

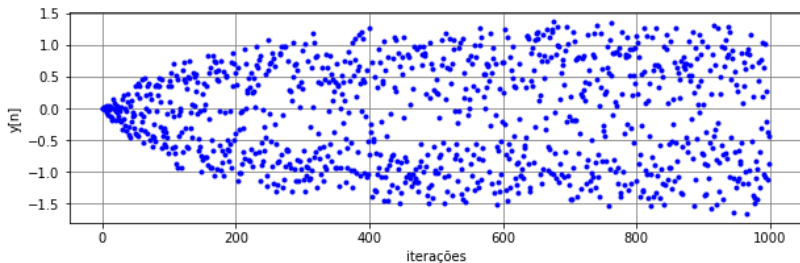
PSI3471 – Fundamentos de Sistemas Eletrônicos Inteligentes
O neurônio de Rosenblatt

Magno T. M. Silva e Renato Candido

Escola Politécnica da USP

1 LMS no problemas das meias-luas

- A saída do algoritmo fica espalhada no intervalo $[-1,5 \ 1,5]$, não havendo uma clara separação em torno do zero.



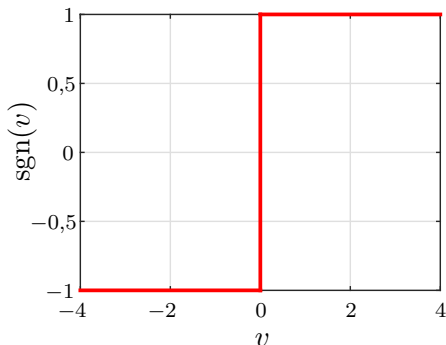
Saída do algoritmo LMS ($\eta = 10^{-4}$ e $M = 2$) durante o treinamento no modo estocástico ($N_t = 1000$, $N_b = 1$ e $N_e = 1$) utilizado no problema de classificação das meias-luas ($r_1 = 10$, $r_2 = 1$ e $r_3 = 6$); taxa de erro de 2,5%.

2 Neurônio de Rosenblatt

- ▶ O perceptron de Rosenblatt força a saída $y(n)$ a assumir valores do conjunto $\{-1, +1\}$
- ▶ Para isso, considera a função

$$\varphi(v) = \text{sgn}(v) = \begin{cases} +1, & v \geq 0 \\ -1, & v < 0 \end{cases},$$

na saída do combinador



3 Neurônio de Rosenblatt

- ▶ vetor dos dados de treinamento:

$$\mathbf{x}(n) = [1 \ x_{1n} \ x_{2n} \ \cdots \ x_{Mn}]^T$$

- ▶ vetor de pesos:

$$\mathbf{w}(n) = [b(n) \ w_1(n) \ \cdots \ w_M(n)]^T$$

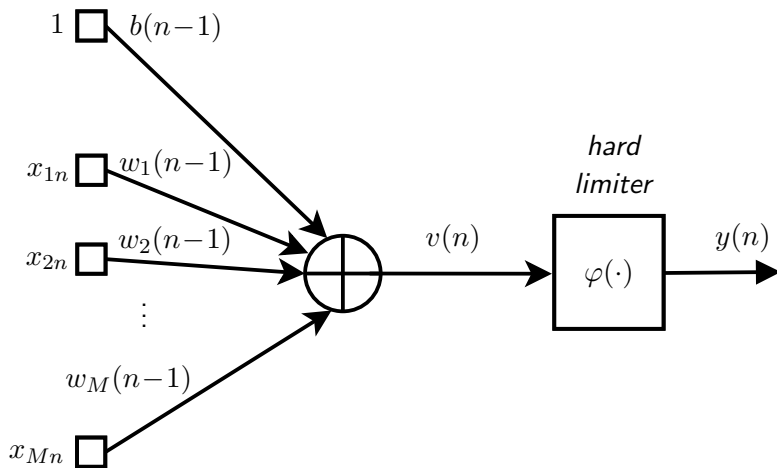
- ▶ saída do combinador linear:

$$v(n) = \mathbf{x}^T(n) \mathbf{w}(n-1)$$

- ▶ saída do perceptron de Rosenblatt:

$$y(n) = \varphi(v(n)) = \text{sgn}(v(n))$$

4 Fluxo de sinal do perceptron de Rosenblatt



5 Algoritmo

- ▶ Os pesos são atualizados para minimizar $e^2(n)$, em que

$$e(n) = d(n) - \varphi(\mathbf{x}^T(n)\mathbf{w}(n-1)) = d(n) - \text{sgn}(v(n)) = d(n) - y(n)$$

- ▶ $e(n)$ assume três valores possíveis:
 - ▶ -2 ou $+2$ quando $d(n) \neq y(n)$
 - ▶ 0 quando $d(n) = y(n)$
- ▶ Como a função sinal não é derivável em todos os pontos, não é possível obter o algoritmo de maneira formal.
- ▶ Note que

$$\frac{\partial e(n)}{\partial \mathbf{w}(n-1)} = -\frac{\partial \text{sgn}(\mathbf{x}^T(n)\mathbf{w}(n-1))}{\partial \mathbf{w}(n-1)} = -\mathbf{x}(n) \text{sgn}'(v(n)) = \begin{cases} \mathbf{0}, & v(n) \neq 0 \\ \nexists, & v(n) = 0 \end{cases}$$

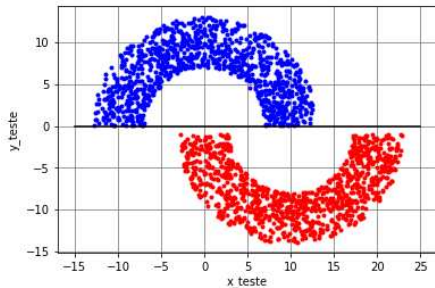
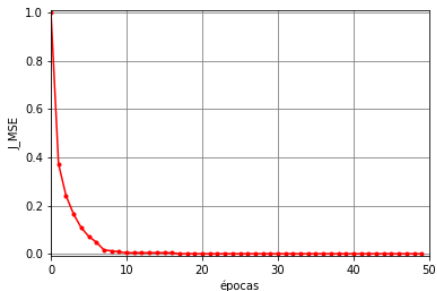
6 Algoritmo

- ▶ Ignorando o fato da derivada não existir para $v(n) = 0$, os pesos não seriam atualizados, pois o vetor gradiente é nulo para $v(n) \neq 0$
- ▶ Utiliza-se a equação de atualização

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \eta e(n) \mathbf{x}(n)$$

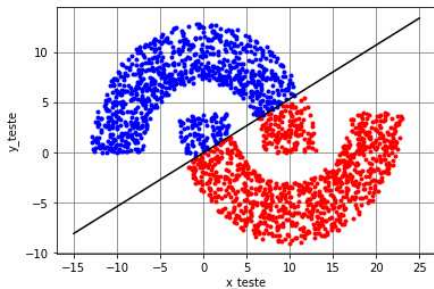
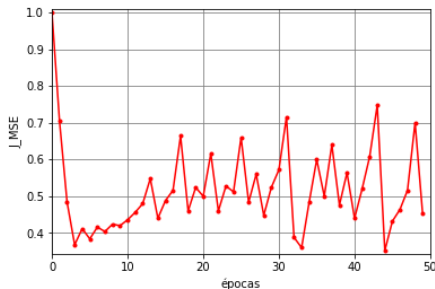
- ▶ Os pesos são atualizados apenas quando $e(n) \neq 0$, ou seja, quando $y(n) \neq d(n)$. Caso contrário, $\mathbf{w}(n) = \mathbf{w}(n-1)$.
- ▶ O passo de adaptação η , também chamado de taxa de aprendizado, é uma constante positiva que deve ser escolhida no intervalo $0 < \eta \leq 1$

7 Rosenblatt nas meias-luas



O problema de classificação das meias-luas ($r_1 = 10$, $r_2 = 1$ e $r_3 = 6$). Função custo ao longo das épocas de treinamento (figura à esquerda); Dados de teste ($N_{teste} = 2000$) e reta de separação das regiões (figura à direita) obtida com o perceptron de Rosenblatt treinado em *batch* ($M = 2$, $\eta = 0,001$, $N_t = 1000$, $N_b = N_t$ e $N_e = 50$); taxa de erro=0%

8 Rosenblatt nas meias-luas



O problema de classificação das meias-luas ($r_1 = 10$, $r_2 = -4$ e $r_3 = 6$). Função custo ao longo das épocas de treinamento (figura à esquerda); Dados de teste ($N_{teste} = 2000$) e reta de separação das regiões (figura à direita) obtida com o perceptron de Rosenblatt treinado em *batch* ($M = 2$, $\eta = 0,001$, $N_t = 1000$, $N_b = N_t$ e $N_e = 50$); taxa de erro=12,8%

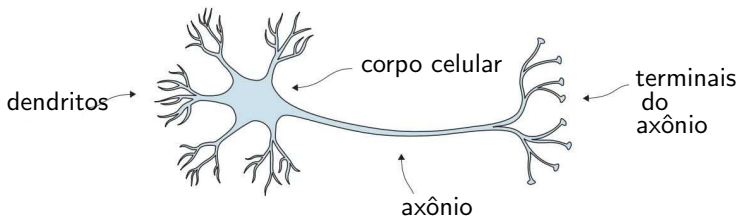
9 Rosenblatt nas meias-luas

- ▶ Como $e(n) \in \{-2, 0, 2\}$, o perceptron de Rosenblatt não sofre de divergência desde que as entradas sejam limitadas.
- ▶ O mesmo não se pode afirmar sobre o algoritmo LMS, pois o sinal de erro não é limitado. Dependendo do valor do passo de adaptação η , o erro e consequentemente os pesos podem divergir.
- ▶ Tanto Rosenblatt como o LMS levam a fronteiras de separação que são retas (ou hiperplanos no caso em que $M > 2$)

10 Neurônio biológico e um pouco de história

- ▶ O médico espanhol **Ramón y Cajál** foi o primeiro a introduzir a **ideia dos neurônios** como unidades básicas do sistema nervoso.
- ▶ Os neurônios são células especializadas na **transmissão de informações na forma de pulsos nervosos**.
- ▶ As ligações entre os neurônios são chamadas de **sinapses**, que tem por função enviar sinais para ocorrer ações no corpo.
- ▶ A taxa dessas **transmissões é baixa** (ms) quando comparada com portas lógicas de silício (ns).
- ▶ Essa taxa “baixa” é compensada pelo número de neurônios (**86 bilhões**) e sinapses (**60 trilhões**) existentes no sistema nervoso humano
- ▶ O resultado é que o cérebro é uma estrutura muito **eficiente**.

11 Neurônio biológico e um pouco de história



- ▶ Atividade caracterizada por **pulsos elétricos** (mV, ms), recebidos de outros neurônios pelos dendritos.
- ▶ Se o sinal acumulado exceder um limiar, **um pulso é enviado aos seus terminais**, que se acoplam a outros neurônios.
- ▶ A computação realizada por um neurônio pode ser resumida na **frequência dos pulsos**.
- ▶ Se houver **poucos pulsos** por unidade de tempo, o neurônio é considerado **pouco ativo**. Se houver **muitos pulsos** por unidade de tempo, haverá **mais estímulos sinápticos**.

12 Neurônio biológico e um pouco de história

- ▶ Redes neurais surgiram para buscar **modelar o cérebro humano**.
- ▶ Nos anos de surgimento das redes neurais (1943-1960), vários pesquisadores se destacaram:
 - ▶ McCulloch e Pitts (1943) por introduzirem a ideia de **redes neurais como máquinas de computação**;
 - ▶ Hebb (1949) por postular a **primeira regra de aprendizagem auto-organizada**;
 - ▶ Rosenblatt (1958) por propor o **perceptron como o primeiro modelo de aprendizagem supervisionada**;
 - ▶ Widrow e Hoff (1960) por propor o **Adaline (*adaptive linear element*)**, que deu origem ao algoritmo LMS.

13 Neurônio biológico e um pouco de história

- ▶ Rosenblatt propôs o **modelo de neurônio artificial**, chamado de perceptron.
- ▶ O neurônio biológico recebe vários estímulos de outros neurônios que chegam por seus dendritos, esses estímulos são então acumulados e se exceder um limiar, o neurônio gera um estímulo no seu axônio que são transmitidos a outros neurônios.
- ▶ No modelo de Rosenblatt:
 - ▶ $\mathbf{x}(n)$: **estímulos**
 - ▶ $v(n)$: **acúmulo dos estímulos**
 - ▶ $v(n) < 0$: **neurônio em repouso**
 - ▶ $v(n) \geq 0$: **um novo estímulo $y(n)$ é gerado**
- ▶ Para representar o neurônio em repouso, talvez fosse mais adequado considerar a função degrau em vez da função sinal. Mas a função sinal é mais adequada para implementação com um circuito analógico.

14 Neurônio biológico e um pouco de história

- ▶ Julho de 1958: um IBM 704, um computador de 5 toneladas que ocupava uma sala, foi alimentado com uma série de cartões perfurados. Após 50 tentativas, o computador aprendeu a distinguir os cartões marcados à esquerda dos cartões marcados à direita.
- ▶ Foi uma demonstração do perceptron de Rosenblatt, a primeira máquina capaz de ter uma ideia original.
- ▶ Na época, Rosenblatt era psicólogo pesquisador e engenheiro de projetos no Laboratório Aeronáutico da Cornell em Buffalo, Nova York.

“As histórias sobre a criação de máquinas com qualidades humanas têm sido fascinantes em ficção científica. No entanto, estamos prestes a testemunhar o nascimento de tal máquina – uma máquina capaz de perceber, reconhecer e identificar seus arredores sem qualquer treinamento ou controle humano”

- ▶ Rosenblatt estava certo, mas levou aproximadamente meio século para vermos isso acontecer.

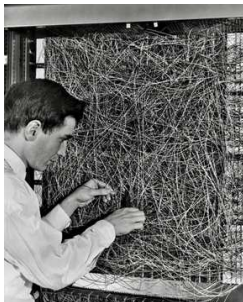
15 Neurônio biológico e um pouco de história

Psychological Review
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN¹

F. ROSENBLATT

Cornell Aeronautical Laboratory



Publicação de Rosenblatt de 1958 (à esquerda) e foto de Rosenblatt e seu perceptron chamado de Mark I em 1960 (à direita).

16 Neurônio biológico e um pouco de história

- ▶ Desde 1960, muita pesquisa foi feita com o objetivo de melhorar o modelo do cérebro humano.
- ▶ Apesar dos inúmeros avanços, ainda estamos longe de termos um sistema que consiga modelar de maneira precisa o cérebro, devido à sua alta complexidade e eficiência.
- ▶ Apesar das redes neurais artificiais serem inspiradas no funcionamento do cérebro, vamos encará-las como **sistemas não lineares** que podem ser aplicados como soluções eficientes em problemas de regressão e classificação.