Risanje velikih grafov s pomočjo linearne algebre

Gašper Ažman

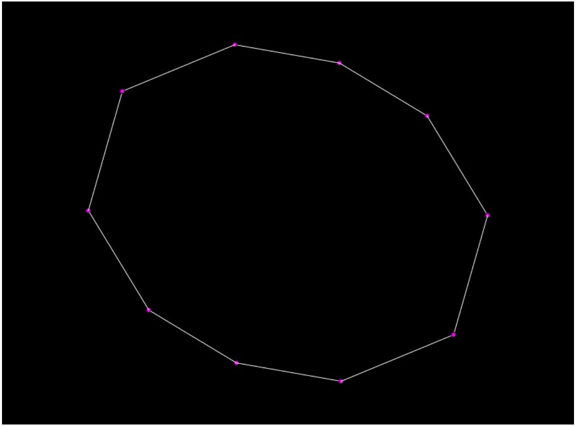
6. junij 2012

Kaj bomo videli?

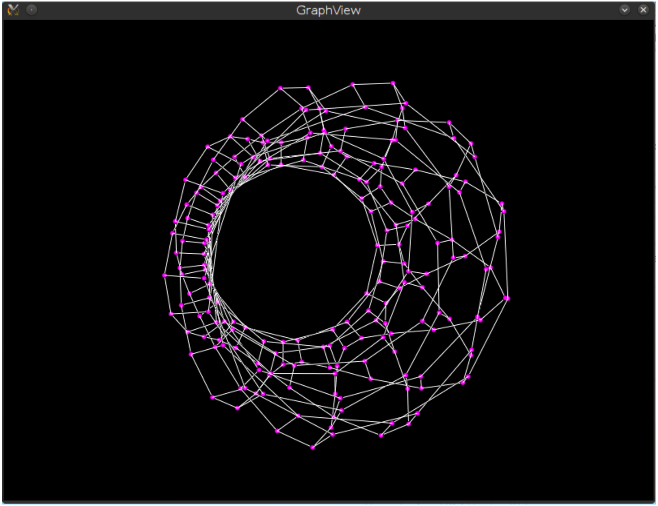
Seznam povezav:

0	1
1	2
2	3
3	4
4	5
5	6
6	7
7	8
8	9
9	0

$\longrightarrow$

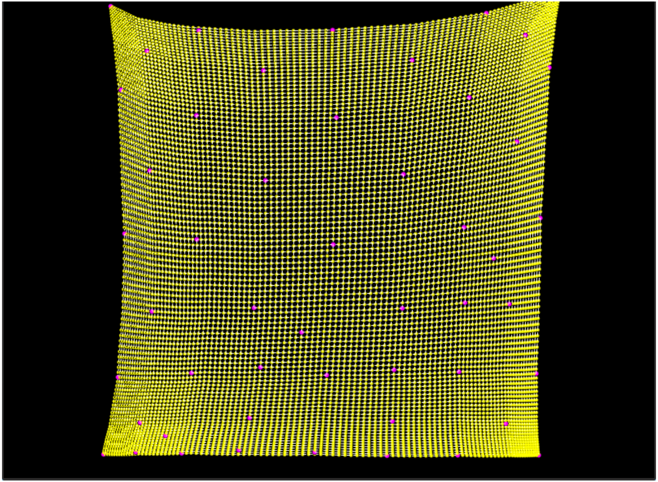


Slika: 10 cikel



Slika: Posplošen petersenov graf 101 37

Mreža 100x100 točk



Slika: Mreža 100 x 100

Osnovna ideja: MDS

Opozorilo: ta algoritem je za velike grafe neuporaben.
Osnovni algoritem:

1. Vhod: seznam povezav povezanega grafa.
2. Izračunamo razdaljno matriko ($O(n^2)$ če z BFS)

$$\begin{bmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ 1 & 2 & 2 & 1 & 0 \end{bmatrix}$$

Osnovna ideja: MDS

Opozorilo: ta algoritem je za velike grafe neuporaben.
Osnovni algoritem:

1. Vhod: seznam povezav povezanega grafa.
2. Izračunamo razdaljno matriko ($O(n^2)$ če z BFS)
3. Dobimo matriko skalarnih produktov B $b_{ij} = \langle x_i, x_j \rangle$. ($O(n^2)$) torej $B = X^T X$.

$$\begin{bmatrix} 0.8 & -0.1 & 0.2 & 0.2 & -0.1 \\ -0.1 & 0.8 & -0.1 & 0.2 & 0.2 \\ 0.2 & -0.1 & 0.8 & -0.1 & 0.2 \\ 0.2 & 0.2 & -0.1 & 0.8 & -0.1 \\ -0.1 & 0.2 & 0.2 & -0.1 & 0.8 \end{bmatrix}$$

Opozorilo: ta algoritem je za velike grafe neuporaben.
Osnovni algoritem:

1. Vhod: seznam povezav povezanega grafa.
2. Izračunamo razdaljno matriko ($O(n^2)$ če z BFS)
3. Dobimo matriko skalarnih produktov B $b_{ij} = \langle x_i, x_j \rangle$. ($O(n^2)$) torej $B = X^T X$.
4. Izračunamo d lastnih parov te matrike. ($O(dn)$)

$$\begin{bmatrix} 0.450415 & -0.625364 & 0.561445 \\ -0.443989 & 0.094448 & 0.291169 \\ -0.447214 & -0.447214 & -0.447214 \\ -0.60556 & -0.360673 & 0.382652 \\ -0.182476 & 0.519533 & 0.503565 \end{bmatrix}$$

Osnovna ideja: MDS

Opozorilo: ta algoritem je za velike grafe neuporaben.
Osnovni algoritem:

1. Vhod: seznam povezav povezanega grafa.
2. Izračunamo razdaljno matriko ($O(n^2)$ če z BFS)
3. Dobimo matriko skalarnih produktov B $b_{ij} = \langle x_i, x_j \rangle$. ($O(n^2)$) torej $B = X^T X$.
4. Izračunamo d lastnih parov te matrike. ($O(dn)$)
5. Uporabimo transponirano matriko izbranih lastnih vektorjev za koordinate (skaliramo s korenom lastne vrednosti)

$$\begin{bmatrix} 0.469256 & -0.651523 & 0.561445 \\ -0.462562 & 0.098399 & 0.291169 \\ -0.465921 & -0.465921 & -0.447214 \\ -0.63089 & -0.37576 & 0.382652 \\ -0.190109 & 0.541266 & 0.503565 \end{bmatrix}$$

Matrika skalarnih produktov

Iščemo tako razporeditev $X = [x_1, \ldots, x_n]$, da bodo razdalje med točkami v $\mathbb{R}^d$ enake razdaljam v grafu.

$$\|x_i - x_j\| = d_{ij}$$

BŠS lahko postavimo izhodišče v baricenter:

$$\sum_{i=1}^{n} x_i = 0$$

Iz $\|x_i - x_j\|^2 = \langle x_i - x_j, x_i - x_j \rangle$ lahko $\langle x_i, x_j \rangle$ izrazimo kot

$$\langle x_i, x_j \rangle = -\frac{1}{2}(d_{ij}^2 - \|x_i\|^2 - \|x_j\|^2),$$

Za vsak j:

$$\begin{aligned}
\sum_{i=1}^{n} d_{ij}^2 &= \sum_{i=1}^{n} (\|x_i\|^2 - 2\langle x_i, x_j\rangle + \|x_j\|^2) \qquad\qquad \sum_{i=1}^{n} x_i = 0 \\
&= \sum_{i=1}^{n} (\|x_i\|^2 + \|x_j\|^2) - \left\langle \left(\sum_{i=1}^{n} x_i\right), x_j \right\rangle \\
&= \sum_{i=1}^{n} (\|x_i\|^2 + \|x_j\|^2) \\
&= \sum_{i=1}^{n} \|x_i\|^2 + n\|x_j\|^2 \\
\Rightarrow \|x_j\|^2 &= \frac{1}{n}\sum_{i=1}^{n} d_{ij}^2 - \frac{1}{n}\sum_{i=1}^{n} \|x_i\|^2
\end{aligned}$$

Seveda velja simetrično tudi za vsak i.

Matrika skalarnih produktov (nadalj. 2)

Če vzamemo enačbo za j in seštejemo po vseh j:

$$\forall j : \sum_{i=1}^{n} d_{ij}^2 = \sum_{i=1}^{n} \|x_i\|^2 + n\|x_j\|^2$$

$$\sum_{j=1}^{n} \sum_{i=1}^{n} d_{ij}^2 = n \sum_{i=1}^{n} \|x_i\|^2 + n \sum_{j=1}^{n} \|x_j\|^2$$

$$\sum_{j=1}^{n} \sum_{i=1}^{n} d_{ij}^2 = 2n \sum_{i=1}^{n} \|x_i\|^2$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} \|x_i\|^2 = \frac{1}{2n^2} \sum_{j=1}^{n} \sum_{i=1}^{n} d_{ij}^2$$

Matrika skalarnih produktov (nadalj. 3)

Prej smo videli:

$$\langle x_i, x_j \rangle = -\frac{1}{2}(d_{ij}^2 - \|x_i\|^2 - \|x_j\|^2),$$

$$\|x_j\|^2 = \frac{1}{n}\sum_{i=1}^{n} d_{ij}^2 - \frac{1}{n}\sum_{i=1}^{n} \|x_i\|^2$$

$$\frac{1}{n}\sum_{i=1}^{n} \|x_i\|^2 = \frac{1}{2n^2}\sum_{j=1}^{n}\sum_{i=1}^{n} d_{ij}^2$$

Od tu vidimo, da

$$\langle x_i, x_j \rangle = -\frac{1}{2}\left(d_{ij}^2 - \frac{1}{n}\sum_{i=1}^{n} d_{ij}^2 - \frac{1}{n}\sum_{i=j}^{n} d_{ij}^2 + \frac{1}{n^2}\sum_{i=1,j=1}^{n} d_{ij}^2 \right)$$

Kako pa to lahko naredimo hitro?

Ideja: izberemo nekaj točk, glede na katere bomo pozicionirali vse ostale. Rečemo jim *pivoti*.

1. Izberemo k pivotov.
2. Izračunamo razdalje od pivotov do vseh točk in dobimo $k \times n$ matriko. ($O(kn)$)
3. Izračunamo $k \times n$ matriko približkov[1] C za $\langle x_v, x_p \rangle$ za vsak pivot p. ($O(kn)$)
4. Konfiguracijo dobimo z izračunom d singularnih vektorjev matrike C. ($O(k^2 n)$)

[1] Zelo podobno kot prej

Razlike v izračunu matrike skalarnih produktov

Dodatna predpostavka: Pivoti so dobro razporejeni po grafu, zato

$$\sum_{p \in P} x_p = 0$$

Po istem postopku kot prej dobimo

$$\langle x_v, x_p \rangle = -\frac{1}{2} \left(d_{vp}^2 - \frac{1}{k} \sum_{p \in P} d_{vp}^2 - \frac{1}{n} \sum_{i=1}^{n} d_{vp}^2 + \frac{1}{nk} \sum_{v \in V, p \in P} d_{vp}^2 \right)$$

Izračun singularnih vektorjev

Imamo matriko $k \times n$, torej je globoka in ozka. Matrika CC^T je zato $n \times n$, kar je preveliko.
Vektorje bomo računali s potenčno metodo:

$$y = \lim_{i \to \infty} \frac{(CC^T)^i x}{\|(CC^T)^i x\|} \qquad = \frac{C \lim_{i \to \infty} \frac{(C^T C)^i x}{\|(C^T C)^i x\|}}{\| \uparrow \|}$$

Matrika $C^T C$ je $k \times k$, torej je majhna in na njej hitro izvajamo potenčno metodo.
Skaliranje končnega vektorja opravimo kar z množenjem s C.

Primeri

Hvala za pozornost!

Vir: C. Pich, *Applications of multidimensional scaling to graph drawing, Doktorsko delo*, univ. v Konstanzu, 2009