4. prednáška: TSP (The Travelling Salesman Problem) - Hopfieldov prístup

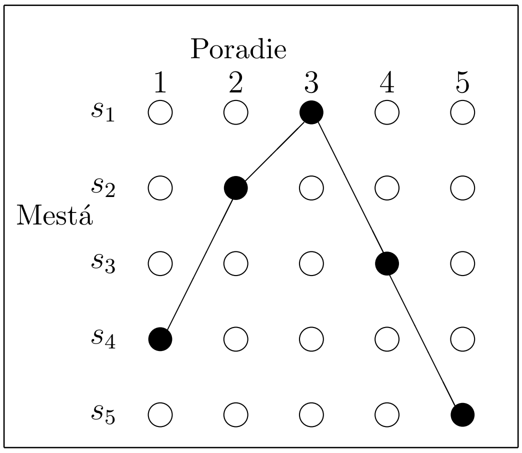
Majme danú množinu n miest a vzdialenosť $d_{i,j}$ pre všetky dvojice miest i, j. Potrebujeme zistiť, v akom poradí má obchodný cestujúci prechádzať mestá tak, aby sa vrátil do mesta, z ktorého vyrazil a zároveň každé mesto na trase navštívil práve jedenkrát. Podmienkou však je, že vzdialenosť, ktorú precestoval, je najkratšia možná. Riešenie tohto problému je prípustné, ak sú splnené nasledujúce obmedzenia :

i. obchodný cestujúci prechádza každým mestom práve jedenkrát,

ii. jeho cesta v grafe tvorí cyklus.

Prípustnosť riešenia je teda daná nájdením hamiltonovskej kružnice a mieru optimality riešenia popisuje súčet vzdialeností medzi mestami na Hamiltonovej kružnici.

Neurónová sieť, ktorá tento problém rieši musí vyjadriť pre dané mesto, ktoré mesto mu predchádza a ktoré mesto za ním nasleduje. Toto je vyjadriteľné v sieti s n^2 neurónmi pri maticovom usporiadaní uvedenom na obr. 1.



Obr. 1. Návrh neurónovej siete pre TSP pri $n = 5$. Kvôli prehľadnosti nie sú zakreslené prepojenia neurónov. Riešenie je zobrazené pomocou čiar. Riešením je postupnosť miest 4-2-1-3-5, prípadne jej cyklická premutácia.

Je zrejmé, že usporiadanie v tomto tvare pomáha pri pochopení riešenia, inak je nepodstatné. Vzťah pre energiu siete je možné upraviť pomocou zavedenia dvoch indexov pre jeden neurón. Jedná sa len o technickú úpravu vzťahu, význam zostane nezmenený.

Teraz sformulujeme túto úlohu pomocou novej sústavy dvojstavových premenných tak, aby hľadanie prípustného riešenia mohlo byť vyjadrené ako minimalizácia funkcie týchto nových premenných. Zadefinujme maticu V typu $n \times n$ s prvkami V_{ij}, ktoré nadobúdajú hodnoty 0 alebo 1, pre $i, j \in \langle 1, n \rangle$. $V_{ij} = 1$ vtedy, keď obchodný cestujúci prechádza mestom i v j-tom kroku. V opačnom prípade $V_{ij} = 0$. Prípustnému riešeniu pôvodnej úlohy teraz zodpovedá stav matice, ktorý sa dá popísať nasledovne :

a) v každom riadku je najviac jedna jednička, t.j. pre x-te mesto platí $V_{xj}V_{xl} = 0$, ak $j \neq l$,

b) v každom stĺpci je najviac jedna jednička, tj. pre x-ty krok obchodného cestujúceho platí $V_{ix}V_{kx} = 0$, ak $i \neq k$,

c) v matici je práve n jedničiek, t.j. $\sum_i^n \sum_j^n V_{ij} = n$,

d) súčet vzdialeností medzi mestami je určený maticou vzdialeností, pričom celková dĺžka cesty je $\frac{1}{2}\sum_i^n \sum_{k\neq i}^n \sum_y^n d_{ki}V_{ky}(V_{i,y-1} + V_{i,y+1})$

Teraz vytvoríme funkcie E_a, E_b, E_c a E_d, ktoré nadobúdajú minimálne hodnoty pri splnení predchádzajúcich podmienok. Funkcie sú vytvárané na základe vzťahov uvedených v podmienkach a)-d).

$$E_a = \frac{A}{2}\sum_x^n \sum_j^n \sum_{l\neq j}^n V_{xj}V_{xl} \tag{1}$$

$$E_b = \frac{B}{2}\sum_i^n \sum_x^n \sum_{i\neq k}^n V_{ix}V_{kx} \tag{2}$$

$$E_c = \frac{C}{2}(\sum_i^n \sum_j^n V_{ij} - n)^2 \tag{3}$$

$$E_d = \frac{D}{2}\sum_i^n \sum_{k\neq i}^n \sum_y^n d_{ki}V_{ky}(V_{i,y-1} + V_{i,y+1}) \tag{4}$$

Vo vzťahoch (1)-(4) A, B, C a D sú voliteľné parametre. Takže riešenie je optimálne ak $E = E_a + E_b + E_c + E_d$ nadobúda svoje minimum. Na tomto mieste je vhodné pripomenúť si, ako vyzerá energetická funkcia pre diskrétny model Hopfieldovej siete pri maticovom očíslovaní neurónov:

$$E_{hop} = -\frac{1}{2}\sum_i^n \sum_j^n \sum_k^n \sum_l^n w_{ij,kl}V_{ij}V_{kl} + \sum_i^n \sum_j^n \Theta_{ij}V_{ij} \tag{5}$$

Prepísaním (1),(2),(3) a (4) na tvar energetickej funkcie (5) dosiahneme to, že tvar "našej" funkcie E bude ekvivalentný tvaru funkcie E_{hop}, ktorá sa počas konvergencie siete minimalizuje (základná vlastnosť Hopfieldovej siete). Po tomto prepísaní nám teda bude stačiť porovnať výsledky s (5) a ľahko vypočítame nastavenie váh a prahov. Nasledujúce úpravy spočívajú v zavedení symbolu Croneckerovo delta, tj. $\delta_{ij} = 1$, ak $i = j$, inak $\delta_{ij} = 0$, čo nám umožní doplniť do niektorých výrazov ďalšie sumy.
Z (1) vyplýva

$$E_a = \frac{A}{2}\sum_i^n \sum_j^n \sum_k^n \sum_l^n \delta_{ik}(1 - \delta_{jl})V_{ij}V_{kl} \tag{6}$$

Výraz $\delta_{ik}(1 - \delta_{jl})$ je rovný 0, ak $i \neq k$ alebo $j = l$.

$$E_a = -\frac{1}{2}\sum_i^n \sum_j^n \sum_k^n \sum_l^n -A\delta_{ik}(1 - \delta_{jl})V_{ij}V_{kl} \tag{7}$$

Tento výraz je vlastne výraz (5), kde

$$w^a_{ij,kl} = -A\delta_{ik}(1 - \delta_{jl}) \tag{8}$$

Analogicky z (2) dostaneme

$$w^b_{ij,kl} = -B\delta_{jl}(1-\delta_{ik}) \tag{9}$$

Z (3) dostaneme

$$E_c = \frac{C}{2}(\sum_i^n \sum_j^n \sum_k^n \sum_l^n V_{ij}V_{kl} - 2n \sum_i^n \sum_j^n V_{ij} + n^2) \tag{10}$$

Po prenásobení výrazu v zátvorke výrazom $\frac{C}{2}$ dostaneme

$$E_c = -\frac{1}{2} \sum_i^n \sum_j^n \sum_k^n \sum_l^n -CV_{ij}V_{kl} + \sum_i^n \sum_j^n -CnV_{ij} + \frac{C}{2}n^2 \tag{11}$$

Pretože posledný člen $\frac{C}{2}n^2$ nezmení polohu minima vyššie uvedenej funkcie E_c, zanedbáme ho a dostávame

$$w^c_{ij,kl} = -C, \qquad \Theta_{ij} = -Cn. \tag{12}$$

Zo (4) vyplýva

$$E_d = \frac{D}{2} \sum_i^n \sum_k^n \sum_y^n d_{ki}V_{ky}V_{i,y-1} + \frac{D}{2} \sum_i^n \sum_k^n \sum_y^n d_{ki}V_{ky}V_{i,y+1}) \tag{13}$$

pričom platí, že $d_{ik} = 0$, ak $i = k$. Zavedením Croneckerovho symbolu δ_{lx} a nahradením indexu y indexom l vo výraze pre E_d dostávame

$$E_d = \frac{D}{2} \sum_i^n \sum_x^n \sum_k^n \sum_l^n d_{ki}\delta_{lx}V_{kl}V_{i,x-1} + \frac{D}{2} \sum_i^n \sum_x^n \sum_k^n \sum_l^n d_{ki}\delta_{lx}V_{kl}V_{i,x+1} \tag{14}$$

$E_d = 0$, ak $l \neq x$.
Položíme $j = x-1$ v prvej časti výrazu a $j = x+1$ v druhej časti a dostaneme

$$= \frac{D}{2} \sum_i^n \sum_j^n \sum_k^n \sum_l^n (d_{ki}\delta_{l,j+1}V_{kl}V_{ij} + d_{ki}\delta_{l,j-1}V_{kl}V_{ij}) \tag{15}$$

$$= -\frac{1}{2} \sum_i^n \sum_j^n \sum_k^n \sum_l^n -Dd_{ki}(\delta_{l,j+1} + d_{ki}\delta_{l,j-1})V_{kl}V_{ij} \tag{16}$$

$$w^d_{ij,kl} = -Dd_{ki}(\delta_{l,j+1} + \delta_{l,j-1}) \tag{17}$$

Takže nakoniec pre nastavenie váh prepojení v sieti medzi neurónmi x_{ij} a x_{kl} bude platiť

$$w_{ij,kl} = w^a_{ij,kl} + w^b_{ij,kl} + w^c_{ij,kl} + w^d_{ij,kl} \tag{18}$$

$$w^d_{ij,kl} = -A\delta_{ik}(1-\delta jl) - B\delta_{jl}(1-\delta_{ik}) - C - Dd_{ki}(\delta_{l,j+1}\delta_{l,j-1}) \tag{19}$$

Po odvodení vyššie uvedených vzťahov možeme uviesť nasledujúci algoritmus:

ALGORITMUS pre TSP:

Krok 1. Priradenie váh prepojeniam

$w_{ij,kl} = -A\delta_{ik}(1-\delta_{jl}) - B\delta_{jl}(1-\delta_{ik}) - C - Dd_{ki}(\delta_{l,j+1} + \delta_{l,j-1})$
pre $1 \le i, j, k, l \le n$,
- A,B,C,D sú zvolené parametre siete (sú meniteľné),

Krok 2. Inicializácia

$V_{ij}(0) = x_{ij}$, pre $1 \le i, j \le n$,

V tejto formule $V_{ij}(t)$ je výstup vrcholu ij v čase $t = 0$ a x_{ij} je náhodná premenná, ktorá nadobúda hodnotu 0 alebo 1.

Krok 3. Iterácia pokiaľ sieť neskovergovala

$V_{ij}(t+1) = g_h[\sum_{k=1}^{n} \sum_{i=1}^{n} w_{ij,kl}.V_{kl}(t)]$
pre $1 \le i, j \le n$.

Krok končí, ak sieť skonvergovala, tj. jej stav sa už nemení. Tu môže dôjsť k zacykleniu siete.

Krok 4. Opakovanie od kroku 2.

Ak došlo k zacykleniu siete, je potrebný nový výpočet začínajúci nastavením nových počiatočných hodnôt siete. Tiež je možné zmeniť parametre siete a nastaviť nové váhy, t.j. začať krokom 1.