Невронни Мрежи

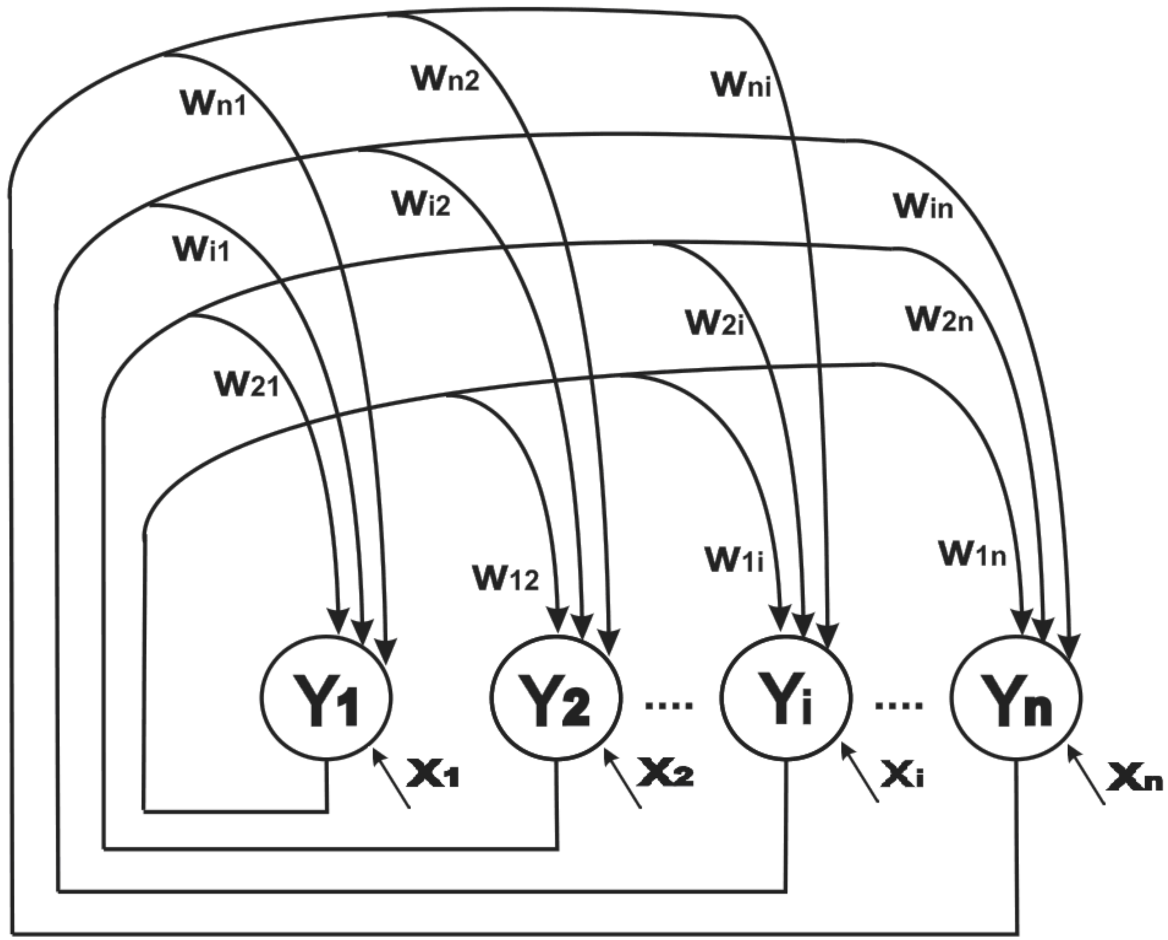
проф. дн В. Младенов

4. Дискретни мрежи на Хопфийлд. Анализ на сходимостта на дискретните мрежи на Хопфийлд. Енергийни функции. Непрекъснати мрежи на Хопфийлд.

Мрежи на Hopfield

- Дискретните мрежи на Hopfield са въведени през 1982 г.
- Непрекъснатите са въведени през 1984 г.
- Еднослойни мрежи, при които всеки неврон е свързан с всички останали.
- Основният принос на Hopfield е разглеждането на невронните мрежи с апарата на динамичните системи и установяването на връзката между тяк.
- Той за първи път въвежда понятията енегийни функции и атрактори в теорията на невронните мрежи.
- Използвани са за решаване на сложни комбинаторни задачи, а не само като асоциативни памети.

Мрежи на Hopfield



Мрежи на Hopfield

- Еднослойна мрежа, всички неврони на която са свързани, като $w_{ij}=w_{ji}$ и $w_{ii}=0$, т.е. матрицата с теглата на мрежата е симетрична с нулеви елементи в главния диагонал.
- Входните сигнали x_1, x_2, …,x_n не са задължителни. В някои случаи входни сигнали може да има, а в други случаи може да няма.
- Входове за невронната мрежа са началните състояния на изходните сигнали на невроните, а изходи са изходните сигнали на невроните след съответен брой итерации. В тези случаи изходните сигнали на невроните, които определят *състоянието на мрежата* на всяка итерация, служат съответно за вход и изход на мрежата.
- Активационните функции на невроните са релейни, като в зависимост от данните (бинарни или биполярни) са в диапазон [0, 1] или [-1, 1].

Мрежи на Hopfield

Матрицата с теглата на мрежата обикновено се изчислява като се използва правилото на Hebb и се нулират диагоналните и елементи.

При биполярни данни

$$w_{ij} = \sum_p s_i(p) s_j(p) \quad i \neq j,$$

$$w_{ii} = 0\,,$$

а при бинарни данни

$$w_{ij} = \sum_p (2s_i(p) - 1)(2s_j(p) - 1) \quad i \neq j,$$

$$w_{ii} = 0\,,$$

като в случая при конструиране на матрицата с теглата всеки бинарен вектор се преобразува в биполярен.

Мрежи на Hopfield

- Важен момент при реализацията на итерационната процедура е асинхронната промяна на изходните сигнали на невроните. Това означава, че на всяка стъпка се променя изходният сигнал само на един неврон, който може да бъде избиран случайно.
- Вероятността за всеки от невроните да бъде избран за промяна на състоянието на дадена стъпка трябва да бъде еднаква.
- Тази нова стойност се използва при промяната на изходния сигнал за следващите неврони.
- Тази асинхронна промяна гарантира сходимостта на промяната на изходните сигнали на невроните на мрежата.

Мрежи на Hopfield

Ако на входа на мрежата е подаден входен вектор **x**$=[x_1\ x_2 \dots x_n]^T$, това означава, че векторът с изходните сигнали на невроните от мрежата в началния момент е **y**$(0)=[y_1(0)\ y_2(0) \dots y_n(0)]^T$ = **x** $= [x_1\ x_2 \dots x_n]^T$.

Процедурата по промяна на теглата на мрежата включва следните стъпки.

а) Подава се входен вектор **x**$=[x_1\ x_2 \dots x_n]^T$,
т.е. **y**$(0)=[y_1(0)\ y_2(0) \dots y_n(0)]^T$ = **x**$=[x_1\ x_2 \dots x_n]^T$.

б) Докато критерия за сходимост не се удоволетвори се изпълняват следните стъпки.

в) Случайно се избира неврон – например невронът с номер i и $x_i= y_i(k)$.

г) Изчислява се претеглената сума за входа на този неврон на стъпка k

$$y_{in\_i} = x_i + \sum_{j=1}^{n} w_{ij} y_j(k)$$

Мрежи на Hopfield

д) Изчислява се изходният сигнал на неврона за стъпка *k+1* като се използва активационната функция

$$y_i(k+1) = f(y_{in\_i}) = \begin{cases} 1, & \text{ako} \quad y_{in\_i} > \theta_i \\ y_i(k), & \text{ako} \quad y_{in\_i} = \theta_i \\ -1, & \text{ako} \quad y_{in\_i} < \theta_i \end{cases}$$

е) Проверява се условието за сходимост.

Векторно-матричната форма за изхода се изразява с

$$y_i(k+1) = f\left(x_i + \mathbf{W}_{i\bullet}.\mathbf{y}(k)\right), \; i=1,2,\ldots,n \;,$$

където $\mathbf{W}_{i\bullet}$ е i-тият ред на матрицата $\mathbf{W}$, а

$$\mathbf{y}(k)=[y_1(k)\; y_2(k)\; \ldots\; y_n(k)]^T.$$

Условието за сходимост на итерационната процедура е спиране на промяната на изходните сигнали на невроните, т.е.

$$y_i(k+1)=y_i(k)\,, \quad i=1,2,\ldots,n.$$

Мрежи на Hopfield - пример

Запазваме
бинарната последователност: $[1\ 1\ 1\ 0]^T$

Извикваме на входа: $\mathbf{x}=[0\ 0\ 1\ 0]^T$ – първите два бита са грешни

$$\mathbf{W} = \begin{bmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{bmatrix}$$

Избираме $Y1$:

$$y_{in\_1} = x_1 + \mathbf{y}^T.\mathbf{W}_{\bullet 1} = 0+1=1$$

$$y_1 = 1$$

$$\mathbf{y}=[1\ 0\ 1\ 0]^T$$

Избираме $Y3$:

$$y_{in\_3} = x_3 + \mathbf{y}^T.\mathbf{W}_{\bullet 3} = 1+1=2$$

$$y_3 = 1$$

$$\mathbf{y}=[1\ 0\ 1\ 0]^T$$

Избираме $Y1$:

$$y_{in\_1} = x_1 + \mathbf{y}^T.\mathbf{W}_{\bullet 1} = 1+2=3$$

$$y_1 = 1,\ \mathbf{y}=[1\ 1\ 1\ 0]^T$$

Избираме $Y4$:

$$y_{in\_4} = x_4 + \mathbf{y}^T.\mathbf{W}_{\bullet 4} = 0+(-2)=-2$$

$$y_4 = 0$$

$$\mathbf{y}=[1\ 0\ 1\ 0]^T$$

Избираме $Y2$:

$$y_{in\_2} = x_2 + \mathbf{y}^T.\mathbf{W}_{\bullet 2} = 0+2=2$$

$$y_2 = 1$$

$$\mathbf{y}=[1\ 1\ 1\ 0]^T$$

⇒ Запазената последователност е правилно извикана (recalled)

Мрежи на Hopfield - сходимост

- Основния въпрос при мрежите на Hopfield е дали са сходящи, т.е. дали при произволен начален вектор, итерационната процедура по промяна на състоянията (изходните сигнали на невроните на мрежата) ще спре.
- Друг не по-малко важен въпрос е, дали дадена мрежа на Hopfield е сходяща към най-близкия вектор, който се съхранява от мрежата. Този въпрос в общия случай няма отговор, защото е свързан с т.нар. *области на привличане* на устойчивите стационарни точки на нелинейната система, която описва динамичния режим (режимът по време на итерациите) на дискретната мрежа на Hopfield.
- Въпросът със сходимостта е решен от Hopfield. Той въвежда т.нар. *енергийна функция* за модела на съответната нел. система:

$$y_i(k+1) = f\left(x_i + \mathbf{W}_{i\bullet}.\mathbf{y}(k)\right),\ \mathrm{i} = 1, 2, \ldots, \mathrm{n}$$

- *Енергийната функция* е *функция от типа на Ляпунов* с тази разлика, че енергийната функция трябва да е ограничена отдолу, докато функциите на Ляпунов трябва да са положително определени.

Мрежи на Hopfield - сходимост

- Най-общо енергийната функция $E(\mathbf{y}(t))$ за дадена система, която описва поведението на дискретната невронна мрежа на Hopfield, е функция от състоянието на мрежата (изходните сигнали на невроните) на всяка стъпка *t*, която трябва да удоволетворява условията:

  -$E(t)$ е ограничена отдолу, т.е съществува константа *c*, такава че $E(t) \geq c$

  -$E(t)$ е монотонно ненарастваща функция, т.е. $\Delta E(t+1) = E(t+1) - E(t) \leq 0$

- Ако за системата съществува енергийна функция, то системата е сходяща, т.е. при кое да е начално състояние (начален вектор) на системата е изпълнено условието за спиране на промяната на състоянието след определен брой стъпки $y_i(k+1)=y_i(k)$, $i=1,2,\ldots,n$.

Мрежи на Hopfield - сходимост

Hopfield показва, че ако

- матрицата с теглата на мрежата е избрана съгласно условията (при биполярни и бинарни данни)

$$w_{ij} = \sum_p s_i(p)s_j(p) \quad i \neq j, \qquad w_{ij} = \sum_p (2s_i(p)-1)(2s_j(p)-1) \quad i \neq j,$$

$$w_{ii} = 0\,, \qquad\qquad w_{ii} = 0\,,$$

- и промяната на състоянията на невроните е асинхронно

за мрежата, която се описва с уравненията

$$y_i(k+1) = f\left(x_i + \mathbf{W}_{i\bullet}.\mathbf{y}(k)\right),\ \mathrm{i} = 1, 2, \ldots, \mathrm{n}$$

то съществува енергийна функция от вида

$$E(t) = E(\mathbf{y}(t)) = -0.5\sum_{i \neq j}\sum_j y_i(t)y_j(t)w_{ij} - \sum_i x_i y_i(t) + \sum_i \theta_i y_i(t)$$

Мрежи на Hopfield - сходимост

Лесно се забелязва, че тъй като x_i, y_i, w_{ij} и θ_i са ограничени, то и *E(t)* е ограничена отдолу. Ако на стъпка *t* за промяната е избран изходният сигнал на *k*-ти неврон, то

$$\Delta y_k(t+1)= y_k(t+1)- y_k(t)$$

и

$$\Delta y_j(t+1)= y_j(t+1)- y_j(t) = 0 \text{ , т.к. } y_j(t+1) = y_j(t) \text{ , } j \neq k \text{ .}$$

Като се отчетат тези зависимости и също така, че $w_{ij}=w_{ji}$, $w_{ii}=0$, то

$$E(t+1)-E(t)=(-0.5\sum_{i\neq j}\sum_{j} y_i(t+1)y_j(t+1)w_{ij} - \sum_i x_i y_i(t+1) + \sum_i \theta_i y_i(t+1))$$
$$-(-0.5\sum_{i\neq j}\sum_{j} y_i(t)y_j(t)w_{ij} - \sum_i x_i y_i(t) + \sum_i \theta_i y_i(t))$$

или $$\Delta E(t+1)=E(t+1)-E(t)=-[\underbrace{\sum_{j\neq k} y_j(t)w_{kj}+x_k-\theta_k}_{y_{in\_k}}]\Delta y_k(t+1)$$

Оттук

$$\Delta E(t+1)=-y_{in\_k}.\Delta y_k(t+1)$$

Мрежи на Hopfield - сходимост

Нека $y_k(t)=1$ и $y_k(t+1)=-1$, тогава $y_{in\_k} < \theta_k$ и $\Delta y_k(t+1) = -2$.
Следователно $\Delta E(t+1)<0$.
Ако $y_k(t)=-1$ и $y_k(t+1)=1$, тогава $y_{in\_k} > \theta_k$ и $\Delta y_k(t+1)= 2$.
Следователно и в този случай $\Delta E(t+1)<0$.
Накрая, ако $y_k(t+1)= y_k(t)$ то $\Delta y_k(t+1)= 0$ и оттук $\Delta E(t+1)=0$.

От горните разглеждания следва, че на всяка стъпка *t* енергийната функция намалява или не се променя. Тъй като *E(t)* е ограничена отдолу, след определен брой стъпки (итерации), енергийната функция ще спре да се променя, т.е. $\Delta E(t+1)=0$ независимо от това кой неврон е избран да си променя състоянието.
Тогава за всеки неврон *k* или $\Delta y_k(t+1)= 0$, а оттам $y_k(t+1)= y_k(t)$ или $y_{in\_k} = \theta_k$, така че и в този случай $y_k(t+1)= y_k(t)$,
т.е. векторът на състоянието спира да се променя и състоянието на невронната мрежа схожда към устойчива стационарна точка наречена още *атрактор*.

Мрежи на Hopfield - сходимост

- Всеки *атрактор* има т.нар. *област на привличане*, която представлява множеството от точки (вектори), които, ако са начално състояние на мрежата, след спиране на итерациите по промяната на състоянието, крайното състояние на мрежата ще е съответният атракторът. В този контекст векторите (образите), които се съхраняват от невронната мрежа на Hopfield, съответстват на устойчивите стационарни точки на системата, с която се описва динамиката на мрежата.
- Броят *P* на векторите, които могат да бъдат съхранени определя капацитета на невронната мрежа. Hopfield е установил емпирично, че $P \approx 0.15.n$, където *n* е броят на невроните. По-прецизните теоретични изследвания показват, че $P \approx n/(2.\log_2 n)$.
- От горните оценки се вижда, че *P/n* намалява с нарастването на *n* т.к. при повече неврони се усилват различните взаимодействия между тях. Съществуват разработки на различни модификации на мрежите на Hopfield с цел увеличаване на капацитета им. В тези случаи за определяне на матрицата с теглата на връзките на мрежата не се използва правилото на Hebb.

Непрекъснати Мрежи на Hopfield

Непрекъснатите мрежи на Hopfield са обобщение на дискретните мрежи и също като тях всички неврони са свързани като $w_{ij}=w_{ji}$ при $i \neq j$ и $w_{ii}=0$. В случая състоянията на мрежата се променят съгласно системата уравнения

$$\frac{du_i}{dt} = \sum_{j=1}^{n} w_{ij} v_j + x_i, \quad i = 1, 2, \ldots, n$$

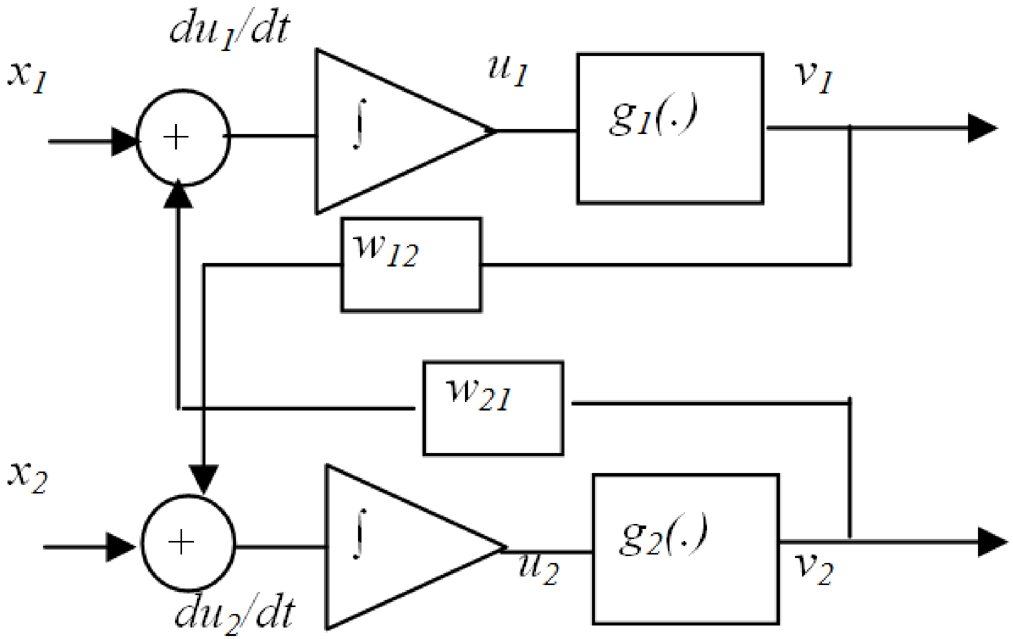
където $v_i = g(u_i)$, $i = 1, 2, \ldots, n$

са изходните сигнали на невроните, а x_i е външният входен сигнал свързан с i-тия неврон.

Функцията g е сигмоидална функция с диапазон [-1, 1] или [0, 1] с цел изходните образи да са биполярни или бинарни и първата и производна да е положителна в целия диапазон на промяна на аргумента и. В непрекъснатия случай, всички неврони променят състоянието си едновременно.

Непрекъснати Мрежи на Hopfield

Двумерен (*n=2*) модел на непрекъсната мрежа на Hopfield е показан на следната фигура



Hopfield показва, че ако теглата на мрежата са избрани, така че $w_{ij}=w_{ji}$ при $i \neq j$ и $w_{ii}=0$, то за невронната мрежа съществува енергийна функция от вида $E = E(t) = -\frac{1}{2}\sum_{ij} v_i w_{ij} v_j - \sum_i x_i v_i$

Непрекъснати Мрежи на Hopfield

Ако се отчете, че v_i, w_{ij} и x_i са ограничени, то и $E(t)$ е ограничена отдолу. Пълната производна на $E(t)$ по отношение на времето е

$$\frac{dE}{dt} = \sum_i \frac{\partial E}{\partial v_i} \cdot \frac{dv_i}{du_i} \cdot \frac{du_i}{dt} = \left(-\sum_{j \neq i} w_{ij} v_j - x_i \right) \cdot \frac{dv_i}{du_i} \cdot \frac{du_i}{dt}$$

където

$$\frac{dv_i}{du_i} = g'(u_i) > 0$$

и

$$\frac{du_i}{dt} = -\frac{\partial E}{\partial v_i} = \sum_{j=1}^{n} w_{ij} v_j + x_i$$

След заместването им се получава, че:

$$\frac{dE}{dt} = -\sum_i g'(u_i) . \left(\sum_{j=1}^{n} w_{ij} v_j + x_i \right)^2 \leq 0$$

Непрекъснати Мрежи на Hopfield

Вижда, че енергийната функция намалява с времето до досигане на някое стационарно състояние, което съответства на такъв

вектор $\mathbf{u}(t)=[u_1(t)\ u_2(t)\ \ldots u_n(t)]^T$, за който $\dfrac{du_i}{dt}=0,\ i=1,2,\ldots,n$.

Това е в сила, тъй като системата уравнения, с която описва режима, е непрекъснатата мрежа на Hopfield е градиентна, т.к. съществува функция

$E=E(t)=E(v_1(t), v_2(t),\ldots, v_n(t))$,

векторът на градиента, за която е равен точно на дясната страна на системата уравнения за режима в мрежата със знак минус. За такива системи, траекториите в пространството на състоянията винаги клонят към някое стационарно състояние при $t\rightarrow\infty$.

Непрекъснати Мрежи на Hopfield

- Едно от най-важните приложения на непрекъснатите мрежи на Hopfield е при решаване на сложни оптимизационни задачи.
- В тези случаи, на база на целевата функция и ограниченията се формира енергийна функция *E*, която зависи от състоянията, с които се моделира множеството от решения на разглежданата задача (с други думи, оптимизационната задача с ограничения по подходящ начин се трансформира в оптимизационна задача без ограничения, като множествата от решенията на двете задачи съвпадат).
- Въз основа на градиента на енергрийната функция се определят десните части на системата и се определят теглата на връзките между невроните от мрежата.
- Най-трудното в тези случаи е да се гарантира, че всяко решение на разглежданата оптимизационна задача съответства на локален минимум на енергийната функция, а оптималното решение съответства на нейния глобален минимум.

Непрекъснати Мрежи на Hopfield

Типичен пример за приложение на непрекъснатите невронни мрежи на Hopfield е решаването на задачата на търговския пътник - тази задача най-общо може да бъде представена по следния начин:

-дадени са *n* града и функцията на разходите за преминаване на разстоянията между всеки от градовете (най-често това са разстоянията между различните градове).

-търси се най-евтиният (в смисъла на функцията на разходите) път за обхождане на всички градове, така че тръгвайки от даден град, след обхождане на всички градове (без повторение) да се стигне до изходния град.

Такъв тип задачи са с голямо практическо приложение.

Съществуват редица оптимизационни алгоритми, които могат да бъдат прилагани при решаването на подобни задачи, но в много от случаите, особено при големи размерности на задачите, те са свързани със сингулярности. За разлика от тези числени алгоритми, невронните мрежи винаги водят до “някакво” решение, което в болшинството от случаите е оптималното решение на разглежданите задачи.