UM MODELO FÍSICO PARA REDES NEURAIS

J. R. Campanha
A. Tancredo
Depto. de Física – UNESP
Rio Claro – SP

Resumo

Recentemente as redes neurais têm despertado um grande interesse entre os físicos, devido à semelhança com sistemas magnéticos denominados vidros de spins. Apresentamos o modelo de uma rede neural do tipo Hopfield de maneira simples, utilizando apenas operações de multiplicação de matrizes para possibilitar que alunos e professores de física se interessem pelo sistema mostrado em sala de aula e o funcionamento de uma rede neural.

I. Introdução

O cérebro humano(1) é constituído de aproximadamente 10^{12} células denominadas neurônios. O neurônio e composto de três partes: corpo ou soma, axônio e dendritos. O axônio é o condutor dos estímulos nervosos; o corpo e os dendritos atuam como propagadores desses estímulos. Os neurônios são conectados uns aos outros através das sinapses. As sinapses desempenham uma função importante porque possibilitam a transmissão de informações entre os neurônios.

O neurônio pode ser ativado quando um estímulo elétrico provoca liberação na junção sináptica de uma substância química, o neurotransmissor, possibilitando que um sinal se propague pelo axônio até o corpo e dendritos, e conseqüentemente a outros neurônios. Os efeitos elétricos que chegam até o neurônio podem ser excitatórios ou inibitórios. Se a soma de todos os potenciais de ação (excitatórios e inibitórios), vindos de todas as sinapses, excede um certo valor limiar (-70 mV) o neurônio dispara um potencial de ação que se propaga ate a junção cinética de outro neurônio. Assim um neurônio responde aos estímulos da forma: *tudo ou nada*.

Ao conjunto de neurônios interligados chamamos de rede neural. A memória e a aprendizagem são exemplos de funções emergentes, que dependem exclusivamente da cooperatividade entre os neurônios.

McCulloch e Pitts[2] em 1943 foram os primeiros a tentar simular o comportamento de uma rede neural supondo os neurônios como simples processadores que podem estar em dois estados possíveis: ativo ou inativo. Posteriormente, em 1954, veio o trabalho de Cragg e Temperley[3] onde fizeram a primeira analogia entre uma rede neural e os estados coletivos de dipolos magnéticos, mas não chamou a atenção o suficiente para o desenvolvimento de novos trabalhos.

Em 1974, Little[4] comparou os neurônios com um sistema de spin de Ising, onde o estado do neurônio **i** no tempo **t** é representado por um spin de Ising $\mathbf{S_i}$, o qual toma o valor +1 se está ativo e -1 se inativo. Esta representação binária reflete a condição do "tudo ou nada" mencionada acima.

O tempo passou e os físicos somente voltaram a se interessar por redes neurais em 1982 quando J.J. Hopfield[5] sugeriu a analogia entre o comportamento da rede neural com certos materiais magnéticos desordenados, conhecidos como vidros de spins[6,7].

Vidros de spins são sistemas magnéticos onde as interações ferromagnéticas e interações antiferromagnéticas são distribuídas aleatoriamente dentro do material. Estudos teóricos de vidros de spins mostram que eles possuem uma rica topologia com muitos mínimos locais cuja energia para um sistema com N spins pode ser descrita pela equação:

$$E = -1/2 \sum_{i,j}^{N} J_{ij} S_i S_j \, , \qquad (1)$$

onde J_{ij} mede o acoplamento entre os spins S_i e S_j e pode ser negativo (interações antiferromagnéticas) ou positivo (interações ferromagnéticas). A existência de uma função energia implica que, à temperatura T = 0, o sistema deve fluir a um mínimo local da energia. Esses mínimos locais são determinadas configurações de spins que minimizam a energia do sistema.

II. Modelo de Hopfield

Vamos supor um sistema com N neurônios, onde o comportamento de cada neurônio seja análogo ao comportamento de um spin. Assim o neurônio terá dois estados, $S_i = +1$ quando está disparando um potencial de ação ou “aceso” e $S_i = -1$ quando está inativo ou “apagado”. (Utilizamos o termo “aceso” procurando dar a idéia de um ponto brilhante na tela de um monitor ou numa lousa cujo fundo é escuro.)

A uma configuração dos N neurônios tomando os valores $S_i = +1$ ou $S_i = -1$ denominamos de padrão, e podemos representá-lo por um vetor. Assim por exemplo:

$$\overset{\omega}{S} = (+1,-1,+1,+1,+1,+1,-1,-1,+1) \tag{2}$$

é um padrão representado por um vetor para um sistema com nove neurônios. Se associarmos esse vetor a um padrão visual que formam letras ou números ele pode ser identificado com o número 4 (ver Fig.1), onde representamos por +1 os pontos da matriz 3 x 3 que estão “acesos” e por -1 os pontos da matriz que estão ”apagados”.

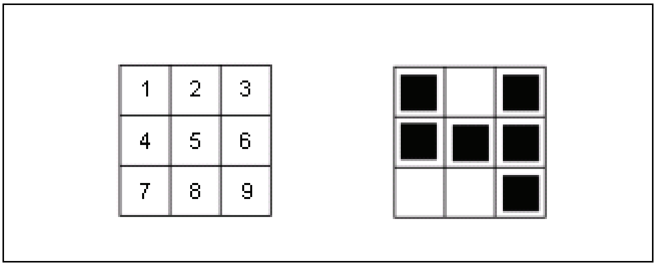


Fig.1

A letra U (Fig.2) pode então, por exemplo, ser representada pelo vetor:

$\vec{S} = (+1,-1,+1,+1,-1,+1,+1,+1,+1)$

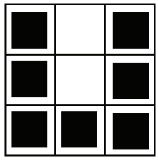


Fig. 2

Podemos estender a dimensão do vetor e construir qualquer figura, como por exemplo uma foto de jornal, apenas com pontos “acesos” e “apagados”.

Para que nossa rede neural funcione de maneira análoga ao conjunto de neurônios que formam o cérebro humano, quanto ao reconhecimento de padrões, é necessário que numa primeira fase a rede “aprenda” um certo número de padrões e numa segunda fase, ao lhe ser apresentada um dos padrões aprendidos (ou um padrão muito parecido com os ensinados), a rede deverá ser capaz de reconhecê-lo.

Para cumprirmos a primeira fase precisamos de uma regra para o aprendizado de nossa rede. Tal regra de aprendizado foi enunciada pela primeira vez pelo psicologista Donald Hebb[(8)] em 1949. A regra estabelece que o aprendizado em sistemas neurais é feito através da modificação seletiva das sinapses entre os neurônios.

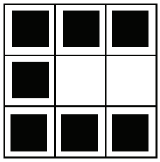
No caso tratado por nós podemos efetuar ligações entre os neurônios cuja regra, baseada em Hebb, é dada por:

$$J_{i\,j} = \sum_{\mu=1}^{p} \xi_i^{\mu}\ \xi_{\mu}^{j}\ , \qquad (3)$$

onde p é o número de padrões aprendidos e $J_{i\,j}$ é o valor da conexão entre o neurônio i e o neurônio j. Os ξ_i^{μ} são as componentes dos padrões aprendidos que assumem os valores ±1. Notamos que com o conjunto de todos os $J_{i\,j}$ podemos construir uma matriz. Essa matriz que denominamos de matriz **J** é simétrica ou $J_{i\,j} = J_{j\,i}$ e também tomamos os elementos da diagonal como nulos ou $J_{i\,i} = 0$, o que reflete o fato de que um neurônio conecta-se a outro neurônio, mas não a si próprio.

Utilizando a regra, ensinamos a nossa rede, numa primeira fase, dois padrões, as letras C e T (Fig.3), ou:

$\vec{C} = (+1,+1,+1,+1,-1,-1,+1,+1,+1)$



$\vec{T} = (+1,+1,+1,-1,+1,-1,-1,+1,-1)$

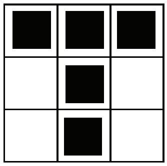


Fig. 3

Com a equação (3) calculamos os elementos J_{ij}, por exemplo:

$$J_{12} = \sum_{\mu=1}^{p=2} \xi_1^{\mu}\ \xi_2^{\mu} = \xi_1^1\ \xi_2^1 + \xi_1^2\ \xi_2^2\,, \tag{4}$$

onde ξ_1^1 é a primeira componente do vetor do primeiro padrão e ξ_1^2 é a primeira componente do vetor do segundo padrão, etc..

Então:

$J_{12} = (+1)(-1) + (+1)(-1) = 2$, logo $J_{21} = 2$.

Assim podemos calcular todos os elementos J_{ij} e alguns exemplos são:

$J_{19} = 0$; $J_{91} = 0$; $J_{16} = -2$; $J_{61} = -2$; etc., e finalmente a matriz J:

$$J = \left| \begin{array}{ccccccccc} 0 & 2 & 2 & 0 & 0 & -2 & 0 & 2 & 0 \\ & 0 & 2 & 0 & 0 & -2 & 0 & 2 & 0 \\ & & 0 & 0 & 0 & -2 & 0 & 2 & 0 \\ & & & 0 & -2 & 0 & 2 & 0 & 2 \\ & & & & 0 & 0 & -2 & 0 & -2 \\ & & & & & 0 & 0 & -2 & 0 \\ & & & & & & 0 & 0 & 2 \\ & & & & & & & 0 & 0 \\ & & & & & & & & 0 \end{array} \right|, \tag{5}$$

Nessa matriz estão contidas todas as informações memorizadas na rede neural.

Podemos então representar algumas conexões de nossa rede como apresentado na Fig.4.

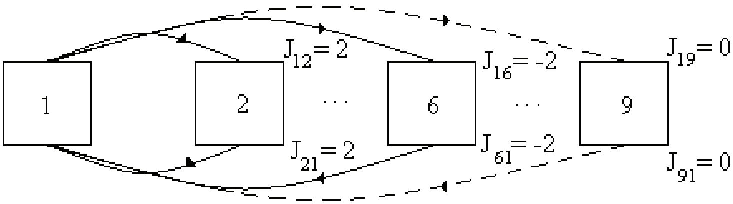


Fig. 4

A busca das informações memorizadas ou a recuperação dos padrões ensinados é feita através de uma dinâmica estabelecida na rede neural. Essa dinâmica foi proposta em 1982 por J.J. Hopfield[(5)], e é dada pela equação:

$$S_i\,(t+1) = sign\left(\sum_j J_{i\,j}\; S_j(t) \right), \tag{6}$$

onde sign (x) é a função dada por:

$$\text{sign (x)} = +1 \text{ se } x > 0 \text{ e sign (x)} = -1 \text{ se } x \le 0 \tag{7}$$

Uma vez estabelecidas as conexões entre os neurônios, isto é, a matriz J e conhecendo a dinâmica da rede neural passamos para a segunda fase na qual apresentaremos um padrão para ser reconhecido pela nossa rede.

Vamos apresentar para nossa rede o vetor (Fig.5):

$$\vec{S}(0) = (+1,+1,+1,+1,-1,-1,-1,+1,+1)$$

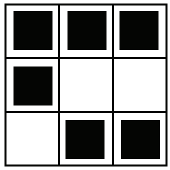


Fig. 5

Então a dinâmica estabelece a evolução de $\vec{S}(0)$, e utilizando a equação (6) obtemos:

$$S_i(1) = sign\left(\sum_j J_{ij}\ S_j(0)\right),$$

logo explicitamente calculando todas as componentes do vetor temos

$$\vec{S}(1) = (+1,+1,+1,+1,-1,-1,+1,+1,+1),$$

Continuando a iteração obtemos:

$$S_i\ (t+1) = S_1(2) = sign\left(\sum_j J_{ij}\ S_j(1)\right),$$

logo

$$\vec{S}(2) = (+1,+1,+1,+1,-1,-1,+1,+1,+1),$$

e assim sucessivamente

$$S_i(3) = sign\left(\sum_j J_{ij}\ S_j(2)\right),$$

logo

$$\vec{S}(3) = (+1,+1,+1,+1,-1,-1,+1,+1,+1).$$

Observamos que partindo de $\vec{S}(0)$ o sistema com sua dinâmica converge logo na primeira iteração para o vetor $\vec{C}$ que é um dos padrões que foi ensinado à nossa rede. A rapidez dessa convergência é devido ao fato de que o vetor $\vec{S}(0)$ é um vetor muito próximo ao vetor $\vec{C}$ ensinado.

Se você tiver prática computacional poderá programar uma rede neural e obter resultados surpreendentes na recuperação dos padrões ensinados, e observara que existe um limite para a quantidade de padrões ensinados em relação ao número de neurônios da rede.

III. Conclusões

Como pudemos observar, a nossa rede procura imitar uma rede de neurônios biológicos, ou, como o nosso cérebro funciona no caso de reconhecimento de padrões. No entanto, a dinâmica cerebral é muito mais complexa. Para sentir esse problema lembramos o fato de que qualquer criança (com algum tempo de aprendizado) re conhece facilmente letras, palavras, sons, etc., já para que um computador reconheça, por exemplo, as letras **A**, **a**, A, a, *A*, *a*, seria necessário programá-lo para "explicar-lhe" que não importa o tamanho, a espessura, a forma, a posição espacial, etc. da letra, pois todas representam o mesmo padrão, isto é, a *letra a*.

Outro fato é a comparação da rede neural com sistemas de vidros de spins[(9)]. Dissemos que os vidros de spins possuem muitos mínimos locais. Esses mínimos podem representar os padrões aprendidos ou a memória de forma que, se apresentarmos um padrão semelhante a um dos padrões aprendidos, haverá uma dinâmica do sistema na qual os padrões ensinados terminam sendo os atratores dessa dinâmica, como vimos na nossa rede onde o atrator seria o vetor que representa a letra C.

O professor poderá mostrar em sala de aula o funcionamento de uma rede neural, onde as regras e conceitos para montagem do sistema são ensinados aos alunos com auxílio deste artigo e na seqüência sugerir que os alunos ensinem as suas redes alguns padrões e tentem recuperá-los através da apresentação de configurações muito próximas às dos padrões ensinados.

IV. Referências Bibliográficas

1. SMITH, C. M. **The brain.** New York: G. P. Putmanns, 1970.

2. McCULLOCH, W. S.; PITTS, W. A. **Bull. Math. Biophys.,** v. 5, p. 115, 1943.

3. CRAGG, B. G.; TEMPERLEY, H. H. V. **Electroencephalog. Clin. Neurophys.,** v. 6, p. 85, 1954.

4. LITTLE, W. A. **Math. Biosci.,** v. 19, p. 101, 1974.

5. HOPFIELD, J. J. **Proc. Nat. Acad. Sci. USA,** v. 79, p. 2554, 1982.

6. TANCREDO, A.; CAMPANHA, J. R. **Ecl. Quim.,** v. 15, p. 51-61, 1990.

7. KIRKPATRICK, S.; SHERRINGTON, D. J. **Phys. Rev. B**, v. 17, p. 4384, 1978.

8. HEBB, D. O. **The organization of behaviour.** New York: Willey,1949.

9. AMIT, D. J.; GUTFREUND, H.; SOMPOLINSKY, H. **Phys. Rev. A,** v. 32, p. 1007, 1985.

Pense e Responda!

É possível achar um ponto *P* no vácuo tal que ao colocar uma carga elétrica nele esta fique em equilíbrio estável, só pelo efeito de forças eletrostáticas? O que se está procurando obter é uma distribuição de cargas e superfícies com diversos potenciais que criem o ponto P com a propriedade descrita acima. Este problema é de interesse para o confinamento e resfriamento de átomos em espectroscopia de altíssima resolução.