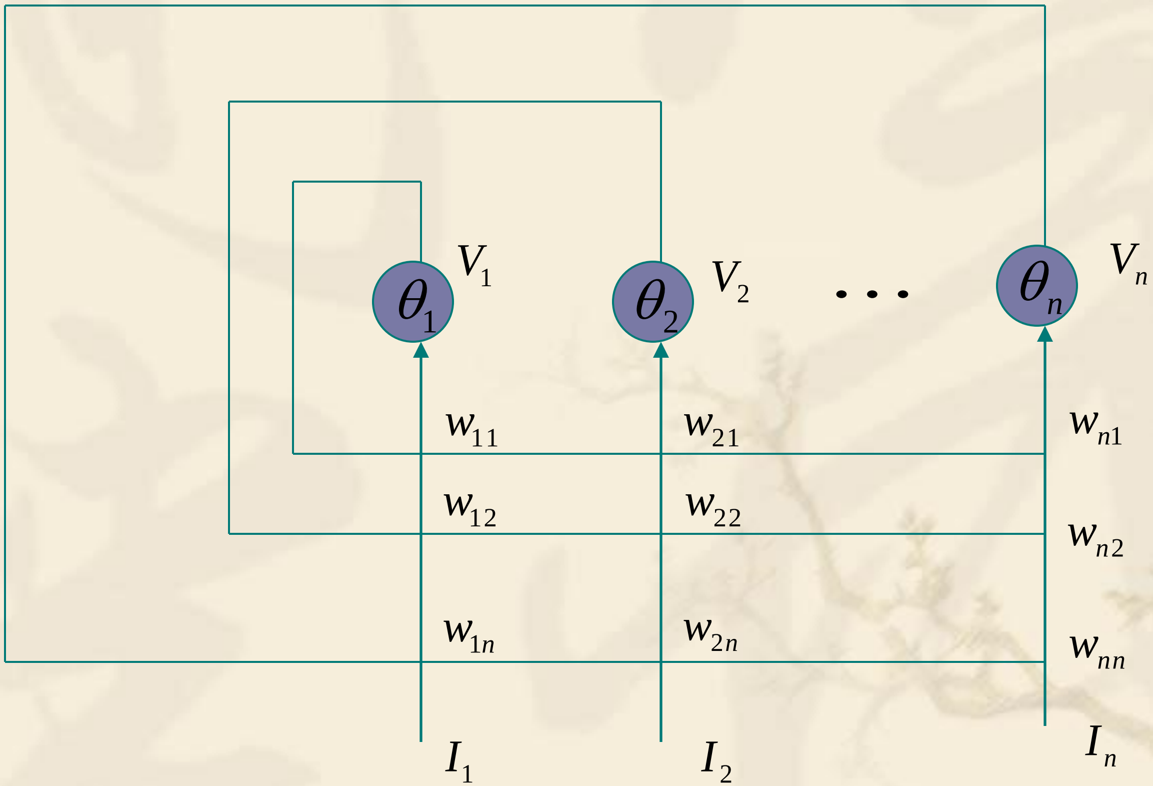
第3章 反馈型人工神经网络

- 动力系统的稳定性
- 离散Hopfield网络的稳定性
- 离散Hopfield网络与联想记忆
- 连续Hopfield网络的稳定性
- 连续Hopfield网络与与组合优化

第3章 反馈型人工神经网络

A. 网络的模型与分类

反馈式网络的结构如下图：



第3章 反馈型人工神经网络

这种网络中的每个神经元的输出都与其它神经元的输入相连接，其输入输出关系如下：

$$\begin{cases} s_i = \sum_{j=0}^{n} w_{ij} V_j + I_i, \qquad (w_{i0} = \theta_i,\ V_0 \equiv -1) \\ x_i = g(s_i) \qquad (x_i = u_i, \text{为神经元的状态}) \\ V_i = f(x_i) \qquad \text{（神经的输出）} \end{cases}$$

注意：在反馈网络中，神经元这一时刻的状态（或输出）与上一时刻的状态（或输出）有关，与前馈网络不同。

函数的选取： $x_i = g(s_i) = s_i$

$$f(x_i) = Sgn(x_i) \leftrightarrow \text{离散型反馈网络}$$

第3章 反馈型人工神经网络

如果通过下列方程给出输入输出关系，则称网络为连续型反馈网络：

$$
\begin{cases}
\dfrac{dx_i}{dt} = -\dfrac{1}{\tau_i} x_i + s_i \\
V_i = f(x_i), \quad f(\bullet)\text{为一个连续单调上升的有界函数。}
\end{cases}
$$

演化过程与轨迹：令 $I = (I_1, I_2, \cdots, I_n)^T \in R^n$

$$X = (x_1, x_2, \cdots, x_n)^T \in R^n, \qquad V = (V_1, V_2, \cdots, V_n)^T \in R^n$$

则在 R^n 空间上状态点 $X(t)$（或输出 $V(t)$）随时间则形成一条轨迹，从这一轨迹的趋势可以判断网络系统的稳定性。对于离散型网络，轨迹是跳跃变化的；对于连续型网络，轨迹是连续变化的。（$X(t) \leftarrow I(t)$ (或$I(t_0)$)）

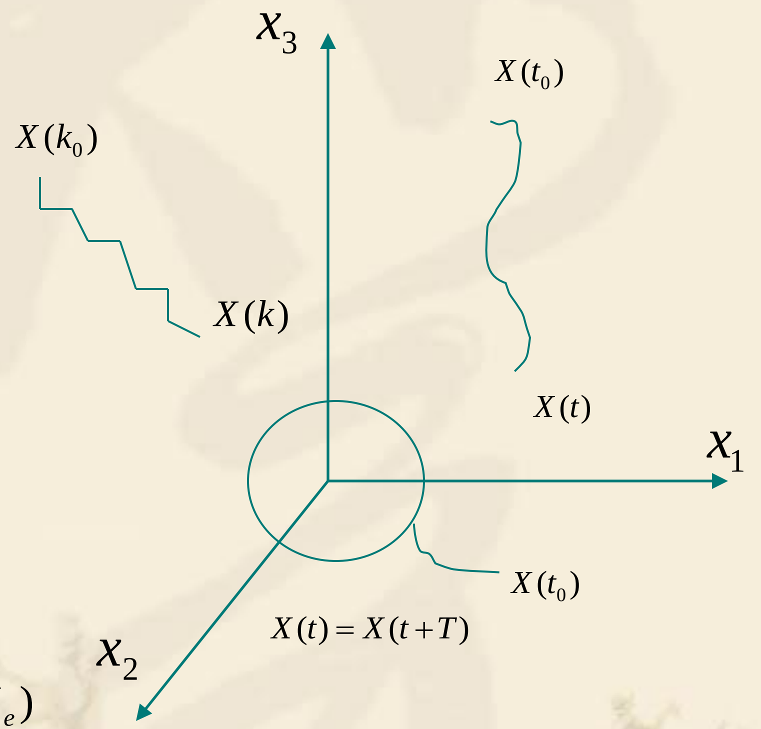
B. 状态轨迹的分类

(1). 稳定性轨迹与稳定点

$$X(t_0) \to X(t) = X(t+\Delta t) = \cdots$$

称网络收敛到一个稳定点

（或平衡点）。



(i). 吸引子 X_e

存在一个 X_e 的 δ 邻域 $N_\delta(X_e)$

当 $X(t_0) \in N_\delta(X_e)$，必有：

$$X(t) \to X_e \quad (t \to \infty)$$

同时，最大的 $N_\delta(X_e)$ 被称为吸引域。

(ii). 不稳定的平衡点 X_{en}

即没有吸引域的平衡点----鞍点或拐点。

(iii). 网络的解----设计者所希望的稳定点。

(iv). 伪稳定点----设计者不希望的稳定点。

(2). 极限环

离散情况：$X(k) = X(k+m)$----m步极限环

连续情况：$X(t) = X(t+T)$----周期为T的极限环

(3). 混沌(Chaos)：不稳定，不发散。

(4). 发散：趋向无穷远点。

第3章 反馈型人工神经网络

C. 联想记忆与组合优化

(1). 联想记忆：从部分信息或含有错误的信息恢复出全部信息模式；

(2). 组合优化：

吸引子 $\leftrightarrow$ 网络能量函数极小点

$\updownarrow$

目标函数

可应用于实际问题，如TSP，二分图问题等。

第3章 反馈型人工神经网络

D. 网络的设计目标

(1). 网络能够达到稳定（收敛）；

(2). 网络的稳定点是实际问题的解（如联想记忆的模式或组合优化的解）；

(3). 吸引域尽可能的大。

3.1 离散Hopfield网络(DHNN)

DHNN是一种单层的，输入输出为二值的反馈式神经网络，在联想记忆和组合优化方面有着重要的应用。当输入向量 I 作为一个初始值时，网络通过反馈演化最终达到稳定状态。

网络的数学描述：

$N=(W,\theta)$, $W=(w_{ij})_{n\times n}$, $w_{ii}=0$ (零对角)，$w_{ij}=w_{ji}$（对称性）

1. 基本公式

令 $V_0(t)\equiv -1$, $w_{i0}=\theta_i$，则

$$x_i(t)=s_i(t)=\sum_{j=0}^{n} w_{ij}V_j(t)+I_i(t), \quad i=1,2,\cdots,n$$

3.1 离散Hopfield网络(DHNN)

$$V_i(t+1) = Sgn(x_i(t)) = \begin{cases} 1, & if \quad x_i(t) \geq 0; \\ -1(0), & if \quad x_i(t) < 0. \end{cases}$$

网络初始值的赋予：

(i). 在 $t = t_0$ 时， 令 $V_i(t_0) = 0$， 则 $x_i(t_0) = I(t_0) - \theta_i$

根据上面的演化公式即可得到

$$V(t_0) \to V(t_0 + 1) \to \cdots$$

(ii). 直接赋予初始状态：

$$V(t_0) = [V_1(t_0), V_2(t_0), \cdots, V_n(t_0)]^T \in \{\pm 1\}^n$$

网络便依此进行演化。这时输入 I 可以看作一个初始推动力（ $I(t) \equiv 0$ ）。

3.1 离散Hopfield网络(DHNN)

网络的运行方式：

(i).异步（或串行）方式

每一时刻，只选取一个神经元进行演化，即若所选取的神经元为第 i 个，则

$$\begin{cases} V_i(t+1) = Sgn(x_i(t)) \\ V_j(t+1) = V_j(t), \quad j \neq i \end{cases}$$

神经元的选取方式可以是随机的，也可以是按某个确定性规则进行。

(ii).同步（或并行）方式

每一时刻，所有神经元都进行演化：

$$V_i(t+1) = Sgn(x_i(t)), \quad j = 1, 2, \cdots, n.$$

3.1 离散Hopfield网络(DHNN)

2. 稳定点

网络从一个初始状态 $t=t_0$， $X(t_0)$ 开始进行演化，经过一个有限的时间，网络的输出不在发生变化，这些输出就为网络的稳定点，在数学上表示为：

$$V(t+1)=V(t)$$

$$V_i(t+1)=Sgn(x_i(t))=Sgn(\sum_{j=0}^{n}w_{ij}V_j(t)+I_i(t))=V_i(t),\quad i=1,2,\cdots,n$$

即得： $V_i(t)x_i(t)\geq 0,\quad i=1,2,\cdots,n$

$$x_i(t)=\sum_{j=0}^{n}w_{ij}V_j(t)+I_i(t)$$

3.1 离散Hopfield网络(DHNN)

假若有一组二元样本 $\{V^1, V^2, \cdots, V^m\}, V^i \in \{\pm 1\}^n$ ，我们要设计一个单层二元反馈网络 $N = (W, \theta)\ \ (\theta = 0)$ （DHNN的一种推广形式），使得这些样本都成为网络的稳定点。则需要：

$$WV^k = \beta^k V^k, \qquad k = 1, 2, \cdots, m$$

$$W \in R^{n \times n}, \qquad \beta^k = Diag[\beta_1^k, \beta_2^k, \cdots, \beta_n^k], \ \ \beta_i^k > 0$$

$$W[V^1, V^2, \cdots, V^m] = [\beta^1 V^1, \beta^2 V^2, \cdots, \beta^m V^m]$$

$$WZ = B, \quad Z = [V^1, V^2, \cdots, V^m], \ \ B = [\beta^1 V^1, \beta^2 V^2, \cdots, \beta^m V^m]$$

$$Z \in R^{n \times m}, \qquad B \in R^{n \times m}, \qquad W \in R^{n \times n}$$

$$Z^T\ W^T = Z^T W = B^T$$

3.1 离散Hopfield网络(DHNN)

令：
$$Z^T=\begin{bmatrix} V_1^1 & V_2^1 & \cdots & V_n^1 \\ V_1^2 & V_2^2 & \cdots & V_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ V_1^m & V_2^m & \cdots & V_n^m \end{bmatrix},\quad W_j=\begin{bmatrix} w_{1j} \\ w_{2j} \\ \vdots \\ w_{nj} \end{bmatrix},\quad B_j=\begin{bmatrix} \beta_j^1 V_j^1 \\ \beta_j^2 V_j^2 \\ \vdots \\ \beta_j^m V_j^m \end{bmatrix}$$

从方程： $Z^T W_j = B_j \quad \Rightarrow \quad W_j=(Z^T)^* B_j$

其中 $(Z^T)^*$ 是 Z^T 的伪逆。

当 $m \le n$ ，且 Z^T 是满秩的，即 $V^1, V^2, \cdots V^m$ 线性无关，那么任取一组 $\beta_j^i > 0$，必存在 W 使得 $V^1, V^2, \cdots V^m$ 为网络 $N=(W,0)$ 的平衡点。

注意：这里存在两个问题：（1）吸引域问题；（2）稳定性问题。

3.1 离散Hopfield网络(DHNN)

3. DHNN的稳定性

所谓网络的稳定性是指它能够从任意一个初始状态演化到一个稳定状态而停止下来，一般稳定系统是由一个能量函数来控制的，下面借用铁磁材料中的哈密顿函数来引入DHNN的能量函数。

在铁磁体中，分子自旋只有两个方向，表示为：

哈密顿函数为： $P_i \in \{\pm 1\}, i = 1, 2, \cdots, n$

$$H = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} J_{ij} P_i P_j - \sum_{i=1}^{n} H_i P_i$$

3.1 离散Hopfield网络(DHNN)

这里 J_{ij} 表示第 i 个分子与第 j 个分子之间的作用，并且是对称的，H_i 是外加的随机场，整个物质的相互作用使得哈密顿函数 H 达到最小。

参照哈密顿函数，我们引入DHNN的能量函数如下：

$$E(V(t)) = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}V_i(t)V_j(t) + \sum_{i=1}^{n}\theta_i V_i(t)$$

显然能量函数是有界的：

$$|E(V(t))| \le \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}|w_{ij}| + \sum_{i=1}^{n}|\theta_i|$$

3.1 离散Hopfield网络(DHNN)

(i).异步方式的稳定性

(a).当网络在异步方式下，满足连接权对称和主对角线元素为零，即 $w_{ij}=w_{ji},\ w_{ii}=0$，若其能量函数能够单调下降，网络必然下降到能量的一个极小点而停止，故网络是稳定的。

定理3.1 DHNN网络在异步方式是稳定的，即从任意初始状态 $V(t_0)$ 出发，网络必然收敛到某个稳定状态 V^*。

证明： 令

$$\Delta V_i(t)=V_i(t+1)-V_i(t)=\begin{cases}0 & if\ \ V_i(t+1)=V_i(t)\\ 2 & if\ \ V_i(t+1)=1, V_i(t)=-1\\ -2 & if\ V_i(t+1)=-1, V_i(t)=1\end{cases}$$

3.1 离散Hopfield网络(DHNN)

根据异步方式的定义，每个时刻只有一个神经元进行演化，现设第 k 个神经元演化，其它神经元不变，能量公式为：

$$\Delta E = E(V(t+1)) - E(V(t))$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}V_i(t+1)V_j(t+1) + \sum_{i=1}^{n}\theta_i V_i(t+1) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}V_i(t)V_j(t) - \sum_{i=1}^{n}\theta_i V_i(t)$$

$$= -\frac{1}{2}\sum_{j=1}^{n} w_{kj}V_j(t)\Delta V_k(t) - \frac{1}{2}\sum_{i=1}^{n} w_{ik}V_i(t)\Delta V_k(t) + \theta_k \Delta V_k(t)$$

$$= -(\sum_{j=1}^{n} w_{kj}V_j(t))\Delta V_k(t) + \theta_k \Delta V_k(t)$$

$$= -\Delta V_k(t)(\sum_{j=1}^{n} w_{kj}V_j(t) - \theta_k)$$

$$= -(V_k(t+1) - V_k(t))(\sum_{j=1}^{n} w_{kj}V_j(t) - \theta_k) \le 0$$

3.1 离散Hopfield网络(DHNN)

由于能量函数随着时间递减，而能量函数又是有界，则必然达到一个极小值。那么该能量函数极小值所对应的神经元输出或着是一个稳定点，或着是一个极限环。下面证明极限环是不存在的。实际上在这种情况下，神经元输出有着实质性变化，但有$\Delta E(t)=0$。在异步方式下，当$\Delta E(t)=0$时，只能有：

$$V_k(t)=1 \quad \Rightarrow \quad V_k(t+1)=-1 \quad \Rightarrow \quad V_k(t+2)=1$$

这时会出现 $\Delta E(t+1)<0$，得出 t 及 $t+1$ 时刻能量函数还未达到极小值，与假设矛盾。故不会出现极限环的情况。因此，网络在异步方式下必收敛到一个稳定状态。 证毕

3.1 离散Hopfield网络(DHNN)

(b).连接权对称、主对角非负的稳定性

定理3.2 在DHNN中，若使得 $w_{ii} \geq 0$ ，在异步方式依然是稳定的。

证明：在这种情况下，我们有：

$$\Delta E = E(V(t+1)) - E(V(t))$$

$$= -\frac{1}{2}\sum_{j\neq k} w_{kj}V_j(t)\Delta V_k(t) - \frac{1}{2}\sum_{i\neq k} w_{ik}V_i(t)\Delta V_k(t) + \theta_k \Delta V_k(t)$$

$$-\frac{1}{2}w_{kk}(V_k^2(t+1) - V_k^2(t))$$

$$= -\sum_{j\neq k} w_{kj}V_j(t)\Delta V_k(t) - w_{kk}V_k(t)\Delta V_k(t) + w_{kk}V_k(t)\Delta V_k(t)$$

$$-\frac{1}{2}w_{kk}(V_k^2(t+1) - V_k^2(t)) + \theta_k \Delta V_k(t)$$

3.1 离散Hopfield网络(DHNN)

$$= -(\sum_{j=1}^{n} w_{kj} V_j(t)) \Delta V_k(t) + \theta_k \Delta V_k(t) + w_{kk} V_k(t) V_k(t+1)$$

$$- w_{kk} V_k(t) V_k(t) - \frac{1}{2} w_{kk} V_k^2(t+1) + \frac{1}{2} w_{kk} V_k^2(t)$$

$$= -\Delta V_k(t)(\sum_{j=1}^{n} w_{kj} V_j(t) - \theta_k) - \frac{1}{2} w_{kk} (V_k(t+1) - V_k(t))^2$$

$$\leq 0$$

故网络必然收敛到一个稳定点。 证毕

定理3.3(Ma, 1997) 在随机异步方式下，连接权非对称、非负，主对角元素为零 $(w_{ij} \geq 0,\ w_{ii} = 0)$ 的DHNN是稳定的。

证明见： Jinwen Ma, The stability of the generalized Hopfield networks in randomly asynchronous mode, **Neural Networks**, vol.10(1997), No.6,pp:1109-1116

3.1 离散Hopfield网络(DHNN)

(ii).同步方式的稳定性

定理3.4 在并行方式下，DHNN（ $w_{ij}=w_{ji}$ ）收敛到一个稳定点，或2步的极限环。

证明：采用向量、矩阵符号，并引入：

$$E(t)=-\frac{1}{2}\sum_i\sum_j w_{ij}V_i(t+1)V_j(t)+\frac{1}{2}\sum_i\theta_i(V_i(t+1)+V_i(t))$$

$$=-\frac{1}{2}V^T(t+1)WV(t)+\frac{1}{2}\theta^T[V(t+1)+V(t)]$$

$$V(t)\in R^n,\quad W\in R^{n\times n},\quad \theta\in R^n$$

则有

3.1 离散Hopfield网络(DHNN)

$$
\begin{aligned}
\Delta E(t) &= E(t) - E(t-1) \\
&= -\frac{1}{2}V^T(t+1)WV(t) + \frac{1}{2}\theta^T[V(t+1)+V(t)] \\
&\quad + \frac{1}{2}V^T(t)WV(t-1) - \frac{1}{2}\theta^T[V(t)+V(t-1)] \\
&= -\frac{1}{2}[V^T(t)W][V(t+1)-V(t-1)] + \frac{1}{2}\theta^T[V(t+1)-V(t-1)] \\
&= -\frac{1}{2}[V^T(t)W-\theta^T][V(t+1)-V(t-1)] \\
&= -\frac{1}{2}[X(t)]^T[V(t+1)-V(t-1)] \\
&\le 0
\end{aligned}
$$

3.1 离散Hopfield网络(DHNN)

若 $V(t)=V(t+1)=V(t-1)$ ，则 $\Delta E(t)=0$ 且网络达到稳定点；

若 $V(t)\neq V(t+1)=V(t-1)$ ，则 $\Delta E(t)=0$ 且网络达到2步的极限环。 证毕

推论3.1 (1). 若 W 是正定的， $\theta=0$ ，则网络必达到稳定收敛；

(2). 若 W 是负定的， $\theta=0$ ，则网络必是周期振荡，极限环的步长为2。

作业：证明推论3.1 。并讨论异步方式的情况。