Submission code: 0VKAYBRNVK

See this and other abstracts on our website: https://home.sbgf.org.br/Pages/resumos.php

Enhancement of seismic events in passive seismic data via dictionary learning

Lucas Aires da Costa Silva (UFRN), Yuri Shalom De Freitas Bezerra (UFRN), Walter Söllner, German Garabito (DPET/CT/UFRN)

Enhancement of seismic events in passive seismic data via dictionary learning

Please, do not insert author names in your submission PDF file

Abstract

Seismic data from passive sources is generally very noisy, with weak sporadic signals of interest obscured by noise. One crucial and difficult task when dealing with such data is improving the signal-to-noise ratio without distorting the signal of interest. Dictionary learning (DL), which is highly effective due to its capability for sparse representation, has already been applied to passive seismic data. One of the challenges of applying DL is finding the optimal configuration of parameters for an effective and efficient solution of the problem. This work aims to determine the optimal parameters for applying the K-SVD and SGK dictionary learning algorithms, which commonly use Discrete Cosine Transform (DCT) starting dictionary. Specifically, the aim is to create a dictionary from the input data itself and, mainly, to create a new dictionary with atoms that represent a local approximation of passive seismic events. The initial tests were performed using synthetic data containing two passive seismic sources. Both DL algorithms (K-SVD and SGK) were applied to this data by using a starting dictionary created from the data itself. Both algorithms were able to attenuate the noise and enhance diffraction events from data with the highest signal-to-noise ratio (SNR). In data with a very low SNR, both algorithms still produced good quality results. Tests are ongoing to implement a kinematically constrained DL algorithm.

Introduction

Passive seismic surveys have many applications in natural resource exploration and engineering, and their success depends on the correct selection and use of the monitoring equipment. Studies of monitoring hydrocarbon production can be used to identify changes in subsurface reservoir fluids and stress behaviour, and to locate areas where the fluids are moving during production. For weak seismic sources, such as those of passive seismic applications, denoising and reconstruction are crucial processing steps to enhance passive seismic signals (Chen *et al.,* 2023). To identify, denoise and enhance seismic signals, different methods are driven to solve this problem. A dictionary learning (DL) method is based on atoms that are small patches learned from the data itself, which can be combined to represent the original signal more clearly and sparsely (Aharon et al., 2006).

Sparse representation has been used for nearly two decades to enhance seismic data. However, these methods often rely on fixed-basis, predefined dictionaries, which can be too rigid and may not always ensure a truly sparse representation. To address this limitation, DL offers a data-driven solution that adapts the dictionary to better satisfy the sparsity condition (Wang et al., 2021). Among the most widely used DL algorithms are K-Singular Value Decomposition (K-SVD) and Sequential Generalized K-means (SGK) (Wang et al., 2021; Chen et al., 2023).

One of the main challenges of DL methods is determining the optimal configuration of parameters to effectively solve tasks and maximize the performance while reducing computational cost. This involves factors such as the number and size of atoms, superposition of data patches, and others specific to the algorithms (K-SVD and SGK). The objective of this work is to investigate efficient strategies for determining the optimal parameters of the K-SVD and SGK algorithms, and, more importantly, to introduce new dictionaries in order to increase their ability to attenuate noise and enhance coherent events from noisy seismic data. We will introduce a new type of dictionary in which the atoms represent local approximations of passive seismic events, i.e. we will implement a kinematically constrained DL algorithm. This work presents the initial results of applying the SKG and KSVD algorithms to synthetic data simulating a passive seismic monitoring study of a hydrocarbon reservoir.

Method

Dictionary Learning (DL), learns a set of simple structures (atoms of the dictionary) that, when combined, can represent complex seismic waves. By representing seismic waves with these atoms, it is possible to reduce the complexity of the data without losing important information, hence facilitating the analysis. The DL method is applied in two steps, the first one is called sparse coding which uses a predefined dictionary as sparse representation. The second step is called dictionary update, in this step the atoms from the dictionary are updated. There are several methods for the dictionary update step, we used the DL methods K-SVD and SGK in this work. According to Chen *et al.* (2023) K-SVD is robust and accurate and SGK is very efficient.

The K-SVD (Aharon et al., 2006) is one of the most widely used DL algorithms, where the atoms are updated individually using singular value decomposition and orthogonal matching pursuit operations, resulting in a high computational cost. The SGK method introduced by *Chen* (2020) is a faster alternative that updates the atoms by taking the arithmetic average of the training signals related to each atom.

Each learned atom can represent a redundant seismic pattern in the dataset. However, without any additional constraints during the dictionary update step, noise or spurious events may also be included in it, which may lead to unsatisfactory results. One way to minimize this issue in streamer data is by using an initial dictionary based on the data itself and constraining the learning to possible surface seismic events (Turquais *et al.*, 2018). In this study, we present preliminary results using an initial dictionary extracted from the data itself and as ongoing research work, the atoms of the dictionary will be constrained to obey the kinematic parameters of passive seismic events.

Results

The synthetic velocity model shown in Figure 1 was built using geological information from the Parnaíba Basin, located in the north-eastern region of Brazil. It was used as a base for elaborating the synthetic passive seismic signal. Geologically speaking it is mainly composed of horizontal layers of sedimentary rocks, with diabase sills that have intruded the sedimetary layers in different levels, thus having the form of apparent folds, as represented with the red band or stripe between 1350 m and 1950 m deep. This diabase sill acts as a sealing rock for natural gas accumulations in sandstone reservoir rocks that occur below the sill. This type of reservoir in the Parnaíba Basin is also being studied for CO2 storage due to its favorable geological features for this activity.

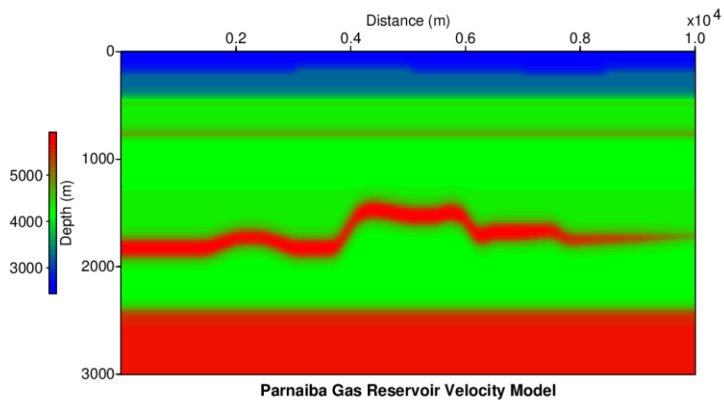

Figure 1: Velocity model inspired by a gas reservoir in the Parnaíba Basin, northeastern Brazil.

We generated synthetic data for two sources located at the base of the diabase sill, using a finite difference numerical scheme to solve the acoustic wave equation. That is, the sources are positioned below the fold-like structure in the central part of the model at the following coordinates: x1 = 4,280 m, z1 = 1,550 m, x2 = 5,950 m, z2 = 1,590 m. The data has 500 traces, each with an interval of 25 m, covering the entire surface of the model. The time sample is 2 ms and the record length is 2 seconds.

The parameters of the dictionary are: Number of atoms (K), size of atoms (L1, L2) and superposition of data patches (S1, S2). In addition, the parameters of the K-SVD and SGK algorithms are: Number of iterations (Niter) and Sparsity level (T). We used for this synthetic data two noise levels, the first with -11.18 dB and the second with −17.20 dB. For both tests we used the same parameters set: S1 = 1; S2 = 1; T = 6; Niter = 10. Figure 2 shows the initial and learned dictionaries for the test with SNR = -11.18 dB.

The results presented in Figure 3 were obtained using the parameters K = 32, L1 = 32 and L2 = 8 for SNR = -11.18 dB. Figure 4 are the results using the parameters K = 32, L1 = 32 and L2 = 16 for SNR = -17.41 dB. In both figures, panels a) and b) show the reference data and the noisy data, c) and d) show the results of K-SVD and SGK algorithms and e) and f) show the errors of both methods in relation to the reference data a), respectively. We can observe that random noise has been effectively removed.

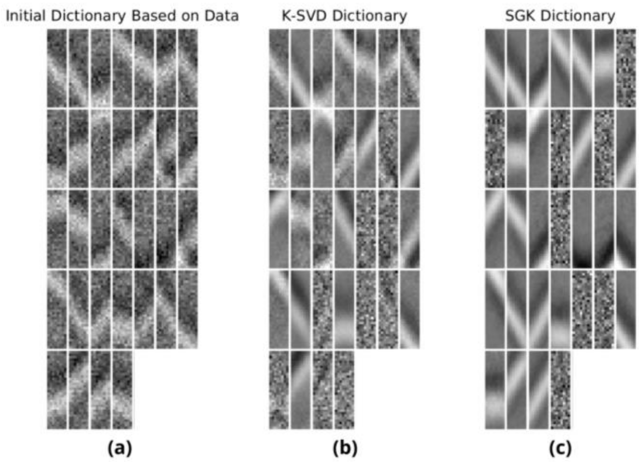

Figure 2: The initial dictionary was created based on the data itself (a) and the learned dictionaries for K-SVD (b) and SGK (c) methods. The parameters are: K = 32, L1 = 32 and L2 = 8.

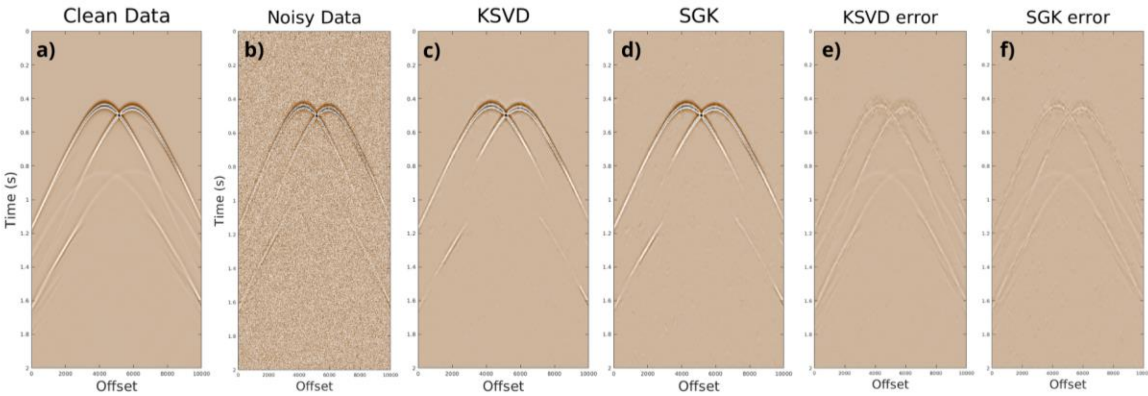

Figure 3: Comparison of 2D passive seismic data denoising. (a) Clean data. (b) Noisy data (SNR = -11.18 dB). (c) Denoising data using KSVD (SNR = 12.41 dB). (d) Denoising data using SGK (SNR = 13.48 dB). (e) KSVD denoising error. (f) SGK denoising error. K = 32, L1 = 32 and L2 = 8.

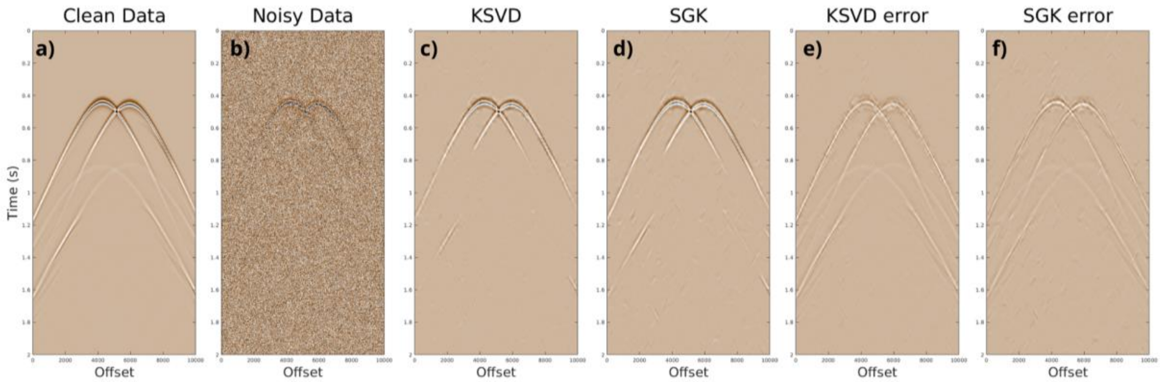

Figure 4: Comparison of 2D passive seismic data denoising. (a) Clean data. (b) Noisy data (SNR = −17.20 dB). (c) Denoising data using KSVD (SNR = 8.36 dB). (d) Denoising data using SGK (SNR = 9.42 dB). (e) KSVD denoising error. (f) SGK denoising error. K = 32, L1 = 32 and L2 = 16.

Conclusion

We present the preliminary results of the successful application of the K-SVD and SGK dictionary learning methods to very noisy synthetic data representing seismic data from passive sources. We applied both algorithms using initial dictionaries extracted from the data itself. The results were quite satisfactory in both tests, which demonstrates the great potential of these DL algorithms for removing random noise and enhancing the signal of interest in seismic data from passive sources. The results confirm that SGK is much faster than K-SVD. Ongoing tests are being performed on constrained dictionary learning using kinematic parameters associated with passive seismic events in order to facilitate the identification and enhance these types of events.

Acknowledgments

We would like to thank FINEP for supporting a research project 01.23.0522.00 FINEP/UFRN/FUNPEC and CNPq for supporting Project 306393/2022-0.

References

AHARON, M.; ELAD, M.; BRUCKSTEIN, A. **K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation.** *IEEE Transactions on Signal Processing*, v. 54, n. 11, p. 4311–4322, 2006.

CHEN, Y.; SAVVAIDIS, A.; FOMEL, S. **Dictionary learning for single-channel passive seismic denoising.** *Seismological Research Letters*, v. 94, n. 6, p. 2840–2851, 2023.

CHEN, Y. **Fast dictionary learning for noise attenuation of multidimensional seismic data,** *Geophysical Journal International*, v. 222, n. 3, p. 1717–1727. 2020,

TURQUAIS, P.; ASGEDOM, E. G.; SÖLLNER W.; GELIUS, L. **Parabolic dictionary learning for seismic wavefield reconstruction across the streamers**. *GEOPHYSICS*, v. 83, p. 1-61, 2018.

WANG, H.; CHEN, W.; ZHANG, Q.; LIU, X.; ZU, S.; CHEN, Y. **Fast dictionary learning for high-dimensional seismic reconstruction.** *IEEE Transactions on Geoscience and Remote Sensing*, v. 59, n. 8, p. 7098-7108, 2021.