PERGAMON

AVAILABLE AT www.ComputerScienceWeb.com POWERED BY SCIENCE DIRECT®

Neural Networks 16 (2003) 349–353

Neural Networks

www.elsevier.com/locate/neunet

2003 Special Issue

Application of the kernel method to the inverse geosounding problem

Hugo Hidalgo[a,*], Sonia Sosa León[b], Enrique Gómez-Treviño[a]

[a]*CICESE, Km. 107 Carr. Tijuana-Eda., Ensenada 22860, Mexico*
[b]*Departamento de Matemáticas, Universidad de Sonora, Hermosillo, Son. 83000, Mexico*

Abstract

Determining the layered structure of the earth demands the solution of a variety of inverse problems; in the case of electromagnetic soundings at low induction numbers, the problem is linear, for the measurements may be represented as a linear functional of the electrical conductivity distribution. In this paper, an application of the support vector (SV) regression technique to the inversion of electromagnetic data is presented. We take advantage of the regularizing properties of the SV learning algorithm and use it as a modeling technique with synthetic and field data. The SV method presents better recovery of synthetic models than Tikhonov's regularization. As the SV formulation is solved in the space of the data, which has a small dimension in this application, a smaller problem than that considered with Tikhonov's regularization is produced. For field data, the SV formulation develops models similar to those obtained via linear programming techniques, but with the added characteristic of robustness.

Keywords: Support vector; Regularization; Kernel method

1. Introduction

To investigate the internal structure of the earth, geophysicists relay mostly on the interpretation of measurements taken on the surface. This applies for deep soundings of hundreds of kilometers as well as for shallow studies of merely a few meters below the surface. The electrical conductivity of rocks is often the property of interest in these types of studies. For this reason, a great amount of electrical techniques had been developed to infer the conductivity structure of the subsurface on the basis of surface measurements.

One of these techniques is based on electromagnetic induction by means of an alternating current that is made to flow in a transmitting coil. This current generates an alternating magnetic field in the surrounding environment, which in turn induces an electromotive force both in the conductive ground and in a receiving coil (Grant & West, 1965). A particular version that works at low induction numbers is of special interest from both theoretical and practical reasons. We exploit here the peculiar theoretical aspect of the technique. It turns out that apparent conductivity, a normalized quantity of the surface measurement is a linear functional of the unknown conductivity of the subsurface (Gómez-Treviño, Esparza, & Méndez-Delgado, 2002). The relationship is in the form of a first order Fredholm equation

$$\sigma_a(r) = F_i\sigma := \int_0^\infty A(r,z)\sigma(z)\mathrm{d}z, \tag{1}$$

where r is the separation between the transmitting and the receiving inductors, z depth ($z = 0$ on the surface). $A(r,z)$ is, for the vertical magnetic dipole case (Esparza & Gómez-Treviño, 1987)

$$A(r,z) = \frac{4zr}{\sqrt{(4z^2 + r^2)^3}}, \tag{2}$$

and for horizontal magnetic dipoles (Gómez-Treviño et al., 2002)

$$A(r,z) = \frac{2}{r} - \frac{4z}{r\sqrt{4z^2 + r^2}}. \tag{3}$$

In order to obtain an estimation of the distribution of the conductivity $\sigma(z)$, an inversion process has to be realized. Although the exact solution for the corresponding inverse problems is known (Esparza & Gómez-Treviño, 1987; Gómez-Treviño et al., 2002), the inversion process is affected by the problem of instability because we have responses for only a few values of r. For this reason the inverse problem is often treated applying some kind of

* Corresponding author. Tel.: +52-646-174-5050; fax: +52-646-175-0593.

E-mail addresses: hugo@cicese.mx (H. Hidalgo), sonia@hades.mat.uson.mx (S. Sosa León), egomez@cicese.mx (E. Gómez-Treviño).

0893-6080/03/$ - see front matter © 2003 Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0893-6080(03)00025-X

regularization (Tenorio, 2001). In the present work, the regularization properties of the support vector (SV) method (Smola, Schölkopf, & Müller, 1998) are considered for the inversion process of the linear operators (2) and (3).

2. The inverse problem and SV regularization

The Tikhonov's regularization method (Tikhonov & Arsenin, 1977) has become the preferred technique to obtain models of the subsurface conductivity distribution from electromagnetic measurements (Constable, Parker, & Constable, 1987; Hidalgo, Marroquín, & Gómez-Treviño, 1998). It is our interest to explore in this direction the use of SV Methods, given the close relationship with regularization (Girosi, 1998; Smola et al., 1998).

The statement of the problem, considering the regularization approach, is as follows: the model proposed for the conductivity distribution can be obtained minimizing the functional

$$I_\lambda(\sigma) = \frac{1}{N}\sum_{i=1}^{N} g(d_i - F_i\sigma) + \frac{\lambda}{2}\|\sigma\|^2_{\mathscr{H}}. \tag{4}$$

In this formulation, $g(\cdot)$ represents the cost function that measures the fitness of the model responses to each data d_i of the N available. The second term correspond to the Tikhonov's regularization functional, a norm in the Hilbert Space $\mathscr{H}$ of the functions. λ is the regularization parameter, used to control the tradeoff between fitness of model's responses to data and model smoothness. When $g(x) = x^2$, the minimizer of Eq. (4) satisfies $F_i^* F_i\sigma + \lambda\sigma = F_i^* d_i$ for each i, where F_i^* is the adjoint operator of F_i.

When the space of models is considered a reproducing kernel Hilbert space (RKHS) (Máté, 1989) with kernel $Q(\cdot,\cdot)$, and the F_i are bounded linear functionals, as in the present case, the minimizer of Eq. (4) is of the form (Kimeldorf & Wahba, 1971)

$$\hat{\sigma}(z) = \sum_{i=1}^{N} \beta_i R(r_i, z). \tag{5}$$

where $R(r_i, z)$ is the representer for F_i in $\mathscr{H}$, i.e. for any σ, $F_i\sigma \equiv \langle R(r_i,\cdot),\sigma\rangle$. The representer can be evaluated in our case as

$$R(r_i, z_0) = F_i Q(z_0,\cdot) = \int_0^\infty A(r_i, u) Q(u, z_0)\,du. \tag{6}$$

The RKHS formulation provides for a very rich structure for the space of models (Girosi (1998), Máté (1989), and Vapnik (1998) for a collection of kernels and their properties). Considering the equivalence between regularization penalizers and the SV method, a Sobolev space may be generated, as those presented in Table 1.

Also, a space of models consisting of functions with Fourier transform of limited band on $[-w, w]$ may be generated using the kernel

$$R(s,t) = \frac{\sin(w(t-s))}{w(t-s)}.$$

The finite and discrete version of the previous representation, consisting of trigonometric polynomials $h(t) = e^{-ikt}$, for $k = 0, \pm 1, \pm 2, \ldots, \pm n$ uses the Fejér kernel

$$R(s,t) = \frac{\sin(n + \frac{1}{2}(t-s))}{\sin\frac{1}{2}(t-s)} \text{ (Máté (1989))}.$$

When $g(s) = s^2$ we obtain the linear least squares problem, studied by Wahba with smoothing splines (Wahba, 1990). Our main interest is in the case of $g(\cdot) = |\cdot|_\varepsilon$, Vapnik's ε-insensitive loss function (Vapnik, 1998)

$$|x|_\varepsilon = \begin{cases} 0, & |x| \le \varepsilon, \\ |x| - \varepsilon, & \text{otherwise.} \end{cases} \tag{7}$$

Considering the equivalence $\|\sigma\|^2_{\mathscr{H}} = \sum_{i,j}\beta_i\beta_j K_{ij}$, with

$$K_{ij} = F_i R_j = \int_0^\infty A(u, r_i) R_j(u)\,du, \tag{8}$$

the original minimization functional is restated as follows (Vapnik, 1998). Considering slack variables ξ_i, ξ_i^*, $i = 1, \ldots, N$, the original problem is now a constrained optimization one, with the objective of finding the parameters β_i that minimize the functional

$$L_\lambda(\beta, \xi, \xi^*) = \frac{1}{N}\sum_{i=1}^{N}(\xi_i + \xi_i^*) + \frac{\lambda}{2}\sum_{i,j}\beta_i\beta_j K_{ij} \tag{9}$$

under the constraints

$$d_i - \sum_j \beta_j K_{ij} \le \varepsilon + \xi_i, \quad \sum_j \beta_j K_{ij} - d_i \le \varepsilon + \xi_i^*,$$
$$\xi_i, \xi_i^* \ge 0, \quad i = 1, 2, \ldots, N. \tag{10}$$

Applying the Lagrangian formulation and obtaining the dual problem, as in Vapnik, 1998 the minimization problem (9) becomes equivalent to maximization of the functional

$$L_\lambda(\alpha, \alpha^*) = -\varepsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N} d_i(\alpha_i - \alpha_i^*) - \frac{1}{2} \times \sum_{i,j}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K_{ij} \tag{11}$$

subject to α_i, $\alpha_i^* \in [0, 1/\lambda N]$, with $\beta_i = \alpha_i - \alpha_i^*$, for $i = 1, 2, \ldots, N$. In formulation (4), we can include the case

Table 1
Some kernels and their representations

Kernel	Norm
$e^{-\lvert x-y\rvert} + e^{-x}e^{-y}$	$\|f\|^2 = \int_0^\infty \lvert f'(x)\rvert^2 + \lvert f(x)\rvert^2 dx$, even functions
$e^{-\lvert x-y\rvert} - e^{-x}e^{-y}$	$\|f\|^2 = \int_0^\infty \lvert f'(x)\rvert^2 + \lvert f(x)\rvert^2 dx$, odd functions
$e^{-\lvert x-y\rvert}[1 + \lvert x-y\rvert]$	$\|f\|^2 = \int_0^\infty (\lvert f''(x)\rvert^2 + 2\lvert f'(x)\rvert^2 + \lvert f(x)\rvert^2) dx$

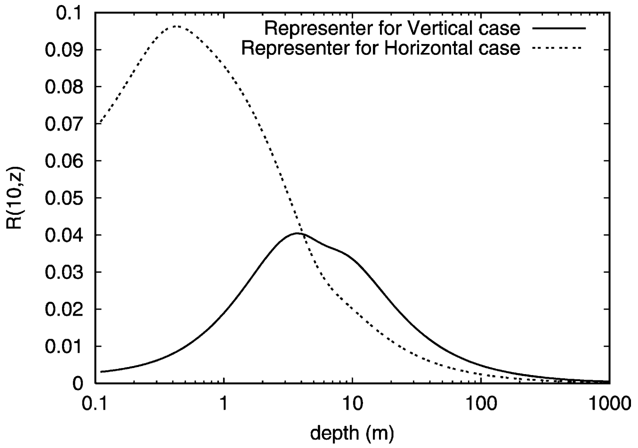

Fig. 1. Representers of the linear functionals for the vertical and horizontal cases, when $r = 10$, and considering a RKHS with kernel $Q(z,u) = \exp((-|z-u|^2)/(2\gamma^2))$ with $\gamma = 0.22$.

of different functionals for the data, and the inversion process can be realized on the vertical or horizontal soundings alone, or in a joint version of both. In the usual approach of regularization, a multilayered model is considered, and the size of the minimization problem depends on the discretization of the model, or regularization mesh. For the case under consideration, the size of the problem is equal to the number of data, usually small, compared to regression problems in other areas. The construction of the model requires, besides the definition of the regularization mesh, the evaluation of the integrals (6) for the representers, and for the elements of K_{ij} in Eq. (8).

3. Application to synthetic data

In this section we attempt the recovery of models when the data are synthetically generated from known 'real' models. The models considered are: $\sigma_1(z) = 1 + 10\delta(10-z)$ and $\sigma_2(z) = 1 + \sin(\pi z/20)$. These models were studied for vertical dipoles in Esparza and Gómez-Treviño (1987), obtaining the analytic inverse for them. First we will concentrate on the representers. In Fig. 1 we show the representers for the functionals (2) and (3), when $r = 10$ and a Gaussian kernel is considered

$$Q(z,u) = \exp\left(\frac{-|z-u|^2}{2\gamma^2}\right),$$

with $\gamma = 0.22$.

It can be observed the difference in depth of penetration for both methods: the representer for the vertical instrument attains its greatest resolution in $[1, 20]$ m, and the horizontal dipole's representer obtains its greatest perception in the interval $[0.2, 2]$ m.

The recovered models for the case of synthetic data generated from the model $\sigma_1(z) = 1 + 10\delta(10-z)$ are presented on Fig. 2. The data used were generated for values of $r = 0.1, 1, 5, 10, 20, 40, 80, 100$, and $1000m$, and a Gaussian Kernel was considered with $\gamma = 0.22$. The regularization parameter considered was $\lambda = 10^{-5}$, and $\varepsilon = 10^{-4}$ for vertical and horizontal cases. Increasing λ smooths the model developed, and deteriorates its form. For the joint inversion we had to increase λ to 0.005 in order to obtain convergence. As for the fitness to data, the mean squared error was of 0.0075 for the joint case, 0.00376 for vertical and 0.0011 S^2 for the horizontal cases. The recovered models contain the structure of the delta, although with some fluctuations at depths before and after the $10m$.

In Fig. 2b we present a comparison with the model developed using Tikhonov's regularization, for the vertical data, with λ estimated using cross-validation. The models are very similar around $10m$ because of the use of a Gaussian kernel, which is equivalent to regularize penalizing all of the derivatives of the model. The model developed with Tikhonov's regularization deteriorates out of the $[1, 100m]$ interval, it is unable to use the data below $5m$ or above $100m$.

Fig. 3 shows the models recovered for synthetic data obtained solving the forward equations for the model $\sigma(z) = 1 + \sin(\pi z/20)$. The inductor's separations employed were

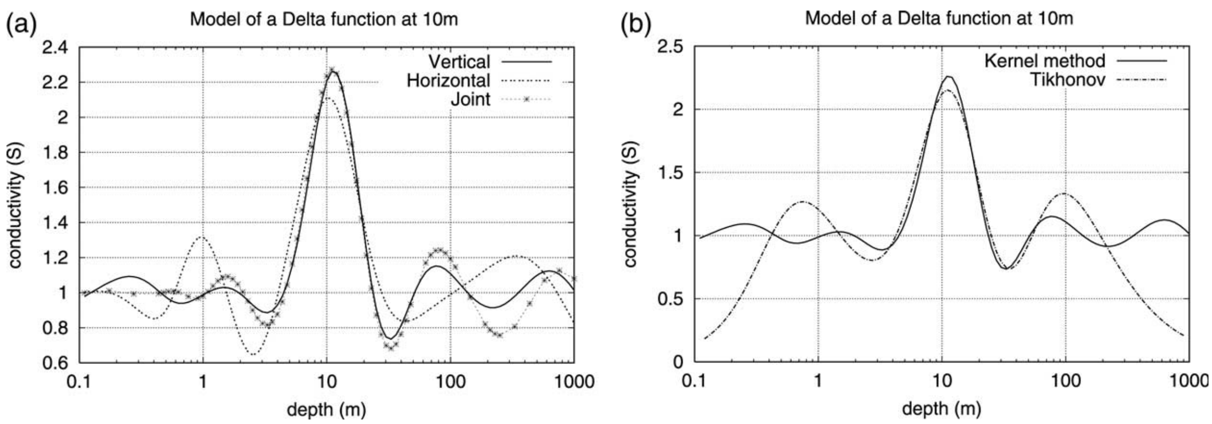

Fig. 2. (a) Model recovered from synthetic data for a delta function at $10m$, and considering a RKHS with kernel $Q(z,u) = \exp(-|z-u|^2/2\gamma^2)$, with $\gamma = 0.22$. (b) Comparison of a previous model (vertical dipoles) with a model obtained using Tikhonov's regularization.

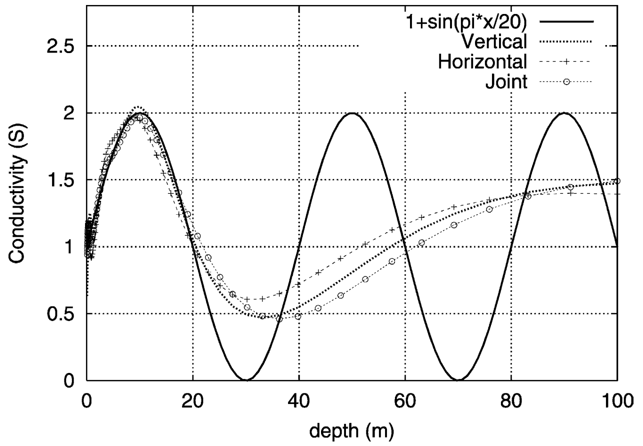

Fig. 3. Model recovered from synthetic data for an oscillating function at $1 + \sin(\pi z/20)$ and considering a RKHS with kernel $Q(z,u) = \exp(-|z-u|^2/2\gamma^2)$, with $\gamma = 0.22$.

$r = 1, 10, 20, 40, 80,$ and $100m$. The model simulates a layered earth situation. The main interest here is on the actual number of layers that can be recovered from the data.

From Fig. 3 we observe that due to the small amount of data available for the inversion process, we can only recover information below $50m$.

4. Application to field data

In this section we present models developed applying the previous procedure to field data from a local study performed in a region of Baja California, México. The soundings were performed at spacings of 10, 20 and 40 m. The resulting models are shown in Fig. 4.

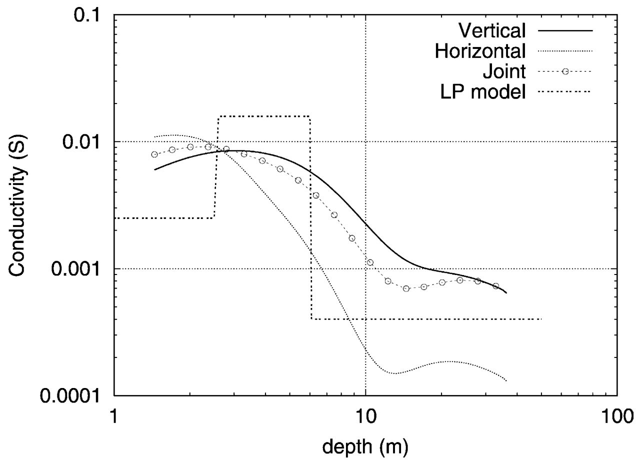

Fig. 4. Model recovered from field data when inversion was performed on vertical, horizontal and joint data. The parameters employed were $\varepsilon = 0.00001$, $\lambda = 5$, and a RKHS with kernel $Q(z,u) = \exp(-|z-u|^2/2\gamma^2)$, with $\gamma = 1$ was considered. The model named 'LP' was obtained by Gómez-Treviño et al. (2002) using LP techniques.

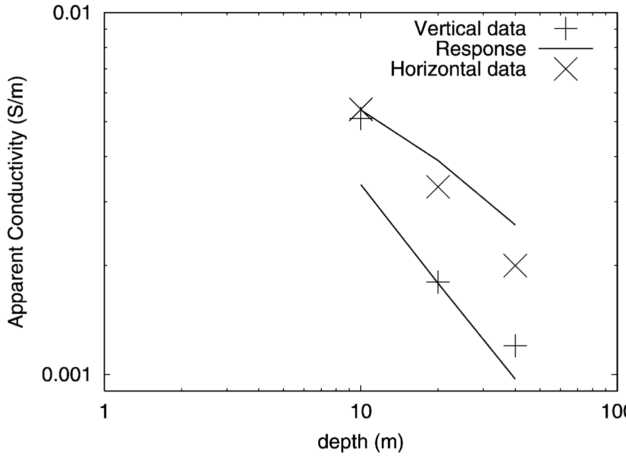

Fig. 5. Responses for vertical and horizontal field data, for the model shown in Fig. 4 (joint inversion).

Basically, the same structure was observed by Gómez-Treviño et al. (2002) with a linear programming (LP) method that enforces a layered structure with minimum vertical derivative norm for the logarithm of conductivity. They recover two-layer structures for the separate vertical and horizontal sounding data, and obtain an extra third layer for the joint inversion. Their three-layer model is also shown in Fig. 4. The fitness to the data is presented in Fig. 5 for our joint inversion procedure. It can be observed that the first vertical datum is not well approximated by the response. The method actually incorporated a big slack variable ξ for that data, considering it as an outlayer. In order to fit that datum it would have to obtain a very complicated model, but the smoothness penalizer does not allow for that, due to the robust fitness function.

5. Conclusions

The regularization properties of the SV formulation were applied to the geosounding inverse problem. The modeling strategy allows to incorporate a priori information about the space function where the models are supposed to belong. The function spaces available with regularization are no longer limited to those obtained from first or second order derivative penalizers. The capabilities of this learning method were evaluated on synthetic and field data, obtaining results better or comparable with those of Tikhonov's or more elaborated inversion procedures.

References

Constable, S. C., Parker, R. L., & Constable, C. G. (1987). Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, *52*, 289–300.

Esparza, F. J., & Gómez-Treviño, E. (1987). Electromagnetic sounding in the resistive limit and the Backus–Gilbert method for estimating averages. *Geoexploration*, *24*, 441–454.

Girosi, F. (1998). An equivalence between sparse approximation and support vector machines. *Neural Computation*, *10*, 1455–1480.

Gómez-Treviño, E., Esparza, F., & Méndez-Delgado, S. (2002). New theoretical and practical aspects of electromagnetic soundings at low induction numbers. *Geophysics*, *67*, 1441–1451.

Grant, F. S., & West, G. F. (1965). *Interpretation theory in applied geophysics*. New York: McGraw-Hill.

Hidalgo, H., Marroquín, J. L., & Gómez-Treviño, E. (1998). Piecewise smooth models for electromagnetic inverse problems. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 556–561.

Kimeldorf, G., & Wahba, G. (1971). Some results on Tchebycheffian spline functions. *Journal on Mathematical Analysis and Applications*, *33*, 82–95.

Máté, L. (1989). *Hilbert space methods in science and engineering*. Bristol, England: Adam Hilger.

Smola, A., Schölkopf, B., & Müller, K. R. (1998). The connection between regularization operators and support vector kernels. *Neural Networks*, *11*, 637–649.

Tenorio, L. (2001). Statistical regularization of inverse problems. *SIAM Review*, *43*, 347–366.

Tikhonov, A., & Arsenin, V. Y. (1977). *Solutions of ill-posed problems*. Washington, DC: V.H. Winston and Sons.

Vapnik, V. (1998). *Statistical learning theory*. London: Wiley.

Wahba, G. (1990). *Spline models for observational data* (*Vol. 59*). *CBMS-NSF regional conference series in applied mathematics*, Society for Industrial and Applied Mathematics.