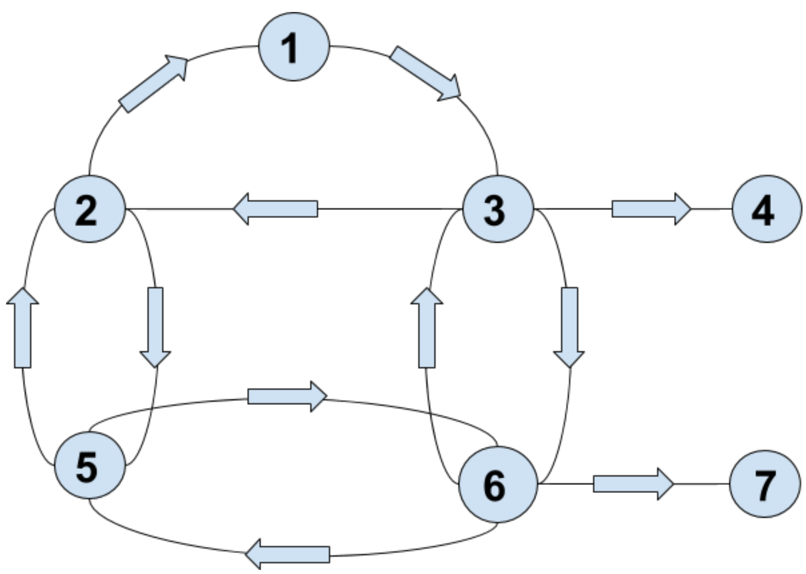
Linjär Algebra
Föreläsning 19

Erik Sjöström

December 14, 2015

1 Sökalgoritmer

Man tänker sig internet som en riktad graf där noden v_i är en webbsida och kanten (v_i, v_j) är länk från webbsida v_i till v_j.



I t.ex. Googles sökalgoritm tänker man sig att sidor med många länkar in eller ut är mer relevanta/viktiga för ett visst ämne/sökfråga.

Exempel 1.1.
Sida ③ ovan är mer relevant än t.ex. ⑦

Övergångsmatrisen för grafen ovan:

$$\mathbf{P} = \begin{bmatrix} 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/3 & 0 & 1/2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 1/3 & 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 1/3 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 1 \end{bmatrix}$$

I en slumpvandring så kommer den stationära fördelningen att beskriva sannolikheten att vara på en viss sida i det långa loppet.

- Vi använde markovkedja för att utföra slumpvandring.

- Varje markovkedja med reguljär övergångsmatris $\mathbf{P}$ har en stationär fördelning. Som är egenvektorn hörande till egenvärdet 1 (det största egenvärdet i $\mathbf{P}$).
- $\mathbf{P}$ är reguljär om det finns något $\mathbf{P}^k$ där alla element > 0.

I ex. ovan, om man kommit till ④ eller ⑦ fastnar man. I slumpvandringen kommer man att få två egenvektorer hörande till egenvärdet 1.
I Googles fall justeras $\mathbf{P}$ enligt följande:

- Om vi hamnat i ⑦ eller ④ kan vi välja vilken nod som helst, och ta oss vidare. Kolumn 4 och 7 i $\mathbf{P}$ blir $\begin{bmatrix}1/7 & 1/7 & ... & 1/7\end{bmatrix}^T$

Vi får då:

$$\mathbf{P} = \begin{bmatrix} 0 & 1/2 & 0 & 1/7 & 0 & 0 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 1/2 & 0 & 1/7 \\ 1 & 0 & 0 & 1/7 & 0 & 1/3 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 0 & 0 & 1/7 \\ 0 & 1/2 & 0 & 1/7 & 0 & 1/3 & 1/7 \\ 0 & 0 & 1/3 & 1/7 & 1/2 & 0 & 1/7 \\ 0 & 0 & 0 & 1/7 & 0 & 1/3 & 1/7 \end{bmatrix}$$

- Även vissa typer av cykler i den underliggande grafen kan göra $\mathbf{P}$ icke-reguljär.

Vår Google-matris ($\mathbf{G}$) är:

$$\mathbf{G} = \overbrace{0.85 \cdot \mathbf{P}}^{1} + \overbrace{0.15[1/n]}^{2}$$

1. Om man är på viss sida, väljer man att klicka enligt grafen.
2. Väljer vilken nod som helst i gruppen.

- Hur beräknas egenvektorn till egenvärdet 1. Vi kan ej lösa:

$$(\mathbf{G} - \lambda \cdot \mathbf{I}) \cdot \vec{v}_1 = \emptyset$$

ty $\mathbf{G}$ är för stor.
Lösning: Potensmetoden

$$\vec{x}_0 = \text{standardfördelning}$$
$$\vec{x}_{k+1} = \mathbf{G} \cdot \vec{x}_k$$
$$k = 0, 1, 2, 3, ...$$

2 Potensmetoden (power method)

Låt $\lambda_1, \lambda_2, ..., \lambda_n$ vara egenvärden till $\mathbf{P}$ och sammanhörande egenvektorer $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$
Om:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ... \geq |\lambda_n|$$

Låt:

$$\vec{x} = c_1 \cdot \vec{v}_1 + c_2 \cdot \vec{v}_2 + ... + c_n \cdot \vec{v}_n$$

Vi har då:

$$\mathbf{G}^k \vec{x} = c_1 \cdot \mathbf{G}^k \cdot \vec{v}_1 + ... + c_n \mathbf{G}^k \cdot \vec{v}_n =$$
$$c_1 \lambda_1^k \cdot \vec{v}_1 + c_2 \lambda_2^k \cdot \vec{v}_2 + ... + c_n \lambda_n^k \cdot \vec{v}_n$$

Antag $c_1 \neq 0$, dela (λ_1^k):

$$\frac{1}{\lambda_1^k} \mathbf{G}^k \cdot \vec{x} = c_1 \vec{v}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^k \cdot c_1 \vec{v}_2 + ... + c_n \cdot \left(\frac{\lambda_n}{\lambda_1}\right)^k \vec{v}_n$$