of the layers (Pakmanesh et al., 2017). Deconvolution is a numerical process of evaluating the unknown input of a linear time-invariant (LTI) system which the output and the response of the signal are known (Selesnick, 2012). The MM algorithm has been presented by (Selesnick, 2012). A difficult minimization problem is solved with this algorithm by optimization condition and convex function. During the solving process, the algorithm monitors to minimize the difference between the observed and desired signal and simultaneously applies the control parameter to the signal (Pakmanesh et al., 2017). In practice, the available output of GPR data can be presented as a convolution of the source wavelet and the earth reflectivity series and additive noise according to Irving and Knight (2006). Therefore, the deconvolution problem is non-invertible, and the use of the direct inverse problem can significantly damage the estimated signal, making the result unusable (Selesnick, 2012). However, to verify the performance of the MM algorithm when compressing and obtaining the reflectivity series of GPR data, the least square deconvolution is performed on the simulated and real data. The ultimate goal of this study is to obtain an image of the GPR data with less possible significant coefficients for the restoration of the anomalies; so that the emphasis is on the use of the L1 norm that is related to sparsity. GPR signal compression can deliver more savings regarding costs and operational time; therefore, in studies of the combination of several different antennas, more data are required by compressing with new processing methods. On the other hand, compression increases the resolution, and in this way, it can lead to a reliable interpretation of GPR sections.

The least square deconvolution

Solving the inverse problem using least-squares deconvolution as a general method is to minimize the energy of x (Selesnick, 2012):

$$minimize\|x\|_2^2 \quad \text{Subject to y=Ax} \qquad (1).$$

In this method, the objective is to get a smoother signal closer to the input with the least-square regularization approach, and there is a tradeoff between smoothness and energy. The solution of $\|x\|_2^2 = \sum_{n=0}^{N-1}|x(n)|^2$ is written below (Selesnick, 2012):

$$x = A^*(AA^*)^{-1}y \qquad (2).$$

In the above equation, A^* is a transpose of the complex conjugate of the matrix A. When y is noisy, the approximate solution of the above minimization is reformed as:

$$x = (A^*A + \lambda I)^{-1}A^*y \qquad (3).$$

where λ is a regularization factor for smoothing the input signal which can be define by the Pareto curve or the generalized cross-validation (GCV) method.

Sparse solution

Supposing that the noisy signal can be written as below:

$$y = (h * x)(n) + w(n) \quad (4).$$

where h(n) is the impulse response of the LTI system, x(n) is sparse signal, w(n) is white Gaussian noise, and * denotes convolution symbol. To solve the problem, it is assumed that the supposed LTI system can be described by a recursive difference equation as written below:

$$y(n) = \sum_k b(k)x(n-k) - \sum_k a(k)y(n-k) \quad (5).$$

The above difference equation can be expressed with a matrix form as following:

$$Ay = Bx \quad (6).$$

where A and B are banded matrixes in the equation (6). The output of this system can be written as (Selesnick, 2012):

$$y = A^{-1}Bx = Hx \qquad (7).$$

where the system matrix of H is equal to:

$$H = A^{-1}B \qquad (8).$$

In the above relation, even though A and B are banded matrices, the inverse matrix (H) is not banded in general. Then the data model (4) can be written in a discretization form as:

$$y = Hx + w \qquad (9).$$

In this case, the approximate sparse solution of inverse problems is to minimize cost function as written below:

$$minimize\|y - Hx\|_2^2 + \lambda\|x\|_1 \qquad (10).$$

This equation is called the basis pursuit de-noising (BPD). In the processing of a large data set which is significant to specify which variables are more valuable, the sparse solution can be productive. Using the L1 norm as a regularization term has the advantage of measuring sparsity. Furthermore, compared with the L2 norm in the least square deconvolution that promotes smoothness, the L1 norm leads to solutions with many zero components, and this helps to determine the relevant variables. The proper regularization parameter $\lambda > 0$ is selected in accordance with the noise sensitivity w(n).

Application

The input of the system is defined with the convolution of the reflectivity model (Fig. 1a) and the source wavelet (Fig. 1b) which is depicted in Fig. 1c according to (Jazayeri et al., 2019). The noisy, sparse signal is considered as the output of the system (Fig. 1d). The contrast between the MM algorithm and the least squares deconvolution of 1D synthetic data has been illustrated in Fig. 1e. As we can see, the least squares solution is noisier and attenuated. On the other hand, the decrease in λ can lead to an even noisier solution. Higher values of λ will increase the attenuation of the solution. An advantage of this comparison lies in the preference of the MM algorithm for which the sparse signal is valid. The RMS values after the use of two algorithms have been reduced by 0.0205 for the MM algorithm and by 0.542 for the least squares method, which confirms the large capacity of the MM algorithm concerning the deconvolution by the least squares.

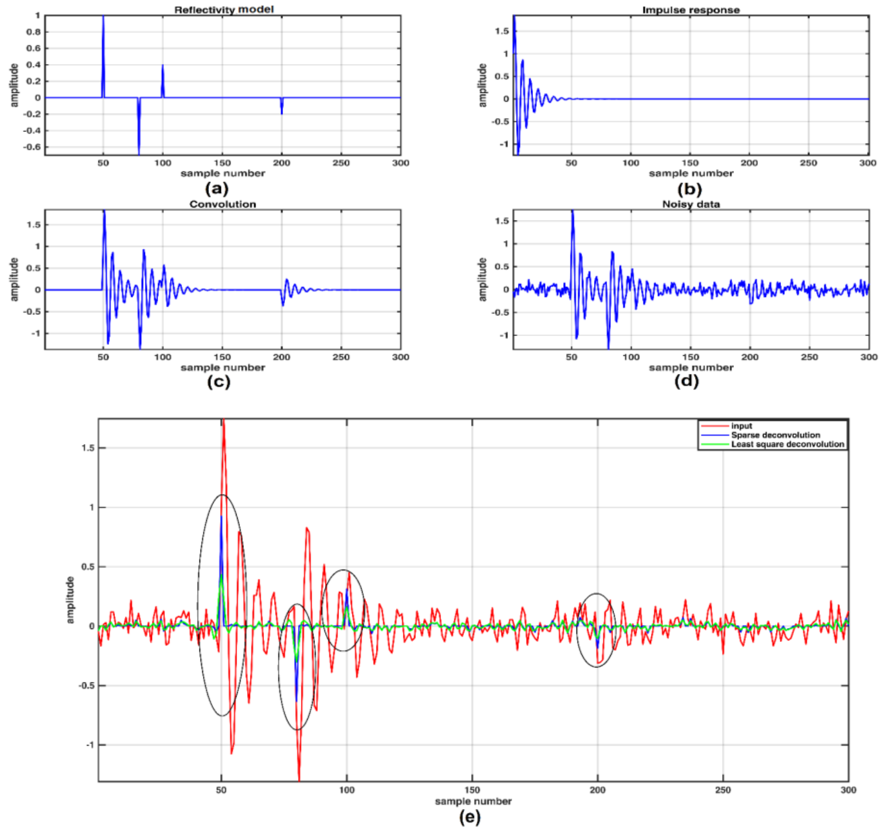

Fig. 1. a). Reflectivity model, b) Source wavelet, c) convolution of Reflectivity model and the source wavelet as an input data, d) noise added data, e) red line is the noisy data, blue line is the output of the MM algorithm with the RMS of 0.0205 and green line is least square deconvolution output with the RMS of 0.542.

For a systematic evaluation of the performance of the proposed methods on GPR data, a synthetic model has been performed using full-wave solver of Maxwell's equations (GPRMax 2.0). The simulated model chosen in this study is considered to consist of three layers with different dielectric and conductivities, where the upper layer is made of concrete (t_1=1m, $\sigma_1 = 0.005\frac{s}{m}, \varepsilon_{r1} = 6$), the middle layer of wet sandy soil (t_2=2m, $\sigma_2 = 0.0001\frac{s}{m}, \varepsilon_{r2} = 25$) and the third layer is a saturated sandy soil (t_3=1m, $\sigma_3 = 0.07\frac{s}{m}, \varepsilon_{r3} = 30$). The ultimate goal of this study is to achieve a series of compact and sparse reflection coefficients, in which the underlying layers are well separated. To test the effects of both methods, a random noise equal to 0.1 of the amplitude of the data was added to the data. The second synthetic trace after application of the MM algorithm and the least squares deconvolution has been shown in Fig. 2. As can be seen, the improvement of the synthetic model is clearly visible with a sparse deconvolution using the MM algorithm, and the layers are also well separated. The RMS of the noisy input was equal to 9.39 and, after application of the two algorithms, to 0.019 for the MM algorithm and 0.1542 for the least squares method, which confirms the large capacity of the MM algorithm with respect to the least squares deconvolution. According to the power spectrum result of Fig. 3, the power of the MM algorithm showed the improvement of the bandwidth compared to the least squares method which indicated the preference of the MM algorithm by providing the satisfactory results for compressing the synthetic example.

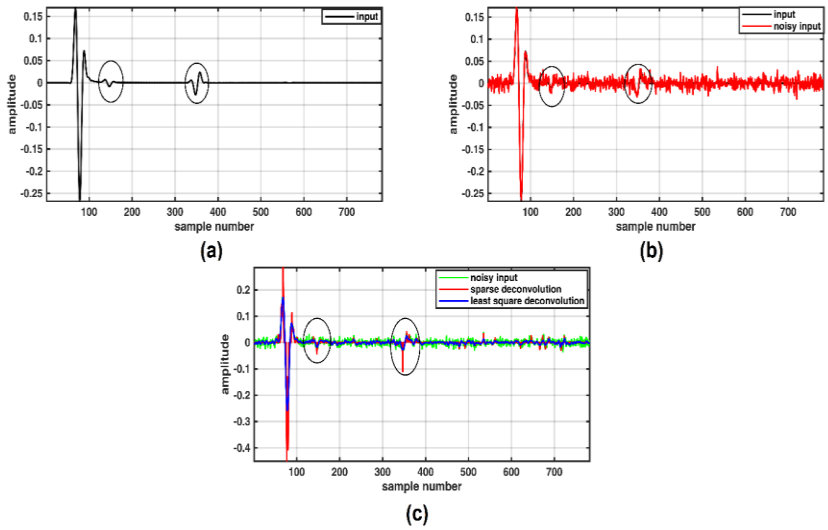

Fig. 2. a) The second trace of synthetic GPR data, b) black line is input, the red line is noisy input, c) the red line is MM algorithm output, the green line is noisy input, and the blue line is least square deconvolution output.

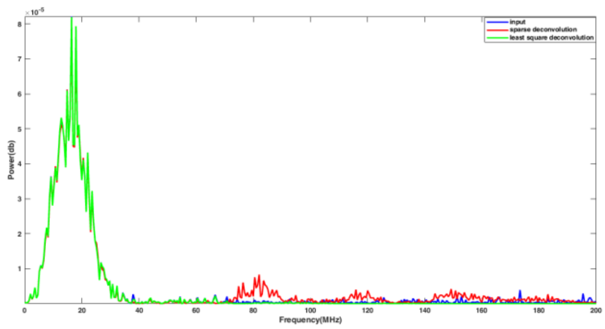

Fig. 3. Power spectra: blue line is input, the red line is the MM algorithm output, and the green line is the least square deconvolution output.

CONCLUSION

The results showed that the existence of L1 and L2 norms in the cost function of the MM algorithm could improve the accuracy and temporal resolution of the GPR data. The layers and anomalies can be detected at an acceptable level; the power spectrum has been improved after using the MM algorithm. Also, whatever the higher frequencies along with the existing ones (which the MM algorithm has been able to maintain them) are retrieved, it leads to more compression. This compression can lead to more proper resolution and clarity than the LSQR method; because the LSQR method using the same parameters is treated as a low-pass filter. The convolution of any mix-phase wavelet and reflectivity series is input, and a spiky sparse signal is an output of the MM algorithm. On the other hand, the LSQR method as a common method cannot increase data compression, and only smoothes the data; because the LSQR method with the same parameters treats as a low-pass filter.

REFERENCES

Bednarczyk, Z., and Szynkiewicz, A., 2015. Applied Engineering Geology Methods for Exemplar Infrastructure Projects in Malopolskie and Podkarpackie Provinces. Springer International Publishing In Engineering Geology for Society and Territory, v.6, pp: 203-210.

Pakmanesh, P., Goudarzi, A., Kourki, M., 2017. Hybrid Sparse Blind Deconvolution: an implementation of SOOT algorithm to real data. Journal of Geophysics and Engineering,128.114.228.241. https://doi.org/10.1088/1742-2140/aaa3d1.

Selesnick, I., 2012. Sparse deconvolution (an MM algorithm), Connexions.

Irving, J., Knight, R., 2006. Numerical modeling of ground-penetrating radar in 2-D using MATLAB. Comput Geosci 32:1247–1258.