



IDENTIFICAÇÃO DE SISTEMAS NÃO-LINEARES ATRAVÉS DE REDES NEURAIIS USANDO O MÉTODO DE LEVENBERG-MARQUARDT

Helder da Silva Camargos – Fundação de Ensino Superior de São João Del Rei – FUNREI-MG / Praça Frei Orlando 170 – Centro CEP: 36307-954 – São João Del Rei – MG
e-mail: heldercamargos@yahoo.com

Evaldo Khater – Fundação de Ensino Superior de São João Del Rei – FUNREI-MG / Praça Frei Orlando 170 – Centro CEP: 36307-954 – São João Del Rei – MG
e-mail: khater@funrei.br

***Resumo.** O propósito desse trabalho consiste na identificação de sistemas dinâmicos não lineares usando seis funções através do algoritmo da retropropagação do erro. Quatro métodos de otimização foram implementados comparando a convergência da rede. Foram eles: Gradiente Decrescente, Gradiente com Momento, Gradiente Conjugado e Levenberg-Marquardt. Os resultados obtidos foram de suma importância para que se pudesse definir o melhor método de otimização e aplicá-lo em sistemas e redes dinâmicas. Através de uma comparação entre os dados obtidos foi possível perceber que a eficácia dos métodos depende da função que está sendo identificada.*

***Palavras – Chaves:** Identificação, Redes Neurais, Otimização, Levenberg-Marquardt.*

1. INTRODUÇÃO

A visão moderna de redes neurais começou em 1940 com o trabalho de Warren McCulloch e Walter Pitts, que mostraram que redes artificiais de neurônios poderiam, em princípio, computar qualquer função lógica ou aritmética. Eles descreveram o cálculo lógico de redes neurais que uniam os estudos da neurofisiologia e a lógica matemática. O trabalho deles é conhecido como a origem do campo de trabalho em redes neurais (Haykin, 1999).

A primeira aplicação prática de redes neurais artificiais foi proposta por Frank Rosenblatt em 1958 que construiu uma rede de perceptrons e demonstrou que uma rede com neurônios tipo lógicos e com conexões modificáveis poderiam ser treinadas para classificar certos conjuntos de padrões. Neste trabalho foi proposto também um algoritmo para ajuste de pesos para o perceptron e provou sua convergência quando os padrões são linearmente separáveis (Khater, 1998).

Em 1982, Hopfield usa a idéia de uma função de energia para formular um novo caminho de entendimento do desempenho computacional por redes recorrentes com conexões sinápticas simétricas. Seu modelo apresentado foi caracterizado como uma coleção de neurônios com suas correspondentes unidades de atraso formando um sistema múltiplo de retroalimentação (Hagan, 1996).

O uso de redes neurais artificiais na identificação de sistemas dinâmicos possui as seguintes vantagens: (i) habilidade de tratar sistemas não lineares: relevante nas aplicações em identificação de sistemas dinâmicos e classificação de padrões; (ii) tolerância a falhas: o conhecimento é distribuído pela RN, mais que em uma simples localização de memória. Uma parte das conexões pode estar inoperante, sem mudanças significativas no desempenho de toda a RN; (iii) adaptabilidade: capacidade da RN em auto-ajustar-se. Os aspectos de aprendizado, auto-

organização, generalização e treinamento estão intrinsecamente ligados a esta característica; (iv) aprendizado: uma RN pode modificar seu comportamento em resposta ao ambiente (Kovacs, 1996).

As identificações de sistemas dinâmicos não lineares, usando o algoritmo de retropropagação do erro, foram desenvolvidas aplicando-se métodos de otimização que pudessem melhorar a convergência da rede proporcionando uma melhor acurácia aos resultados. O dinamismo aplicado a arquitetura da rede foi desenvolvido realimentando-a à cada iteração e os resultados permitiram uma comparação com a rede estática. Tais resultados foram de suma importância para que se pudesse definir o melhor método de otimização e aplicá-lo em sistemas e redes dinâmicas. Através de uma comparação entre os dados obtidos foi possível perceber que a eficácia dos métodos depende da função que está sendo identificada.

2. MATERIAL E MÉTODO UTILIZADO

Para identificação de sistemas não-lineares usando o algoritmo da retropropagação do erro, fez-se, primeiramente, o estudo da influência dos parâmetros da rede neural artificial através de sucessivos treinamentos. Para que isso fosse possível, realizou-se os ajustes dos parâmetros da planta a ser identificada. Foram selecionadas 6 funções não lineares de acordo com seu modo crescente de complexidade. Foram elas (Narendra,1991):

$$y(k+1) = 0,3y(k) + 0,6y(k-1) + f[u(k)] \quad (1)$$

onde

$$f(u) = 0.6\sin(\pi u) + 0.3\sin(3\delta u) + 0.1\sin(5\delta u)$$

$$y(k+1) = (y(k)y(k-1)(y(k)+2,5))/(1+y^2(k)+y^2(k-1)) + u(k) \quad (2)$$

$$y(k+1) = y(k)/(1+y^2(k)) + u^3(k) \quad (3)$$

$$y(k+1) = 0,8y(k) + (u - 0,8)u(u + 0,5) \quad (4)$$

$$y(k+1) = (y(k)y(k-1))(y(k)+2,5)/(1+y^2(k)+y^2(k-1)) + u(k) \quad (5)$$

$$y(k+1) = (5y(k)y(k-1))/(1+y^2(k)+y^2(k-1)+y^2(k-2)) + u + 0.8u(k-1) \quad (6)$$

onde $u = \sin\left(\frac{2\pi k}{25}\right)$ exceto na Eq. (3) onde $u = \sin\left(\frac{2\pi k}{250}\right) + \sin\left(\frac{2\pi k}{25}\right)$ (7)

Foram feitas simulações dos modelos matemáticos não lineares, acima descritos, na identificação de sistemas. A definição de uma arquitetura de rede adequada não é uma tarefa trivial. Surgiu, então, a necessidade do estudo de métodos numéricos de otimização que pudessem fornecer maior acurácia nos resultados. Dessa forma foram estudados quatro métodos: Gradiente, Gradiente com Momento, Gradiente Conjugado e Levenberg-Marquardt.

A Eq. (6), de complexidade maior, foi utilizada para a realização do treinamento da rede neural dinâmica.

Para uma melhor comparação da complexidade dos sistemas treinados, foi utilizada uma única arquitetura para as 6 redes. Adotou-se a seguinte configuração:

- 1 vetor de entrada (k = 1x100) com valores distribuídos no intervalo 1 a 100.
- 10 neurônios na camada de entrada
- 20 neurônios na camada oculta.
- 1 neurônio na camada de saída.

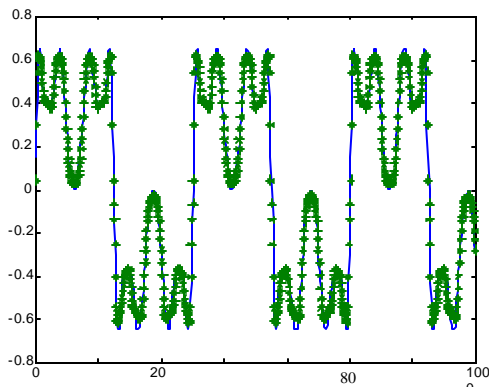
Como critério de parada usou-se um erro de 10^{-3} ou 5000 iterações para ambas funções.

Dada a entrada da rede neural, juntamente com o modelo de referência, o próximo passo foi realização da realimentação da rede neural a partir da resposta da mesma ao término de cada iteração definindo assim um sistema múltiplo de alimentação da rede neural dinâmica (Haykin, 1999).

3. RESULTADOS

Estão apresentados abaixo os resultados da identificação das funções estudadas. Para a Eq. (1), foram realizados 4 treinamentos utilizando os métodos de otimização Gradiente, Gradiente com Momento, Gradiente Conjugado e Levenberg-Marquardt. Para as demais funções, Eq. (2) à Eq. (6), de maior complexidade, foi realizado o treinamento utilizando somente o método de otimização Levenberg-Marquardt. Nos gráficos, a saída desejada da rede neural treinada está representada pela linha contínua, já a saída real está representada por asterisco. As figuras (1) à (4) representam a Eq. (1). As figuras (5) à (10) representam as demais equações como indicado nos gráficos.

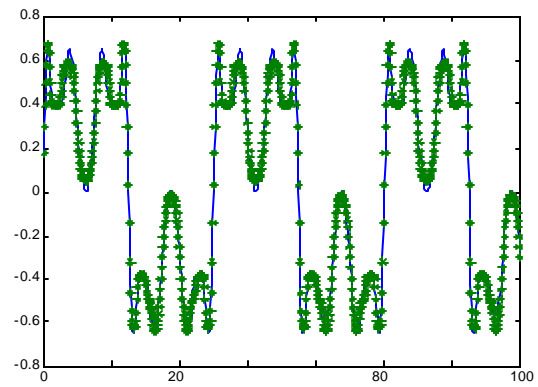
Saídas



k

Figura 1. Método Gradiente

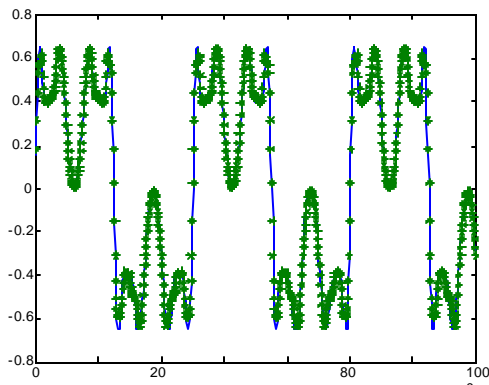
Saídas



k

Figura 2. Método Gradiente com Momento

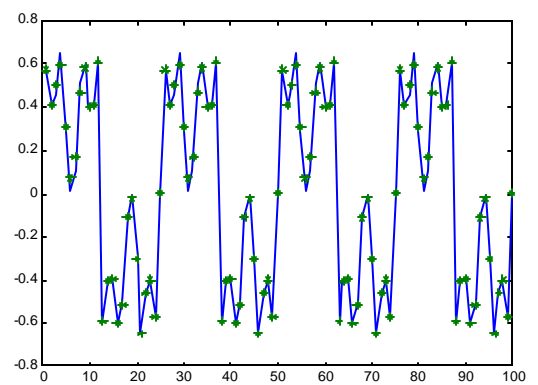
Saídas



k

Figura 3. Método Gradiente Conjugado

Saídas



k

Figura 4. Método Levenberg-Marquardt

LEGENDA: — saída desejada

*** saída real

Saídas

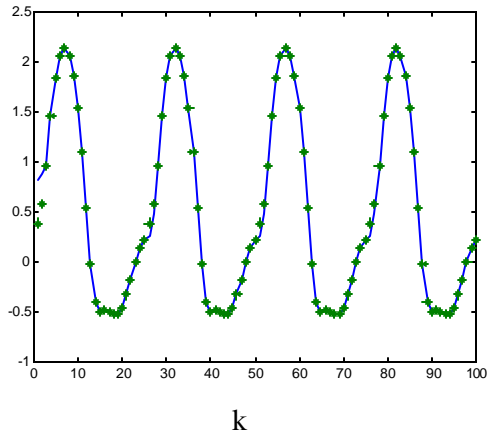


Figura 5. Identificação da Eq. (2)

Saídas

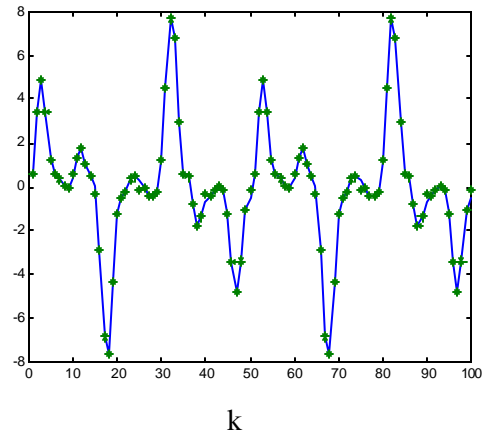


Figura 6. Identificação da Eq. (3)

Saídas

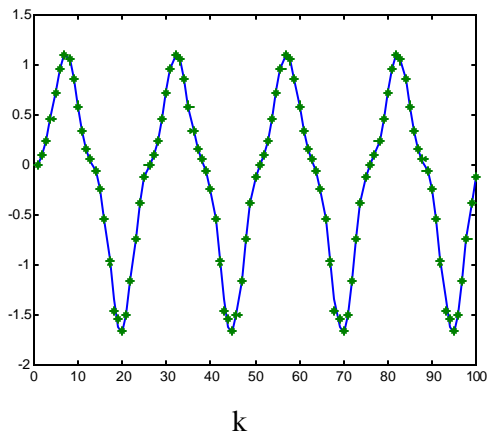


Figura 7. Identificação da Eq. (4)

Saídas

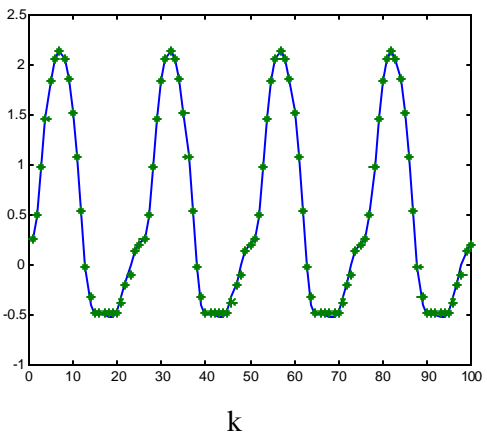


Figura 8. Identificação da Eq. (5)

Saídas

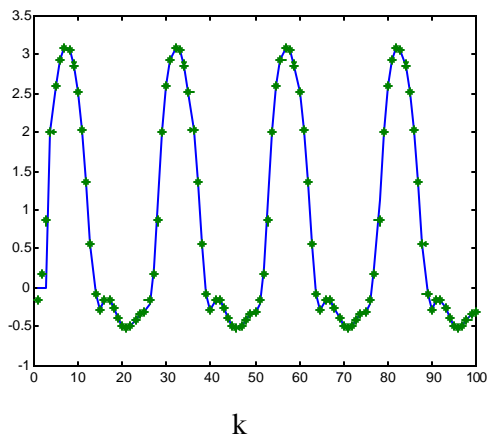


Figura 9. Identificação da Eq. (6)
(Rede Estática)

Saídas

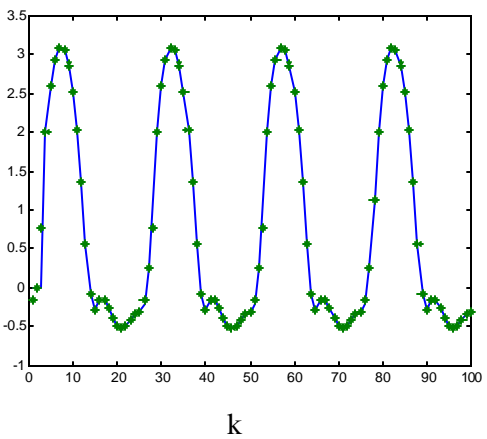


Figura 10. Identificação da Eq. (6)
(Rede Dinâmica)

LEGENDA: — saída desejada *** saída real

As tabelas abaixo resumem os principais resultados obtidos com a identificação das funções não lineares:

Tabela 3. Resultado numérico do treinamento das funções

	MÉTODO	ITERAÇÕES	TEMPO PROC.	ERRO
FUNÇÃO (1)	Gradiente	5000	32'15''	0,00108
	Gradiente Mom.	939	9'42''	0,00099
	Gradiente Conjug.	54	1'03''	0,00098
	Levenberg-Mar.	4	0'10''	0,00058
FUNÇÃO (2)	Levenberg-Mar.	152	4'42''	0,00099
FUNÇÃO (3)	Levenberg-Mar.	1919	56'51''	0,01450
FUNÇÃO (4)	Levenberg-Mar.	119	5'40''	0,00091
FUNÇÃO (5)	Levenberg-Mar.	5000	2h 27'30''	0,00937
FUNÇÃO (6)	Levenberg-Mar.	5000	9h 43'34''	0,01051

O resultado a seguir refere-se a rede neural dinâmica desenvolvida utilizando a Eq. (6):

Tabela 2. Resultado do treinamento da Eq. (6) através da realimentação da entrada

	MÉTODO	ITERAÇÕES	TEMPO PROC.	ERRO
FUNÇÃO (6)	Levenberg-Mar	5000	9h 25'59''	0,00901

4. DISCUSSÃO DOS RESULTADOS

Para a identificação dos sistemas propostos foram realizados os treinamentos das redes neurais utilizando os 4 métodos de otimização. Tais métodos foram aplicados na Eq. (1) sendo que nas demais utilizou-se somente o método de Levenberg – Marquardt. Observando os gráficos e as tabelas foi possível perceber que os resultados variam bastante de acordo com a função e com o método numérico de otimização.

Na identificação da Eq. (1) com o método do gradiente foram realizadas 5000 iterações num tempo de 32 min e 15seg e o erro estipulado não foi alcançado, como pode ser visto na tabela (1). Com isso, houve a necessidade de se usar um método mais eficaz. O resultado dado pelo método do gradiente com momento é parecido com o resultado alcançado pelo método do gradiente na identificação da mesma equação. Observando-se a curva da rede treinada e a da saída desejada na Fig. (1) e comparando-a com a Fig. (2) é possível perceber a semelhança entre as curvas. No entanto, o método do gradiente com momento apresenta uma ligeira vantagem no que diz respeito ao tempo, como pode ser visto na Tab. (1).

É na identificação dessa Eq. (1) que se percebe a vantagem do método gradiente conjugado em relação ao dois métodos já mencionados. Esse fato é confirmado pela Tab. (1) que mostra que o erro atingiu também 10^{-3} como era previsto num menor número de iterações e menor tempo.

O método de Levenberg-Marquardt na identificação da equação (1) mostrou-se o mais eficiente. Observando a identificação na Fig. (4) percebe-se que as curvas estão superpostas em quase todo o intervalo da função. Os resultados observados na Tab. (1) mostram que este método ofereceu o melhor desempenho na identificação da Eq. (1), reduzindo o tempo e o n° de iterações em relação aos outros métodos.

A seguir, realizou-se novamente, o treinamento da Eq. (6), porém dinamicamente. O resultado obtido mostrou que a rede convergiu para o erro estipulado e num menor intervalo de tempo ao contrário do que ocorreu com a mesma função treinada estaticamente como pode ser visto pela Tab. (2).

5. CONCLUSÃO

Redes neurais têm aplicações em problemas práticos de engenharia, em áreas como reconhecimento de padrões, processamento de sinais e sistemas de controle, sendo possível usá-las para resolver problemas industriais complexos.

Nessa pesquisa pode-se perceber a flexibilidade que essas redes possuem, podendo ser adaptadas de acordo com a complexidade da função. Para se fazer essa adaptação é necessário encontrar um equilíbrio entre a suposição e a lógica.

Verificou-se também que o método do gradiente apresenta bons resultados na identificação de funções simples. No entanto, quando se deseja implementar uma função mais complexa, deve-se buscar métodos mais robustos, que possam melhorar a convergência da rede neural treinada. Posteriormente as funções foram treinadas com o método de Levenberg-Marquardt e este mostrou-se o mais adequado para a implementação da função não linear, reduzindo o tempo e o n° de iterações. A rede neural dinâmica desenvolvida nesse trabalho, mostrou um menor custo operacional para funções mais complexas, como foi feito utilizando a Eq. (6).

Os resultados obtidos nesse trabalho proporcionaram incentivos para a continuidade da análise do Método Levenberg-Marquardt em outros modelos de identificação. O vasto campo das redes neurais artificiais proporcionam estudos avançados na busca do aperfeiçoamento na identificação e controle de sistemas dinâmicos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- Hagan, T. M., Demuth B. H., Beale, M. Neural Network Design. PWS Publishing Company, 1996.
- Haykin, S. Neural Networks: A Comprehensive Foundation. Prentice-Hall, 1999, 696p.
- Khater, Evaldo. Controle de vibração torcional em sistemas rotativos usando redes neurais multicamadas. Campinas: Universidade Estadual de Campinas (UNICAMP), 1998. 126p. Tese (doutorado).
- Kovács, L. Z. Redes Neurais Artificiais: Fundamentos e Aplicações. 2º edição; Collegium Cognitio, 1996.
- Narendra, S. K., Parthasaraty, K. Identification and Control of Dynamical Systems Using Neural Networks. IEEE Trans. on Neural Networks, v.1, n.1, march 1990.

TITLE: Identification of Nonlinear Systems by Neural Networks Using the Levenberg-Marquardt Method.

Helder da Silva Camargos – Fundação de Ensino Superior de São João Del Rei – FUNREI-MG / Praça Frei Orlando 170 – Centro CEP: 36307-954 – São João Del Rei – MG
e-mail: heldercamargos@yahoo.com

Evaldo Khater – Fundação de Ensino Superior de São João Del Rei – FUNREI-MG / Praça Frei Orlando 170 – Centro CEP: 36307-954 – São João Del Rei – MG
e-mail: khater@funrei.br

Abstract. The intention of this work consists of the identification of nonlinear dynamic systems using six functions through of the backpropagation algorithm. Four methods of optimization had been implemented comparing the convergence of the network. They had been they: Decreasing Gradient, Gradient with Moment, Gradient Conjugate and Levenberg-Marquardt. The gotten results had been of utmost importance so that if it could define the best method of optimization and apply it in systems and dynamic networks. Through a matching between the gotten data it was possible to perceive that the effectiveness of the methods depends on the function that is being identified.

Key-Words. Identification, Neural Networks, Optmization, Levenberg-Marquardt.