Aplicações de Autovalores e Autovetores

Prof. Alexandre Lymberopoulos

Instituto de Matemática e Estatística
Universidade de São Paulo

1. Page Ranking

2. Cadeias de Markov
   - Biologia
   - Market Share

3. Índice de acessibilidade em Redes

4. Referências Bibliográficas

Buscas na WWW

- Objetivo: ordenar por “importância” o resultado de uma busca.

Buscas na WWW

- Objetivo: ordenar por “importância” o resultado de uma busca.
- O que define a importância de uma página?

Buscas na WWW

- Objetivo: ordenar por "importância" o resultado de uma busca.
- O que define a importância de uma página?
- A importância de uma página é proporcional à importância das páginas que apontam para ela.

Buscas na WWW

- Objetivo: ordenar por “importância” o resultado de uma busca.
- O que define a importância de uma página?
- A importância de uma página é proporcional à importância das páginas que apontam para ela.
- O método que veremos é creditado a Kendall e Wei, em meados da década de 1950, ver [3, 6].

Buscas na WWW

- Objetivo: ordenar por “importância” o resultado de uma busca.
- O que define a importância de uma página?
- A importância de uma página é proporcional à importância das páginas que apontam para ela.
- O método que veremos é creditado a Kendall e Wei, em meados da década de 1950, ver [3, 6].
- Os inventores do Google descrevem o método utilizado por eles (semelhantes ao aqui apresentado) com algum detalhe em [1].

Buscas na WWW

- Objetivo: ordenar por "importância" o resultado de uma busca.
- O que define a importância de uma página?
- A importância de uma página é proporcional à importância das páginas que apontam para ela.
- O método que veremos é creditado a Kendall e Wei, em meados da década de 1950, ver [3, 6].
- Os inventores do Google descrevem o método utilizado por eles (semelhantes ao aqui apresentado) com algum detalhe em [1].
- Este modelo pode ser usado para "rankear" competidores que se enfrentam diretamente, veja [7].

O Modelo

- Seja $x_i, 1 \leq i \leq n$, a importância i-ésimo site.

O Modelo

- Seja $x_i, 1 \leq i \leq n$, a importância i-ésimo site.
- A hipótese do modelo é que a importância de um site é proporcional à importância dos sites que apontam para ele.

O Modelo

- Seja $x_i, 1 \leq i \leq n$, a importância i-ésimo site.
- A hipótese do modelo é que a importância de um site é proporcional à importância dos sites que apontam para ele.
- Isto nos leva a um sistema de equações do tipo

$$\left\{ \begin{array}{lcl} x_1 & = & K(x_2 + x_{14} + x_{541}) \\ x_2 & = & K(x_1 + x_{23} + x_{541} + x_{1023}) \\ & \vdots & \\ x_n & = & K(x_{25} + x_{133}) \end{array} \right. , \qquad (1)$$

O Modelo

- Seja $x_i, 1 \leq i \leq n$, a importância i-ésimo site.
- A hipótese do modelo é que a importância de um site é proporcional à importância dos sites que apontam para ele.
- Isto nos leva a um sistema de equações do tipo

$$\left\{ \begin{array}{lcl} x_1 & = & K(x_2 + x_{14} + x_{541}) \\ x_2 & = & K(x_1 + x_{23} + x_{541} + x_{1023}) \\ & \vdots & \\ x_n & = & K(x_{25} + x_{133}) \end{array} \right. , \tag{1}$$

- onde K é a constante de proporcionalidade que está multiplicada, em cada linha, pela soma das importâncias dos sites que apontam para x_i.

O Modelo

- Cada linha do sistema em (1) pode ser ressсrita como

$$\sum_{j=1}^{n} a_{ij} x_i = \frac{1}{K} x_i \iff Ax = \frac{1}{K} x. \tag{2}$$

O Modelo

- Cada linha do sistema em (1) pode ser ressctrita como

$$\sum_{j=1}^{n} a_{ij} x_i = \frac{1}{K} x_i \iff Ax = \frac{1}{K} x. \tag{2}$$

- Estamos procurando então um autovetor de coordenadas não negativas associada ao autovalor $\frac{1}{K}$ da matriz A.

O Modelo

- Cada linha do sistema em (1) pode ser ressrita como

$$\sum_{j=1}^{n} a_{ij}x_i = \frac{1}{K}x_i \iff Ax = \frac{1}{K}x. \tag{2}$$

- Estamos procurando então um autovetor de coordenadas não negativas associada ao autovalor $\frac{1}{K}$ da matriz A.
- Tal autovetor sempre existe devido ao teorema de Perron.

O Modelo

- Cada linha do sistema em (1) pode ser ressrita como

$$\sum_{j=1}^{n} a_{ij} x_i = \frac{1}{K} x_i \iff Ax = \frac{1}{K} x. \tag{2}$$

- Estamos procurando então um autovetor de coordenadas não negativas associada ao autovalor $\frac{1}{K}$ da matriz A.
- Tal autovetor sempre existe devido ao teorema de Perron.
- Uma vez determinado este autovetor, sua maior coordenada é a coordenada do site mais importante, a segunda maior é a do segundo site mais importante e assim sucessivamente.

O Modelo

- Cada linha do sistema em (1) pode ser ressrcrita como

$$\sum_{j=1}^{n} a_{ij} x_i = \frac{1}{K} x_i \iff Ax = \frac{1}{K} x. \tag{2}$$

- Estamos procurando então um autovetor de coordenadas não negativas associada ao autovalor $\frac{1}{K}$ da matriz A.
- Tal autovetor sempre existe devido ao teorema de Perron.
- Uma vez determinado este autovetor, sua maior coordenada é a coordenada do site mais importante, a segunda maior é a do segundo site mais importante e assim sucessivamente.
- Basta então listar os sites de acordo com a ordem decrescente de importância, o que poupa muito tempo nas buscas.

O Modelo

- Cada linha do sistema em (1) pode ser ressscrita como

$$\sum_{j=1}^{n} a_{ij} x_i = \frac{1}{K} x_i \iff Ax = \frac{1}{K} x. \tag{2}$$

- Estamos procurando então um autovetor de coordenadas não negativas associada ao autovalor $\frac{1}{K}$ da matriz A.
- Tal autovetor sempre existe devido ao teorema de Perron.
- Uma vez determinado este autovetor, sua maior coordenada é a coordenada do site mais importante, a segunda maior é a do segundo site mais importante e assim sucessivamente.
- Basta então listar os sites de acordo com a ordem decrescente de importância, o que poupa muito tempo nas buscas.
- No caso de uma busca no Google a matriz A é gigante!

Um exemplo

- Suponha que temos uma liga de basquete com $n > 10$ times. Como encontrar os “top 10”?

Um exemplo

- Suponha que temos uma liga de basquete com $n > 10$ times. Como encontrar os "top 10"?
- Podemos aplicar o modelo de "page ranking"!

Um exemplo

- Suponha que temos uma liga de basquete com $n > 10$ times. Como encontrar os “top 10”?
- Podemos aplicar o modelo de “page ranking”!
- Se x_i é a força do i-ésimo time, podemos assumir que ela é proporcional à soma das forças dos times que i venceu.

Um exemplo

- Suponha que temos uma liga de basquete com $n > 10$ times. Como encontrar os "top 10"?
- Podemos aplicar o modelo de "page ranking"!
- Se x_i é a força do i-ésimo time, podemos assumir que ela é proporcional à soma das forças dos times que i venceu.
- Isto nos leva a um sistema do tipo (2) onde $a_{ij} = 1$ se i venceu j e $a_{ij} = 0$ em caso contrário.

Um exemplo

- Suponha que temos uma liga de basquete com $n > 10$ times. Como encontrar os "top 10"?
- Podemos aplicar o modelo de "page ranking"!
- Se x_i é a força do i-ésimo time, podemos assumir que ela é proporcional à soma das forças dos times que i venceu.
- Isto nos leva a um sistema do tipo (2) onde $a_{ij} = 1$ se i venceu j e $a_{ij} = 0$ em caso contrário.
- Exemplo: Determine o ranking para 6 times cuja matriz A é dada por

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

Definições

- Uma cadeia de Markov é um modelo para um sistema que admite n estados diferentes e o estado num dado instante k depende somente do estado imediatamente anterior, ou seja, o instante $k - 1$.

Definições

- Uma cadeia de Markov é um modelo para um sistema que admite n estados diferentes e o estado num dado instante k depende somente do estado imediatamente anterior, ou seja, o instante $k-1$.
- A matriz $T = (p_{ij})$ cujo elemento p_{ij} é a probabilidade de sairmos do estado i para o estado j no instante seguinte é chamada *matriz de transição do modelo*.

Definições

- Uma cadeia de Markov é um modelo para um sistema que admite n estados diferentes e o estado num dado instante k depende somente do estado imediatamente anterior, ou seja, o instante $k-1$.
- A matriz $T = (p_{ij})$ cujo elemento p_{ij} é a probabilidade de sairmos do estado i para o estado j no instante seguinte é chamada *matriz de transição do modelo*.
- Fato: a soma dos elementos de cada coluna de T é 1.

Definições

- Uma cadeia de Markov é um modelo para um sistema que admite n estados diferentes e o estado num dado instante k depende somente do estado imediatamente anterior, ou seja, o instante $k-1$.
- A matriz $T=(p_{ij})$ cujo elemento p_{ij} é a probabilidade de sairmos do estado i para o estado j no instante seguinte é chamada *matriz de transição do modelo*.
- Fato: a soma dos elementos de cada coluna de T é 1.
- T é dita *regular* se alguma potência de T possui todos os elementos positivos.

Definições

- Uma cadeia de Markov é um modelo para um sistema que admite n estados diferentes e o estado num dado instante k depende somente do estado imediatamente anterior, ou seja, o instante $k-1$.
- A matriz $T = (p_{ij})$ cujo elemento p_{ij} é a probabilidade de sairmos do estado i para o estado j no instante seguinte é chamada *matriz de transição do modelo*.
- Fato: a soma dos elementos de cada coluna de T é 1.
- T é dita *regular* se alguma potência de T possui todos os elementos positivos.
- Se $x^{(k)}$ é um vetor com as probabilidades de cada estado ocorrer no instante k então $Tx^{(k)}$ é o vetor com as probabilidades de cada estado ocorrer no instante $k+1$, ou seja,

$$x^{(k+1)} = Tx^{(k)}. \qquad (3)$$

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então
  1. $\lim_{n \to \infty} T^n = A$, onde A é uma matriz com todas as colunas iguais, chamemo-as de u.

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então
  1. $\lim_{n\to\infty} T^n = A$, onde A é uma matriz com todas as colunas iguais, chamemo-as de u.
  2. O vetor u acima é um vetor de probabilidade (todos positivos e somando 1).

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então
  1. $\lim_{n\to\infty} T^n = A$, onde A é uma matriz com todas as colunas iguais, chamemo-as de u.
  2. O vetor u acima é um vetor de probabilidade (todos positivos e somando 1).
- Se T é a matriz de transição regular de alguma cadeia de Markov, A e u são como acima então

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então
  1. $\lim_{n\to\infty} T^n = A$, onde A é uma matriz com todas as colunas iguais, chamemo-as de u.
  2. O vetor u acima é um vetor de probabilidade (todos positivos e somando 1).
- Se T é a matriz de transição regular de alguma cadeia de Markov, A e u são como acima então
  1. Para todo vetor de probabilidade x temos $\lim_{n\to\infty} T^n x = u$.

Alguns resultados importantes

- Se T é a matriz de transição regular de alguma cadeia de Markov então
  1. $\lim\limits_{n\to\infty} T^n = A$, onde A é uma matriz com todas as colunas iguais, chamemo-as de u.
  2. O vetor u acima é um vetor de probabilidade (todos positivos e somando 1).
- Se T é a matriz de transição regular de alguma cadeia de Markov, A e u são como acima então
  1. Para todo vetor de probabilidade x temos $\lim\limits_{n\to\infty} T^n x = u$.
  2. O vetor u acima é o único tal que $Tu = u$.

Genética

- Pesquisadores em uma universidade descobriram que vacas com genótipo AA produzem leite de qualidade melhor que as com outros genótipos. Esses pesquisadores desejam descobrir a fração de descendentes dessas vacas com o genótipo AA. Se são cruzados somente vacas de genótipo AA com as de outros genótipos quais são as probabilidades de obtermos cada um dos genótipos (AA, Aa, ou aa)?

Genética

- Pesquisadores em uma universidade descobriram que vacas com genótipo AA produzem leite de qualidade melhor que as com outros genótipos. Esses pesquisadores desejam descobrir a fração de descendentes dessas vacas com o genótipo AA. Se são cruzados somente vacas de genótipo AA com as de outros genótipos quais são as probabilidades de obtermos cada um dos genótipos (AA, Aa, ou aa)?
- Suponhamos que as probabilidades de propagação dos genes são uniformes.

Genética

- Pesquisadores em uma universidade descobriram que vacas com genótipo AA produzem leite de qualidade melhor que as com outros genótipos. Esses pesquisadores desejam descobrir a fração de descendentes dessas vacas com o genótipo AA. Se são cruzados somente vacas de genótipo AA com as de outros genótipos quais são as probabilidades de obtermos cada um dos genótipos (AA, Aa, ou aa)?
- Suponhamos que as probabilidades de propagação dos genes são uniformes.
- Com isso temos uma matriz de transição de genes dada por

$$A = \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Genética

- Pesquisadores em uma universidade descobriram que vacas com genótipo AA produzem leite de qualidade melhor que as com outros genótipos. Esses pesquisadores desejam descobrir a fração de descendentes dessas vacas com o genótipo AA. Se são cruzados somente vacas de genótipo AA com as de outros genótipos quais são as probabilidades de obtermos cada um dos genótipos (AA, Aa, ou aa)?
- Suponhamos que as probabilidades de propagação dos genes são uniformes.
- Com isso temos uma matriz de transição de genes dada por

$$A = \begin{bmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

- Partindo de proporções iguais de cada genótipo, o que obtemos após dois anos? E a longo prazo?

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:
- 20% dos que comeram no restaurante chinês vão ao mexicano, 20% ficam em casa e 30% vão à pizzaria no próximo jantar;

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:
- 20% dos que comeram no restaurante chinês vão ao mexicano, 20% ficam em casa e 30% vão à pizzaria no próximo jantar;
- 10% dos que comeram no restaurante mexicano vão à pizzaria, 25% ficam em casa e 25% vão ao chinês no próximo jantar;

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:
- 20% dos que comeram no restaurante chinês vão ao mexicano, 20% ficam em casa e 30% vão à pizzaria no próximo jantar;
- 10% dos que comeram no restaurante mexicano vão à pizzaria, 25% ficam em casa e 25% vão ao chinês no próximo jantar;
- 20% dos que comeram em casa vão ao restaurante chinês, 25% vão ao mexicano e 30% vão à pizzaria no próximo jantar;

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:
- 20% dos que comeram no restaurante chinês vão ao mexicano, 20% ficam em casa e 30% vão à pizzaria no próximo jantar;
- 10% dos que comeram no restaurante mexicano vão à pizzaria, 25% ficam em casa e 25% vão ao chinês no próximo jantar;
- 20% dos que comeram em casa vão ao restaurante chinês, 25% vão ao mexicano e 30% vão à pizzaria no próximo jantar;
- 30% dos que comeram na pizzaria jantam em casa, 30% vão ao chinês e 10% vão ao mexicano.

Comida chinesa, mexicana, pizza ou caseira?

- Numa pequena cidade as pessoas tem quatro opções de jantar: um restaurante chinês, um mexicano, uma pizzaria ou jantar em casa.
- Suponha que:
- 20% dos que comeram no restaurante chinês vão ao mexicano, 20% ficam em casa e 30% vão à pizzaria no próximo jantar;
- 10% dos que comeram no restaurante mexicano vão à pizzaria, 25% ficam em casa e 25% vão ao chinês no próximo jantar;
- 20% dos que comeram em casa vão ao restaurante chinês, 25% vão ao mexicano e 30% vão à pizzaria no próximo jantar;
- 30% dos que comeram na pizzaria jantam em casa, 30% vão ao chinês e 10% vão ao mexicano.
- Se numa noite a clientela é igualmente dividida, qual a proporção de clientes em cada restaurante depois de duas noites? E depois de muito tempo?

Aplicações

- Aplicações na física relativa a História e Geografia.

Aplicações

- Aplicações na física relativa a História e Geografia.
- Em Geografia: consideramos redes de transporte envolvendo ferrovias, rodovias, linhas aéreas e marítimas.

Aplicações

- Aplicações na física relativa a História e Geografia.
- Em Geografia: consideramos redes de transporte envolvendo ferrovias, rodovias, linhas aéreas e marítimas.
- Estas redes são representadas por grafos.

Aplicações

- Aplicações na física relativa a História e Geografia.
- Em Geografia: consideramos redes de transporte envolvendo ferrovias, rodovias, linhas aéreas e marítimas.
- Estas redes são representadas por grafos.
- Em História: a supremacia de Moscou é atribuída por historiadores russos à sua posição estratégica em relação às rotas comerciais russas.

Aplicações

- Aplicações na física relativa a História e Geografia.
- Em Geografia: consideramos redes de transporte envolvendo ferrovias, rodovias, linhas aéreas e marítimas.
- Estas redes são representadas por grafos.
- Em História: a supremacia de Moscou é atribuída por historiadores russos à sua posição estratégica em relação às rotas comerciais russas.
- Técnicas recentes da teoria de grafos foram usadas para por à prova tal teoria.

Aplicações

- Aplicações na física relativa a História e Geografia.
- Em Geografia: consideramos redes de transporte envolvendo ferrovias, rodovias, linhas aéreas e marítimas.
- Estas redes são representadas por grafos.
- Em História: a supremacia de Moscou é atribuída por historiadores russos à sua posição estratégica em relação às rotas comerciais russas.
- Técnicas recentes da teoria de grafos foram usadas para por à prova tal teoria.
- Com essas técnicas Moscou ficou na sexta posição, quando consideramos as rotas comerciais russas existentes no século 12. (Ver [4, 5]).

Exemplos

1. Considere as seguintes redes de rodovias (de mão dupla) em cada um dos grupos de cidades:

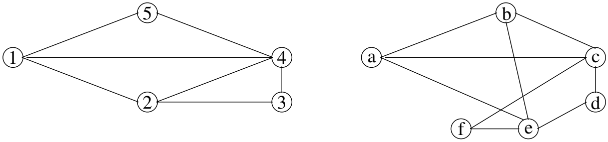

Figura: Uma rede rodoviária

Exemplos

1. Considere as seguintes redes de rodovias (de mão dupla) em cada um dos grupos de cidades:

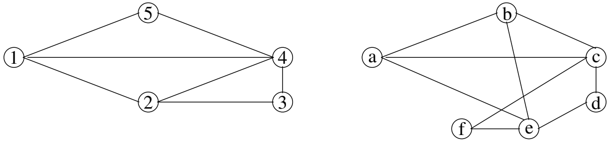

Figura: Uma rede rodoviária

- Construa as matrizes de adjacências e aumentada do grafo da figura 1.

Exemplos

1. Considere as seguintes redes de rodovias (de mão dupla) em cada um dos grupos de cidades:

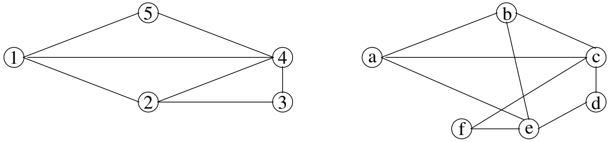

Figura: Uma rede rodoviária

- Construa as matrizes de adjacências e aumentada do grafo da figura 1.
- Calcule o índice de Gould (ver [2]) de cada cidade, ou seja cada coordenda do autovetor do autovalor dominante sobre a soma delas.

Exemplos

1. Considere as seguintes redes de rodovias (de mão dupla) em cada um dos grupos de cidades:

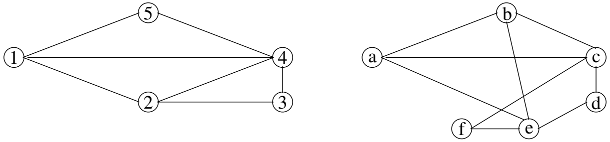

Figura: Uma rede rodoviária

- Construa as matrizes de adjacências e aumentada do grafo da figura 1.
- Calcule o índice de Gould (ver [2]) de cada cidade, ou seja cada coordenda do autovetor do autovalor dominante sobre a soma delas.
- Decida qual das cidades têm maior acessibilidade em cada grupo.

Sergey Brin and Lawrence Page.
The anatomy of a large-scale hypertextual web search engine.
Computer Networks and ISDN Systems, 30:107–117, 1998.

P. R. Gould.
On the geographical interpretation os eigenvectors.
Transactions of the Institute of British Geographers, 42:53–86, Dec. 1967.

M. G. Kendall.
Further contributions to the theory of paired comparisons.
Biometrics, 11:43, 1955.

Forrest Pitts.
A graph theoretical approach to historical geography.
The Professional Geography, 17(5):15–20, Sept. 1965.

Philip Straffin.
Linear algebra in geography: Eigenvectors of networks.
Mathematics Magazine, 53(5):269–276, Nov. 1980.

T. H. Wei.
The algebraic foundations of ranking theory.
PhD thesis, Cambridge, 1952.

H. Wilf.
Searching the web, using eigenvectors, April 2001.
http://www.math.upenn.edu/ wilf/website/KendallWei.pdf.