

Ciągły model Hopfielda

Jacek Kluska

Politechnika Rzeszowska

Ciągły model Hopfielda

- 1 N - neuronów o wejściach x_1, \dots, x_N i wyjściach y_1, \dots, y_N
- 2 $w_{ii} = 0, \forall i$ - charakter sprzężeń zwrotnych
- 3 $w_{ij} = w_{ji}, \forall i, j$ - symetria

Część **liniowa** sieci

$$\begin{bmatrix} n_1 \\ \vdots \\ n_N \end{bmatrix} = \begin{bmatrix} 0 & w_{12} & \dots & w_{1N} \\ w_{21} & 0 & \dots & w_{2N} \\ \vdots & \vdots & \dots & \vdots \\ w_{N1} & w_{N2} & \dots & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} + \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}$$

Część **nieliniowa** sieci

$$y_i = f(n_i) = \frac{1}{2} (1 + \tanh(\beta n_i)) = \frac{1}{1 + e^{-2\beta n_i}}, \quad y_i \xrightarrow{\beta \rightarrow \infty} \mathbf{1}(n_i)$$

Model elektryczny sieci Hopfielda

$$\mathbf{C} \frac{d}{dt} \mathbf{n}(t) = -\mathbf{G} \mathbf{n}(t) + \mathbf{W} \mathbf{y}(t) + \mathbf{i}$$
$$\mathbf{y}(t) = \mathbf{f}[\mathbf{n}(t)]$$

\mathbf{i} - wektor stałych wymuszeń, ($\mathbf{i} = \mathbf{x} = \text{const}$)

$\mathbf{C} = \text{diag}(C_1, \dots, C_N)$, $\mathbf{G} = \text{diag}(G_1, \dots, G_N)$

$$\mathbf{W} = \begin{bmatrix} 0 & w_{12} & \dots & w_{1N} \\ w_{21} & 0 & \dots & w_{2N} \\ \vdots & \vdots & \dots & \vdots \\ w_{N1} & w_{N2} & \dots & 0 \end{bmatrix}$$

1. Bez utraty ogólności można przyjąć $\mathbf{C} = \mathbf{I}$ (macierz jednostkowa).
2. Punkt równowagi $\mathbf{n} = \mathbf{n}^*$ (o ile istnieje), spełnia:

$$-\mathbf{G} \mathbf{n}^* + \mathbf{W} \mathbf{f}(\mathbf{n}^*) + \mathbf{i} = \mathbf{0}$$

Przykład (N=3)

$$C_1 \frac{dn_1(t)}{dt} = -G_1 n_1(t) + w_{12} f_2(n_2(t)) + w_{13} f_3(n_3(t)) + i_1$$

$$C_2 \frac{dn_2(t)}{dt} = -G_2 n_2(t) + w_{21} f_1(n_1(t)) + w_{23} f_3(n_3(t)) + i_2$$

$$C_3 \frac{dn_3(t)}{dt} = -G_3 n_3(t) + w_{31} f_1(n_1(t)) + w_{32} f_2(n_2(t)) + i_3$$

Funkcja energetyczna (Lapunova)

$$E(\mathbf{y}) = -\frac{1}{2}\mathbf{y}^T \mathbf{W} \mathbf{y} - \mathbf{i}^T \mathbf{y} + \sum_{i=1}^N G_i \int_0^{y_i} f_i^{-1}(z) dz$$

f_i^{-1} - funkcja odwrotna do f_i

Przykład

$$\begin{aligned} E(y_1, y_2, y_3) = & -[w_{12}y_1(t)y_2(t) + w_{13}y_1(t)y_3(t) + w_{23}y_2(t)y_3(t)] \\ & - [i_1y_1(t) + i_2y_2(t) + i_3y_3(t)] \\ & + G_1h_1(y_1(t)) + G_2h_2(y_2(t)) + G_3h_3(y_3(t)) \end{aligned}$$

Uwaga:

$$h_i(y_i) = \int_0^{y_i} f_i^{-1}(z) dz < 0$$

Wyznaczenie całki z funkcji odwrotnej do funkcji aktywacji: $h(y)$

$$h(y) = \int_0^y f^{-1}(z) dz$$

$$y = f(n) = \frac{1}{1 + \exp(-2n)} \quad \Rightarrow \quad f^{-1}(z) = \frac{1}{2} \ln \left(\frac{z}{1-z} \right)$$

$$h(y) = \int_0^y f^{-1}(z) dz = \frac{1}{2} \ln y^y (1-y)^{(1-y)} < 0$$

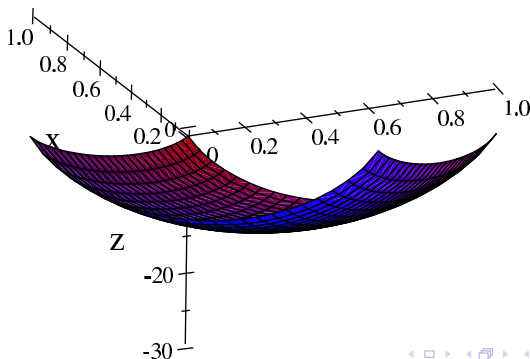
$$\frac{d}{dy} \int_0^y f^{-1}(z) dz = \frac{1}{2} \ln \left(\frac{y}{1-y} \right) = f^{-1}(y)$$

Jak wygląda funkcja energetyczna ?

Przykład. $N = 2$, $f(n) = \frac{1}{1 + \exp(-2n)}$,

$$\mathbf{W} = \begin{bmatrix} 0 & 2 \\ 2 & 0 \end{bmatrix}, \mathbf{i} = \begin{bmatrix} 15 \\ 6 \end{bmatrix}, \mathbf{G} = \begin{bmatrix} 8 & 0 \\ 0 & 28 \end{bmatrix}$$

$$E(y_1, y_2) = -2y_1y_2 - 15y_1 - 6y_2 + 4 \ln y_1^{y_1} (1 - y_1)^{(1-y_1)} + 14 \ln y_2^{y_2} (1 - y_2)^{(1-y_2)}$$



Pochodna funkcji energetycznej

$$E(\mathbf{y}) = -\frac{1}{2}\mathbf{y}^T \mathbf{W} \mathbf{y} - \mathbf{i}^T \mathbf{y} + \sum_{i=1}^N G_i \int_0^{y_i} f_i^{-1}(z) dz$$

Obliczenia

$$\frac{dE(\mathbf{y}(t))}{dt} = \sum_{i=1}^N \frac{\partial E}{\partial y_i} \frac{dy_i}{dt} = (\nabla_{\mathbf{y}} E)^T \frac{d\mathbf{y}}{dt}$$

$$\nabla_{\mathbf{y}} [\mathbf{y}^T \mathbf{W} \mathbf{y}] = (\mathbf{W} + \mathbf{W}^T) \mathbf{y}$$

$$\nabla_{\mathbf{y}} (\mathbf{i}^T \mathbf{y}) = \mathbf{i}$$

$$\frac{\partial}{\partial y_i} \left(\int_0^{y_i} G_i f_i^{-1}(z) dz \right) = G_i n_i = G_i f_i^{-1}(y_i)$$

$$\nabla_{\mathbf{y}} \left(\sum_{i=1}^N G_i \int_0^{y_i} f_i^{-1}(z) dz \right) = \mathbf{G} \mathbf{n}$$

$$\begin{aligned}\nabla_{\mathbf{y}} E &= \nabla_{\mathbf{y}} \left(-\frac{1}{2} \mathbf{y}^T \mathbf{W} \mathbf{y} - \mathbf{i}^T \mathbf{y} + \sum_{i=1}^N G_i \int_0^{y_i} f_i^{-1}(z) dz \right) \\ &= -\frac{1}{2} (\mathbf{W} + \mathbf{W}^T) \mathbf{y} - \mathbf{i} + \mathbf{G} \mathbf{n} \\ &= -(-\mathbf{G} \mathbf{n} + \mathbf{W} \mathbf{y} + \mathbf{i}) = -\mathbf{C} \frac{d\mathbf{n}}{dt}\end{aligned}$$

Zmiany \mathbf{n} zachodzą w kierunku przeciwnym do $(\nabla_{\mathbf{y}} E)$.

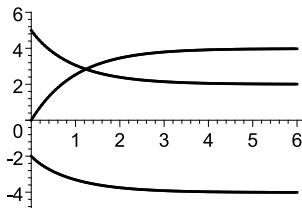
$$\frac{dE}{dt} = (\nabla_{\mathbf{y}} E)^T \frac{d\mathbf{y}}{dt} = - \left(\mathbf{C} \frac{d\mathbf{n}}{dt} \right)^T \frac{d\mathbf{y}}{dt}$$

$$y_i = f_i(n_i) - \text{rosnąca} \Rightarrow \left. \frac{dE}{dt} \right|_{\mathbf{n}=\mathbf{n}(t)} \leq 0$$

$$\begin{aligned}\dot{n}_1 &= -n_1 + 1/\left(1 + e^{-2n_2}\right) + 1/\left(1 + e^{-2n_3}\right) + 1 \\ \dot{n}_2 &= -n_2 + 1/\left(1 + e^{-2n_1}\right) + 1/\left(1 + e^{-2n_3}\right) - 6 \\ \dot{n}_3 &= -n_3 + 1/\left(1 + e^{-2n_1}\right) + 1/\left(1 + e^{-2n_2}\right) + 3\end{aligned}$$

Punkt równowagi: $(n_1^*, n_2^*, n_3^*) = (2.000, -4.018, 3.982)$

$$y_1^* = f_1(n_1^*) = 0.98201, y_2^* = f_2(n_2^*) = 0.00032, y_3^* = f_3(n_3^*) = 0.99965$$



- ❶ Podczas ewolucji sieci Hopfielda, energia maleje aż do momentu, gdy

$$\frac{dE}{dt} = 0, \quad \frac{dy_i}{dt} = 0$$

- ❷ Nie ma oscylacji ani cykli **granicznych**; atraktor jest punktem **równowagi stabilnej**, takim, że:

$$-\mathbf{G}\mathbf{n}^* + \mathbf{W}\mathbf{f}(\mathbf{n}^*) + \mathbf{i} = \mathbf{0}$$

- ❸ W przestrzeni wyjść (y_1, \dots, y_N) , atraktor \mathbf{y}^* znajduje się zwykle **blisko wierzchołka** hipersześcianu $[0, 1]^N$.
- ❹ Model jest "**synchroniczny**", tzn. wszystkie równania różniczkowe są rozwiązywane **jednocześnie**.
- ❺ Model jest **deterministyczny** - inaczej jest w klasycznym dyskretnym modelu Hopfielda.

Realizacje techniczne ciągłej sieci Hopfielda

Realizacja **programowa**.

$$\frac{dn_i(k)}{dt} \cong \frac{n_i(k+1) - n_i(k)}{\Delta t}, \quad k = 0, 1, 2, \dots$$

$$\begin{aligned} \frac{1}{\Delta t} \mathbf{C} [\mathbf{n}(k+1) - \mathbf{n}(k)] &= -\mathbf{G}\mathbf{n}(k) + \mathbf{W}\mathbf{y}(k) + \mathbf{i} \\ \mathbf{y}(k) &= \mathbf{f}[\mathbf{n}(k)] \end{aligned}$$

Równanie rekurencyjne

$$\begin{aligned} \mathbf{n}(k+1) &= \mathbf{A}_0\mathbf{n}(k) + \mathbf{A}_1\mathbf{y}(k) + \mathbf{b} \\ \mathbf{y}(k) &= \mathbf{f}[\mathbf{n}(k)] \end{aligned}$$

$$\mathbf{A}_0 = (\mathbf{I} - \Delta t \mathbf{C}^{-1} \mathbf{G}), \quad \mathbf{A}_1 = \Delta t \mathbf{C}^{-1} \mathbf{W}, \quad \mathbf{b} = \Delta t \mathbf{C}^{-1} \mathbf{i},$$

Przykład - Problem komiwojażera

Dane: N - miast - punktów w przestrzeni R^2 .

Macierz odległości

$$\mathbf{D} = \{d_{pq}\} = \begin{bmatrix} 0 & d_{12} & \dots & d_{1N} \\ d_{21} & 0 & \dots & d_{2N} \\ \vdots & \vdots & \dots & \vdots \\ d_{N1} & d_{N2} & \dots & 0 \end{bmatrix}, \quad d_{pq} - \text{odległość miasta } p \text{ od } q$$

Przykład

(1): (3, 7), (2): (7, 0), (3): (4, 0), (4): (7, 4), (5): (0, 3)

$d_{12} = \|(3, 7) - (7, 0)\| = \sqrt{65}, \dots$

$$\mathbf{D} = \begin{bmatrix} 0.0 & 8.1 & 7.1 & 5.0 & 5.0 \\ 8.1 & 0.0 & 3.0 & 4.0 & 7.6 \\ 7.1 & 3.0 & 0.0 & 5.0 & 5.0 \\ 5.0 & 4.0 & 5.0 & 0.0 & 7.1 \\ 5.0 & 7.6 & 5.0 & 7.1 & 0.0 \end{bmatrix}$$

Problem

Znaleźć najkrótszą trasę łączącą N miast. Każde miasto musi być odwiedzane dokładnie jeden raz.

Stopień trudności

Liczba tras: $(N - 1)!/2$

N	$\frac{(N - 1)!}{2}$
$N = 3$	1
$N = 4$	3
$N = 10$	181 440
$N = 30$	4.4×10^{30}
$N = 50$	3.0×10^{62}
$N = 100$	4.7×10^{155}

Pełne przeszukiwanie ?

Zał. Sprawdzenie 1 trasy trwa 1[ns]. $N = 22$:

$$\frac{\frac{(22 - 1)!}{2}}{365 * 24 * 60 * 60 * 10^9} = 810.04 \text{ [lat]}$$

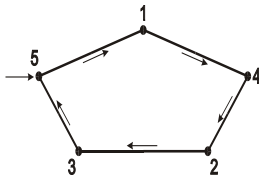
Nie znaleziono jeszcze żadnego algorytmu dokładnego rozwiązania w czasie wielomianowym (**NP-trudne** zadanie optymalizacji kombinatorycznej).

Założenia

N^2 - **neuronów**, każdy generuje sygnał $y_{pi} \in \{0, 1\}$,
 $p = 1, \dots, N, i = 1, \dots, N$

Neurony mają **podwójne** wskaźniki: y_{pi} - neuron wskazujący, czy miasto p zajmuje i -tą pozycję na trasie komiwojażera” $y_{pi} = 1$ - tak, $y_{pi} = 0$ - nie.

Przykład $N = 5$



$\mathbf{Y} = \{y_{pi}\} =$

	1	2	3	4	5
1	0	1	0	0	0
2	0	0	0	1	0
3	0	0	0	0	1
4	0	0	1	0	0
5	1	0	0	0	0

\iff trasa 5-1-4-2-3

Wskaźniki cząstkowe ("funkcje kary")

$$\mathbf{Y} = \{y_{pi}\} =$$

	1	2	3	4	5	
1	0	1	0	1	1	←
2	0	0	0	1	0	
3	0	0	0	0	1	
4	0	0	1	0	0	
5	1	0	0	0	0	

Liczba takich sytuacji, że miasto p , ($p = 1, \dots, N$) zostało odwiedzone na i -tym etapie podróży i ponownie na j -tym etapie podróży:

$$k_1(p) = \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N y_{pi} y_{pj}$$

Przykład $p = 1, [0 \ 1 \ 0 \ 1 \ 1], k_1(1) = 3$

Kara za ponowne odwiedzanie tych samych miast ($E_1 = \sum_{p=1}^N k_1(p)$):

$$E_1(y_{11}, \dots, y_{NN}) = \frac{1}{2} \sum_{p=1}^N \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N y_{pi} y_{pj}$$

Wskaźniki cząstkowe ("funkcje kary") - c.d.

$$\mathbf{Y} = \{y_{pi}\} = \begin{array}{c|ccccc} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & 1 & 0 & \mathbf{1} & \mathbf{1} \\ \hline 2 & 0 & 0 & 0 & \mathbf{1} & 0 \\ \hline 3 & 0 & 0 & 0 & 0 & \mathbf{1} \\ \hline 4 & 0 & 0 & 1 & 0 & 0 \\ \hline 5 & 1 & 0 & 0 & 0 & 0 \\ \hline \end{array}$$

↑ ↑

Liczba takich sytuacji, że miasta p i q , ($p, q = 1, \dots, N$) zostały jednocześnie odwiedzone na i -tym etapie:

$$k_2(i) = \frac{1}{2} \sum_{p=1}^N \sum_{\substack{q=1 \\ q \neq p}}^N y_{pi} y_{qi}$$

Kara za "rozdwojenie" komiwojażera ($E_2 = \sum_{i=1}^N k_2(i)$):

$$E_2(y_{11}, \dots, y_{NN}) = \frac{1}{2} \sum_{i=1}^N \sum_{p=1}^N \sum_{\substack{q=1 \\ q \neq p}}^N y_{pi} y_{qi}$$

Wskaźniki cząstkowe ("funkcje kary") - c.d.

Jeżeli miasto p , ($p, q = 1, \dots, N$) zostało odwiedzone na i -tym etapie podróży, to $y_{pi} = 1$. **Kara za brak odwiedzenia miasta lub wielokrotne odwiedzanie**

$$E_3(y_{11}, \dots, y_{NN}) = \frac{1}{2} \left(\sum_{p=1}^N \sum_{i=1}^N y_{pi} - N \right)^2$$

Długość trasy: zsumowanie d_{pq}

$d_{pq}y_{pi}y_{qi}$ - uwzględnianie tego składnika jest zbędne

$d_{pp}y_{pm}y_{pn} = 0$ - uwzględnianie tego składnika jest zbędne

$d_{pq}y_{pi}y_{q,i+1} = d_{pq} \Leftrightarrow$ kolejność odwiedzania miast: $p \rightarrow q$

$d_{pq}y_{pi}y_{q,i-1} = d_{pq} \Leftrightarrow$ kolejność odwiedzania miast: $q \rightarrow p$

$$E_4(y_{11}, \dots, y_{NN}) = \frac{1}{2} \sum_{p=1}^N \sum_{\substack{q=1 \\ q \neq p}}^N \sum_{i=1}^N d_{pq}y_{pi} (y_{q,i+1} + y_{q,i-1})$$

$$y_{q,N+1} := y_{q1}, \quad y_{q0} := y_{qN}$$

Uwaga: Indeksy są obliczane modulo N .

Funkcja energetyczna (Lapunova)

$$\begin{aligned} E &= A_1 E_1 + A_2 E_2 + A_3 E_3 + A_4 E_4 + \sum_{p=1}^N \sum_{i=1}^N \frac{1}{\tau} \int_{0.5}^{y_{pi}} f^{-1}(z) dz \\ &= A_1 \frac{1}{2} \sum_{p=1}^N \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N y_{pi} y_{pj} + A_2 \frac{1}{2} \sum_{i=1}^N \sum_{p=1}^N \sum_{\substack{q=1 \\ q \neq p}}^N y_{pi} y_{qi} \\ &\quad + A_3 \frac{1}{2} \left(\sum_{p=1}^N \sum_{i=1}^N y_{pi} - N \right)^2 + A_4 \frac{1}{2} \sum_{p=1}^N \sum_{\substack{q=1 \\ q \neq p}}^N \sum_{i=1}^N d_{pq} y_{pi} (y_{q,i+1} + y_{q,i-1}) \\ &\quad + \sum_{p=1}^N \sum_{i=1}^N \frac{1}{\tau} \int_{0.5}^{y_{pi}} f^{-1}(z) dz \end{aligned}$$

f^{-1} - funkcja odwrotna do funkcji aktywacji, $A_1, A_2, A_3, A_4, \tau > 0$.

Rozwiązanie problemu komiwojażera

Przypomnienie: $\frac{d}{dt}\mathbf{n}(t) = \mathbf{C}^{-1}\nabla_{\mathbf{y}}E \implies$ **minimalizacja** E

$$E = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N w_{ij} y_i y_j - \sum_{i=1}^N c_i y_i + \sum_{i=1}^N G_i \int_0^{y_i} f_i^{-1}(z) dz$$

jest równoważna z **rozwiązaniem** N -równań różniczkowych

$$\frac{dn_i(t)}{dt} = -G_i n_i(t) + \sum_{j=1}^N w_{ij} \cdot f_j[n_j(t)] + c_i, \quad y_i = f_i(n_i), \quad i = 1, \dots, N$$

Wniosek: **Minimalizacja** funkcji energetycznej N^2 -zmiennych y_{pi}

$$E = A_1 E_1 + A_2 E_2 + A_3 E_3 + A_4 E_4 + \sum_{p=1}^N \sum_{i=1}^N \frac{1}{\tau} \int_{0.5}^{y_{pi}} f^{-1}(z) dz$$

jest równoważna z **rozwiązaniem** N^2 -równań różniczkowych

$$\frac{dn_{pi}}{dt} = - \left. \frac{\partial E}{\partial y_{pi}} \right|_{y=y(t)=f[n(t)]}, \quad p, i = 1, \dots, N$$

Rozwiązanie problemu komiwojażera - c.d.

$$\frac{dn_{pi}}{dt} = - \left. \frac{\partial E}{\partial y_{pi}} \right|_{y=y(t)=f[n(t)]}, \quad p, i = 1, \dots, N$$



$$\begin{aligned} \frac{dn_{pi}}{dt} = & -\frac{1}{\tau}n_{pi} - A_1 \sum_{j \neq i} y_{pi} - A_2 \sum_{q \neq p} y_{qi} - A_3 \left(\sum_q \sum_j y_{qj} - N \right) \\ & - A_4 \sum_{q \neq p} d_{pq} (y_{q,i+1} + y_{q,i-1}) \\ y_{pi} = & f(n_{pi}) \end{aligned}$$

Używa się N^2 -neuronów.

Przykład (Hopfield, Tank, 1985)

$N = 10$, (181 440 dróg)

1. $n_{pi}(0) = \frac{1}{N} + \varepsilon \text{rand}(\cdot)$
2. $A_1 = A_2 = A_4 = 500, A_3 = 200, \beta = 50$
3. $E'_3 = \frac{1}{2} \left(\sum_{p=1}^N \sum_{i=1}^N y_{pi} - (N + \sigma) \right)^2, \sigma = 5$

50% uzyskanych wyników reprezentowało jedną z dwóch najkrótszych dróg.

Wady ciągłego modelu Hopfielda:

- utykanie w minimach lokalnych
- brak uniwersalnych metod definiowania wag w funkcji energii,
- duża wrażliwość sieci na dobór tych wag.