Advanced Studies in Medical Sciences, Vol. 5, 2017, no. 1, 63 - 69
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/asms.2017.784

Back Propagation Neural Network (BPNN) as Tracing Method to Trace Physionet EMG Signals: a Case Study

Jyh-Woei Lin

Department of Electrical Engineering, Southern Taiwan University of Science and Technology, No. 1, Nan-Tai Street, Yungkang Dist., Tainan City, Taiwan

Abstract

A new Modified Levenberg–Marquardt Algorithm (M-LMA) different from Levenberg–Marquardt Algorithm (LMA) was used to trace the Physionet EMG signals in tanning back propagation neural network (BPNN). M-LMA and LMA were simultaneously used to minimize back propagation errors in tanning BPNN to trace the Physionet EMG signals under the same learning rates of 0.1. Resulted shown M-LMA was better then LMA in training BPNN and could be as a better tracing method in this case.

Keywords: Modified Levenberg–Marquardt Algorithm (M-LMA); Levenberg–Marquardt Algorithm (LMA); Physionet EMG signals, Back Propagation Neural Network (BPNN)

1. Introduction

The physical concept of Artificial Neural Network (ANN) related to Biological Neuron Network was introduced in a study [13]. Back-propagation neural network (BPNN) [2, 6, 16, 17, 19, 21] will be performed by the classical Levenberg–Marquardt Algorithm (LMA) [7, 14] to decide desired minimum error in correcting back-propagation errors [1, 4, 8; 9, 10, 11, 15, 18, 19, 20]. LMA is a function of an independent variable. The computed processing of LMA to update weight and bias is not a parallel distribute processing (PDP) as a human neuron [3, 12, 13]. In this paper, BPNN is used to trace the Physionet EMG signals [23] with the modified Levenberg–Marquardt Algorithm (M-LMA).

2. Levenberg–Marquardt Algorithm (LMA)

A nonlinear real-valued function of f be composed of x. The parameter Y_t is defined as a target output [22] corresponding to the parameters x_t and y_t. That is, $Y_k = f(x_t)$. After approaching a nonlinear real-valued function, such as when applied to some simulated problems, the function f will generate the co-domain Y_k i.e., $Y_k = f(x_k)$, $\mathrm{k} = 1,2....$. The codomain is defined as the initial codomain when k is 1. Therefore, limits of $\delta x \to 0$ can be considered. Let x^o best satisfies the real-valued functional relation f that must be determined to minimize $\varepsilon^T \varepsilon$ with the error $\varepsilon = Y_k - Y_t$, where T denotes the transposition of a vector. For small $\|\delta x\|$, where $\|.\|$ denotes the Euclidean norm, a Taylor-series expansion leads to the approximation

$$f(x + \delta x) \approx f(x) + J\delta x; \tag{1}$$

where J is the Jacobian matrix, $\dfrac{\partial f(x)}{\partial x}$; using this nonlinear iterative method then produces a series of (x_1), (x_2), (x_3)…, which finally converge toward a local minimizer $Y^o = f(x^o)$; that is, optimized output corresponding to the optimized x^o and. Therefore $\|Y_k - f(x+\delta x)\| \approx \|Y_k - f(x) - J\delta x\| = \|\varepsilon - J\delta x\|$ is estimated. The element of δx is used for a nonlinear least squares problem whose minimum is attained when $\varepsilon - J\delta x$ is orthogonal to the column space of J and $J^T(\varepsilon - J\delta x) = \delta x(uI)$, where I is the identity matrix. Then,

$$\delta x(uI) + J^T J\delta x = J^T \varepsilon \tag{2}$$

Thus formula (2) becomes

$$(I_\mu + H_\mu)(\delta x) = J^T \varepsilon, \tag{3}$$

where $H_a = J^T J$ is the approximated Hessian matrix [5, 18] and $I_\mu = uI$, u is considered to be the learning rate.

3. Modified Levenberg–Marquardt Algorithm (M-LMA)

Now let a nonlinear real-valued function of f be composed of x and y, which belong to independent variables. The parameter Y_t is defined as a target output corresponding to the parameters x_t and y_t. That is, $Y_k = f(x_t, y_t)$. After approaching a true curve, such as when applied to some simulated problems, the function f will generate the co-domain Y_k i.e., $Y_k = f(x_k, y_k), \mathrm{k} = 1,2....$. The codomain is defined as the initial codomain when k is 1. If $\delta x \to 0, \delta y \to 0$, then it is implied that $\delta x \delta y \to 0$. On the contrary, if $\delta x \delta y \to 0$, it is uncertain if $\delta x \to 0, \delta y \to 0$. Therefore, limits of $\delta x \to 0, \delta y \to 0$ can be considered. Let x^o and y^o best satisfy the real-valued curve functional relation f that must be determined to minimize $\varepsilon^T \varepsilon$ with the error $\varepsilon = Y_k - Y_t$, where T denotes the transposition of a vector. For small $\|\delta x, \delta y\|$, where $\|.\|$ denotes the Euclidean norm, a Taylor-series expansion leads to the approximation

$$f(x+\delta x, y+\delta y) \approx f(x, y) + J\delta x\delta y; \tag{4}$$

where J is the Jacobian matrix, $\dfrac{\partial f(x,y)}{\partial x \delta y}$; using this nonlinear iterative method then produces a series of (x_1, y_1), (x_2, y_2), (x_3, y_3)..., which finally converge toward a local minimizer $Y^o = f(x^o, y^o)$; that is, optimized output corresponding to the optimized x^o and y^o. Therefore $\|Y_k - f(x+\delta x, y+\delta y)\| \approx \|Y_k - f(x,y) - J\delta x\delta y\| = \|\varepsilon - J\delta x\delta y\|$ is estimated. The elements of δx and δy are used for a nonlinear least squares problem whose minimum is attained when $\varepsilon - J\delta x\delta y$ is orthogonal to the column space of

J and $J^T(\varepsilon - J\delta x\delta y) = \delta x\delta y(uI)$, where I is the identity matrix. Then,

$$\delta x\delta y(uI) + J^T J\delta x\delta y = J^T\varepsilon \tag{5}$$

Thus formula (4) becomes

$$(I_\mu + H_\mu)(\delta x\delta y) = J^T\varepsilon, \tag{6}$$

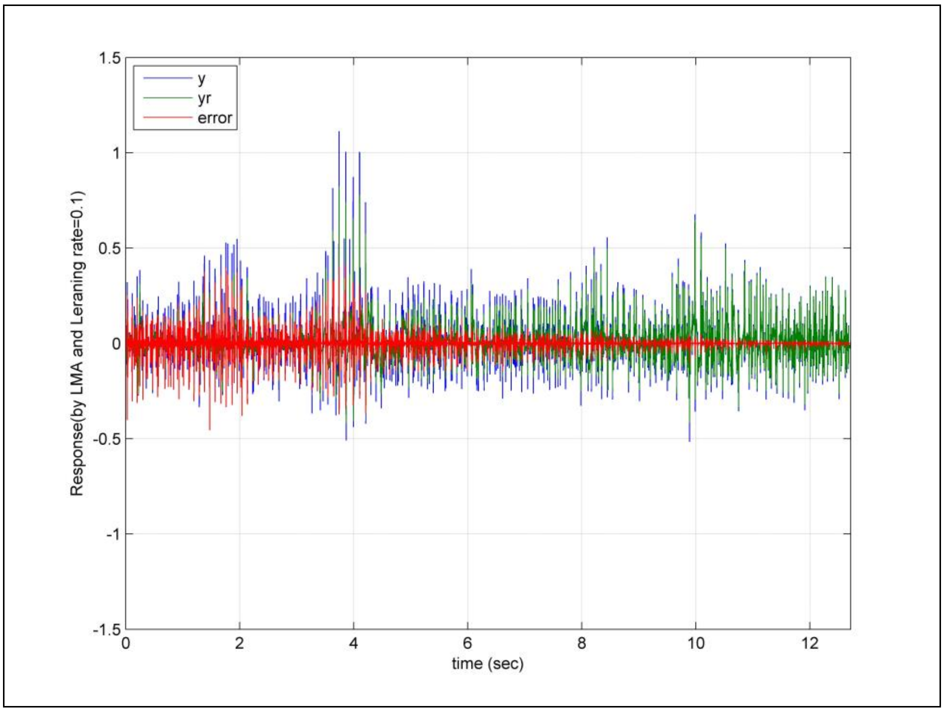
where $H_a = J^T J$ is the approximated Hessian matrix and $I_\mu = uI$, u is considered to be the learning rate when simultaneously adjusting or updating the

x, y values. Formulas (4), (5), and (6) belong to M-LMA. In this work, the parameters x and y are considered as the weight and bias of the BPNN [6, 16, 21]. Therefore, the updated processing of weight and bias in BPNN becomes a PDP, simulated more like the processing of a biological neuron using M-LMA.

4. Tests and Results

In this section, Physionet EMG signals will be traced by BPNN with both LMA and M-LMA of 2 hidden layers with 10 neurons in each layer to update the weight and bias with the same learning rates of 0.1 to minimize back propagation Errors.

Figure.1 shows the results of LMA and M-LMA. When M-LMA was used to the Physionet EMG signals and then the traced results of M-LMA were better than using LMA.



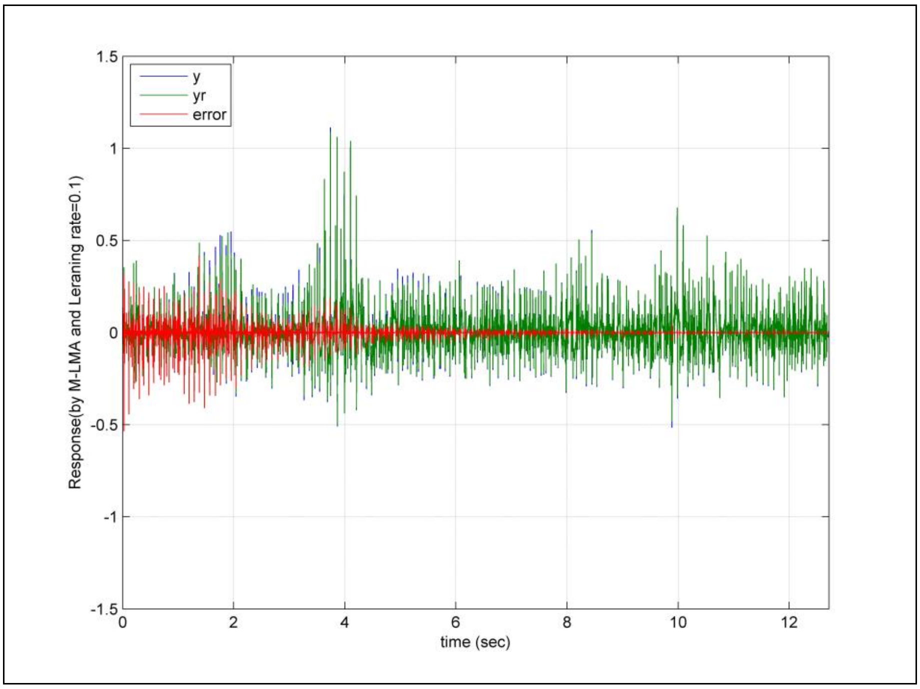


Fig.1 This figure shows the traced results of LMA and M-LMA to the Physionet EMG signals with the same learning rate of 0.1. The traced errors of M-LMA are smaller.

5. Conclusion

The M-LMA is better than LMA to trace the Physionet EMG signals with the same learning rate of 0.1. The M-LMA has been shown smaller traced errors in training BPNN in this case.

Acknowledgements. The authors are grateful to the support of Department of Electrical Engineering, Southern Taiwan University of Science and Technology.

References

[1] C. Alippi, *Intelligence for Embedded Systems: A Methodological Approach*, Springer International Publishing, 2014.
https://doi.org/10.1007/978-3-319-05278-6

[2] A, Bird, Kuhn, Nominalism, and Empiricism, *Philosophy of Science*, **70** (2003), no. 4, 690–719. https://doi.org/10.1086/378860

[3] D. Ferrier, *The Functions of the Brain*, London: Smith, Elder and Company, 1876. https://doi.org/10.1037/12860-000

[4] S. Finsterle and M. B. Kowalsky, A truncated Levenberg–Marquardt algorithm for the calibration of highly parameterized nonlinear models, *Computers & Geosciences*, **37** (2011), no. 6, 731-738.
https://doi.org/10.1016/j.cageo.2010.11.005

[5] X. Fu, S. Li, M. Fairbank, D. C. Wunsch, E. Alonso, Training Recurrent Neural Networks With the Levenberg–Marquardt Algorithm for Optimal Control of a Grid-Connected Converter, *IEEE Transactions on Neural Networks and Learning Systems*, **26** (2015), no. 9, 1900-1912.
https://doi.org/10.1109/tnnls.2014.2361267

[6] K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biological Cybernetics*, **36** (1980), no. 4, 193–202. https://doi.org/10.1007/bf00344251

[7] M. T. Hagan and M. B. Menhaj, Training feedforward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, **5** (1994), no. 6, 989-993. https://doi.org/10.1109/72.329697

[8] M. T. Hagan, H. B. Demuth, M. Beale, *Neural Network Design*, PWS Publishing Co. Boston, MA, USA, 1997.

[9] F.M. Ham and I. Kostanic, *Principles of Neurocomputing for Science and Engineering*, McGraw-Hill, 2001.

[10] S. Haykin, *Neural Networks and Learning Machines*, (3rd Edition), Prentice Hall, 2008.

[11] I. C. F. Ipsen, C. T. Kelley and S. R. Pope, Rank-Deficient Nonlinear Least Squares Problems and Subset Selection, *SIAM J. Numer. Anal.*, **49** (2011), no. 3, 1244–1266. https://doi.org/10.1137/090780882

[12] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Quarterly of Applied Mathematics*, **2** (1944), 164–168.
https://doi.org/10.1090/qam/10666

[13] J. W. Lin, Artificial Neural Network Related to Biological Neuron Network: A Review, *Advanced Studies in Medical Sciences*, **5** (2017), no. 1, 55 – 62.
https://doi.org/10.12988/asms.2017.753

[14] D. W. Marquardt, An Algorithm for Least-Squares Estimation of Nonlinear

Parameters, *Journal of the Society for Industrial and Applied Mathematics*, **11** (1963), no. 2, 431–441. https://doi.org/10.1137/0111030

[15] I. Manolis and A. Lourakis, A Brief Description of the Levenberg-Marquardt Algorithm Implemened by levmar, Institute of Computer Science, Foundation for Research and Technology – Hellas (FORTH), Vassilika Vouton, P.O. Box 1385, GR 711 10, Heraklion, Crete, Greece, 2005.

[16] W. McCulloch and W. Pitts, A Logical Calculus of Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics*, **5** (1943), no. 4, 115–133. https://doi.org/10.1007/bf02478259

[17] M. Mohri, A. Rostamizadeh and A. Talwalkar, Foundations of Machine Learning, MIT Press, 2012, pp.432, ISBN: 9780262018258

[18] L. S. H. Ngia and J. Sjoberg, Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm, *IEEE Transactions on Signal Processing*, **48** (2000), no. 7, 1915-1927. https://doi.org/10.1109/78.847778

[19] J. Pickies, Ground Truth, *The Social Implications of Geographic Information Systems*, Guilford Press, 1994.

[20] R Pourbagher and S. Y. Derakhshandeh, Application of high-order Levenberg–Marquardt method for solving the power flow problem in the ill-conditioned systems, *IET Generation, Transmission & Distribution*, **10** (2016), no. 12, 3017-3022. https://doi.org/10.1049/iet-gtd.2016.0064

[21] N. Rochester, J.H. Holland, L.H. Habit, and W.L. Duda, Tests on a cell assembly theory of the action of the brain, using a large digital computer, *IEEE Transactions on Information Theory*, **2** (1956), no. 3, 80–93. https://doi.org/10.1109/tit.1956.1056810

[22] D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*, Vol. 1, Cambridge, MA, MIT Press, 1986.

[23] A. Salman, E. G. Allstot, A. Y. Chen, A. M. R. Dixon, D. Gangopadhyay and D. J. Allstot, Compressive sampling of EMG bio-signals, *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, Rio de Janeiro (2011), 2095-2098. https://doi.org/10.1109/iscas.2011.5938011

Received: August 21, 2017; Published: October 7, 2017