

Exercício Computacional 2: Adaline

Aluno: *Pedro Vinícius A. B. de Venâncio*

Disciplina: *Redes Neurais Artificiais (EEE950)* – Professor: *Antônio de Pádua Braga*
Data: *Setembro de 2019*

Introdução

Este relatório tem como objetivo desenvolver¹ a rede linear do tipo *adaline*, tendo em vista sua importância histórica e suas diversas aplicações.

Adaline

O modelo *adaline* tem grande semelhança com o *perceptron*, sendo, portanto, adaptado à problemas de regressão. Sua principal diferença é que a ativação do neurônio ocorre por meio da função identidade $f(z) = z$, isto é, sua saída é exatamente a soma ponderada do vetor de entradas $\hat{y} = \sum_j w_j x_j$ [1].

```
# Activation function
activation <- function(X, W){
  z = t(X %*% t(W))
  return(z)
}
```

Problema de Regressão

Um problema de regressão consiste em aproximar uma função de mapeamento das variáveis de entrada para uma variável de saída contínua. Como um modelo para esse fim prevê um valor real qualquer, o desempenho do modelo é comumente avaliado com cálculos de erros das previsões.

Assim como no problema de classificação, deseja-se avaliar o modelo em dois problemas: um primeiro com base de dados sintética e outro com base de dados reais. Os dados sintéticos foram gerados pela equação linear $y = x + b + \mathcal{N}(0, \sigma^2)$, onde tem-se um ruído Gaussiano com variância σ^2 . Os parâmetros da reta foram definidos como $m = 5$ e $n = 10$ e a variância do ruído como $\sigma^2 = 25$.

```
# Straight line
slope <- 5
intercept <- 10
Y <- slope*X[,2] + intercept

# Adding white Gaussian noise to linear data
```

¹Todas as implementações foram feitas em linguagem R.

```
noise <- rnorm(n_samples, mean = 0, sd = 25)
Y <- Y + noise
```

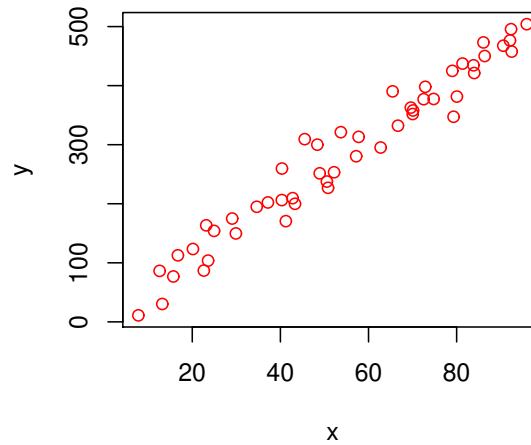


Figura 1: Base sintética de dados lineares.

Os parâmetros utilizados para o treinamento do modelo foram $\eta = 0.001$ e $max_epochs = 25000$, sendo a divisão dos dados similar, em proporção, a do problema anterior. Através da função linear gerada (Figura 2), corrobora-se que o *adaline* foi capaz de se ajustar aos dados através da Regra Delta, obtendo um erro quadrático médio de 6.67×10^{-2} .

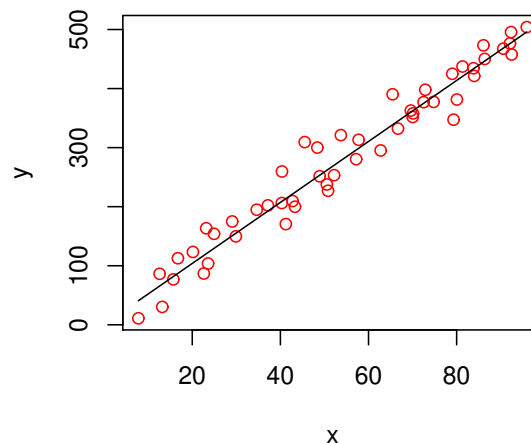


Figura 2: Função linear ajustada aos dados pelo *adaline*.

A fim de investigar a correlação entre a variância do ruído presente nos dados de entrada e o erro quadrático médio da função ajustada à medida que o número de amostras cresce, considera-se um gráfico de erro por número de amostras (Figura 3). Dada uma mesma partição de treino e teste, os erros foram calculados a cada 20 amostras no

intervalo de 0 a 10000, onde os valores extremos foram substituídos pela mediana dos erros para melhor visualização da curva.

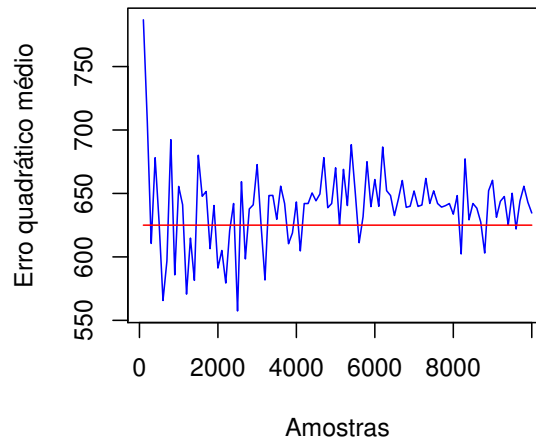


Figura 3: Erro quadrático médio por número de amostras.

É possível notar que à proporção que o número de amostras aumenta, tem-se que o erro quadrático médio tende à variância do ruído $\sigma^2 = 625$ (reta horizontal em vermelho).

Por fim, deseja-se avaliar o *adaline* para problemas de regressão múltipla, isto é, mais de duas variáveis explicativas. Para tal, um base de dados sobre consumo de energia, disponível no arquivo *BUILDING1paraR.DT*, foi considerada. No total são 14 *features* (7 binárias e 7 contínuas) para prever três funções de saída: *Energy*, *Hot Water* e *Cold Water*.

A taxa de aprendizagem para ajustar a função *Energy* (Figura 4) foi de $\eta = 0.001$ e o número máximo de épocas foi de $max_epochs = 15000$.

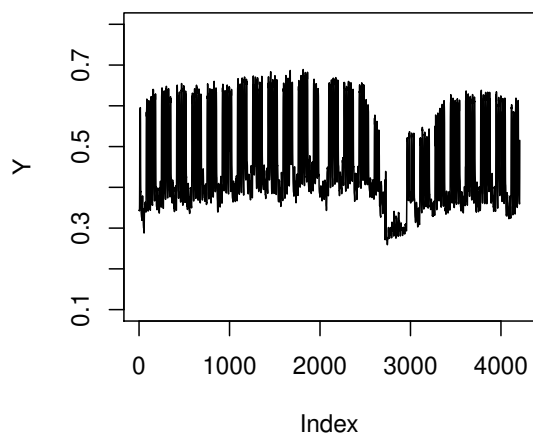


Figura 4: Função *Energy* original.

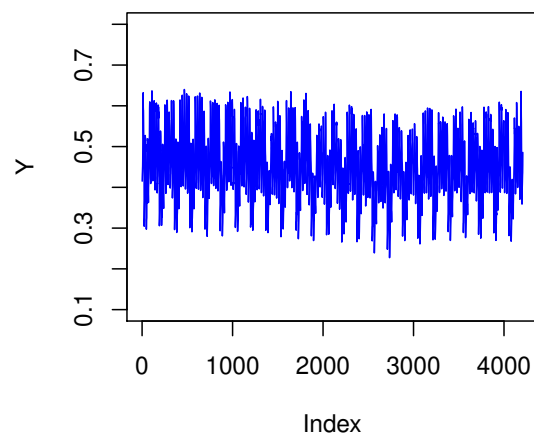


Figura 5: Função ajustada pelo *adaline*.

Ainda que o erro quadrático médio tenha sido relativamente baixo 4.8×10^{-3} e exista uma similaridade bastante evidente entre os valores extremos da função ajustada e da função original, percebe-se que a aproximação não conseguiu a suavização na faixa de 2700 a 3000 (Figura 5). Diante da sobreposição das funções na Figura 6, nota-se também que, ao atingir um pico, a aproximação não decai tanto quanto a função original.

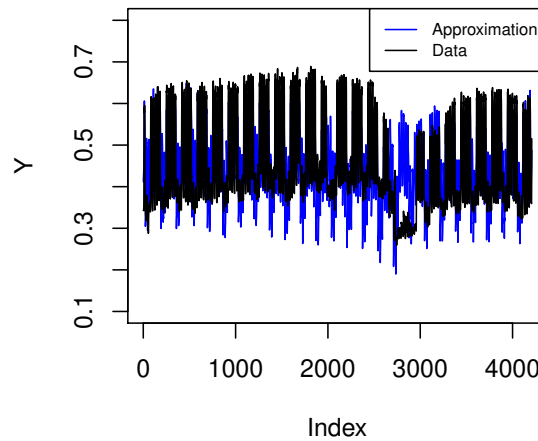


Figura 6: Sobreposição da função *Energy* com sua respectiva aproximação.

Para ajustar a segunda função, *Hot Water* (Figura 7), utilizou-se a mesma taxa de aprendizagem $\eta = 0.001$ e um número máximo de épocas substancialmente menor $max_epochs = 10000$.

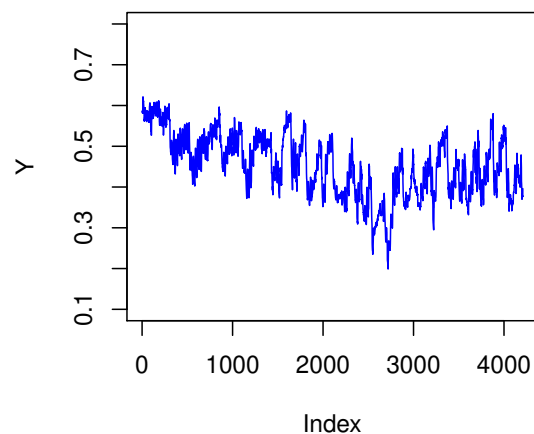
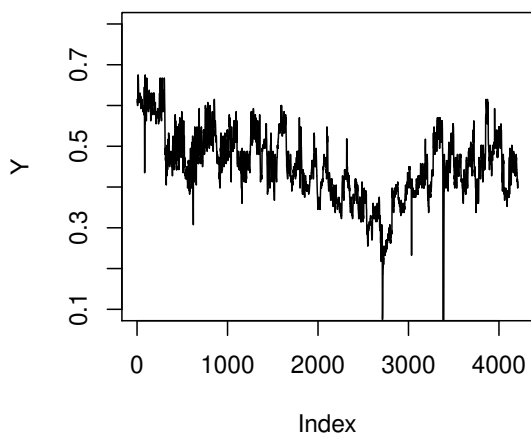


Figura 7: Função *Hot Water* original.

Figura 8: Função ajustada pelo *adaline*.

Ao contrário do caso anterior, a função ajustada pelo *adaline* para a saída *Hot Water* (Figura 8) foi visualmente melhor. Em conformidade com tal percepção visual de melhoria, o erro quadrático médio apresentou um valor três vezes mais baixo (1.4×10^{-3}),

embora a aproximação não tenha modelado alguns dos mínimos locais mais expressivos. A sobreposição da função original e sua respectiva aproximação ratificam as observações apresentadas (Figura 9).

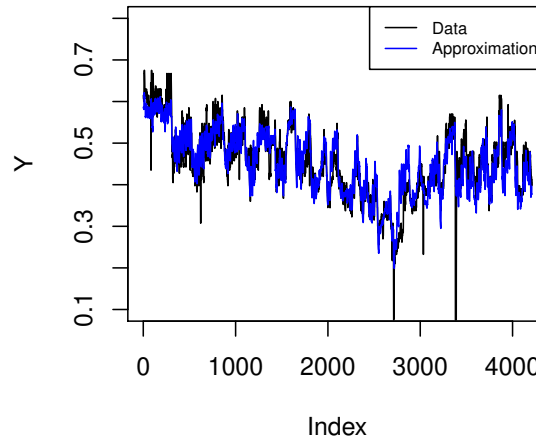


Figura 9: Sobreposição da função *Hot Water* com sua respectiva aproximação.

Finalmente, a taxa de aprendizagem para o ajuste da função *Cold Water* (Figura 10) foi de $\eta = 0.001$ e o número máximo de épocas foi de $max_epochs = 14000$.

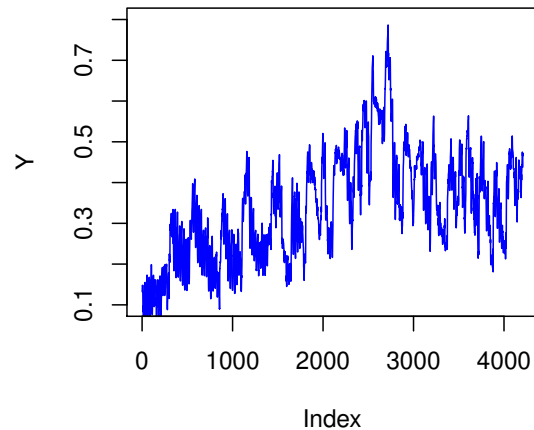
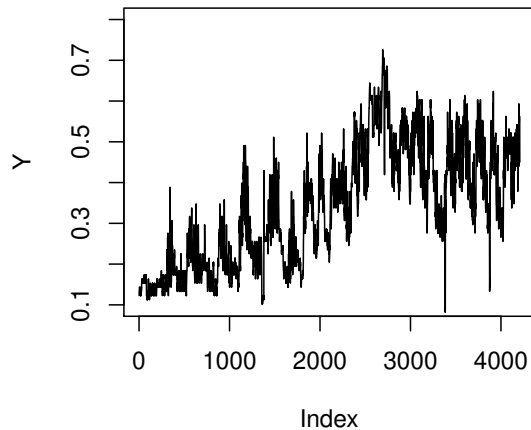


Figura 10: Função *Cold Water* original.

Figura 11: Função ajustada pelo *adaline*.

O erro quadrático médio foi de 4×10^{-3} , sendo, portanto, ligeiramente menor que o da função *Energy* ajustada e consideravelmente maior que o da função *Hot Water*. Todavia, percebe-se, para esse último caso, uma boa similaridade no comportamento da aproximação com sua respectiva função original. Isso torna-se mais claro com a sobreposição das curvas presentes na Figura 12, ainda que os extremos entre 3000 e 4000 sejam mais atenuados na função obtida pelo *adaline* do que na saída desejada.

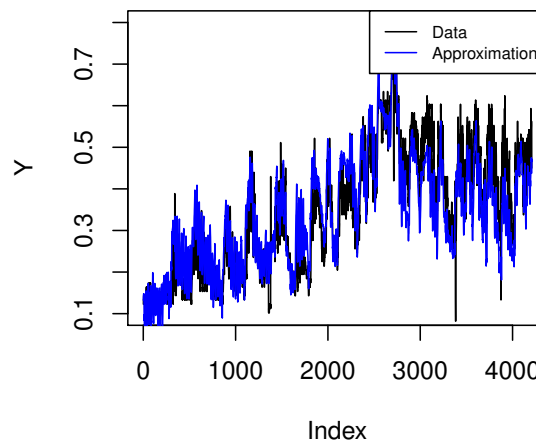


Figura 12: Sobreposição da função *Cold Water* com sua respectiva aproximação.

Conclusão

Apesar de hoje ser mais comum utilizar outros modelos de redes neurais mais complexos, o *adaline* permite uma compreensão clara do funcionamento de uma rede neural em termos matemáticos para problemas de regressão, sendo uma excelente introdução.

Referências

- [1] A. d. P. Braga, *Redes neurais artificiais: teoria e aplicações*. Livros Técnicos e Científicos, 2000.