On the Statistical Efficiency of LMS Algorithms

Bernard Widrow
ISL, Department of Electrical Engineering, Stanford University, Stanford CA, USA
e-mail: widrow@stanford.edu

The statistical efficiency of a learning algorithm applied to the adaptation of a given set of variable weights is defined as the ratio of the quality of the converged solution to the amount of data used in training the weights. Statistical efficiency is computed by averaging over an ensemble of learning experiences. A high quality solution is very close to optimal, while a low quality solution corresponds to noisy weights and less than optimal performance.

In this work, two gradient descent adaptive algorithms are compared, the LMS algorithm and the LMS/Newton algorithm. LMS is simple and practical, and is used in many applications worldwide. LMS/Newton is based on Newton's method and the LMS algorithm. LMS/Newton is optimal in the least squares sense. It maximizes the quality of its adaptive solution while minimizing the use of training data. Many least squares adaptive algorithms have been devised over the years, but no other least squares algorithm can give better performance, on average, than LMS/Newton.

LMS is easily implemented, but LMS/Newton, although of great mathematical interest, cannot be implemented in most practical applications. Because of its optimality, LMS/Newton serves as a benchmark for all least squares adaptive algorithms. When the eigenvalues of the input signal's autocorrelation matrix are all equal, LMS is identical to LMS/Newton. When the eigenvalues are not equal, then these algorithms behave differently.

The performances of LMS and LMS/Newton can be compared, and it is found that under many circumstances, both algorithms provide equal performance. For example, when both algorithms are tested with statistically nonstationary input signals, their average performances are equal. When adapting with stationary input signals and with random initial conditions, their respective learning times are on average equal. However, under worst-case initial conditions, the learning time of LMS can be much greater than that of LMS/Newton, and this is the principal disadvantage of the LMS algorithm.

The strong points of LMS are ease of implementation and optimal performance under important practical conditions. For these reasons, the LMS algorithm has enjoyed very widespread application. It is used in almost every modem for channel equalization and echo cancelling. Furthermore, it is related to the famous backpropagation algorithm used for training neural networks.

The full version of this paper can be found in [1].

References

[1] B. Widrow and M. Kamenetsky, "Statistical efficiency of adaptive algorithms," *Neural Networks*, vol. 16, no. 5–6, pp. 735–744, June–July 2003.

0-7803-8622-1/04/$20.00 ©2004 IEEE