UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO
Departamento de Ciências de Computação

SCC-5809 - Capítulo 7 Redes Associativas

João Luís Garcia Rosa[1]

[1]SCC-ICMC-USP - joaoluis@icmc.usp.br

2011

Sumário

Sumário

Neurodinâmica

- *Neurodinâmica*: RNAs vistas como sistemas dinâmicos não-lineares (SDNL), com ênfase no problema da *estabilidade*.
- A estabilidade de um SDNL é característica de todo o sistema.
- "A presença de estabilidade sempre implica alguma forma de coordenação entre as partes individuais do sistema." [1]
- A estabilidade mencionada aqui é no sentido do *método direto de Lyapunov* de 1892, usado para análise de estabilidade de sistemas lineares e não-lineares, invariante e variante no tempo: aplicável diretamente para RNA!

Neurodinâmica determinística

- O estudo da neurodinâmica pode ser dividido em:
  1. *Neurodinâmica determinística*: na qual o modelo de RNA tem um comportamento determinístico descrito por um conjunto de *equações diferenciais não-lineares* que definem a evolução exata do modelo como uma função do tempo.
  2. *Neurodinâmica estatística*: na qual o modelo de RNA é perturbado pela presença de ruído. Neste caso, deve-se tratar com *equações diferenciais não-lineares estocásticas* que expressam a solução em termos probabilísticos. A combinação da estocasticidade e não-linearidade torna-a mais difícil de tratar.
- Este capítulo será restrito à neurodinâmica determinística.

Sistemas Dinâmicos

- Para proceder com o estudo da neurodinâmica, é necessário um *modelo matemático estado-espaço*.
- Conjunto de *variáveis de estado* cujos valores devem conter informação suficiente para predizer a evolução futura do sistema.
- Sejam $x_1(t)$, $x_2(t)$, ... $x_N(t)$ as variáveis de estado de um sistema dinâmico não-linear onde o tempo contínuo t é a *variável independente* e N é a *ordem* do sistema.
- Para facilitar a notação, um *vetor de estados* N-por-1 $\mathbf{x}(t)$ contém essas variáveis.

Sistemas Dinâmicos

- A dinâmica de sistemas dinâmicos não-lineares pode ser representada na forma de um sistema de equações diferenciais de primeira ordem:

$$\frac{d}{dt}x_j(t) = F_j(x_j(t)), \quad j = 1, 2, ..., N \tag{1}$$

onde $F(\cdot)$ é uma função não-linear. Posto em forma compacta:

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{F}(\mathbf{x}(t)) \tag{2}$$

onde a função não-linear **F** é um vetor em que cada elemento opera em um elemento correspondente do vetor de estados

$$\mathbf{x}(t) = [x_1(t), x_2(t), ..., x_N(t)]^T \tag{3}$$

Sistemas Dinâmicos

- Um sistema dinâmico não-linear para o qual a função vetor $\mathbf{F}(\mathbf{x}(t))$ não dependa *explicitamente* do tempo t, como na equação 2, é chamado de *autônomo*.
- Independentemente da forma exata da função não-linear $\mathbf{F}(\cdot)$, o vetor de estados $\mathbf{x}(t)$ deve variar com o tempo t; caso contrário $\mathbf{x}(t)$ é constante e o sistema não é dinâmico.

  Um sistema dinâmico é um sistema cujo estado varia com o tempo.

- Pode-se pensar em $d\mathbf{x}/dt$ como um vetor "velocidade", não em termos físicos, mas num sentido abstrato.
- Então, de acordo com a equação 2, pode-se referir à função vetor $\mathbf{F}(\mathbf{x})$ como um campo vetor velocidade ou simplesmente como um *campo vetor*.

Sumário

Estabilidade de estados de equilíbrio

- Considere um sistema dinâmico autônomo descrito pela equação 2.
- Um vetor constante $\bar{\mathbf{x}} \in \mathcal{M}$ é um *estado de equilíbrio (estacionário)* do sistema se a seguinte condição for satisfeita:

$$\mathbf{F}(\bar{\mathbf{x}}) = \mathbf{0} \tag{4}$$

  onde $\mathbf{0}$ é o vetor nulo.
- O vetor velocidade $d\mathbf{x}/dt$ desaparece no estado de equilíbrio $\bar{\mathbf{x}}$, e portanto a função constante $\mathbf{x}(t) = \bar{\mathbf{x}}$ é uma solução da equação 2.
- Ainda, por causa da propriedade da unicidade de soluções, nenhuma outra curva pode passar através do estado de equilíbrio $\bar{\mathbf{x}}$, também referido como *ponto singular*, significando que no caso de um ponto de equilíbrio a trajetória irá se degenerar no próprio ponto.

Definições de estabilidade

- No contexto de um sistema dinâmico não-linear autônomo com estado de equilíbrio $\bar{\mathbf{x}}$, as definições de estabilidade e convergência são
  1. O estado de equilíbrio $\bar{\mathbf{x}}$ é uniformemente estável se para qualquer positivo ϵ existe um positivo δ tal que a condição
  $$\| \mathbf{x}(0) - \bar{\mathbf{x}} \| < \delta \tag{5}$$
  implica
  $$\| \mathbf{x}(t) - \bar{\mathbf{x}} \| < \epsilon \tag{6}$$
  para todo $t > 0$.
  Esta definição estabelece que uma trajetória do sistema pode ser feita para ficar dentro de uma vizinhança pequena do estado de equilíbrio $\bar{\mathbf{x}}$ se o estado inicial $\mathbf{x}(0)$ é próximo de $\bar{\mathbf{x}}$.

Definições de estabilidade

- Definições de estabilidade e convergência (cont.)

  2. O estado de equilíbrio $\bar{\mathbf{x}}$ é convergente se existe um positivo δ tal que a condição

  $$\| \mathbf{x}(0) - \bar{\mathbf{x}} \| < \delta \tag{7}$$

  implica que

  $$\mathbf{x}(t) \to \bar{\mathbf{x}} \;\; \textit{quando} \;\; t \to \infty \tag{8}$$

  Se o estado inicial $\mathbf{x}(0)$ de uma trajetória é suficientemente próximo do estado de equilíbrio $\bar{\mathbf{x}}$, então a trajetória descrita pelo vetor de estados $\mathbf{x}(t)$ aproximar-se-á de $\bar{\mathbf{x}}$ quando o tempo t se aproximar do infinito.

Definições de estabilidade

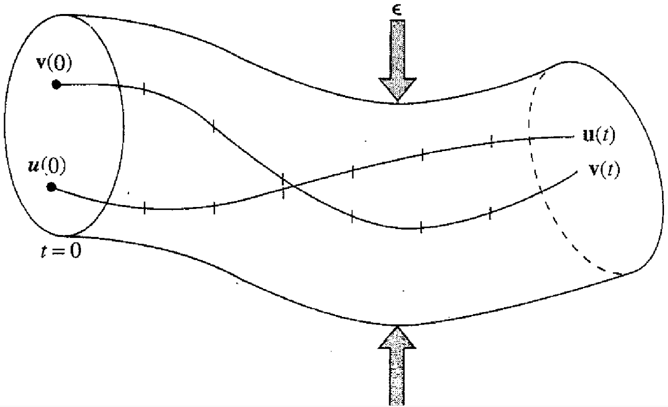
- Definições de estabilidade e convergência (cont.)
  3. O estado de equilíbrio $\bar{\mathbf{x}}$ é assintoticamente estável se for estável e convergente.
     Estabilidade e convergência são propriedades independentes. Apenas quando ambas são satisfeitas, tem-se a estabilidade assintótica.
  4. O estado de equilíbrio $\bar{\mathbf{x}}$ é assintoticamente estável ou globalmente assintoticamente estável se for estável e todas trajetórias do sistema convergem para $\bar{\mathbf{x}}$ quando o tempo t aproxima-se do infinito.
     O sistema não pode ter outros estados de equilíbrio e é necessário que toda trajetória permaneça limitada para todo tempo $t > 0$.

Exemplo

- Seja uma solução $\mathbf{u}(t)$ do sistema dinâmico não-linear descrito pela equação 2 que varia com o tempo t como indicado na figura 1.
- Para a solução $\mathbf{u}(t)$ ser uniformemente estável, é necessário que $\mathbf{u}(t)$ e qualquer outra solução $\mathbf{v}(t)$ fiquem próximas para os mesmos valores de t.
- Esse tipo de comportamento é conhecido como *correspondência isócrona* das duas soluções $\mathbf{u}(t)$ e $\mathbf{v}(t)$.
- A solução $\mathbf{u}(t)$ é convergente pois para toda outra solução $\mathbf{v}(t)$ para qual $\| \mathbf{v}(0) - \mathbf{u}(0) \| \leq \delta(\epsilon)$ no tempo $t = 0$, as soluções $\mathbf{v}(t)$ e $\mathbf{u}(t)$ convergem para um estado de equilíbrio quando t se aproxima do infinito.

Exemplo

Figure: Noção de estabilidade (convergência) uniforme de um vetor de estados [2].



Teoremas de Lyapunov

- Os teoremas de Lyapunov sobre a estabilidade e a estabilidade assintótica da equação de estado-espaço (equação 2) descrevendo um sistema dinâmico não-linear autônomo com vetor de estados $\mathbf{x}(t)$ e estado de equilíbrio $\bar{\mathbf{x}}$:
  1. O estado de equilíbrio $\bar{\mathbf{x}}$ é estável se em uma vizinhança pequena de $\bar{\mathbf{x}}$ existe uma função definida positiva $V(\mathbf{x})$ tal que sua derivada com respeito ao tempo seja semidefinida negativa nessa região.
  2. O estado de equilíbrio $\bar{\mathbf{x}}$ é assintoticamente estável se em uma vizinhança pequena de $\bar{\mathbf{x}}$ existe uma função definida positiva $V(\mathbf{x})$ tal que sua derivada com respeito ao tempo seja definida negativa nessa região.
- Uma função escalar $V(\mathbf{x})$ que satisfaz esses requisitos é chamada de *função de Lyapunov* para o estado de equilíbrio $\bar{\mathbf{x}}$.

Teoremas de Lyapunov

- Esses teoremas requerem que $V(\mathbf{x})$ seja uma função definida positiva.
- Uma função $V(\mathbf{x})$ é *definida positiva* no estado de espaço $\mathscr{L}$ se, para todo $\mathbf{x}$ em $\mathscr{L}$, satisfaça:
  1. A função $V(\mathbf{x})$ tem derivadas parciais contínuas com respeito aos elementos do vetor de estados $\mathbf{x}$.
  2. $V(\bar{\mathbf{x}}) = 0$.
  3. $V(\mathbf{x}) > 0$ se $\mathbf{x} \neq \bar{\mathbf{x}}$.
- Dado que $V(\mathbf{x})$ é uma função de Lyapunov, de acordo com o teorema 1 o estado de equilíbrio $\bar{\mathbf{x}}$ é estável se

$$\frac{d}{dt}V(\mathbf{x}) \leq 0 \quad para\, \mathbf{x} \in \mathscr{U} - \bar{\mathbf{x}} \tag{9}$$

onde $\mathscr{U}$ é uma pequena vizinhança ao redor de $\bar{\mathbf{x}}$.

Teoremas de Lyapunov

- Além disso, de acordo com o teorema 2, o estado de equilíbrio $\bar{\mathbf{x}}$ é assintoticamente estável se

$$\frac{d}{dt}V(\mathbf{x}) < 0 \quad para\ \mathbf{x} \in \mathscr{U} - \bar{\mathbf{x}} \tag{10}$$

- Os teoremas de Lyapunov podem ser aplicados sem ter que resolver a equação estado-espaço do sistema.
- Infelizmente, os teoremas não indicam como achar uma função de Lyapunov.
- Em muitos casos, a função energia pode servir.
- A inabilidade de achar uma função de Lyapunov adequada não prova a instabilidade do sistema.
- A existência de tal função é suficiente mas não necessária para estabilidade.

Sumário

Modelos de Neurodinâmica

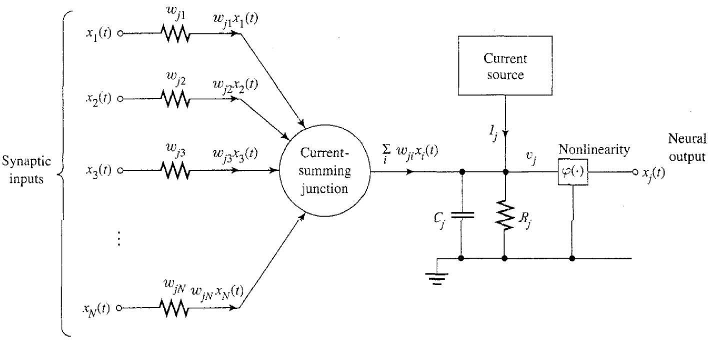
- Os sistemas de interesse possuem quatro características:
  1. *Um grande número de graus de liberdade*. O córtex humano é um sistema distribuído e altamente paralelo com 100 bilhões de neurônios, com cada neurônio modelado por uma ou mais variáveis de estado.
  2. *Não-linearidade*. Essencial para criar uma máquina de computação universal.
  3. *Dissipação*. Caracterizado pela convergência do volume estado-espaço para um dispositivo de pequena dimensionalidade ao passar do tempo.
  4. *Ruído*. Em neurônios reais, ruído na membrana é gerado nas junções sinápticas.

Modelo aditivo

- Considere o modelo dinâmico sem ruído de um neurônio mostrado na figura 2.
- Em termos físicos, os pesos sinápticos w_{j1}, w_{j2}, ..., w_{jN} representam *condutâncias* e as N entradas respectivas $x_1(t)$, $x_2(t)$, ..., $x_N(t)$ representam *potenciais*.
- Essas entradas são aplicadas a uma *junção somadora de corrente* caracterizada como:
  - Baixa resistência de entrada
  - Unidade de ganho de corrente
  - Alta resistência de saída
- Esse modelo pode ser visto como aproximação do circuito de modelo de linha de transmissão distribuída de um neurônio dendrítico biológico.
- A natureza "passa-baixa" do circuito RC pode ser justificado pelo fato de que uma sinapse biológica **é** um filtro passa-baixa.

Modelo aditivo

Figure: Modelo aditivo de um neurônio [2].



Modelo aditivo

- Portanto, a junção age como um nó somatório para as correntes de entrada.
- A corrente total que flui *em direção* ao nó de entrada do elemento não-linear (função de ativação) é

$$\sum_{i=1}^{N} w_{ji} x_i(t) + I_j \tag{11}$$

onde o primeiro termo da soma é devido ao estímulo $x_1(t)$, $x_2(t)$, ..., $x_N(t)$ que agem nos pesos sinápticos (condutâncias) w_{j1}, w_{j2}, ..., w_{jN}, respectivamente, e o segundo termo é devido à fonte de corrente I_j representando um *bias* aplicado externamente.

Modelo aditivo

- Seja $v_j(t)$ o campo local induzido na entrada de uma função de ativação não-linear $\varphi(\cdot)$.
- Expressa-se a corrente total que flui *para fora* do nó de entrada do elemento não-linear:

$$\frac{v_j(t)}{R_j} + C_j \frac{dv_j(t)}{dt} \tag{12}$$

  onde o primeiro termo é devido à resistência de vazamento R_j e o segundo à capacitância de vazamento C_j.
- Da lei de Kirchoff, sabe-se que a corrente total que flui na direção de um nó de um circuito elétrico é zero:

$$C_j \frac{dv_j(t)}{dt} + \frac{v_j(t)}{R_j} = \sum_{i=1}^{N} w_{ji} x_i(t) + I_j \tag{13}$$

Modelo aditivo

- O termo capacitivo $C_j \frac{dv_j(t)}{dt}$ da equação 13 é a forma mais simples de adicionar dinâmica (memória) ao modelo de um neurônio.
- Dado o campo local induzido $v_j(t)$, pode-se determinar a saída do neurônio j usando a relação não-linear

$$x_j(t) = \varphi(v_j(t)) \tag{14}$$

- O modelo RC descrito na equação 13 é chamado de *modelo aditivo* para diferenciar de modelos multiplicativos onde w_{ji} é dependente de x_i.
- *Neurodinâmica clássica*: $x_i(t)$ aplicado ao neurônio j pelo neurônio adjacente i é uma função que varia lentamente no tempo t.

Modelo aditivo

- Agora, considere uma *rede recorrente* consistindo de uma interconexão de N neurônios, onde cada um tem o mesmo modelo matemático das equações 13 e 14.
- Ignorando os atrasos de tempo de propagação interneurônios, define-se a dinâmica da rede pelo seguinte *sistema de equações diferenciais de primeira-ordem acopladas*:

$$C_j \frac{dv_j(t)}{dt} = -\frac{v_j(t)}{R_j} + \sum_{i=1}^{N} w_{ji} x_i(t) + I_j, \quad j = 1, 2, ..., N \qquad (15)$$

- Assume-se que $\varphi(\cdot)$ é uma função contínua e portanto diferenciável. Função logística (mais comum):

$$\varphi(v_j) = \frac{1}{1 + exp(-v_j)}, \quad j = 1, 2, ..., N \qquad (16)$$

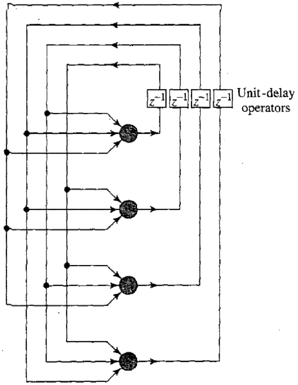
Sumário

Rede de Hopfield

- A *rede* (ou *modelo*) *de Hopfield* consiste de um conjunto de neurônios e um conjunto correspondente de unidades de atraso, formando um *sistema de retro-alimentação (feedback) de loops múltiplos*, como ilustrado na figura 3.
- O número de *loops* de *feedback* é igual ao número de neurônios.
- Basicamente, a saída de cada neurônio é retro-alimentada, via unidade de atraso, a cada um dos outros neurônios da rede.
- Ou seja, não há auto-retro-alimentação.

Rede de Hopfield

Figure: Grafo arquitetural de uma rede de Hopfield com $N = 4$ neurônios [2].



Rede de Hopfield

- Reconhecendo que $x_i(t) = \varphi_i(v_i(t))$, pode-se re-escrever a equação 15

$$C_j \frac{d}{dt} v_j(t) = -\frac{v_j(t)}{R_j} + \sum_{i=1}^{N} w_{ji} \varphi_i(v_i(t)) + I_j, \quad j = 1, ..., N \quad (17)$$

- Faz-se as seguintes assunções:
  1. A matriz de pesos sinápticos é *simétrica*:

$$w_{ji} = w_{ij} \quad para\ todo\ i\ e\ j \quad (18)$$

  2. Cada neurônio tem sua própria ativação *não-linear* - daí o uso de $\varphi(\cdot)$ na equação 17.
  3. A *inversa* da função de ativação não-linear existe, portanto pode-se escrever:

$$v = \varphi_i^{-1}(x) \quad (19)$$

Ganho

- Seja a função sigmoide $\varphi_i(v)$ definida pela tangente hiperbólica:

$$x = \varphi_i(v) = \tanh\left(\frac{a_i v}{2}\right) = \frac{1 - exp(-a_i v)}{1 + exp(-a_i v)} \qquad (20)$$

  que tem uma inclinação de $a_i/2$ na origem:

$$\frac{a_i}{2} = \left.\frac{d\varphi_i}{dv}\right|_{v=0} \qquad (21)$$

  Refere-se a a_i como o *ganho* do neurônio i.

- A relação inversa saída-entrada da equação 19 pode ser re-escrita como

$$v = \varphi_i^{-1}(x) = -\frac{1}{a_i}\log\left(\frac{1-x}{1+x}\right) \qquad (22)$$

Função energia (de Lyapunov)

- A forma *padrão* da relação inversa saída-entrada para um neurônio de ganho unitário

$$\varphi^{-1}(x) = -\log\left(\frac{1-x}{1+x}\right) \tag{23}$$

- Pode-se re-escrever a equação 22 em termos dessa relação padrão

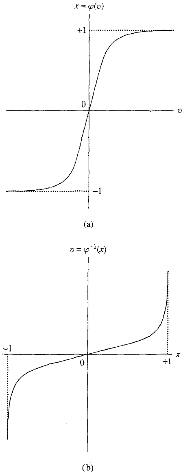
$$\varphi_i^{-1}(x) = \frac{1}{a_i}\varphi^{-1}(x) \tag{24}$$

- A figura 4 mostra a não-linearidade sigmoidal padrão $\varphi(v)$ (a) e sua inversa $\varphi^{-1}(x)$ (b).
- A função energia (de Lyapunov) da rede de Hopfield da figura 3 é

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ji}x_i x_j + \sum_{j=1}^{N}\frac{1}{R_j}\int_0^{x_j}\varphi_j^{-1}(x)dx - \sum_{j=1}^{N} I_j x_j \tag{25}$$

Função energia (de Lyapunov)

Figure: (a) Não-linearidade sigmoidal padrão e (b) sua inversa [2].



Função energia (de Lyapunov)

- A função energia E da equação 25 pode ter um *cenário* complicado com muitos mínimos.
- A dinâmica da rede é descrita por um mecanismo que procura evitar esses mínimos.
- Então, diferenciando E com respeito ao tempo

$$\frac{dE}{dt} = -\sum_{j=1}^{N}\left(\sum_{i=1}^{N} w_{ji}x_i - \frac{v_j}{R_j} + I_j\right)\frac{dx_j}{dt} \tag{26}$$

- Por causa da equação 17, a quantidade dentro dos parênteses da equação 26 é reconhecida como $C_j {}^{dv_j}/_{dt}$. Simplificando

$$\frac{dE}{dt} = -\sum_{j=1}^{N} C_j \left(\frac{dv_j}{dt}\right)\frac{dx_j}{dt} \tag{27}$$

Função energia (de Lyapunov)

- Reconhece-se agora a relação inversa que define v_j em termos de x_j.
- O uso da equação 19 em 27

$$\frac{dE}{dt} = -\sum_{j=1}^{N} C_j \left[\frac{d}{dt}\varphi_j^{-1}(x_j)\right]\frac{dx_j}{dt} = -\sum_{j=1}^{N} C_j \left(\frac{dx_j}{dt}\right)^2 \left[\frac{d}{dx_j}\varphi_j^{-1}(x_j)\right] \tag{28}$$

- Da figura 4b vê-se que a relação saída-entrada inversa $\varphi_j^{-1}(x_j)$ é uma função monotônica crescente da saída x_j. Portanto segue que

$$\frac{d}{dx_j}\varphi_j^{-1}(x_j) \geq 0 \;\; \textit{para todo } x_j \tag{29}$$

Função energia (de Lyapunov)

- Nota-se também que

$$\left(\frac{dx_j}{dt}\right)^2 \geq 0 \;\; para\ todo\ x_j \tag{30}$$

- Assim, todos os fatores que fazem a soma do lado direito da equação 28 são não negativos.
- Em outras palavras, para a função energia E definida na equação 25 tem-se

$$\frac{dE}{dt} \leq 0 \tag{31}$$

- Da definição da equação 25, nota-se que a função E é limitada:
  1. A função energia E é uma função de Lyapunov do modelo de Hopfield contínuo.
  2. O modelo é estável de acordo com o teorema 1 de Lyapunov.

Espaço de estados

- Em outras palavras, a evolução do tempo do modelo de Hopfield contínuo descrito pelo sistema de equações diferenciais de primeira ordem não-lineares (equação 17) representa uma trajetória no espaço de estados, que busca os mínimos da função energia (de Lyapunov) E e para em tais pontos fixos.
- Da equação 28 nota-se também que a derivada ${}^{dE}/_{dt}$ desaparece apenas se

$$\frac{d}{dt}x_j(t) = 0 \quad \textit{para todo } j \tag{32}$$

- Pode-se escrever

$$\frac{dE}{dt} < 0 \quad \textit{exceto em um ponto fixo} \tag{33}$$

Função energia (de Lyapunov)

- A equação 33 provê a base para o seguinte teorema:

Theorem

A função energia (de Lyapunov) E de uma rede de Hopfield é uma função do tempo monotonicamente decrescente.

- A rede de Hopfield é globalmente assintoticamente estável.

Sumário

Hopfield discreto e contínuo

- A rede de Hopfield pode ser operada no modo contínuo ou discreto.
- O modo contínuo é baseado no modelo aditivo.
- O modo discreto é baseado no modelo de McCulloch-Pitts.
- Pode-se estabelecer um relacionamento entre os estados estáveis do modelo contínuo com aqueles do modelo discreto redefinindo a relação entrada-saída para um neurônio, satisfazendo duas características de simplificação:
  1. A saída de um neurônio tem os valores assintóticos:
  $$x_j = \begin{cases} +1 & para\ v_j = \infty \\ -1 & para\ v_j = -\infty \end{cases}$$
  2. O ponto médio da função de ativação de um neurônio fica na origem:
  $$\varphi_j(0) = 0 \tag{34}$$
- Da mesma forma, pode-se fazer o *bias* $I_j = 0$ para todo j.

Hopfield discreto e contínuo

- Na formulação da função energia E para um modelo de Hopfield contínuo, os neurônios podem ter auto-*loops*.
- Um modelo discreto, não precisa de ter auto-*loops*.
- Para simplificar: $w_{jj} = 0$ para todo j em ambos modelos.
- Assim, pode-se redefinir a função energia de um modelo de Hopfield contínuo dado na equação 25 como:

$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1, i\neq j}^{N} w_{ji}x_i x_j + \sum_{j=1}^{N}\frac{1}{R_j}\int_0^{x_j}\varphi_j^{-1}(x)dx \qquad (35)$$

- A função inversa $\varphi_j^{-1}(x)$ é definida pela equação 24. Re-escrevendo a equação 35:

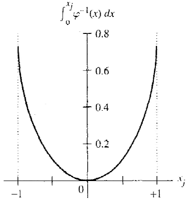
$$E = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1, i\neq j}^{N} w_{ji}x_i x_j + \sum_{j=1}^{N}\frac{1}{a_j R_j}\int_0^{x_j}\varphi^{-1}(x)dx \qquad (36)$$

Hopfield discreto e contínuo

- A integral

$$\int_0^{x_j} \varphi^{-1}(x)dx \qquad (37)$$

tem a forma padrão plotada na figura abaixo [2].



- Se o ganho a_j do neurônio j torna-se infinitamente grande, ou seja, a não-linearidade sigmoidal aproxima-se da forma linear (*hard limiter*), o segundo termo da equação 36 torna-se muito pequeno.

Hopfield discreto e contínuo

- No caso limite ($a_j = \infty$ para todo j) os máximos e mínimos do modelo de Hopfield contínuo tornam-se idênticos aos do modelo discreto.
- Nesse caso, a função energia (de Lyapunov) é definida simplesmente por

$$E = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} w_{ji} x_i x_j \qquad (38)$$

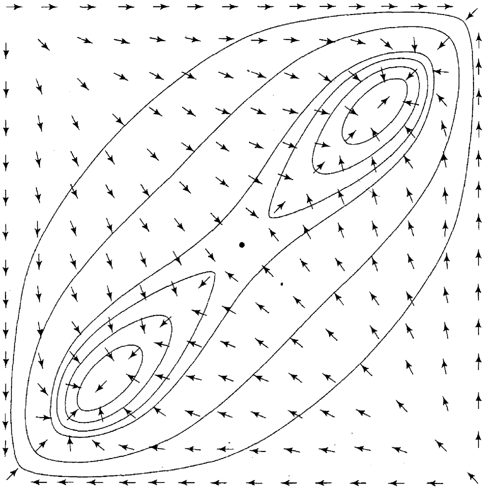
onde o j-ésimo estado do neurônio $x_j = \pm 1$.

- Conclui-se que os únicos pontos estáveis do modelo de Hopfield determinístico, contínuo e de ganho muito alto correspondem aos pontos estáveis do modelo de Hopfield estocástico discreto.

Hopfield discreto e contínuo

- Quando cada neurônio j tem um ganho a_j grande porém finito, o segundo termo da equação 36 torna-se uma contribuição importante à função energia do modelo contínuo.
- A contribuição é grande e positiva perto das superfícies, bordas e cantos do hipercubo unitário que define o espaço de estados do modelo.
- Por outro lado, a contribuição é desprezível nos pontos removidos da superfície.
- A função energia tem seus máximos nos cantos mas os mínimos estão no interior do hipercubo.
- A figura 7 mostra o *mapa de contorno da energia* para um modelo de Hopfield contínuo com dois neurônios.

Hopfield discreto e contínuo [2]



Hopfield discreto e contínuo

- Na figura 7, as saídas dos dois neurônios definem as duas coordenadas do mapa.
- Os cantos inferior esquerdo e superior direito representam mínimos estáveis para o caso limite de ganho infinito.
- Estados instáveis estão nos outros dois cantops.
- As setas mostram o movimento do estado, que geralmente não é perpendicular aos contornos de energia.
- O fluxo para os pontos fixos (isto é, mínimos estáveis) podem ser interpretados como solução à minimização da função energia E definida na equação 25.

Sumário

Hopfield como Memória associativa

- *Memória endereçável pelo conteúdo* (MEC).
- Nessa aplicação, conhece-se os pontos fixos da rede *a priori* pois correspondem aos padrões a serem armazenados.
- Entretanto, os pesos sinápticos da rede que produzem os pontos fixos desejados são desconhecidos.
- O problema é determiná-los.
- A função básica de uma memória endereçável pelo conteúdo é recuperar um padrão armazenado na memória, em resposta à apresentação da versão incompleta ou com ruídos desse padrão.

Hopfield como Memória associativa

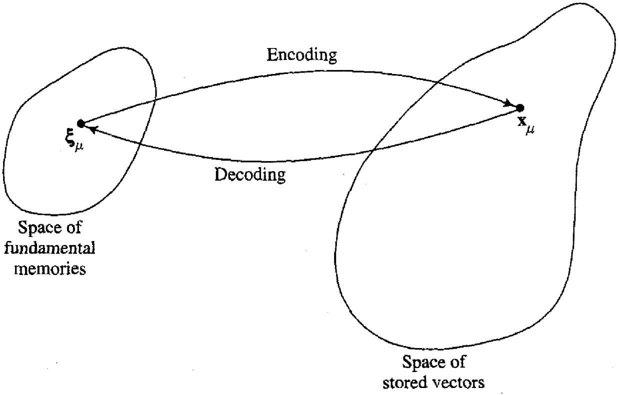
Suponha que um item armazenado na memória seja "H.A. Kramers & G.H. Wannier Physi Rev. 60, 252 (1941)." Uma memória endereçável pelo conteúdo geral seria capaz de recuperar esse item inteiro da memória baseado em informação parcial suficiente. A entrada "& G.H. Wannier (1941)" pode ser suficiente. Uma memória ideal poderia tratar erros e recuperar essa referência mesmo a partir da entrada "Wannier, (1941)." [3]

- A essência da MEC é mapear uma memória fundamental ξ_μ em um ponto fixo (estável) $\mathbf{x}_\mu$ de um sistema dinâmico, como ilustrado na figura 5.
- Matematicamente:

$$\xi_\mu \rightleftharpoons \mathbf{x}_\mu \tag{39}$$

Hopfield como Memória associativa

Figure: Codificação-decodificação realizado por uma rede recorrente [2].



Hopfield como Memória associativa

- Com o modelo de Hopfield usando o neurônio formal de McCulloch-Pitts, cada neurônio tem dois estados determinados pelo nível do campo local induzido que age no mesmo ($x_i = +1$ ou $x_i = -1$).
- Para uma rede com N neurônios, o *estado* da rede é definido pelo vetor

$$\mathbf{x} = [x_1, x_2, ..., x_N]^T \tag{40}$$

- Com $x_i = \pm 1$, o estado do neurônio i representa um *bit* de informação.
- O campo local induzido v_j do neurônio j é definido por

$$v_j = \sum_{i=1}^{N} w_{ji} x_i + b_j \tag{41}$$

onde b_j é um *bias* fixo aplicado externamente ao neurônio j.

Hopfield como Memória associativa

- Assim, o neurônio j modifica seu estado x_j de acordo com a *regra determinística*

$$x_j = \begin{cases} +1 & se\ v_j > 0 \\ -1 & se\ v_j < 0 \end{cases}$$

  re-escrita na forma compacta $x_j = sgn[v_j]$, onde *sgn* é a função *signum*.
- Se $v_j = 0$ assume-se que o neurônio j permanece no estado anterior.
- Duas fases na operação da rede de Hopfield discreta como uma MEC: a fase de armazenamento e a fase de recuperação.

Armazenamento

1. *Fase de armazenamento*:
   - Deseja-se armazenar um conjunto de M vetores N-dimensionais (palavras binárias), denotados por $\{\xi_\mu | \mu = 1, 2, ..., M\}$.
   - Esses vetores, chamados de *memórias fundamentais*, representam os padrões a serem memorizados pela rede.
   - Seja $\xi_{\mu,i}$ o i-ésimo elemento da memória fundamental ξ_μ, classe $\mu = 1, 2, ..., M$.
   - De acordo com a *regra do produto externo* (generalização do *postulado de aprendizado de Hebb*), o peso sináptico do neurônio i para o j é definido por

$$w_{ji} = \frac{1}{N} \sum_{\mu=1}^{M} \xi_{\mu,j} \xi_{\mu,i} \tag{42}$$

   A razão para usar $1/N$ como constante de proporcionalidade é simplificar a descrição matemática da recuperação da informação.

Armazenamento

1. *Fase de armazenamento* (cont.):
   - A regra de aprendizado da equação 42 é uma computação "única".
   - Na operação normal da rede de Hopfield, faz-se

$$w_{ii} = 0 \quad \textit{para todo } i \tag{43}$$

     que significa que os neurônios não têm auto-retro-alimentação.
   - Seja **W** a matriz de pesos sinápticos N-por-N da rede.
   - Pode-se combinar as equações 42 e 43 na forma matriz:

$$\mathbf{W} = \frac{1}{N} \sum_{\mu=1}^{M} \xi_\mu \xi_\mu^T - M\mathbf{I} \tag{44}$$

     onde $\xi_\mu \xi_\mu^T$ representa o produto externo do vetor ξ_μ com ele mesmo e **I** denota a matriz identidade.

Armazenamento

1. *Fase de armazenamento* (cont.):
   - A partir dessas equações de definição dos pesos sinápticos/matriz de pesos, pode-se reconfirmar o seguinte:
     - A saída de cada neurônio na rede é retro alimentado a todos os outros neurônios.
     - Não há auto-retro-alimentação na rede (isto é, $w_{ii} = 0$).
     - A matriz de pesos da rede é simétrica, como mostrado por (veja equação 18):

$$\mathbf{W}^T = \mathbf{W} \tag{45}$$

Recuperação

2 *Fase de recuperação*:

- Durante a fase de recuperação, um vetor N-dimensional ξ_{sonda}, é imposto à rede de Hopfield como seu estado.
- O vetor sonda tem elementos iguais a ± 1.
- Representa uma versão incompleta ouy ruidosa de uma memória fundamental da rede.
- A recuperação da informação acontece de acordo com uma *regra dinâmica* na qual cada neurônio j da rede *aleatoriamente*, mas numa taxa fixa, examina o campo local induzido v_j (incluindo qualquer *bias* b_j não nulo) aplicado a ele.
- Se, em algum instante de tempo, v_j é maior que zero, o neurônio j mudará seu estado para +1 ou permanecerá em seu estado.
- Se v_j é menor que zero, j mudará seu estado para -1 ou permanecerá em seu estado.
- Se $v_j = 0$, j permanecerá em seu estado anterior.

Condição de estabilidade

2 *Fase de recuperação* (cont.):

- A atualização de estados é portanto determinística, mas a seleção de um neurônio para realizar a atualização é aleatória.
- O procedimento de atualização *assíncrono* (serial) é continuado até que não haja mais mudanças.
- Ou seja, começando pelo vetor sonda **x**, a rede produz um vetor de estados invariante no tempo **y** cujos elementos individuais satisfazem a *condição de estabilidade* ou *condição de alinhamento*:

$$y_i = sgn\left(\sum_{i=1}^{N} w_{ji} y_i + b_j\right), \quad j = 1, 2, ..., N \tag{46}$$

ou, na forma matricial,

$$\mathbf{y} = sgn(\mathbf{W}\mathbf{y} + \mathbf{b}) \tag{47}$$

onde **W** é a matriz de pesos sinápticos da rede e **b** é o *vetor bias* aplicado externamente.

Condição de estabilidade

2 *Fase de recuperação* (cont.):

- O vetor de estados **y** que satisfaz a condição de estabilidade é chamado de *estado estável* ou *ponto fixo* do espaço de estados do sistema.
- Pode-se afirmar que a rede de Hopfield sempre convergirá para um estado estável, quando a operação de recuperação for realizada *assincronamente*.

Sumário

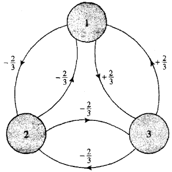
Exemplo [2]

- Para ilustrar o comportamento emergente do modelo de Hopfield, considere a rede da figura 3a, que consiste de três neurônios.
- A matriz de pesos da rede é

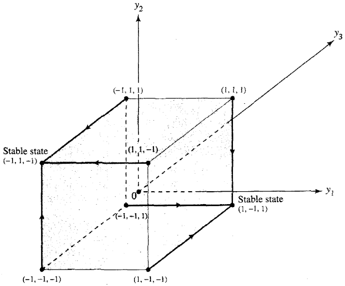
$$\mathbf{W} = \frac{1}{3}\begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix} \tag{48}$$

  que é legítima, pois satisfaz as condições das equações 43 e 45.
- O *bias* aplicado a cada neurônio é zero.
- Com três neurônios, há $2^3 = 8$ possíveis estados.
- Desses estados, apenas (1,-1,1) e (-1,1,-1) são estados estáveis, pois satisfazem a condição de alinhamento da equação 47.

Exemplo [2]


1
2
3
$-\frac{2}{3}$
$+\frac{2}{3}$

(a)


y_2
y_3
y_1
(−1, 1, 1)
(1, 1, 1)
(1, 1, −1)
Stable state
(−1, 1, −1)
0
(−1, −1, 1)
Stable state
(1, −1, 1)
(−1, −1, −1)
(1, −1, −1)

Exemplo [2]

- Para o vetor de estados (1, -1, 1) tem-se

$$\mathbf{Wy} = \frac{1}{3}\begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix}\begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} +4 \\ -4 \\ +4 \end{bmatrix} \tag{49}$$

- Aplicando o *hard limiter*:

$$sgn[\mathbf{Wy}] = \begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix} = \mathbf{y} \tag{50}$$

- Similarmente, para o vetor de estados (-1,1,-1):

$$\mathbf{Wy} = \frac{1}{3}\begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix}\begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} -4 \\ +4 \\ -4 \end{bmatrix} \tag{51}$$

- Aplicando o *hard limiter*:

$$sgn[\mathbf{Wy}] = \begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix} = \mathbf{y} \tag{52}$$

Exemplo [2]

- O mapa do fluxo da figura 3b exibe simetria com repeito aos dois estados estáveis da rede, que é o resultado de deixar um neurônio em seu estado prévio se o campo local induzido que age nele é exatamente zero.
- Se a rede da figura 3a está no estado inicial (1,1,1), (-1,-1,1) ou (1,-1,-1), ela convergirá para o estado estável (1,-1,1) após uma iteração.
- Se o estado inicial for (-1,-1,-1), (-1,1,1) ou (1,1,-1), ela convergirá para o segundo estado estável (-1,1,-1).
- A rede possui duas memórias fundamentais, (1,-1,1) e (-1,1,-1), representando os dois estados estáveis.
- A aplicação da equação 44 leva à matriz de pesos sinápticos

$$\mathbf{W} = \frac{1}{3}\begin{bmatrix} +1 \\ -1 \\ +1 \end{bmatrix}\begin{bmatrix} +1, & -1, & +1 \end{bmatrix} + \frac{1}{3}\begin{bmatrix} -1 \\ +1 \\ -1 \end{bmatrix}\begin{bmatrix} -1, & +1, & -1 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{53}$$

Exemplo [2]

$$\mathbf{W} = \frac{1}{3}\begin{bmatrix} 0 & -2 & +2 \\ -2 & 0 & -2 \\ +2 & -2 & 0 \end{bmatrix} \quad (54)$$

- Que correspondem aos pesos sinápticos mostrados na figura 3a.
- Examinando o mapa de fluxos da figura 3b, a capacidade de correção de erro da rede de Hopfield pode ser vista:
  1. Se o vetor sonda ξ_{sonda} aplicado à rede for igual a (-1,-1,1), (1,1,1) ou (1,-1,-1), a saída resultante é a memória fundamental (1,-1,1). Cada um desses valores representa um único erro, comparado com o padrão armazenado.
  2. Se o vetor sonda ξ_{sonda} aplicado à rede for igual a (1,1,-1), (-1,-1,-1) ou (-1,1,1), a saída resultante é a memória fundamental (-1,1,-1). De novo, cada um desses valores representa um único erro, comparado com o padrão armazenado.

Bibliografia I

[1] W. R. Ashby
Design for a brain.
2nd. edition. New York: Wiley, 1960.

[2] S. Haykin
Neural networks - a comprehensive foundation.
2nd. edition. Prentice Hall, 1999.

[3] J. J. Hopfield
Neural networks and physical systems with emergent collective computational abilities.
Proceedings of the National Academy of Sciences, USA, vol. 79, pp. 2554–2558, 1982.