Lineární asociativní síť

- Lineární asociativní síť (LAS)
- Hebbův zákon
- Učení LAS podle Hebbova zákona
- Pseudohebbovská adaptace LAS

Asociativní sítě - obecně

Cílem je uchovat množinu vzorů $\{(\vec{x}_k, \vec{d}_k) \mid k = 1, \dots, p\}$ tak, aby platilo následující:

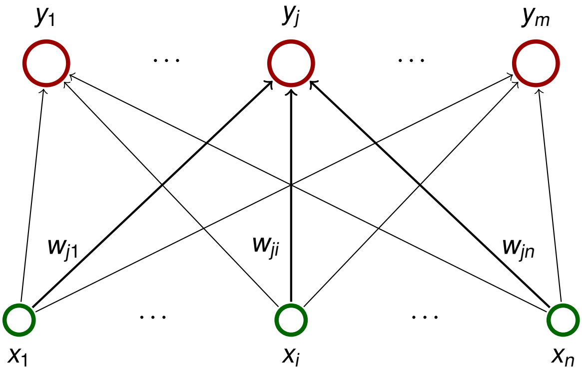
Po předložení nového vstupu $\vec{x}$, který je „blízko“ některému $\vec{x}_k$ bude výstup sítě roven (nebo alespoň blízko) $\vec{d}_k$.

Zejména by síť měla mít schopnost *reprodukce*:
Pro vstup $\vec{x}_k$ by měla dát výstup $\vec{d}_k$.

V připadě tvz. *autoasociativní* sítě předpokládáme $\vec{x}_k = \vec{d}_k$.

Lineární asociativní síť

Organizační dynamika: Jednovrstvá síť, neurony bez biasů.



Aktivní dynamika: Síť počítá funkci z $\mathbb{R}^n$ do $\mathbb{R}^m$ danou

$$y_j = \sum_{i=1}^{n} w_{ji} x_i \qquad \text{pro } j = 1, \ldots, m$$

Lineární asociativní síť - matice

Označme

- $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$
- váhy tvoří matici W danou

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mn} \end{pmatrix}$$

Aktivní dynamiku potom lze zapsat takto:

$$\vec{y}(W, \vec{x}) = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = W \cdot \vec{x}$$

Občas budu psát jen $\vec{y}(\vec{x})$ pokud bude W jasné z kontextu.

LAS - adaptivní dynamika

Dána množina $\mathcal{T}$ **tréninkových vzorů** tvaru

$$\left\{ \left(\vec{x}_k, \vec{d}_k \right) \;\middle|\; k = 1, \ldots, p \right\}$$

kde každé $\vec{x}_k = \begin{pmatrix} x_{k1} \\ \vdots \\ x_{kn} \end{pmatrix}$ je vstupní vektor a každé $\vec{d}_k = \begin{pmatrix} d_{k1} \\ \vdots \\ d_{km} \end{pmatrix}$ je očekávaný výstup sítě.

V případě *autoasociativní* paměti předpokládáme $\vec{d}_k = \vec{x}_k$ pro $k = 1, \ldots, p$.

Cílem je nalézt váhy W takové, že $\vec{d}_k = \vec{y}(W, \vec{x}_k) = W \cdot \vec{x}_k$.

Adaptace podle Hebbova zákona

Hebbův zákon: *When an axon of cell A is near enough to excite cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.*

Zákon formuloval neuropsycholog Donald Hebb v knize „The Organization of Behavior“ z roku 1949.

Jinými slovy: *Cells that fire together, wire together.*

Formulace používaná v umělých NS:
Změna váhy spoje mezi dvěma neurony je úměrná jejich souhlasné aktivitě.

Hebb se snažil vysvětlit podmíněné reflexy: Současná aktivita/pasivita presynaptického neuronu (příčina) a postsynaptického neuronu (reakce) posiluje/zeslabuje synaptickou vazbu.

Hebbovská adaptivní dynamika LAS

Počítáme posloupnost matic vah $W^{(0)}, W^{(1)}, \ldots, W^{(p)}$ kde

$$W^{(k)} = \begin{pmatrix} w_{11}^{(k)} & \cdots & w_{1n}^{(k)} \\ \vdots & \ddots & \vdots \\ w_{m1}^{(k)} & \cdots & w_{mn}^{(k)} \end{pmatrix}$$

- Na začátku nastav $w_{ji}^{(0)} = 0$.
- V k-tém kroku (zde $k = 1, \ldots, p$) je síti předložen k-tý vzor a váhy se adaptují podle Hebbova zákona:

$$w_{ji}^{(k)} = w_{ji}^{(k-1)} + d_{kj} x_{ki}$$

Hebbovská adaptivní dynamika LAS

Zapsáno maticově

- $W^{(0)}$ je nulová matice.
- V k-tém kroku (zde $k = 1, \ldots, p$) vypočti:

$$W^{(k)} = W^{(k-1)} + \vec{d}_k \vec{x}_k^T$$

kde

$$\vec{d}_k \vec{x}_k^T = \begin{pmatrix} d_{k1}x_{k1} & \cdots & d_{k1}x_{kn} \\ \vdots & \ddots & \vdots \\ d_{km}x_{k1} & \cdots & d_{km}x_{kn} \end{pmatrix}$$

Hebbovská adaptivní dynamika LAS

Výsledná matice:

$$W = W^{(p)} = \sum_{k=1}^{p} \vec{d}_k \vec{x}_k^T = DX^T$$

kde

$$D = \begin{pmatrix} d_{11} & \cdots & d_{p1} \\ \vdots & \ddots & \vdots \\ d_{1m} & \cdots & d_{pm} \end{pmatrix}$$

$$X = \begin{pmatrix} x_{11} & \cdots & x_{p1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \cdots & x_{pn} \end{pmatrix} \qquad X^T = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{p1} & \cdots & x_{pn} \end{pmatrix}$$

V autoasociativním případě dostaneme $W = XX^T$.

Schopnost reprodukce

Chceme aby po předložení tréninkového vstupu $\vec{x}_k$ byl výstup $W\vec{x}_k$ roven $\vec{d}_k$.

Hodnota funkce sítě pro r-tý vzor je

$$\vec{y}(\vec{x}_r) = W\vec{x}_r = \sum_{k=1}^{p} \vec{d}_k \left(\vec{x}_k^T \vec{x}_r\right)$$

Předpokládejme, že je množina vektorů $\{\vec{x}_1, \dots, \vec{x}_p\}$ ortonormální, tj.

$$\vec{x}_k^T \vec{x}_r = \begin{cases} 1 & k = r \\ 0 & k \neq r \end{cases}$$

Pak $W\vec{x}_k = \vec{d}_k$.

Asociace

Uvažme vstup: $\vec{x}_r + \vec{u}$ kde norma $\left\|\vec{u}\right\|$ je malá.

Chyba sítě pro r-tý vzor perturbovaný vektorem $\vec{u}$:

$$E_r(\vec{u}) = \left\|\vec{y}(\vec{x}_r + \vec{u}) - \vec{d}_r\right\| = \left\|W\vec{x}_r + W\vec{u} - \vec{d}_r\right\| = \left\|W\vec{u}\right\|$$

Pokud $\left\|\vec{d}_r\right\| = 1$ pro každé $r = 1, \ldots, p$, pak pro každé r platí

$$E_r(\vec{u}) \leq n \left\|\vec{u}\right\|$$

($\left\|\vec{d}_r\right\| = 1$ například platí v autoasociativním případě, protože máme ortonormální vstupy)

Tedy pro vstupy blízké vzorům síť odpovídá přibližně požadovaným výstupem.

Schopnost reprodukce - nelineární aktivace

Přepokládejme **pro tento a následující slajd**, že aktivační funkce není identita, ale $\sigma(\xi) = \begin{cases} 1 & \xi \geq 0 \\ -1 & \xi < 0 \end{cases}$

Předpokládejme, že $\vec{x}_k \in \{-1, 1\}^n$ a $\vec{d}_k \in \{-1, 1\}^m$.

Pak hodnota funkce sítě pro r-tý vzor je

$$\vec{y}(\vec{x}_r) = sgn(W\vec{x}_r) = sgn\left(\vec{d}_r + \sum_{k \neq r} \vec{d}_k \frac{\left(\vec{x}_k^T \vec{x}_r\right)}{\left(\vec{x}_r^T \vec{x}_r\right)}\right)$$

Nyní $\vec{d}_k = \vec{y}(\vec{x}_k)$ pro každé k pokud

$$\left|\sum_{k \neq r} \vec{d}_r \frac{\left(\vec{x}_k^T \vec{x}_r\right)}{\left(\vec{x}_r^T \vec{x}_r\right)}\right| < 1$$

(tato hodnota se nazývá *přeslech*)

Ukázka LAS s nelineární aktivací

Experiment: vzorové vektory z $\{-1, 1\}^{10}$ byly postupně ukládány Hebbovským učením. Poté byl síti předložen nový vektor a očekávala se jeho asociace s nejbližším vzorovým vektorem vzhledem k Hammingově vzdálenosti.

- po každém uložení vzorového vektoru byla síť otestována na velkém množství vektorů, jejichž Hamingova vzdálenost od vzorů byla 0 – 4.
- tabulka udává procento správně asociovaných vektorů z dané H. vzdálenosti od některého vzoru (řádky) ve chvíli, kdy síť byla naučena daný počet vzorů (sloupce)

H	1	2	3	4	5	6	7	8	9	10
0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
1	100.0	100.0	90.0	85.0	60.0	60.0	54.3	56.2	45.5	33.0
2	100.0	86.7	64.4	57.2	40.0	31.8	22.5	23.1	17.0	13.3
3	100.0	50.0	38.6	25.4	13.5	8.3	4.8	5.9	3.1	2.4
4	100.0	0.0	9.7	7.4	4.5	2.7	0.9	0.8	0.3	0.2

Pseudohebbovská adaptace

- Alternativní způsob učení LAS
- Nevyžaduje ortonormalitu vzorů (jen lineární nezávislost)
- Poněkud se vzdaluje od biologické motivace

Pseudohebbovská adaptace

Uvážíme *autoasociativní* případ:

$$\mathcal{T} = \left\{ \left(\vec{x}_k, \vec{x}_k\right) \quad \middle| \quad k = 1, \ldots, p \right\}$$

Předpokládejme, že množina $\{\vec{x}_1, \ldots, \vec{x}_p\}$ je lineárně nezávislá.

Potom $p \leq n$ a $\{\vec{x}_1, \ldots, \vec{x}_p\}$ je bází vektorového prostoru V_p, který je podprostorem $\mathbb{R}^n$.

Idea: Během adaptace budeme postupně konstruovat *ortogonální* bázi $\{\vec{z}_1, \ldots, \vec{z}_p\}$ prostoru V_p pomocí Gram-Schmidtova *ortogonalizačního* procesu.

Pseudohebbovská adaptace

Počítáme posloupnost matic vah $W^{(0)}, W^{(1)}, \dots, W^{(p)}$.

- Na začátku nastav $w_{ji}^{(0)} = 0$.
- V k-tém kroku (zde $k = 1, \dots, p$) je síti předložen k-tý vzor a váhy se adaptují takto:

$$\vec{z}_k = \vec{x}_k - W^{(k-1)} \cdot \vec{x}_k$$

$$W^{(k)} = W^{(k-1)} + \frac{\vec{z}_k \vec{z}_k^T}{\vec{z}_k^T \vec{z}_k}$$

Pseudohebbovská adaptace - reprodukce

Označme výslednou matici

$$W = W^{(p)} = \sum_{k=1}^{p} \frac{\vec{z}_k \vec{z}_k^T}{\vec{z}_k^T \vec{z}_k}$$

Chceme $W\vec{x}_k = \vec{x}_k$ pro každé $k = 1, \dots, p$.

Tvrzení

$\{\vec{z}_1, \dots, \vec{z}_k\}$ je ortogonální bází vektorového prostoru V_p (tj. prostoru s bází $\{\vec{x}_1, \dots, \vec{x}_k\}$).

Důsledek

Pro $\vec{x} \in V_p$ platí $W\vec{x} = \vec{x}$. Zejména $W\vec{x}_k = \vec{x}_k$, tedy síť má schopnost reprodukce.

Pseudohebbovská adaptace - asociace

Uvažme vstup: $\vec{x}_r + \vec{u}$ kde norma $\left\|\vec{u}\right\|$ je malá.

Chyba sítě pro r-tý vzor perturbovaný vektorem $\vec{u}$:

$$E_r(\vec{u}) = \left\|\vec{y}(\vec{x}_r + \vec{u}) - \vec{x}_r\right\| = \left\|W\vec{x}_r + W\vec{u} - \vec{x}_r\right\| = \left\|W\vec{u}\right\|$$

Pak pro každé $r = 1, \ldots, p$ platí

$$E_r(\vec{u}) \leq n\left\|\vec{u}\right\|$$

Tedy pro vstupy blízké vzorům síť odpovídá přibližně požadovaným výstupem.

Pseudohebbovská adaptace - pseudoinverze

$W^{(p)} = XX^{+}$ kde $X^{+} = (X^{T}X)^{-1}X^{T}$ je *Moore-Penroseova pseudoinverze* matice X.

Matice $(X^{T}X)^{-1}$ existuje, protože jsou sloupce matice X lineárně nezávislé.

Matice $W = XX^{+}$ je řešením rovnice $WX = X$ (tj. přesně problém autoasociace)

Pseudohebbovská adaptace - heteroasociativní

Uvážíme *heteroasociativní* případ:

$$\mathcal{T} = \left\{ \left(\vec{x}_k, \vec{d}_k\right) \quad \middle| \quad k = 1, \ldots, p \right\}$$

Předpokládejme, že množina $\{\vec{x}_1, \ldots, \vec{x}_p\}$ je lineárně nezávislá.

Hledáme matici W takovou, že pro $k = 1, \ldots, p$ platí $W\vec{x}_k = \vec{d}_k$.
Tj. $WX = D$ kde

$$X = \begin{pmatrix} x_{11} & \cdots & x_{p1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \cdots & x_{pn} \end{pmatrix} \qquad D = \begin{pmatrix} d_{11} & \cdots & d_{p1} \\ \vdots & \ddots & \vdots \\ d_{1m} & \cdots & d_{pm} \end{pmatrix}$$

Matice $W = DX^+ = D(X^T X)^{-1} X^T$ řeší rovnici $WX = D$

Pseudohebbovská adaptace - heteroasociativní

Poznámka: Problém asociace lze řešit, i když $X^T X$ není invertibilní (tj. bez předpokladu nezávislosti $\{\vec{x}_1, \dots, \vec{x}_p\}$).

Vždy existuje unikátní matice (tzv. pseudoinverze), která splňuje následující:

- $XX^+X = X$
- $X^+XX^+ = X^+$
- X^+X a XX^+ jsou symetrické

Matice $W = DX^+$ potom minimalizuje

$$\|WX - D\|^2 = \sum_{j=1}^{m} \sum_{k=1}^{p} (y_j(W, \vec{x}_k) - d_{kj})^2$$

Tedy matici W lze počítat pomocí gradientního sestupu.

Pseudohebbovská adaptace - heteroasociativní

W lze počítat také pomocí pseudohebbovské adaptace.

Počítáme dvě posloupnosti matic vah $W'^{(0)}, W'^{(1)}, \ldots, W'^{(p)}$ a $W^{(0)}, W^{(1)}, \ldots, W^{(p)}$

- Na začátku nastav $w_{ji}^{(0)} = 0$.
- V k-tém kroku (zde $k = 1, \ldots, p$) je síti předložen k-tý vzor a váhy se adaptují takto:

$$\vec{z}_k = \vec{x}_k - W'^{(k-1)} \cdot \vec{x}_k$$

$$W'^{(k)} = W'^{(k-1)} + \frac{\vec{z}_k \vec{z}_k^T}{\vec{z}_k^T \vec{z}_k}$$

$$W^{(k)} = W^{(k-1)} + \frac{(\vec{d}_k - W^{(k-1)} \vec{x}_k) \vec{z}_k^T}{\vec{z}_k^T \vec{z}_k}$$

Výsledná matice je $W = W^{(p)}$.

Tj. $\vec{z}_k$ se počítají v *autoasociatvní* síti a poté se použijí k výpočtu $W^{(k)}$

Hopfieldova síť

- Definice
- Energetická funkce

Hopfieldova síť

Autoasociativní síť.

Organizační dynamika:

- úplná topologie, tj. každý neuron je spojen s každým
- všechny neurony jsou současně vstupní i výstupní
- označme $\xi_1, \dots, \xi_n$ vnitřní potenciály a $y_1, \dots, y_n$ výstupy (stavy) jednotlivých neuronů
- označme w_{ji} *celočíselnou* váhu spoje od neuronu $i \in \{1, \dots, n\}$ k neuronu $j \in \{1, \dots, n\}$.
- žádný neuron nemá bias a přepokládáme $w_{jj} = 0$ pro každé $j = 1, \dots, n$.

Hopfieldova síť

Adaptivní dynamika: Dána tréninková množina

$$\mathcal{T} = \{\vec{x}_k \mid \vec{x}_k = (x_{k1}, \ldots, x_{kn}) \in \{-1, 1\}^n, k = 1, \ldots, p\}$$

Adaptace probíhá podle Hebbova zákona (stejně jako u LAS). Výsledná konfigurace je

$$w_{ji} = \sum_{k=1}^{p} x_{kj} x_{ki} \qquad 1 \leq j \neq i \leq n$$

Všimněte si, že $w_{ji} = w_{ij}$, tedy matice vah je symetrická.

Adaptaci lze vidět jako hlasování vzorů o vazbách neuronů: $w_{ji} = w_{ij}$ se rovná rozdílu mezi počtem souhlasných stavů $x_{kj} = x_{ki}$ neuronů i a j a počtem rozdílných stavů $x_{kj} \neq x_{ki}$.

Hopfieldova síť

Aktivní dynamika: Iniciálně jsou neurony nastaveny na vstup $\vec{x} = (x_1, \dots, x_n)$ sítě, tedy $y_j^{(0)} = x_j$ pro každé $j = 1, \dots, n$.

V t-tém kroku aktualizujeme neuron j, který splňuje $t = \tau \cdot (n - 1) + j - 1$ takto:

nejprve vypočteme vnitřní potenciál

$$\xi_j^{(t-1)} = \sum_{i=1}^{n} w_{ji} y_i^{(t-1)}$$

a poté

$$y_j^{(t)} = \begin{cases} 1 & \xi_j^{(t-1)} > 0 \\ y_j^{(t-1)} & \xi_j^{(t-1)} = 0 \\ -1 & \xi_j^{(t-1)} < 0 \end{cases}$$

Pozn. τ je počet period v nichž se aktualizovaly všechny neurony.

Hopfieldova síť - aktivní dynamika

Výpočet končí v kroku t^* pokud se síť nachází (poprvé) ve *stabilním* stavu, tj.

$$y_j^{(t^*+n)} = y_j^{(t^*)} \qquad (j = 1, \ldots, n)$$

Věta

Za předpokladu symetrie vah, výpočet Hopfieldovy sítě skončí pro každý vstup.

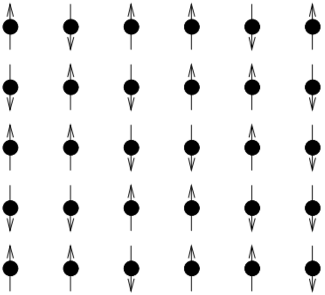
Z toho plyne, že Hopfieldova síť počítá funkci z $\{-1, 1\}^n$ do $\{-1, 1\}^n$ (která závisí na hodnotách vah neuronů).

Označme $\vec{y}(W, \vec{x}) = (y_1^{(t^*)}, \ldots, y_n^{(t^*)})$ hodnotu funkce sítě pro vstup $\vec{x}$ a matici vah W. Dále označme $y_j(W, \vec{x}) = y_j^{(t^*)}$ složku hodnoty funkce sítě, která odpovídá neuronu j.

Pokud bude W jasné z kontextu, budu psát jen $y(\vec{x})$ a $y_j(\vec{x})$

Fyzikální analogie

Jednoduché modely magnetických materiálů připomínají Hopfieldovu síť.



- atomické magnety poskládané do mřížky
- každý magnet může mít pouze jednu ze dvou orientací (v Hopfieldově síti $+1$ a -1)
- orientaci každého magnetu ovlivňuje jednak vnější magnetické pole (vstup sítě), jednak magnetické pole ostatních magnetů (závisí na jejich orientaci)
- synaptické váhy modelují vzájemnou interakci magnetů

Energetická funkce

Energetická funkce E přiřazuje každému stavu sítě $\vec{y} \in \{-1, 1\}^n$ potenciální energii danou

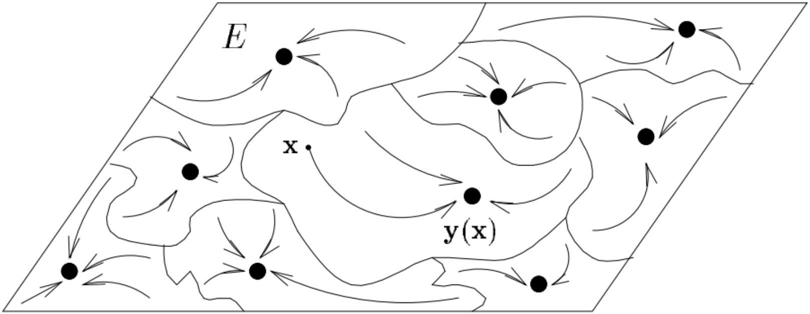
$$E(\vec{y}) = -\frac{1}{2} \sum_{j=1}^{n} \sum_{i=1}^{n} w_{ji} y_j y_i$$

- velké (kladné) $w_{ji} y_j y_i$ je stabilní a malé (záporné) $w_{ji} y_j y_i$ nestabilní
- stavy s nízkou energií jsou stabilní (málo neuronů „chce“ změnit svůj stav), stavy s vysokou energií jsou nestabilní

V průběhu výpočtu se energie snižuje: $E(\vec{y}^{(t)}) \geq E(\vec{y}^{(t+1)})$, stav $\vec{y}^{(t^*)}$ odpovídá lokálnímu minimu funkce E.

Poznámka: Stav $y_j^{(t)}$ neuronu j odpovídá opačnému znamínku gradientu E v bodě $\vec{y}^{(t-1)}$: $\frac{\partial E}{\partial y_j}(\vec{y}^{(t-1)}) = -\sum_{i=1}^{n} w_{ji} y_i^{(t-1)}$

Energetická plocha



Obr. 3.4: Energetická plocha.