ADAPTIBILITY OF LEVENBERG-MARQUARDT ALGORITHM TO DIFFERENT SIGNAL MODELS

Miroslav ŠTĚPÁN, Doctoral Degree Programme (2)
Dept. of Telecommunications, FEEC, VUT
E-mail: stepan@feec.vutbr.cz

Martin MAREŠ, Doctoral Degree Programme (1)
Dept. of Telecommunications, FEEC, VUT
E-mail: mares@kn.vutbr.cz

Supervised by: Karel Bartušek

ABSTRACT

This article is focused on nonlinear signal processing using Levenberg – Marquardt (LM) algorithm [1] [2]. We introduce possibility of using different signal models to fit input data, which increases usability and performance of algorithm in different fields of study and research. In principle each nonlinear algorithm uses a definition of signal model and derivation of this function. This definition is advisable to declare in a separate function, which is called in iterations to compute cost function between model and data and evaluate new set of parameters. Basically this theory is not bounded to specific nonlinear algorithm, but it is applicable to arbitrary nonlinear algorithm, such as Gauss – Newton or pure Newton method. At the end of subscription some observations have been made on exponential signal fitting with linear function, fitting on noisy signal and the basis for nonlinear estimation of parameters from biomedical signals.

1 INTRODUCTION

LM algorithm is a widespread nonlinear tool for parameter estimation of signal data. It is often used for solving problems raised from linear least squares problem tasks and mainly as a basis for more sophisticated algorithms for biomedical signal processing such as VARPRO or AMARES as well [3]. LM is a powerful tool for signal processing not bounded to a specific research discipline. In biomedicine it handles with signals of Nuclear magnetic resonance (NMR), we can find it in neural network domain processing or everywhere the optimization process is needed. In biomedicine is often used Lorentz (Gauss, Voight) model to fit signal or simply exponential model for special cases [4]. Due to this broad use, naturally different description of handled signals is needed. It makes sense to work on adaptability of LM to different signal models.

2 BASIC THEORY

Observed signals in biomedicine can be well modeled using Lorentz function [5], which describes exponentially damped complex sinusoids:

$$M(\overline{A},\overline{\alpha},\overline{\omega},\overline{\varphi},t) = \sum_{i=1}^{k} A_i.\exp(\alpha_i t).\exp(j\omega_i t + \varphi_i)\,, \tag{1}$$

where $\boldsymbol{A}$ stands to amplitudes, $\boldsymbol{\alpha}$ represents damping factors, $\boldsymbol{\omega}$ represents an angular frequencies and $\boldsymbol{\varphi}$ means initial phases. In signal we can find k harmonical components. To simplify annotation it is convenient to pick components related to same harmonical component to one vector $\boldsymbol{x}$. So we search a set of $\boldsymbol{x}_k$ vectors to describe our signal.

Philosophy of Levenberg – Marquardt (LM) algorithm is based on solving lest squares problem, which is defined as:

$$F(\overline{x},t) = \frac{1}{2}\sum_{i=1}^{m}(y_i - M(\overline{x},t_i))^2 = \frac{1}{2}\left\| f(\overline{x})\right\|^2 = \frac{1}{2} f(\overline{x})^T f(\overline{x})\,, \tag{2}$$

where $F(\boldsymbol{x})$ is cost function, and expression to be squared is called residuum. Residuum represents the difference between signal data $\boldsymbol{y}$ and model M, which is dependent on set of estimated parameters $\boldsymbol{x}$, and time vector $\boldsymbol{t}$ of course. Dividing by 2 is only for convenience and has no effect on solution. In absence of noise and correct fitting model it is possible to reach negligible or zero residuum (as well as cost function), which is the goal of algorithm procedure. If we mark residuum as $\boldsymbol{f}$ function, for close neighborhood we can write using Taylor expansion for linear approximation (accepting only first two elements):

$$f(\overline{x}+\overline{h}) \cong f(\overline{x}) + J_f(\overline{x})\overline{h}\,, \tag{3}$$

where vector $\boldsymbol{h}$ determines direction to new set of parameters in next iteration. $\boldsymbol{J}$ matrix represents Jacobean, which contains the first partial derivatives of the residual components:

$$(J_f(\overline{x}))_{ij} = \frac{\partial f_i}{\partial x_j}(\overline{x})\,. \tag{4}$$

All nonlinear algorithms trying to secure descent direction $\boldsymbol{h}$ on cost function $\boldsymbol{F}$ to solution $\boldsymbol{x}^*$, which minimizes it. Change to new set of parameters is reflected on cost function using eq.(1), (2) and Taylor expansion for linear approximation again:

$$F(\overline{x}+\overline{h}) \cong \frac{1}{2} f^T f + h^T J_f^T f + \frac{1}{2}\overline{h}^T J_f^T J_f \overline{h} = F(\overline{x}) + \overline{h}^T J_f^T f + \frac{1}{2}\overline{h}^T J_f^T J_f \overline{h}\,. \tag{5}$$

It is not intended to explain detailed theory about LM algorithm, which is also not necessary, lot of contributions has been published [1] [2] [5]. Set of equations is used to demonstrate requirements for a model. From eq. (1) and (3) is obvious that the model is included in residuum and Jacobean both, so it is necessary to adjust definition of a model function there.

3 USING OF DIFFERENT SIGNAL MODULES

LM uses definition of model in separate function, which can looks like eq.(6):

$$M = \begin{bmatrix} f(x_1,t_1) & \cdots & f(x_k,t_1) \\ \vdots & & \vdots \\ f(x_1,t_n) & \cdots & f(x_k,t_n) \end{bmatrix}, \qquad (6)$$

it is obvious, that estimated parameters create columns. Of course in each iteration are parameters replaced and new fitting curve is obtained. Matrix form for derivation of fitting function can look like this formula:

$$M' = \begin{bmatrix} f'(x_1,t_1) & \cdots & f'(x_k,t_1) \\ \vdots & & \vdots \\ f'(x_1,t_n) & \cdots & f'(x_k,t_n) \end{bmatrix}. \qquad (7)$$

Derivation of fitting curve is necessary to evaluate a Jacobian., which is mainly used for searching descend direction to satisfy convergence criteria.

3.1 ADAPTATION LM TO LINEAR FITTING CURVE

Linear model of fitting curve is very simple. This group of signals related to nonlinear fitting tasks produces the only one linear curve, because the sum of linear estimated curves always produces one curve, no matter of initial number of fitting curves:

$$f_{est}(k,q,t) = \sum_{i=1}^{n} k_i.t + q_i = k'.t + q'. \qquad (8)$$

Interpretation of the final fitting curve is to approximate data in least square sense. This result secures the smallest quadratic difference between linear model and data as is shown on the left of obr.1.

3.2 ADAPTATION LM TO EXPONENTIAL FITTING CURVE

This group of tasks is very common, often used and one of the first solved nonlinear problems. It rises from linear least squares tasks, which reflect solving initial amplitude as linear parameter and damping (omega in argument) factor as nonlinear parameter building and determining cost function F. Estimating parameters of exponentials in signal is helpful in biomedicine to determine chemical components in sample. Estimation of exponential curve from noisy signal is showing on obr.1 on the right.

3.3 ADAPTATION LM TO BIOMEDICAL MODELS

Means to fit observed data with Lorenz (Gauss, Voight) model from eq (1).This is challenging task still in progress [6], lot of modifications and improved sophisticated algorithms have raised [7]. To this account simulations have been observed and are outlined on obr.2. On the left are plotted original data consisted of four decayed sinusoids. For fitting procedure Lorentz function has been chosen and applied successfully, as is obvious from obr.2. on right, where residuum is plotted and is negligible in –12 order.

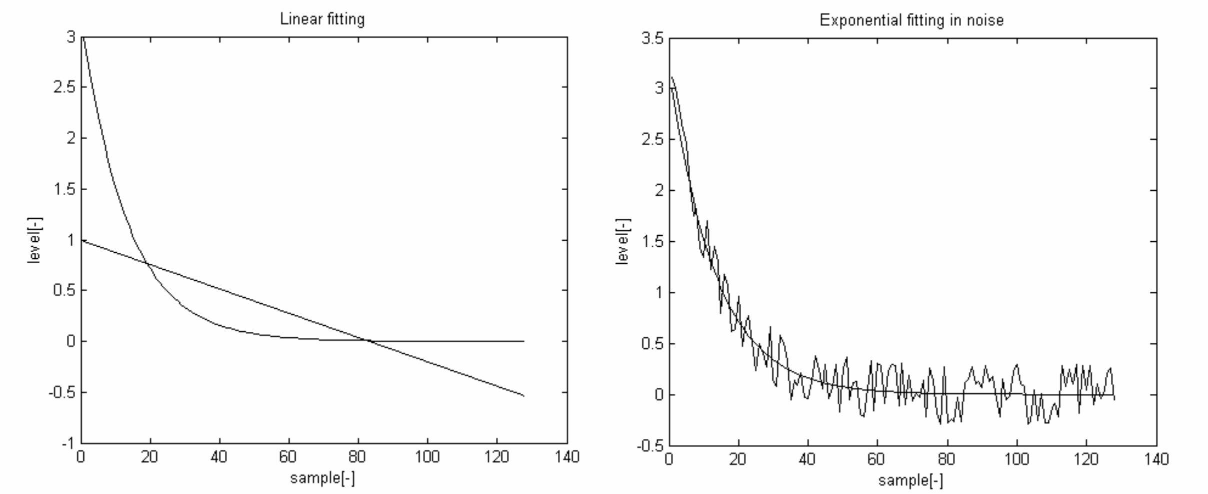

Fig. 1: *Simple fitting tasks, fitting ideal single exponential curve with linear curve on left, fitting with adequate model of noisy data on the right*

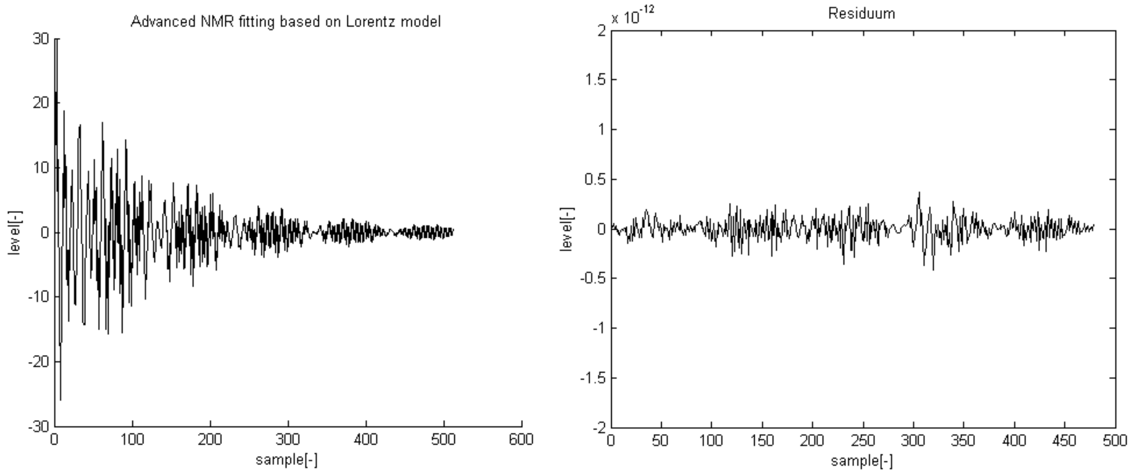

Fig. 2: *Biomedical data containing four decayed sinusoids and fitted with Lorentz model on the left, residuum on the right*

Fitting biomedical signals leads to quantitation, which means converting estimated parameters to biochemical quantities. This processing leads to efficient tumor detections, brain tissues and heart abnormalities.

4 CONCLUSION

This article was aimed to increase the usability of nonlinear algorithms, namely Levenberg – Marquardt [8]. It has been outlined that using proper definition of signal model allow use in different fields of study, starting with optimization techniques and continuing to biomedical applications, where nonlinear fitting is the most used. For speech processing certain constrains exist, there is difficult use of algorithm in real time applications. To secure the reusability of algorithm in advance separate implementation of fitting function and Jacobian is recommended.

ACKNOWLEDGEMENTS

The paper has been prepared as a part of the solution of FRVŠ project No. 1644/G1/2003 named "Implementace nelineárních algoritmů pro odhad parametrů signálů".

REFERENCES

[1] Marquardt, D.: An Algorithm for Least-Squares Estimation of Nonlinear Parameters, 1963, SIAM Journal Applied Math, Vol. 11, pp. 431-441.

[2] Levenberg, K.: A Method for the Solution of Certain Problems in Least-Squares, 1944, Quarterly Applied Math. 2, pp. 164-168.

[3] Vanhamme, L., Boogaart, A., Huffel, S.: Improved Method for Accurate and Efficient Quantification of MRS Data with Use of Prior Knowledge, 1997, J. of Magn. Resonance, pp.35-43.

[4] Nielsen, H. B.: Damping Parameter in Marquardt's Method, 1999, IMM, DTU report IMM-REP-1999-05, document available at http://www.imm.dtu.dk/~hbn/publ/TR9905.ps.Z.

[5] Hong, Y.: Signal Processing for Magnetic Resonance Imaging and Spectroscopy, 2002, ISBN 0-8247-0653-6.

[6] Kumaresan, R., Tufts, D. W.: Estimating the parameters of exponentially damped sinusoids and pole-zero modeling in noise, 1982, IEEE Trans. Acoust., Speech, Signal Process., ASSP-30:833-840.

[7] Li, G., Madsen, K.: Robust Nonlinear Data Fitting, 1988, Numerical analysis, Longman Scientific and Technical, London.

[8] Nielsen, H. B.: Separable Nonlinear Least Squares, 2000, IMM, DTU report IMM-REP-2000-01, document available at http://www.imm.dtu.dk/~hbn/publ/TR0001.ps.Z.