3.1 离散Hopfield网络(DHNN)

4. 外积型DHNN权的设计

(i).外积型设计权的基本公式($\theta = 0$)

设 $S=\{V^1,V^2,\cdots,V^m\}$ 是一组要被记忆的二元样本，其中 $V^k=[v_1^k,v_2^k,\cdots,v_n^k]^T\in\{\pm1\}^n$， $k=1,2,\cdots,m$ 。我们根据Hebb学习律进行如下设计DHNN的权：

$$\begin{cases} w_{ij}=\eta\sum_{k=1}^{m} v_i^k v_j^k \quad if \quad i\neq j \\ w_{ii}=0 \quad if \quad i=j \end{cases}$$

当 $v_i^k\in\{0,1\}$，我们可以用一个 $\{0,1\}$ 到 $\{\pm1\}$ 的一个变换得： $w_{ij}=\eta\sum_{k=1}^{m}(2v_i^k-1)(2v_j^k-1)$

3.1 离散Hopfield网络(DHNN)

用矩阵表示为：

$$W=\eta([V^1,V^2,\cdots,V^m][V^1,V^2,\cdots,V^m]^T-mU)$$

其中U为单位矩阵。显然有：$w_{ij}=w_{ji},w_{ii}=0$

(ii).外积法中稳定点的性质

若S中的样本模式两两相互正交，则按外积法形成的权矩阵W满足（$\eta=1$）：

$$WV^k=(n-m)V^k \qquad k=1,2,\cdots,m$$

从而，在正交情况下，当$m<n$，每个样本模式都被网络所记忆，即为网络的稳定点。

3.1 离散Hopfield网络(DHNN)

如果 S 中的样本模式不是正交的，那么一般无法保证每个样本模式为稳定点，但当 m 相对于 n 较小时，样本是随机的，则能够以较大的概率保证每个样本模式为稳定点，这就是记忆容量问题。从表达式上看，我们有

$$v'^k_j = Sgn(\sum_{i=1}^{n} w_{ij} v_i^k) = Sgn(s_j + n_j)$$

其中 $s_j = n v_j^k$， $n_j = \sum_{l=1, l\neq k}^{m} \sum_{i=1}^{n} v_i^l v_j^l v_i^k$

并且噪声部分 n_j 是一个零均值并方差为

$\sigma^2 = (m-1)n$ 的随机变量。如果 $m < n+1$ ，网络以很高的概率能够记忆 $V^1, V^2, \cdots, V^m$ 。

3.1 离散Hopfield网络(DHNN)

1）若V^k 是Hopfield网络的一个稳定点，且 $WV^k \neq 0$ ，则 $-V^k$ 也是网络的稳定点。

2）用 $d_H(V^l,V^k)$ 表示两个向量模式之间的汉明距离，若 $d_H(V^l,V^k)=1$ 或 $d_H(V^l,V^k)=n-1$ ，那么当V^l 是网络的稳定点，V^k 则不是网络的稳定点。

3）设m个所记忆样本模式 $V^1,V^2,\cdots,V^m$ 满足：

$$\alpha n \leq d_H(V^i,V^j) \leq (1-\alpha)n, \qquad \forall i \neq j, \quad \text{其中} 0 < \alpha < \frac{1}{2}$$

3.1 离散Hopfield网络(DHNN)

若满足 $m \le M = \left\lfloor \frac{1}{1-2\alpha} \right\rfloor$，则当n足够大时，$V^1, V^2, \cdots, V^m$ 为外积型设计的权所形成的网络的稳定点。

4）给定一组正交样本模式 $V^1, V^2, \cdots, V^m$，以及另一个模式 X，如果有 $d_H(X, V^l) < \frac{n-m}{2m}$，则 X 将被吸引到 V^l。

5）若 $V^1, V^2, \cdots, V^m$ 是网络的稳定点，X 是由 $V^1, V^2, \cdots, V^m$ 的线性组合形成的模式，那么 X 也是网络的稳定点。（给网络带来了许多的伪稳定点。）

3.1 离散Hopfield网络(DHNN)

例3.1 考虑 $n=5, m=3$ 的情况，其中：

$$V^1=[1,1,1,1,1]^T \quad V^2=[1,-1,-1,1,-1]^T \quad V^3=[-1,1,-1,-1,-1]^T$$

根据外积规则得：

$$W=\begin{bmatrix} 0 & -1 & 1 & 3 & 1 \\ -1 & 0 & 1 & -1 & 1 \\ 1 & 1 & 0 & 1 & 3 \\ 3 & -1 & 1 & 0 & 1 \\ 1 & 1 & 3 & 1 & 0 \end{bmatrix}, \quad \theta=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

网络有四个稳定点 $V^1, V^2, V^3, V^4=-V^2$ 。

网络在两种方式下的运行情况：

3.1 离散Hopfield网络(DHNN)

在串行方式下：

10个初始状态收敛到V^1 ；

8个初始状态收敛到V^2 ；

8个初始状态收敛到V^3 ；

6个初始状态收敛到V^4 ；

在并行方式下：

收敛到 V^1,V^2,V^3,V^4 的初始状态数分别为8，1，2，1。其余20初始状态都收敛到极限环。

3.1 离散Hopfield网络(DHNN)

外积设计的特点：

(a). 存在伪稳定点，所存储的样本也可能不稳定点；

(b). V^l 是稳定点，其反模式$-V^l$也为稳定点；

©. 所要求记忆的样本之间的汉明距离越大，被记忆的样本的数量越大；

(d). 对于正交的记忆样本，每个样本不仅能够收敛到自己，而且还存在一个吸引域，其半径为：

$$R_H = [\frac{n-m}{2m}]$$

3.1 离散Hopfield网络(DHNN)

5.其它权设计方法

(i). 伪逆法

考虑 m 个样本模式 $V^1,V^2,\cdots,V^m$，将它们按顺序排成一个矩阵：

$$X=[V^1,\cdots,V^k,\cdots,V^m]\in R^{n\times m}$$

对应于 X，设神经元根据联想记忆所要求的对应输出可以写成一个矩阵 $Y=Sgn(WX)$。我们直接设计：$W=YX^*$，$\theta=0$，其中 X^*为 X 的伪逆，且满足：

$$X^*=(X^TX)^{-1}X^T$$

在联想记忆中，我们希望 $Y=X$，即每个样

3.1 离散Hopfield网络(DHNN)

本模式都是网络的稳定点，则

$$W=Y(X^TX)^{-1}X^T=X(X^TX)^{-1}X^T$$

这时所得到的 W 是对称的。当 $Y\neq X$ ，所形成的联想称为异联想，即从一个模式按内容地址形式恢复出另外一个模式。显然，自联想和异联想都可以应用于模式识别中。

(ii).正交化的权设计方法

对于样本模式 $V^1,V^2,\cdots,V^m$，正交化权的设计的计算公式如下：

3.1 离散Hopfield网络(DHNN)

令 $Y=[V^1-V^m,V^2-V^m,\cdots,V^{m-1}-V^m]$ 并对其进行奇异值分解得 $Y=PAQ^T$

$$Y=[Y_1,Y_2,\cdots,Y_{m-1}]^T \qquad \in R^{n\times(m-1)}$$

$$P=[P_1,P_2,\cdots,P_n]^T \qquad \in R^{n\times n}$$

$$Q=[Q_1,Q_2,\cdots,Q_{m-1}]^T \qquad \in R^{(m-1)\times(m-1)}$$

$$A=diag[\lambda_1,\cdots,\lambda_k,0,\cdots,0]$$

设 $\{P_1,P_2,\cdots,P_k\}$ 为 Y 的列向量空间的正交基，而 $\{P_{k+1},P_{k+2},\cdots,P_n\}$ 为其补空间的正交基，则

$$W^+=[w_{ij}]^+_{n\times n}=\sum_{i=1}^{k}P_iP_i^T \qquad \in R^{n\times n}$$

$$W^-=[w_{ij}]^-_{n\times n}=\sum_{i=k+1}^{n}P_iP_i^T \qquad \in R^{n\times n}$$

3.1 离散Hopfield网络(DHNN)

网络权设计为

$$W_\tau = W^+ - \tau W^-,$$

其中 $\tau > -1$ 为一个参数，并且网络的阈值向量为

$$\theta_\tau = V^m - W_\tau V^m$$

正交化设计的优点：

（1）对称性。

（2）所有样本点都是稳定点。

对于任一个样本模式 $V^l(l \neq m)$ ，由于 $V^l - V^m$ 是 Y 矩阵的一个列向量，自然属于维数为 k 的列向量空间。这样便存在一组系数 $\beta_1, \beta_2, \cdots, \beta_k$ 使得

3.1 离散Hopfield网络(DHNN)

$$V^l - V^m = \sum_{i=1}^{k} \beta_i P_i, \quad \text{或} \quad V^l = \sum_{i=1}^{k} \beta_i P_i + V^m$$

对于$\{P_1, P_2, \cdots, P_k\}$ 中任一 P_i ，则有

$$W_\tau P_i = W^+ P_i - \tau W^- P_i = P_i \Rightarrow W_\tau (V^l - V^m) = (V^l - V^m)$$

因此，将 V^l 输入给网络，我们得

$$Sgn(W_\tau V^l + \theta_\tau) = Sgn(W_\tau (V^l - V^m) + V^m)$$
$$= Sgn(V^l - V^m + V^m) = Sgn(V^l) = V^l$$

则 V^l 是网络的稳定点。另外，我们再考虑 V^m ，

$$Sgn(W_\tau V^m + \theta_\tau) = Sgn(V^m) = V^m$$

3.1 离散Hopfield网络(DHNN)

故所有模式样本都是网络的稳定点。

（3）参数τ 可调性使得网络可以减少伪稳定点，扩大样本模式的吸引域。

当 $V'=Sgn(W_\tau V'+\theta_\tau)$ ，总可调整 τ 使得

$$V'\neq Sgn(W_\tau V'+\theta_\tau)$$

这样便可减少一个伪稳定点。在串行 方式下，当伪稳定点的个数减少时，样本模式（真稳定点）的吸引域将增大。

3.1 离散Hopfield网络(DHNN)

(iii). 感知机学习算法

对于一组样本模式 $S=\{V^1,V^2,\cdots,V^m\}$ ，我们根据每个样本模式成为网络稳定点的条件对广义DHNN（ $N=(W,\theta),\ w_{ii}=0$ ）中各单元的权值根据感知机学习算法进行学习获得。即按下面目标学习：

$$W=[W_1,W_2,\cdots,W_n]^T,\ W_i=[w_{i1},\cdots,w_{ii},\cdots,w_{in}]^T \leftrightarrow \text{神经元}i$$
$$V^l=[v_1^l,v_2^l,\cdots,v_n^l]\in\{\pm 1\}^n,\qquad l=1,2,\cdots,m$$

对于广义DHNN中神经元 i 所要求的权和阈值和目标值为：

3.1 离散Hopfield网络(DHNN)

$$\tilde{W}_i = [w_{i1}, \cdots, w_{i,i-1}, w_{i,i+1}, \cdots, w_{in}]^T, \ \theta_i$$

$$\tilde{V}^l = [v_1^l, \cdots, v_{i-1}^l, v_{i+1}^l, \cdots, v_n^l]^T \ \rightarrow \ v_i^l, \quad l = 1,2,\cdots,m, \text{ 即}$$

$$Sgn(\sum_{j=1}^{n} w_{ij} v_j^l - \theta_i) = v_i^l, \ l = 1,2,\cdots,m \quad \Rightarrow$$

$$Sgn(WV^l - \theta) = V^l, \quad l = 1,2,\cdots,m$$

根据广义Hopfield网络在随机串行方式下的稳定性，用这种方法所得到的网络是稳定的，可应用于联想记忆。当任一个目标不是线形可分的，那么说明这一组样本是矛盾的，不可能同时被广义DHNN所记忆，当然同样不可能被DHNN所记忆。

3.1 离散Hopfield网络(DHNN)

(iv). 目标感知机学习算法（OPLA）（Ma, 1999）

对于一组样本$S=\{V^1,V^2,\cdots,V^m\}$，采用感知机学习算法可以得到广义DHNN使得每个样本为其稳定点，但无法保证它们有一个理想的吸引域，为此建立了DHNN。下面我们介绍这一算法。

1） 合理的目标

对于样本模式$V^1,V^2,\cdots,V^m$，最大吸引半径：

$$h[k]=\lfloor(\min\{d_H(V^k,V^l):l=1,\cdots,k-1,k+1,\cdots,m\}-1)/2\rfloor \quad k=1,2,\cdots,m$$

吸引半径的目标值应满足：

$$t(k)\le h[k],\qquad k=1,2,\cdots,m$$

3.1 离散Hopfield网络(DHNN)

2） 参数选取

$\alpha \leftrightarrow$ 学习率；$b \leftrightarrow$ 绝对权值的上界；$\delta \leftrightarrow$ 小的正实数。

设在 t 时刻选取样本 V^k 进行学习， 对第 i 个神经元的权和阈值进行修正， 即对 $(\tilde{W}_i, \theta_i)$ 进行修正， 学习律如下：

$$\Delta w_{ij} = \alpha (v_i^k - u_i^k) v_j^k \quad (j = 1, \cdots, i-1, i+1, \cdots, n)$$

$$\Delta \theta_i = \alpha (v_i^k - u_i^k)$$

其中

$$u_i^k = Sgn(\sum_{j=1}^{n} w_{ij} v_j^k + \theta_i - v_i^k (t_k b + \delta)) \qquad (w_{ii} \equiv 0)$$

3.1 离散Hopfield网络(DHNN)

在修正中，如果修正后的权值满足上界约束，即 $|w_{ij}+\Delta w_{ij}|<b$ ，则接受这一修正，并继续进行，否则，保持权值不变。θ_i 始终按上面规则进行修正，直到收敛。 OPLA的推导见参考文献[3]，经模拟实验证明，该算法优于外积法和其它的一些方法。

3.1 离散Hopfield网络(DHNN)

6. 记忆容量讨论

网络按某种方式能够存储的样本的数量称为记忆容量。显然，记忆容量与吸引半径的要求是矛盾的。

(i). 外积型设计的DHNN的记忆容量

对于样本组 $S=\{V^1,V^2,\cdots,V^m\}$，当这些样本是随机选取时，Hopfield 用模拟的方法估计出网络的记忆容量为 $0.13\sim0.15n$ 。下面我们从数学上对这一结果进行推导。根据前面的记号，我们有：

3.1 离散Hopfield网络(DHNN)

$$\tilde{V}_j^l = Sgn(s_j + n_j)$$

$$s_j = nV_j^l, \qquad n_j = \sum_{k \neq l} \sum_{i=1}^{n} v_i^k v_j^k v_i^l$$

当 $|n_j| < |s_j|$ ， $\tilde{v}_j^l = v_j^l$ ，即 v_j^l 达到平稳。

在随机样本的情况，且当 n 充分大时，则：

$$n_j \sim N(0, \sigma^2), \quad \sigma^2 = n(m-1)$$

我们进一步考虑 $\tilde{v}_j^l \neq v_j^l$ 的概率，即错误概率。

若 $v_j^l = -1 < 0$ ，则仅需要 $n_j > -nv_j^l$

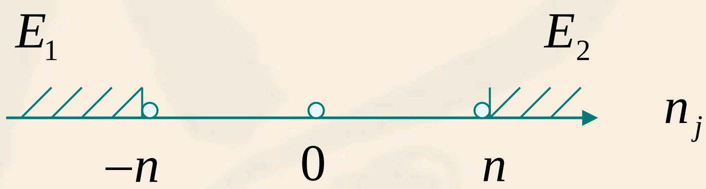
$$P(E_1) = \int_{-nv_j^l}^{+\infty} \varphi(x)dx, \qquad \text{其中}\varphi(x)\text{为正态分布}N(0,\sigma^2)\text{的概率密度函数。}$$

3.1 离散Hopfield网络(DHNN)

若 $v_j^l = 1 > 0$ ，则需要 $n_j < -nv_j^l$

$$P(E_2) = \int_{-\infty}^{-nv_j^l} \varphi(x)dx$$

在 $v_j^l = -1,\ s_j = -n$，我们来计算



$$P(E_1) = \int_n^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{n}{\sqrt{n(m-1)}}}^{+\infty} e^{-\frac{u^2}{2}} du = \int_{\sqrt{\frac{n}{m-1}}}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = \phi(\sqrt{\frac{n}{m-1}})$$

$$= 1 - \Phi(\sqrt{\frac{n}{m-1}}) = \Phi(-\sqrt{\frac{n}{m-1}})$$

假设在 V^l 中出错分量个数服从泊松分布，则 V^l 中的出错分量个数的平均值（泊松分布的参数）：

$$\lambda = n\phi(\sqrt{\frac{n}{m-1}})$$

3.1 离散Hopfield网络(DHNN)

假设 $\beta(0<\beta<1)$ 为所要求的正确率，则有

$$\beta = e^{-n\phi(\sqrt{\frac{n}{m-1}})} \Rightarrow -\frac{\ln\beta}{n} = \phi(\sqrt{\frac{n}{m-1}}) = \frac{\alpha}{n} \qquad (\alpha = -\ln\beta)$$

$$\phi^{-1}(\frac{\alpha}{n}) = \sqrt{\frac{n}{m-1}} \Rightarrow m = \frac{n}{[\phi^{-1}(\frac{\alpha}{n})]^2} + 1$$

当 $\sqrt{\frac{n}{m-1}} > 3.5$ 时，$\phi^{-1}(\frac{\alpha}{n}) \approx \sqrt{2\ln\frac{n}{\alpha}}$，因此得

$$m = \frac{n}{2\ln\frac{n}{\alpha}} + 1$$

若 $\alpha = 1$ 时，则进一步有 $m = \frac{n}{2\ln n} + 1$，这近似地有： $m \approx 0.13 \sim 0.15n$ （注意：$\beta = e^{-1}$）

3.1 离散Hopfield网络(DHNN)

(ii). 广义Hopfield网络的记忆容量（Ma, 1999）

考虑广义DHNN在感知机学习算法下的记忆容量（显然这也是广义DHNN的最大记忆容量），为此先引进记忆概率序列：

$$P(m,n)=P(\{S=\{V^1,V^2,\cdots,V^m\}\text{是可记忆的。}\})$$

显然 $P(m,n)$ 是随 m 递减的。

定义3.1　称一个整数函数 $C(n)$ 是广义DHNN的渐近记忆容量，如果满足下列两个条件：

(1). $\lim_{n\to\infty} P(m,n)=1, \quad if \;\; m\le C(n);$

(2). $\lim_{n\to\infty} \inf P(m,n)<1, \quad if \;\; m>C(n).$

3.1 离散Hopfield网络(DHNN)

若 $C(n)$ 仅满足第一个条件，称其为渐近记忆容量的下界；若 $C(n)$仅满足第二个条件，称其为渐近记忆容量的上界。

定理3.4(Ma, 1999) 假设 $P(m,n)$为广义DHNN的记忆概率序列，则有：

$$\lim_{n\to\infty} P(n-1,n)=1, \qquad \lim_{n\to\infty} \inf P(2n,n) \le \frac{1}{2}$$

证明见[4]。

注意：定理3.4给出了广义DHNN渐近记忆容量的下界、上界分别为 $n-1$, $2n$。

3.1 离散Hopfield网络(DHNN)

注：合理的数学定义应按下列方式进行：

渐近记忆容量 $C(n)$ 的定义为：

$$\lim_{m\to\infty} P(m,n)=\begin{cases}1, & if \quad m<C(n);\\ 0, & if \quad m>C(n).\end{cases}$$

3.1 离散Hopfield网络(DHNN)

(iii). 提高记忆容量和吸引域的方法

(1). 高阶关联Hopfield网络

$$y_i = Sgn(\sum_l T_l(i))$$

$$T_1(i) = \sum_j w_{ij} v_j, \qquad T_2(i) = \sum_j \sum_k w_{ijk} v_j v_k$$

$$\cdots\cdots$$

$$T_p(i) = \sum_{j_1} \cdots \sum_{j_p} w_{j_1 \cdots j_p} v_{j_1} \cdots v_{j_p}$$

$$w_{ijk} = \sum_{\mu=1}^{m} v_i^\mu v_j^\mu v_k^\mu, \quad i, j, k = 1, 2, \cdots, n$$

高阶关联提高了网络的性能，使之能够储存更多的样本模式，并由于多重外积加强了网络的稳定性，则能够扩大样本的吸引域。

3.1 离散Hopfield网络(DHNN)

(2). 为了增大某个样本的吸引域，我们可以直接加强这个样本，采用如下：

$$W=\sum_{\mu=1}^{m}P_i(V^i(V^i)^T-U),\qquad P_i>0,\ \sum_{\mu=1}^{m}P_i=1$$

(3). 非对称δ 学习方法

$$W'=\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}\begin{bmatrix}w_{11} & w_{12} & \cdots & w_{1n}\\ w_{21} & w_{22} & \cdots & w_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ w_{n1} & w_{n2} & \cdots & w_{nn}\end{bmatrix}$$

$$=\begin{bmatrix}\lambda_1 w_{11} & \lambda_1 w_{12} & \cdots & \lambda_1 w_{1n}\\ \lambda_2 w_{21} & \lambda_2 w_{22} & \cdots & \lambda_2 w_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ \lambda_n w_{n1} & \lambda_n w_{n2} & \cdots & \lambda_n w_{nn}\end{bmatrix}\qquad(\lambda_i>0)$$

3.1 离散Hopfield网络(DHNN)

显然，这里 W' 是非对称的，即非对称Hopfield网络同样可以使样本成为稳定点。故可以采用前面的感知机学习算法及目标感知机学习算法得到所要求的广义Hopfield网络。

3.1 离散Hopfield网络(DHNN)

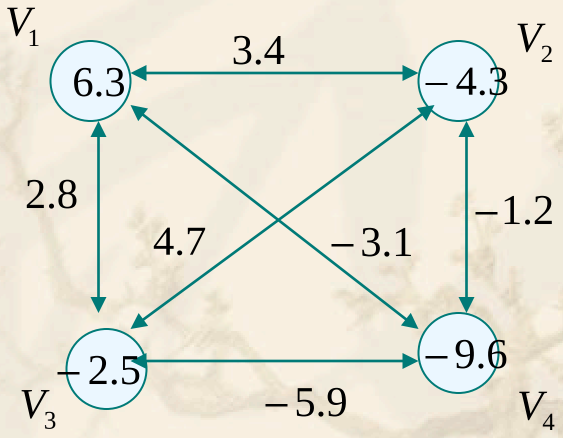
作业：

1. 证明DHNN的能量函数

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

在演化时所产生的减小量 $|\Delta E(t)|$ 有非零下界。

2. 如右图所示的DHNN，计算每个状态的能量函数，求出其稳定点和异步、同步方式下达到个稳定点初始状态的个数。



参考资料

[1] Jinwen Ma, Simplex Memory Neural Networks, **Neural Networks**, vol.10(1997), No.1, pp:25-29.

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

[3] Jinwen Ma, The object perceptron learning algorithm on generalised Hopfield networks for associative memory, **Neural Computing & Applications**, Vol.8(1999), No.1,pp:25-32.

[4] Jinwen Ma, The asymptotic memory capacity of the generalized Hopfield networks , **Neural Networks**, vol.12(1999), No.9, pp:1207-1212.