

The Use of the Levenberg–Marquardt and Variable Projection Curve-Fitting Algorithm in Intravoxel Incoherent Motion Method for DW-MRI Data Analysis

Roberta Fusco · Mario Sansone · Antonella Petrillo

Received: 8 January 2015 / Published online: 4 March 2015
© Springer-Verlag Wien 2015

Abstract The objective of this study was to evaluate the performances of different algorithms for diffusion parameters estimation in intravoxel incoherent motion method for diffusion-weighted magnetic resonance imaging (DW-MRI) data analysis. Traditionally, the method of non-linear least squares analysis by means of Levenberg–Marquardt algorithms has been used to estimate the parameters obtained from exponential decay data. In this study, we evaluated the Variable Projection curve-fitting algorithm and the performance of two non-linear regression methods when single and multiple starting points were used. Analysis was done on simulation data to which different amounts of Gaussian noise had been added. The performance of two non-linear regression methods was compared using the residual sum of squares and the number of failures in data fitting. We conclude that the VarPro algorithm is superior to the LM algorithm for curve fitting in intravoxel incoherent motion method for DW-MRI data analysis.

1 Introduction

Diffusion-weighted magnetic resonance imaging (DW-MRI) shows promise as an imaging biomarker for treatment response in a variety of clinical tumor types [1–9]. Routine in almost all preclinical and clinical scanners, diffusion maps can be

R. Fusco and M. Sansone contributed equally to this work.

R. Fusco (✉) · A. Petrillo

Division of Radiology, Department of Diagnostic Imaging, Radiant and Metabolic Therapy, Istituto Nazionale Tumori Fondazione Giovanni Pascale, IRCCS, via Mariano Semmola, 80131 Naples, Italy
e-mail: r.fusco@istitutotumori.na.it

M. Sansone

Department of Electrical Engineering and Information Technologies, University “Federico II” of Naples, via Claudio, 21, 80132 Naples, Italy

generated from a minimum of two images acquired at low (b value ~ 100 s/mm²) and high (b value ~ 1000 s/mm²) diffusion weightings.

Depending on the acquisition parameters, apparent diffusion coefficient (ADC) is known to reflect variable combinations of diffusion and perfusion effects. The concept of the IVIM method (intravoxel incoherent motion) initially described by Le Bihan et al. [10, 11] has the potential to measure both true molecular diffusion and incoherent motion of water molecules in the capillary network, known as pseudodiffusion. The pulse sequence used in this method is made sensitive to the motion of blood through the application of magnetic field gradients in a manner analogous to the measurement of diffusion coefficient [10, 11]. In the simplest possible model for analysis of IVIM data, the water in the tissue of interest is described as being in either one of two compartments: intravascular or extravascular. Extravascular water moves by ordinary diffusion, and the process can be assigned a diffusion coefficient with dimension mm²/s. Intravascular water not only moves via diffusion but also moves with the bulk flow of blood. Blood flow displaces intravascular spins over much greater distances per unit time than diffusion. If the movement of these spins through vessels is modeled as random or incoherent motion, then this process can be assigned a pseudodiffusion coefficient that also has dimension of mm²/s. Overall, there are three pieces of information available from this model of the IVIM experiment: f the fraction of spins within the volume of interest that are within flowing blood, D_p the pseudodiffusion coefficient of those intravascular flowing spins, and D_t the diffusion coefficient of the extravascular non-flowing spins.

Then, assuming monoexponential signal attenuation with b value, the apparent diffusion coefficient (ADC) can be calculated analytically. Instead, assuming biexponential signal attenuation with b value, tissue pure diffusion (D_t), pseudodiffusion (D_p) and perfusion fraction (f) can be calculated analytically.

Traditionally, the method of non-linear least squares (NLLS) analysis by means of Levenberg–Marquardt algorithms has been used to estimate the parameters obtained from exponential decay data [12].

In this study, we evaluated the Variable Projection algorithms and the performance of two non-linear regression methods when single and multiple starting points were used. Analysis was done on simulation data to which different amounts of Gaussian noise had been added.

The performance of two non-linear regression methods was compared using the residual sum of squares in data fitting.

2 Materials and Methods

2.1 IVIM

The simple, two compartment model described above can be expressed as [10, 11]

$$S(b)/S_0 = (1 - f)e^{-bD_t} + fe^{-bD_p}, \quad (1)$$

where $S(b)$ is the echo amplitude with diffusion gradients on at amplitude b and S_0 is the echo amplitude with diffusion gradients off. The parameters f , D_t , and D_p are, respectively, perfusion fraction, tissue pure diffusion and pseudodiffusion coefficient.

2.2 Levenberg–Marquardt Algorithm

Typically, non-linear regression of tracer kinetics models involves the minimization of the cost functional:

$$S(\vartheta) = \sum_{i=1}^N (y(i) - S_b(i, \vartheta))^2 = \|y - S_b\|_2 \tag{2}$$

where N denotes the number of points of curve, $y = [y(1), \dots, y(N)]^T$ represents the measured data and $\theta = (f; D_p; D_t)$. A widely used approach for estimating the optimum θ is the Levenberg–Marquardt (LM) algorithm, which is based on an approximation of the Hessian of $S_b(\theta)$. An exhaustive description of the algorithm is beyond the scope of the present paper and the reader is referred to [12, 13]. LM has shown to be a good solution for a number of non-linear regression problems and is implemented in a number of commercial packages. To start a minimization, the user has to provide an initial guess for θ . As the cost function surface could have many local minima in the parameter space, the algorithm is not guaranteed to converge to the global minimum unless the starting estimate is close to it. To improve the convergence of LM in DCE-MRI scenario, Ahearn et al. [13] proposed a multiple starting point approach which has been used in our simulation study.

2.3 VARIable PROjection Algorithm

Rearranging Eq. (1), the $S(b)/S_0 - e^{-bD_t}$ is the product of f and a non-linear function of D_t and D_p :

$$f(D_p; D_t; b) = S(b)/S_0 - e^{-bD_t} = f(e^{-bD_p} - e^{-bD_t}) \tag{2}$$

Letting $f(D_p; D_t; b)$, the cost functional becomes:

$$S(b)/S_0 - e^{-bD_t} = \|y - f(D_p; D_t; b)f\|_2 \tag{3}$$

Therefore, a separable non-linear LS algorithm known as Variable Projection (VarPro) can be used to calculate the diffusion parameters [12]. If we knew, the estimate of the non-linear parameters D_p and D_t , the estimate of the linear parameter f could be obtained by (solving a linear LS problem):

$$f = e^{-bD_t} + y + f(D_p; D_t; b)^+ \tag{4}$$

where $f(D_p; D_t; b)^+$ is the Moore–Penrose generalized inverse of $f(D_p; D_t; b)$. Therefore, a new cost functional can be constructed:

$$S(b)/S_0 - e^{-bD_t} = \|y - f(D_p; D_t; b)f(D_p; D_t; b)^+ y\|_2 \tag{5}$$

An exhaustive description of the algorithm is beyond the scope of the present paper and the interested reader is referred to [12]. Implementations of the VarPro algorithm are available in commercial packages.

2.4 Simulation Data

Computer simulated amplitude data were analyzed. The simulated data were generated from b values 0, 50, 100, 150, 200, 300, 400, 600, 800 s/mm² using Eq. (1). The ranges of values for f , D_p , D_t were (0.01, 0.30), (0.001, 0.030), (0.0001, 0.0025) with step 0.05, 0.005 and 0.0005, respectively. These values for b , f , D_p , D_t correspond to those used in a study of NLLS analysis of IVIM reported by Pekar et al. [14]. Gaussian noise was then added at simulated data varying in the range (1.7, 2.3) with step 0.1. To evaluate the performances of the different algorithms examined, the following procedure has been followed: $S_b(\theta)$ curves were simulated for several values of the parameters θ ; noise has been added on simulated curves; per each noisy curves parameters were estimated using all the algorithms. Per each simulated curve, 100 noisy curves have been obtained using random gaussian noise: correspondingly, per each algorithm and per each parameter, 100 estimates have been calculated (Monte Carlo Simulation). The value S_0 was imposed to 200, considering an estimation performed on real data. Each data set was analyzed using both NLLS and VarPro algorithms. For each simulation, we fitted the data using a single search start point (SSSP) in the middle of parameter space and multiple search start points (MSSP), i.e., the first starting point for each search was in the center of parameter space. Additional starting points were then defined at the center of each quadrant of parameter space. For each simulation, we fitted the data using each of the height points described above and selected the best fit of the height as the final result.

2.5 Goodness of Fit

Finding the best fit of a model to data involves the minimization of a merit function. This is usually described by the sum of the squares of the differences between the data points and the model estimated points, the residual sum of squares (RSS):

$$R^2 = \sum_{i=1}^N (y(i) - S_b(i))^2 \quad (2)$$

where N denotes the number of points of curve, y_b are the experimental data and the $S(b)$ are the data obtained by model fitting, a higher R^2 value corresponds to greater discrepancy (worse fit) between the data and the model.

3 Results

Figure 1 shows results of R^2 goodness-of-fit test for S_b curves obtained with LM algorithm with SSSP versus LM algorithm with MSSP (a) and VarPro algorithm

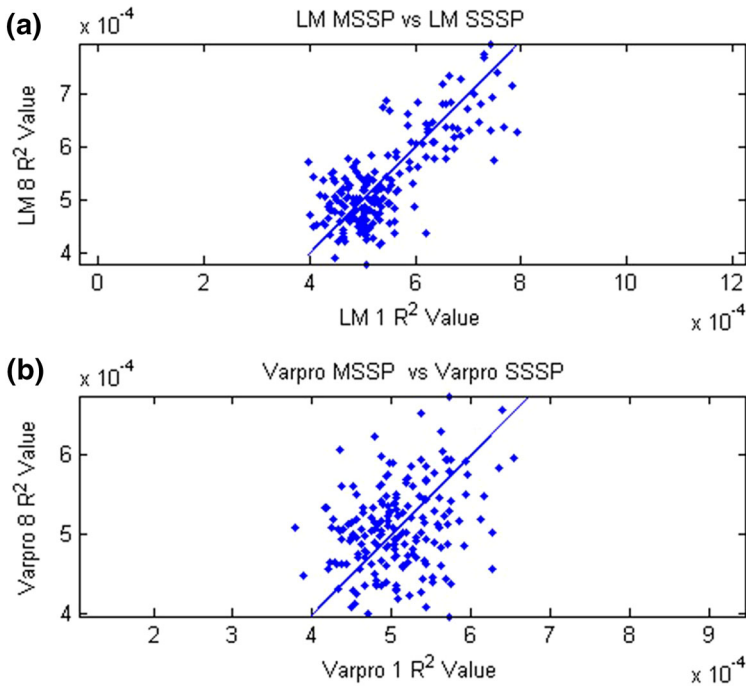


Fig. 1 R^2 goodness-of-fit test for S_b curves obtained with LM algorithm with SSSP versus LM algorithm with MSSP (a) and VarPro algorithm with SSSP versus VarPro algorithm with MSSP (b). Straight lines indicate equal goodness of fit. The points above lines denote cases in which LM or VarPro algorithm with SSSP gave better fit, points below lines denote cases in which LM or VarPro algorithm with MSSP gave better fit

with SSSP versus VarPro algorithm with MSSP (b). Straight lines indicate equal goodness of fit. The points above lines denote cases in which LM or VarPro algorithm with SSSP gave better fit; points below lines denote cases in which LM or VarPro algorithm with MSSP gave better fit. Both VarPro and LM with SSSP showed equivalent results than VarPro and LM with MSSP. Figure 2 shows results of R^2 goodness-of-fit test for S_b curves obtained with VarPro algorithm versus LM algorithm. Straight lines indicate equal goodness of fit. The points above lines denote cases in which LM algorithm gave better fit; points below lines denote cases in which VarPro algorithm gave better fit: (a) SSSP and (b) MSSP methods. VarPro algorithm showed a better fitting in comparison of LM algorithm both for SSSP and for MSSP.

Table 1 reports the comparison of LM and VarPro algorithm with SSSP versus MSSP: the number of simulated curves that showed better fitting of LM with MSSP versus SSSP was 55.6 % and the number of simulated curves that showed better fitting of VarPro with MSSP versus SSSP was 54.4 %.

Table 2 reports the comparison of LM and VarPro algorithm: the number of simulated curves that showed better fitting of VarPro versus LM with SSSP was 73.3 % and the number of simulated curves that showed better fitting of VarPro

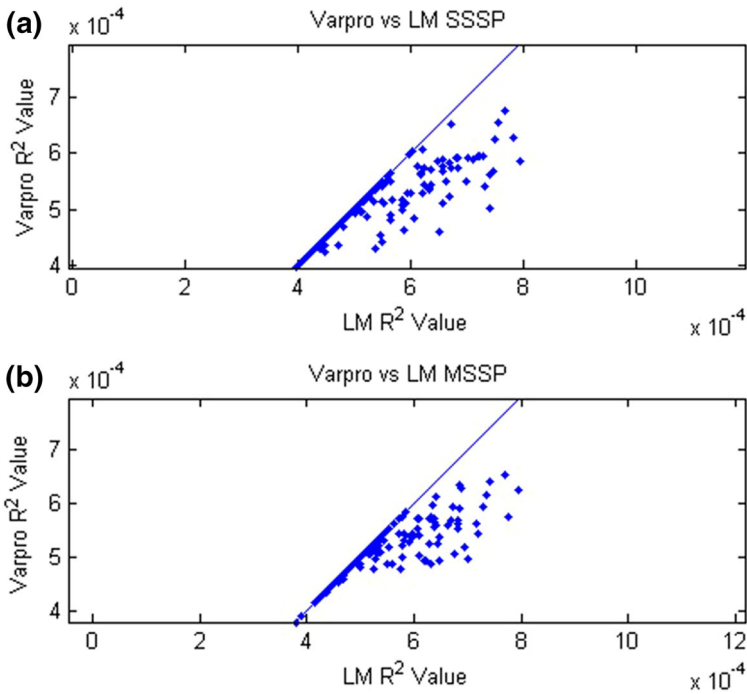


Fig. 2 R^2 goodness-of-fit test for S_b curves obtained with VarPro algorithm versus LM algorithm. *Straight lines* indicate equal goodness of fit. The *points above lines* denote cases in which LM algorithm gave better fit, *points below lines* denote cases in which VarPro algorithm gave better fit: **a** SSSP and **b** MSSP methods

Table 1 Comparison of LM and VarPro algorithm with SSSP versus MSSP: number of simulated curves that showed better fitting of LM and VarPro with SSSP versus MSSP using RSS

Fraction of curves $R_1^2 < R_2^2$	MSSP versus SSSP
LM	0.56
VarPro	0.54

n total number of simulated curves

Table 2 Comparison of LM and VarPro algorithm: number of simulated curves that showed better fitting of VarPro versus LM using RSS

Fraction of curves $R_1^2 < R_2^2/n$	VarPro versus LM
SSSP	0.73
MSSP	0.60

n total number of simulated curves

versus LM with MSSP was 60.0 %. The median \pm standard deviation R^2 values for LM with SSSP, LM with MSSP, VarPro with SSSP and VarPro with MSSP were, respectively: $3.3298e-004 \pm 4.7658e-005$, $3.3739e-004 \pm 5.7141e-005$; $3.2641e-004 \pm 2.4145e-005$, $3.2600e-004 \pm 2.4180e-005$.

4 Discussion and Conclusion

In this study, we evaluated the Variable Projection algorithms and the performance of two non-linear regression methods when single and multiple starting points were used to estimate diffusion parameters of intravoxel incoherent motion method for DW-MRI data analysis. Analysis was done on simulation data to which different amounts of Gaussian noise had been added. The performance of two non-linear regression methods was compared using the residual sum of squares in data fitting.

In a recent paper [15] were reported the results about a comparison of three different curve-fitting methods for intravoxel incoherent motion (IVIM) analysis in breast cancer: a direct estimation of D_t , D_p and f (Method 1); an estimation of D first and then D^* and f (Method 2); an estimation of D and f first and then D^* (Method 3). Among the three biexponential methods, Method 1 best described most of the pixels (63.20 % based on R^2). Their conclusions were that IVIM-derived parameters differ depending on the calculation methods.

Our group in a previous paper [16] evaluated the performances of different algorithms for tracer kinetics parameters estimation in breast Dynamic Contrast Enhanced-MRI. We considered four algorithms: two non-iterative algorithms based on impulsive and linear approximation of the Arterial Input Function, respectively; and two iterative algorithms widely used for non-linear regression (Levenberg–Marquardt, LM and Variable Projection, VarPro). The results of this study showed that the accuracy of all the methods depends on the specific value of the parameters. The methods are in general biased: however, VarPro showed small bias in a region of the parameter space larger than the other methods; moreover, VarPro showed better performances with respect to LM and non-iterative algorithms.

To the best of our knowledge, no paper is present in the research literature that reports the finding of VarPro algorithm to diffusion parameter estimation by DW-MRI data.

Our findings showed that both VarPro and LM with SSSP give equivalent results than VarPro and LM with MSSP. Moreover, VarPro algorithm showed a better fitting in comparison of LM algorithm both for SSSP and for MSSP. The number of simulated curves that showed better fitting of VarPro versus LM with SSSP was 73.3 % and the number of simulated curves that showed better fitting of VarPro versus LM with MSSP was 60.0 %.

Therefore, we conclude that the VarPro algorithm is superior to the LM algorithm for curve fitting in intravoxel incoherent motion method for DW-MRI data analysis.

Conflict of Interest All authors have no conflict of interest to be disclosed.

References

1. M. Sumi, M. Van Cauteren, T. Sumi, M. Obara, Y. Ichikawa, T. Nakamura, *Radiology* **263**(3), 770–777 (2012)
2. A.M. Chow, D.S. Gao, S.J. Fan, Z. Qiao, F.Y. Lee, J. Yang, K. Man, E.X. Wu, *J. Magn. Reson. Imaging* **36**(1), 159–167 (2012)
3. S. Rheinheimer, B. Stieltjes, F. Schneider, D. Simon, S. Pahernik, H.U. Kauczor, P. Hallscheidt, *Eur. J. Radiol.* **81**(3), e310–e316 (2012)
4. T.J. Re, A. Lemke, M. Klauss, F.B. Laun, D. Simon, K. Grünberg, S. Delorme, L. Grenacher, R. Manfredi, R.P. Mucelli, B. Stieltjes, *Magn. Reson. Med.* **66**(5), 1327–1332 (2011)
5. E.E. Sigmund, G.Y. Cho, S. Kim, M. Finn, M. Moccaldi, J.H. Jensen, D.K. Sodickson, J.D. Goldberg, S. Formenti, L. Moy, *Magn. Reson. Med.* **65**(5), 1437–1447 (2011)
6. H. Chandarana, V.S. Lee, E. Hecht, B. Taouli, E.E. Sigmund, *Invest. Radiol.* **46**(5), 285–291 (2011)
7. A. Lemke, F.B. Laun, M. Klauss, T.J. Re, D. Simon, S. Delorme, L.R. Schad, B. Stieltjes, *Invest. Radiol.* **44**(12), 769–775 (2009)
8. M. Iima, D. Le Bihan, R. Okumura, T. Okada, K. Fujimoto, S. Kanao, S. Tanaka, M. Fujimoto, H. Sakashita, K. Togashi, *Radiology* **260**(2), 364–372 (2011)
9. B.A. Hoff, T.L. Chenevert, M.S. Bhojani, T.C. Kwee, A. Rehemtulla, D. Le Bihan, B.D. Ross, C.J. Galbán, *Magn. Reson. Med.* **64**(5), 1499–1509 (2010)
10. D. Le Bihan, *Radiology* **249**(3), 748–752 (2008)
11. D. Le Bihan, E. Breton, D. Lallemand, M.L. Aubin, J. Vignaud, M. Laval-Jeantet, *Radiology* **168**(2), 497–505 (1988)
12. G.A.F. Seber, C.J. Wild, *Nonlinear Regression* (Wiley, New York, 1989)
13. T.S. Ahearn, R.T. Staff, T.W. Redpath, S.I. Semple, *Phys. Med. Biol.* **50**(9), N85–N92 (2005)
14. J. Pekar, C.T. Moonen, P.C. van Zijl, *Magn. Reson. Med.* **23**, 122–129 (1992)
15. S. Suo, N. Lin, H. Wang, L. Zhang, R. Wang, S. Zhang, J. Hua, J.J. Xu, *Magn. Reson. Imaging* (2014). doi:[10.1002/jmri.24799](https://doi.org/10.1002/jmri.24799)
16. R. Fusco, M. Sansone, A. Petrillo, *IJCCE* **5**(4), 2278–4209 (2014)