Parametric dictionary learning for extracting basis functions that resemble local seismic waveforms

Hongling Chen and Mauricio D Sacchi
Department of Physics, University of Alberta

Summary

Convolutional Dictionary Learning (CDL) can represent signals and images via the convolution of sparse coefficients and a dictionary. The dictionary elements represent universal signals that can model different images, whereas the coefficients are inherent to one image. Estimating the coefficients and the dictionary from a set of observed signals is like a blind deconvolution problem where we aim to concurrently represent a signal via the convolution of two unknown signals. Classical Convolutional Dictionary Learning provides data-dependent dictionaries that might not resemble typical waveforms observed in seismic records. We propose a new Convolutional Dictionary Learning algorithm where we add a parametric constraint to enforce simplicity on the filters. In other words, we restrict each filter to include one single waveform parametrizable via a second-order traveltime curve and a seismic wavelet. The learned dictionary is composed of linear and parabolic events that can adapt to observed seismic waveforms and resemble local Radon transform basis functions.

Theory

CDL entails estimating basis functions that can represent seismic data. For this purpose, we denote the seismic data as a 2-dimensional signal that depends on time and a spatial coordinate

$$D(t,x) = \sum_k A_k(t,x) * \Phi_k(t,x), \tag{1}$$

where $A_k = 1 \ldots N$ represent the 2D unknown coefficients. Similarly, $\Phi_k = 1 \ldots N$ are unknown basis functions (Heide et al., 2015). The task of CDL is to simultaneously estimate the coefficients and the basis functions from the data. The symbol * is used to indicate 2D convolution. The coefficients are considered dynamic variables that can vary from one part of the dataset to another. In contrast, the basis functions are t-x invariant universal templates that can model complex waveforms at any position in the data. The unknown coefficients and basis functions can be estimated following the procedure introduced by Veshki and Vorobyov (2022).

Or contribution expands the work of Veshki and Vorobyov (2022). Specifically, we seek basis functions of the form

$$\phi\_k(x,t) = w(t - t_{0_k} - p\_k\left(x - x_{0\_k}\right) + q\_k\left(x - x_{0\_k}\right)^2, \tag{2}$$

where $w(t)$ is an unknown wavelet and p_k, q_k, x_{0k} are the kinematic parameters necessary to describe the basis functions (Bakulin et al., 2020). The algorithm for Parametric CDL (PCDL) is described as follows:

1. Optimize $||D - \sum_k A_k * \Phi_k||_2^2 + \mu \sum_k |A_k|_1$ with respect to filter coefficients using a sparsity promoting algorithm. In this stage, we minimize the *l1*-norm of the filter coefficients to solve for

sparse solutions. The *l2*-norm or error function ensures data fidelity. The latter is controlled by the trade-off parameter μ .

2. Update basis functions $\Phi_k = 1 \dots N$ via the alternating direction method of multipliers (ADMM)
3. A delay-and-sum estimator and a smooth wavelet estimation algorithm are used to compute the wavelet and the second-order traveltime kinematic parameters.
4. For each element of the dictionary, we adjust the phase of the wavelet via a phase rotation that maximizes its kurtosis.
5. Repeat until convergence.

Results

Figure 1 shows dictionaries (basis functions) obtained from a real dataset from the Gulf of Mexico. In this example, we used 100 common-receiver-gathers to estimate the dictionaries. Figure 1a corresponds to basis functions extracted via the classical CDL method (Liu and Lu. 2021). Figure 1b shows the basis function retrieved with PCDL. We observe that in PCDL, each dictionary element contains one waveform. In addition, they resemble the responses of local Radon transform operators (Sacchi et al., 2005), making them more attractive for processing seismic data. The basis functions learned via PCDL are associated with dip and curvature; hence, they directly relate to interpretable parameters. A geophysicist can identify these parameters and interpret them. On the contrary, dictionaries extracted via CDL are not indexed by interpretable parameters such as dip/curvature. This point is essential. Popular tools, such as the parabolic Radon transform, have a simple interpretation (e.g., coefficients represent residual moveout at far offset) and are straightforward to use.

(a) (b)

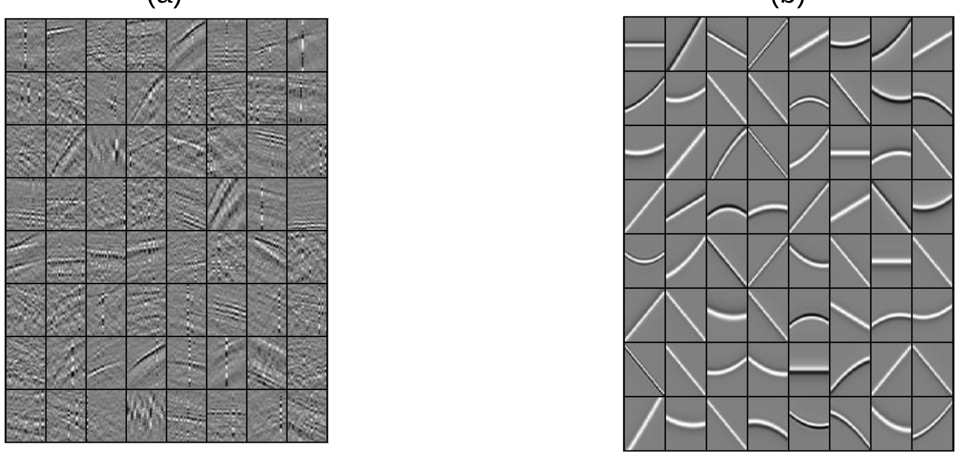

Figure 1. a) Dictionary estimated via classical Convolutional Dictionary Learning (CDL) algorithm. B). Dictionary estimated via the Parametric CDL (PCDL) algorithm proposed in the paper. Both dictionaries were extracted from marine seismic gathers.

The learned dictionaries were also used for trace interpolation. In this numerical experiment, we use sparse reconstruction to recover missing traces. The problem involves estimating the coefficients that represent the available data via the learned basis. Then, once those coefficients are found, they are used to estimate the unknown traces via expression (1). This process is exemplified by Figure 2.

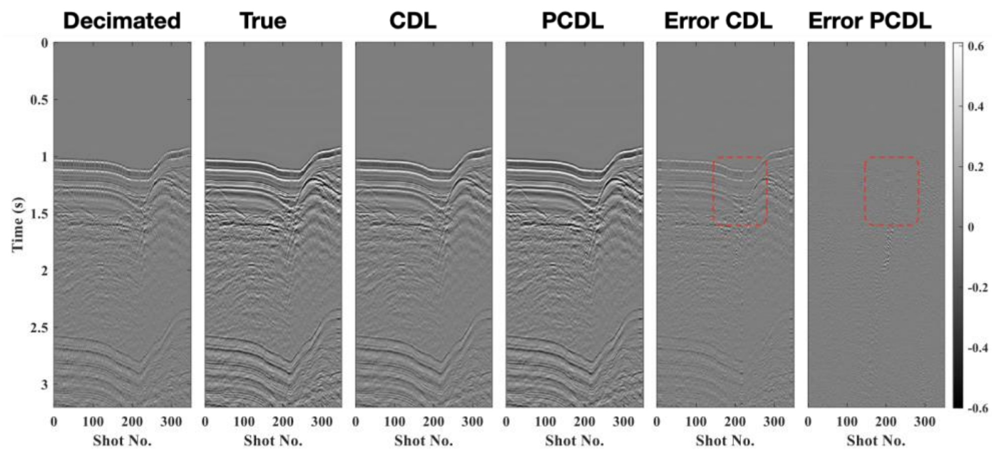

Figure 2. Comparison of CDL and PCDL. The basis functions in Figure 1 were used to reconstruct marine seismic data. In the red boxes, it is possible to appreciate how PCDL can cope with steep complex dips.

Conclusions

Convolutional Dictionary Learning (CDL) can estimate spatio-temporal basis functions to represent seismic data. However, the calculated waveforms do not resemble single simple waveforms. They do not have a strong processing interpretability (which one to choose? what does a given waveform means?). On the other hand, Parametric Convolutional Dictionary Learning (PCDL) yields basic functions that consist of a single waveform per dictionary element. Moreover, associated with them are interpretable parameters such as dip and curvature. Indeed, they look more like local waveforms present in real seismic data. Basis functions extracted via PCDL can be saved and adopted for seismic data regularization and denoising.

Acknowledgements

We thank the sponsor of the Signal Analysis and Imaging Group at the University of Alberta for supporting this research. Hongling Chen also wishes to thank the support from China Scholarship Council.

References

A. Bakulin, I. Silvestrov, M. Dmitriev, D. Neklyudov, M. Protasov, K. Gadylshin, and V. Dolgov, “Nonlinear beamforming for enhancement of 3D prestack land seismic data,” Geophysics, vol. 85, no. 3, pp. V283– V296, 2020

F.Heide, W.Heidrich, and G.Wetzstein,“Fast and flexible convolutional sparse coding,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

Z. Liu and K. Lu, “Convolutional sparse coding for noise attenuation in seismic data,” Geophysics, vol. 86, no. 1, pp. V23–V30, 2021.

M. D. Sacchi, D. J. Verschuur, and P. M. Zwartjes, Data reconstruction by generalized deconvolution, 2005, pp. 1989–1992.

F. G. Veshki and S. A. Vorobyov, “Efficient ADMM-based algorithms for convolutional sparse coding,” IEEE Signal Processing Letters, vol. 29, pp. 389–393, 2022.