Seismic Inversion by Principal Components Analysis and Neural Networks

Oswaldo Ludwig Júnior and Wilson M. Figueiró, CPGG-UFBA, Federal University of Bahia, Salvador, Brazil.

Abstract

This paper describes how Artificial Neural Networks can be employed to solve seismic inverse problems. The main objective is to obtain one-dimensional Earth velocity models from seismic waveform data. Precedent works are discussed and a new approach is suggested. It means, the Principal Component Analysis is adopted to avoid redundant information that is originated due to the Common Middle Point gather. A secondary objective is to reduce the total number of network parameters and increase the generalization skill on the neural network.

Introduction

The geophysical problems are characterized by the attempt to have an image of subsurface geologic structures using data measured on observation surface. The inaccessible character of Earth interior causes problems like the indirect estimation of physical and geologic properties of rocks deeply located. Difficulty like representation of a geologic reality by a model mathematically and computationally tractable, presence of noise in the data and weak relationship between measured data and models parameters, implicates in problems such as: existence, ambiguity and stability of solution.

The deterministic approach to solve this class of problems works well for simple, experimental, controlled and unrealistic models. The deterministic treatment of these problems has also a place as linear steps inside a larger non-linear procedure. Generally, this approach is applied to study; to test or to revels new ideas in a particular aspect of specific situations. For example, the Gauss-Newton method applied to the study of the reflector slope (Figueiró and Goldin, 2005). Iterative procedures like that can work well for moderate non-linear problems, but out of that deterministic reign, there is space for applications of non-deterministic methods, such as: geostatistics, genetic algorithm, chaos, geological feelings and so on. Among such methods, this work is particularly interested in one called Artificial Intelligence (AI). It points some failures in applications of AI to solve inverse seismic problems presented in literature. In addition, it show an alternative strategy to overcome such difficulties and to solve seismic inverse problems.

The inverse geophysical problem consists of using a set of observations to infer subsurface parameters. A mathematical model is necessary in order to do the relationship between the measure data and the unknown parameters. Some authors (Calderón-Macías and Sen, 1993), (An and Moon, 1993), (Röth and Tarantola, 1992) propose the use of Artificial Neural Network (ANN) to modeling this relationship.

Note that a supervised ANN is nothing else than a function that relates an input space with an output space (Haykin,1999). Equations (1) and (2) illustrate a function for a feedforward multi-layer ANN with one hidden layer and a linear output layer respectively:

$$y_h = \varphi(W_1x+b_1) \tag{1}$$

and

$$y = W_2y_h+b_2 \tag{2}$$

where $\varphi(.)$ is a usual activation function as hyperbolic tangent or sigmoid; W_1, W_2 are the synaptic weights matrix, b_1 and b_2 are the bias vectors; x is the input vector; y is the output vector and y_h is the output vector of the ANN hidden layer.

To achieve a satisfactory performance, the ANN must be adjusted (i.e. the synaptic weights must be optimized).

The ANN synaptic weights adjust is done by the use of training examples. The simulation of a synthetic wavelet (i.e. source) applied to a synthetic model gives a seismogram. Some data pairs consisting of a synthetic seismogram and the parameter vector of the synthetic model is used to training the ANN.

The proposed problem

This work is to suggest an automated tool for the seismic inversion problem. The problem scope is limited to obtaining 1D Earth velocity models. However, this tool may be extended to 2D or 3D Earth velocity model.

The idea is to give a seismogram to the ANN and receive a velocity model of each homogeneous and isotropic layer separated by horizontal interfaces. Figure 1 (Dos Santos, 2002) shows an element of such family of models.

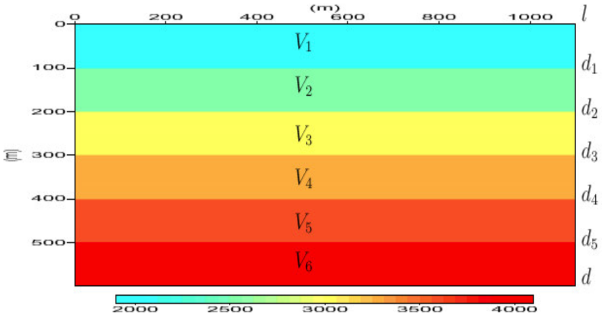

Figure 1 - 1D Earth velocity model of homogeneous isotropic layers separated by horizontal interfaces.

Notice that the input vector and the output vector of ANN have unalterable dimension. So, the dimension of the seismogram matrix and the total number of the model parameters must be predefined according to the neural architecture adopted. To solve this limitation, the sample frequency of the seismic trace may be changed to result a vector of predetermined length.

Synthetic models for ANN training

In order to training the ANN a set of synthetic seismograms related to different models must be generated.

The source (i.e. wavelet) for the synthetic training examples must be close to the real wavelet. The authors suggest the use of the gaussian second derivatives, according to Equation (3). Figure 2 (Cunha, 1997) illustrates this function.

$$f(t) = [1 - 2(\pi f_p t)^2] e^{-(\pi f_p t)^2} \qquad (3)$$

where f_p is the peak frequency.

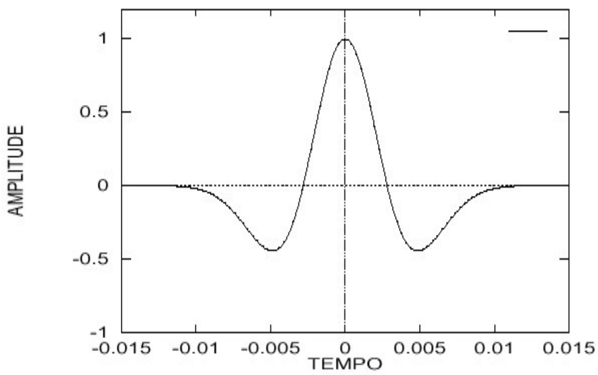

Figure 2 – The source function has a gaussian second derivatives form.

The seismic traces can be calculated based on the finite difference method (FDM) applied to the acoustic wave equation. Assuming that the Earth has an acoustic behavior, that allows us to implement, in a numerical way, a seismic modeling employing regular nets in models representing 2-D geological media. Second derivatives of the wave equation can be obtained by Taylor Series expansion of fourth order to the space and of second for the time (Mufti et al., 1996)

This procedure makes possible to generate synthetic seismograms related to synthetic models. Figure 3 (Dos Santos, 2002) illustrates an example of this seismogram. Additive noise can be added to the seismograms in order to approximate the synthetic seismic trace to the real seismic trace.

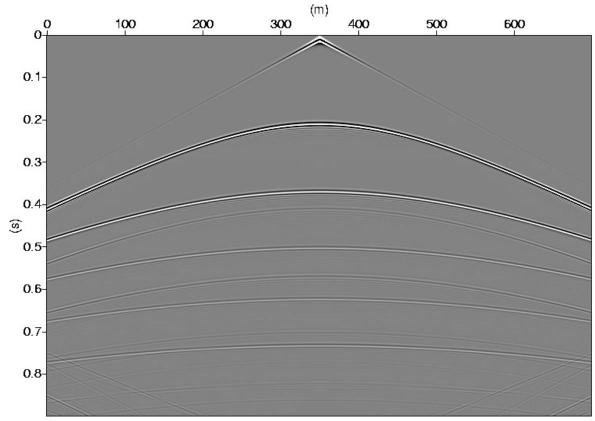

Figure 3 – An example of synthetic seismograms related to a homogeneous isotropic layers separated by horizontal interfaces model.

Adequate ANN architecture

An ANN Feedforward Multilayer with one nonlinear hidden layer and a linear output layer is enough to present the universal approximation skill (Hornik et al., 1990). So, this ANN is adequate for the geophysical inversion task. Figure 4 illustrates this neural architecture.

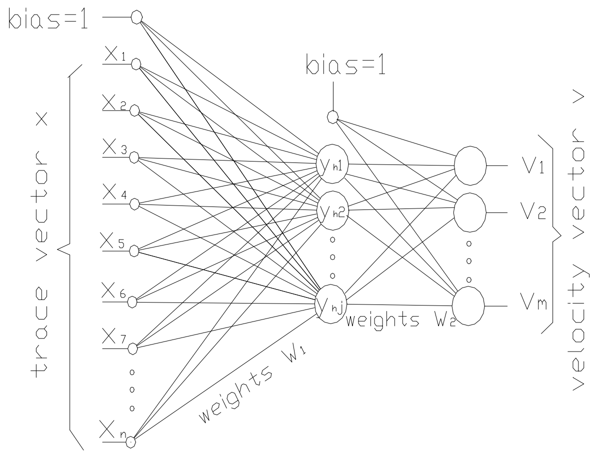

Figure 4 - Flowchart of the adequate ANN.

Equations (1) and (2) presents the mathematical model for this ANN.

Notice that a matrix S of dimension n x m, where n is the number of elements of each seismic trace and m is the total number of seismic traces, represents the seismogram. However, the ANN input must be a vector. The solution is to concatenate all the columns of matrix S in order to compose a vector x with $d=n \cdot m$ dimensions.

The optimization of the ANN parameters demands a scalar function of error energy J. An adequate approach is to consider de Equation (4).

$$J = (v - \hat{v})^T (v - \hat{v}) \qquad (4)$$

where v is the target velocity vector and $\hat{v}$ is the ANN output vector.

Some gradient optimization method can be employed to minimize *J*, conform Equation 5. Another approach is a global optimization method to example of Genetic Algorithm, where the fitness function may be based on *J*.

$$w[n+1] = w[n] - \alpha \nabla J[n] \quad (5)$$

where *w* is a synaptic weight vector, *n* is the iteration, $\alpha \in (0,\ 1]$ is the learning coefficient and ∇J is the gradient of *J* in relation to *w*.

Normally, the total number of hidden neurons is obtained empirically by testing the performance of some types of ANN architectures.

Pre-processing method

Notice that, in geophysical applications, usually vector *x* has a great dimension. The direct application of the vector *x*, as the input data of the ANN, implicates in a great number of adjustable parameters on the ANN.

ANN with a great number of adjustable parameters demand a great training data set to present generalization skill (Freeman and Skapura, 1991). The practice is to use at least ten examples for each adjustable parameter.

The work of (Calderón-Macías and Sen, 1993) they uses a seismogram matrix with dimension 128 x 16 (i.e. 16 seismic traces with 128 positions). Then, the vector *x* has 2048 positions. This vector is directly applied to an ANN with 28761 adjustable parameters. The practice suggests a training set of 287610 examples to reach an adequate generalization skill. This training set is unavailable in practical applications.

Fortunately, due to the Common Middle Point gather, the coordinates of vector *x* has high covariance (i.e. vector *x* has redundant information). So, it is possible to use of Principal Component Analysis (PCA) in order to reduce the dimension of the vector *x*.

The PCA method (Haykin, 1999) allows the coordinates covariance elimination by a change of basis.

To applied PCA method, the first step is the calculation of the covariance matrix *C*. In a discrete approach, each element of this matrix represents the covariance $s^2(x_i, x_j)$ between two coordinates of the vector *x*, according to Equation (6).

$$s^2(x_i, x_j) = \sum_{n=1}^{N} x_i[n] x_j[n] - \frac{\sum_{n=1}^{N} x_i[n] \sum_{n=1}^{N} x_j[n]}{N} \quad (6)$$

where x_i and x_j are coordinates of vector *x* and *N* is the number of examples.

The second step is the determination of the eigenvectors v_n of *C* and matrix *P*, whose columns are these eigenvectors, according to Equation (7).

$$P = \{v_1, v_2, ..., v_n\} \quad (7)$$

Using matrix *P* to change the *C* basis one can obtain the diagonal matrix *D* of eigenvalues of *C*, according to Equation (8).

$$CP = PD \Rightarrow P^{-1}CP = D \quad (8)$$

Note that matrix *D* (i.e., *C* in the new basis) does not present covariance. This diagonal matrix has only variance $s^2(x_i, x_i)$. The *D* matrix is the covariance matrix of vector *x* in the basis *P*.

One to conclude by this fact that vector *x* in the basis *P* has no covariance among its coordinates. In other words, this vector in basis *P* has no redundant information.

Figure 5 illustrates the basis change effect in a set of vectors $x \in \Re^2$. After the change to basis *P*, the covariance is zero and the coordinate x_{1P} justifies almost the total of the variance (i.e., almost all the information).

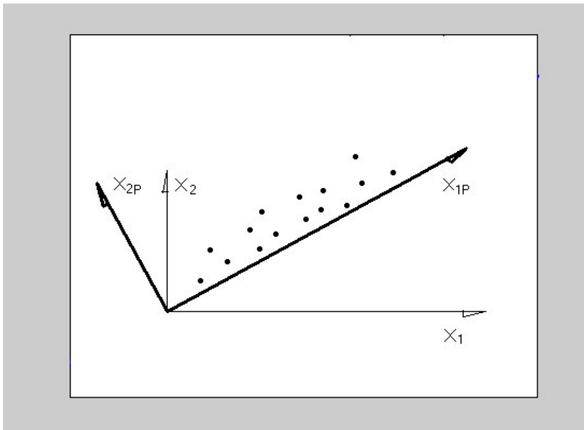

Figure 5 - Basis change effect.

After the basis change, the resultant vector x_P has coordinates with low variance (i.e. low eigenvalue in *D*). These coordinates can be dispensed due to its low entropy or information level. So, vector x_p can be truncated, maintaining only the high variance coordinates.

The ANN can receives the truncated vector x_p instead of vector *x*. So, the PCA method reduces the total number of adjusted parameters of the ANN and consequently reduces the training data set to feasible values. This tool makes possible a real application of ANN to seismic inverse problems.

Precedent works

In the work of (Calderón-Macías and Sen, 1993) the authors have evaluated the application of ANN to the solution of geophysical inverse problems. More specifically, ANNs was used to obtain 1D acoustic velocity models with 6 layers from synthetic seismogram matrix compose by 16 seismic traces with 128 positions.

Results show that the trained networks can learn to relate waveform seismic data with velocity models. Figure 6

(Calderón-Macías and Sen, 1993) illustrates an example of a synthetic seismogram and the velocity model used in the training section.

In this work velocity predictions for the top layers (layers 2 and 3) were more accurate than those for the bottom layers (layers 4, 5 and 6). The authors relate this result to the fact that the shallow layers have more moveout than deep layers.

The work of (Röth and Tarantola, 1992) show that an ANN trained using data with 10% uncorrelated noise added performs better than a network trained with noise-free data for solving a similar problem.

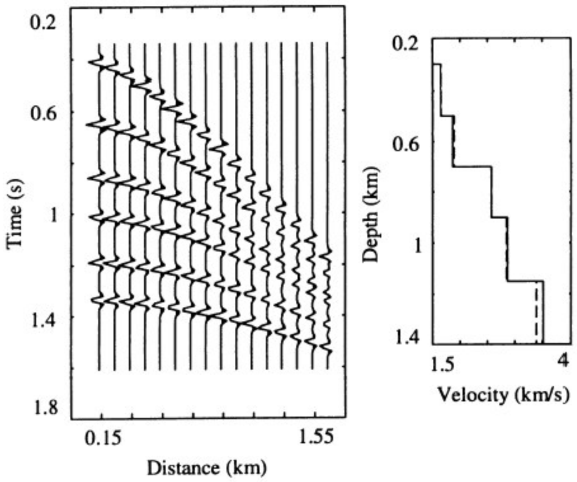

Figure 6 - Synthetic seismograms and a velocity model

Conclusions

Precedent works show that ANN directly applied to inversion problems reproduces training data set perfectly, but it can be seen that there is a considerable difference between the training data set and the testing data set.

In these works, more input patterns should be given to the network in order to reduce errors in the testing set. In case of the testing data set, the network computed a model that produces seismograms that have several mismatches in travel time compared to the ones computed with the true model. This is because the ANN has a high number of adjustable parameters in relation to the number of training data set. This fact may implicate in an underdetermined problem (i.e. more than one solution for the synaptic weights that makes the global error equal to zero). In this case, no training method can do an adequate fit to the net and the ANN does not possess the generalization skill (i.e. skill to predict data not seen in the training stage).

In applications of geophysical inversion the ANN has a high number of adjustable parameters due to the high dimension of the input data. The PCA preprocessing method, suggested in the present paper, can reduce the input vector dimension without significant loss of information. So, the authors believe the training data set may be available for real implementations.

Acknowledgements

We present our Thanks to the CPGG-UFBA (Center for Research in Geophysics and Geology of Federal University of Bahia) and to FTC (Faculdade de Tecnologia e Ciências).

References

An P. and Moon W.M. 1993. Reservoir characterization using feedforward neural networks. 63rd SEG meeting, Washington, DC, USA, Expanded Abstracts, 258–262.

Calderon-Macias C. and Sen M.K. 1993. Geophysical interpretation by artificial neural systems: a feasibility study. 63rd SEG meeting, Washington, DC, USA, Expanded Abstracts, 254–257.

Cunha, P. E. M., 1997. Estratégias Eficientes para Migração Reversa no Tempo Pré-empilhamento 3-D em Profundidade pelo Método das Diferenças Finitas. Dissertação de Mestrado, Pós-Graduação em Geofísica, Instituto de Geociências da UFBA, Salvador, Brasil.

Dos Santos, R. H. M., 2002. Modelagem Acústica Bidimensional Usando Diferentes Parametrizações do Campo de Velocidade. Dissertação de Mestrado, Pós-Graduação em Geofísica, Instituto de Geociências da UFBA, Salvador, Brasil.

Figueiró, W. M., and Goldin, S. V., 2005. Curved interface slope estimation using seismic reflection travel time data. Submited to the Journal of Seismic Exploration.

Freeman J.A. and Skapura D.M. 1991. Neural Networks: Algorithms, Applications, and Programming Techniques. Addison-Wesley Publishing Co.

Haykin, Simon 1999. Redes Neurais: Princípios e Prática, 2a. Edição, Bookman Companhia Editora, São Paulo.

Hornik, K., M. Stinchcombe and H. White. 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, Neural Networks, vol 3, pp.551-560.

Mufti, I; Pita, J. A. e Huntlay, R. W. 1996. Finite difference depth migration of exploration scale 3-D seismic data, Geophysics, 61: 776 –794.

Röth G. and A. Tarantola, 1992, Neural networks and inversion of seismic data, submitted to Journal of Geophysical Research.