

Inteligência Computacional

Rede neural: Adaptive linear element(Adaline)

Gabriel Tambara Rabelo
Departamento de engenharia elétrica
Faculdade de Tecnologia - UnB
Brasília, Brasil
gtambararab@gmail.com

Abstract—Artificial neural networks like Adaline have the ability to identify dynamical systems. This identification is conditioned depending on certain design parameters of the identification model that are covered in the article, their typical characteristics of a sub-optimal identification are also presented.

Keywords: adaline; intelligence; artificial; neural; network.

Resumo—Redes neurais artificiais como a Adaline possuem a capacidade de identificação de sistemas dinâmicos. Essa identificação é condicionada dependendo de certos parâmetros de projeto do modelo de identificação que são abordados no artigo, e suas características típicas de uma identificação sub-ótima também são apresentadas.

Palavras-chave: adaline; inteligência; artificial; rede; neural.

I. INTRODUÇÃO

No contexto de redes neurais artificiais, o perceptron de Rosenblatt, criado por aquele que o nomeia, em 1957, abordou o aprendizado supervisionado, objetivando facilitar a obtenção de resultados nas redes neurais. Contudo, há um modelo de rede que também traz novidades e é pertinente para a história dos algoritmos deste modelo.

Diante desta ferramenta, o presente artigo detalha um estudo realizado visando esclarecer os fundamentos e questões pertinentes quanto ao uso da rede Adaline, bem como elucida as suas possíveis vulnerabilidades.

II. FUNDAMENTAÇÃO TEÓRICA

A. O semelhante Perceptron de Rosenblatt

O modelo de rede neural artificial denominado Perceptron de Rosenblatt foi um dos modelos desenvolvidos por Rosenblatt baseados nos modelos anteriores de neurônios, criados por McCulloch e Pitts, este neurônio era capaz de somar diversos sinais diferentes, cada qual com seu respectivo peso, e então tratar o resultado com uma função de ativação. O formato do neurônio pode ser observado pela Fig.[1].

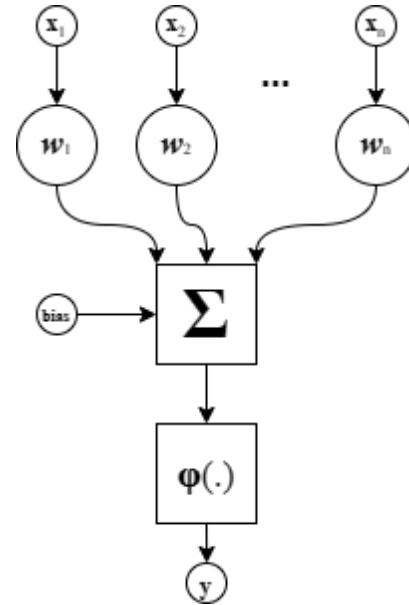


Figura 1: Modelo de neurônio de McCulloch e Pitts

Foi evidente a sua utilização como classificador para modelos lineares, contudo, essa limitação é razoavelmente problemática tendo em vista a característica da maioria dos problemas que envolvem identificação de sistemas na realidade, já que não são lineares.

Os estudos no desenvolvimento de modelos de redes neurais não pararam, tanto referente ao neurônio quanto referente à topologia da rede. Um dos modelos que se sobressaiu foi o modelo de Perceptron de Rosenblatt, que utiliza a ideia geral do neurônio de McCulloch, porém com a denominada Regra Delta, ou *Delta Rule*. Essa regra é responsável por adaptar os pesos de cada entrada do neurônio conforme um erro associado à saída do neurônio e o valor de referência esperado para uma entrada pré-determinada, ou seja, durante a etapa de treinamento do neurônio. As equações que bem descrevem essa dinâmica são apresentadas pelas Eqs.[1,2]. Aqui μ é definida como a taxa de aprendizado que pode ser regulada e δx o próprio erro a corrigir os pesos w . d seria a referência e y a saída real do sistema.

$$\delta = d - y \quad (1)$$

$$w_{novo} = w_{velho} + \mu \delta x \quad (2)$$

B. ADALINE: Adaptive linear element

Porém, em 1960, Widrow e Hoff retomaram a frente nas pesquisas sobre os modelos de neurônio e desenvolveram um modelo muito semelhante ao neurônio de Rosenblatt. Neste novo modelo, percebe-se a presença da *Delta rule* porém sob uma diferente forma. Ela então passa a tomar a forma de uma LMS (*Least Mean Square*), ou método dos mínimos quadrados, modelo matemático reconhecido como uma técnica de otimização através da soma de quadrados. As equações que descrevem este novo modelo são apresentadas pelas Eqs.[1,3,4]. Nesse cenário percebe-se que se utiliza o sinal de saída s do somador do neurônio como elemento para se gerar um erro ϵ , que também será parâmetro para a correção do peso w do neurônio.

$$\epsilon = d - s = d - \sum (wp + bias) \quad (3)$$

$$w_{novo} = w_{velho} + \frac{\mu \epsilon \delta x}{\sum x^2} \quad (4)$$

C. LMS: Método dos mínimos quadrados e Gradient Descent

Objetivando-se descobrir o comportamento de uma função arbitrária com base nos sinais de entrada, na referência, e nas suas saídas, utiliza-se o seguinte algoritmo baseado na variável estatística denominada esperança, de forma a minimizá-la.

$$E(e^2) = E(d - y)^2 = E(d - xw)^2$$

$$E(e^2) = E(d^2) - 2E(dx)w + wE(x^2)w$$

Neste cenário, objetiva-se minimizar os valores dos pesos, sendo estes as variáveis a serem otimizadas. O modelo geral de otimização busca o menor valor possível de modo que busca-se zerar a derivada com base na variável de otimização. Contudo, para matrizes não quadradas, a derivada, que exige a busca pela inversa da matriz, não é acessível. Portanto, faz-se necessário o uso da matriz pseudo-inversa nos casos que convierem.

Portanto, segue da derivada parcial que:

$$w_{novo} = w - c_k \nabla_w E(e^2) (E(x^2))^{-1}$$

Aqui, caso a variável c_k seja igual a meio, denomina-se o método de Newton. Porém, assumindo que a esperança do quadrado dos valores de entrada seja a matriz identidade, o sistema toma a forma de um novo modelo denominado método do gradiente, que vai buscar mínimos locais dentro do espaço onde a função for localizada. Nesse cenário, o algoritmo se resume para o modelo iterativo definido pela Eq.[5].

$$w_{k+1} = w_k + 2c_k ex \quad (5)$$

É evidente que, como um método iterativo e estocástico, o resultado ótimo nunca é alcançado, porém, ainda é capaz de

alcançar mínimos locais, que satisfaz a maioria dos contextos de aplicação de redes neurais artificiais.

D. Situação problema

Usando a ADALINE com um algoritmo de treinamento adaptativo, busca-se identificar uma função de transferência IIR de segunda ordem que varia no tempo. Sistemas discretos no tempo podem ser implementados de duas formas: "FIR- Resposta Finita ao Impulso ou "IIR- Resposta Infinita ao Impulso. Uma estrutura FIR não possui realimentação, fazendo com que a resposta a um impulso decaia para zero após um determinado número de atrasos (períodos de clock) equivalente à ordem do sistema FIR. Por outro lado, um sistema IIR possui realimentação, resultando em uma resposta ao impulso que tende a zero indefinidamente.

Neste artigo, pretende-se através de um estudo aproximar um sistema IIR de segunda ordem usando um filtro adaptativo FIR Adaline. A abordagem FIR assegura a estabilidade do sistema resultante, um requisito fundamental em sistemas adaptativos. No entanto, é importante destacar que uma taxa de aprendizagem inadequada também pode levar o sistema à instabilidade, o que será abordado.

A situação problema envolve a seguinte função de transferência $G(s)$, representada pela Eq.[6].

$$G(s) = \frac{1}{(s^2 + 0.2s + 1)} \quad (6)$$

Essa função é mantida até o instante de 80 segundos após o início da amostragem, quando a função toma a forma representada pela Eq.[7].

$$G(s) = \frac{3}{(s^2 + 2s + 1)} \quad (7)$$

III. LABORATÓRIO

Para realizar a identificação dos sistemas, é necessário um conjunto de sinais de entrada de ordem caótica suficiente grande para que o sistema de identificação possa entender todas as suas características para além do apresentado, se possível. Nesse contexto, foi utilizado um sinal de entrada binário pseudorandômico com faixas de frequência de 0 a 0.01 Hz e outro sinal de 0 a 0.05Hz. Os sinais podem ser observados pelas Figs.[2,3].

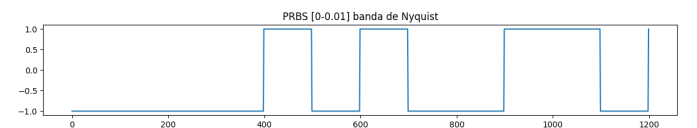


Figura 2: PRBS com frequência de Nyquist até 0.01Hz

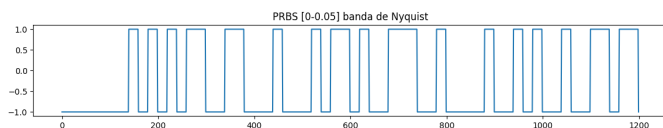


Figura 3: PRBS com frequência de Nyquist até 0.05Hz

De início, é observado o comportamento de uma simples discretização do sistema e sua resposta, via segurador de ordem zero. O resultado é demonstrado pela Fig.[4].

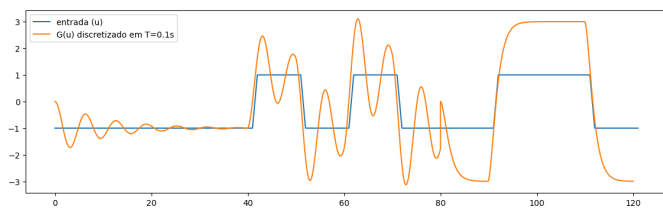


Figura 4: Resposta do sistema $G(s)$ discretizado

Utilizando o algoritmo Adaline para identificar o sistema, com a Regra de Delta, pode-se observar o resultado representado pela Fig.[5].

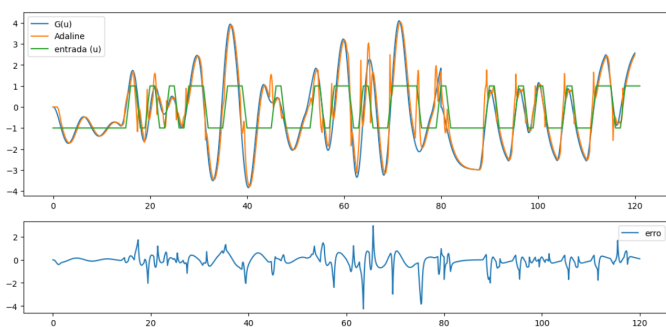


Figura 5: Resposta Adaline

Aqui pode-se perceber como o comportamento da saída acompanha muito bem o perfil da entrada. Este resultado não foi trivial de se encontrar tendo em vista que foram necessárias várias iterações entre variações das variáveis de entrada do sistema, sendo elas o coeficiente de aprendizado μ e os delays do sistema. Os valores que geraram tal resultado foram um delay de 8 e uma taxa de 0.05. Taxas maiores foram responsáveis por levar o sistema à instabilidade, bem como um valor muito grande de delay é responsável por aumentar consideravelmente o ruído percebido no sistema.

Também foi explorado o cenário com pesos fixos, o que levou o sistema à tomar um comportamento muito semelhante ao da entrada, porém com ganhos a serem corrigidos. O resultado pode ser observado pela Fig.[6].

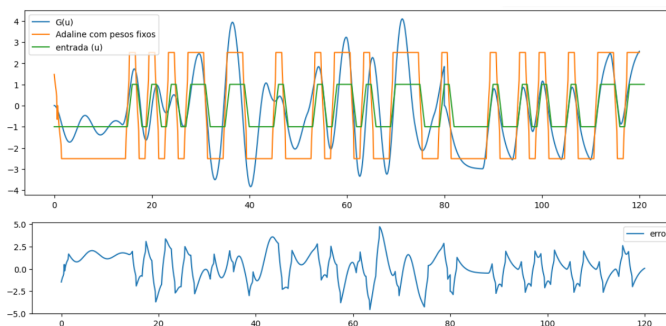


Figura 6: Resposta Adaline com pesos fixos

Por fim, foi realizada uma nova fixação de pesos com base em todos os dados amostrais de uma vez. Este experimento resultou no cenário onde a saída do sistema replica com boa característica a entrada do sistema. O resultado é percebido pela Fig.[7]. Também percebe-se que ao final de cada degrau do sinal de entrada, observa-se um pequeno desvio de sinal da resposta ADALINE, provavelmente causado pelo seu comportamento inesperado próximo a discontinuidades, contornada pelo modelo de pseudoinversa.

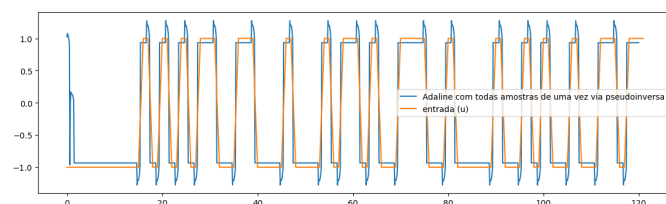


Figura 7: Resposta Adaline com pseudoinversa

IV. CONCLUSÃO

As redes neurais artificiais possuem diversas aplicações, dentre elas a identificação de sistemas. Foi discutido neste artigo a aplicação de uma rede específica denominada ADALINE, capaz de utilizar de da busca por *gradient descent* pela otimização dos pesos de uma rede neural baseada no modelo de neurônio de Rosenblatt. O experimento de avaliação de identificação de um sistema de resposta ao impulso infinito IIR, levou a conclusões positivas sobre o uso da ADALINE, porém questões como ruído e pequenos distúrbios no sinal foram observados.

Para fins de trabalhos futuros no tópico, recomenda-se uma exploração de diferentes ferramentas como o Matlab, tendo em vista que o Python foi incapaz, até o presente momento, de oferecer ferramentas adequadas para o manuseio de sistemas discretos tão bem quanto de sistemas contínuos, o que concederia uma grande gama de observações interessantes que poderiam ter sido feitas.

REFERÊNCIAS

- [1] S. Haykin, "Neural Networks and Learning Machines," 3rd ed., 2009.
- [2] Adaline neural networks: the origins of gradient descent Neural Networks Series — Chapter 2 - Adriano Vinhas