

PREVISÕES BAYESIANAS E MODELOS GRÁFICOS

Ralph dos Santos Silva

<http://www.ralphsilva.org/previsoesbayesianasemodelosgraficos.html>

Instituto de Matemática
Centro de Ciências Matemáticas e da Natureza
Universidade Federal do Rio de Janeiro

Modelos gráficos não direcionados

Nesta aula, veremos os **modelos gráficos não direcionados**.

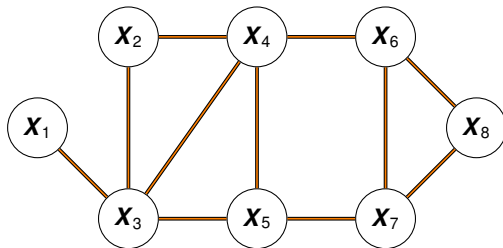


Figura: Um exemplo de modelo gráfico não direcionado.

Note que as arestas são não direcionadas.

Agora, as propriedades de independência condicional são diferentes.

Suponha que em um grafo não direcionado existam 3 conjuntos de vértices, A , B e C , tal que a propriedade de independência condicional seja

$$A \perp B | C.$$

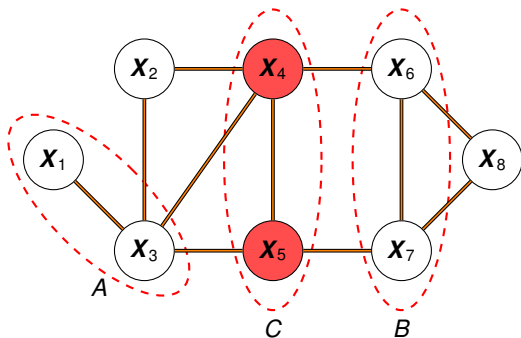


Figura: Independência condicional em um modelo gráfico não direcionado.

Para que $(A \perp B|C)$, precisamos verificar todos os caminhos que conectam os vértices em A aos vértices em B .

Se todos os caminhos passam por um ou mais vértices em C , então todos os caminhos estão **bloqueados**. Portanto, a propriedade de independência condicional se aplica.

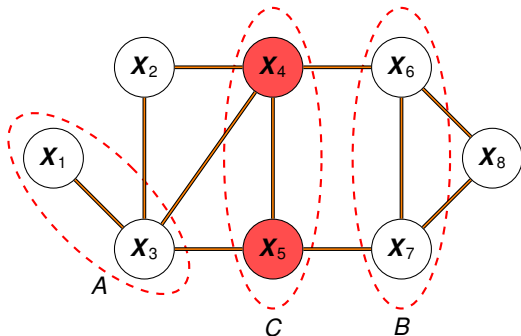


Figura: Independência condicional em um modelo gráfico não direcionado.

Fatorização

Expressaremos a distribuição conjunta $f(\mathbf{x})$ como um produto de funções definidas sobre um conjunto de variáveis que são **locais ao grafo**.

Se considerarmos dois vértices $\mathbf{X}_i$ e $\mathbf{X}_j$ que não são conectados por uma aresta, então estas variáveis devem ser condicionalmente independentes.

Isto advém do fato que não existe caminho direto entre os vértices e todos os outros caminhos passam por vértices que foram observados.

Esta propriedade de independência condicional pode ser expressa como

$$f(\mathbf{x}_i, \mathbf{x}_j | \mathbf{x}_{\setminus \{i,j\}}) = f(\mathbf{x}_i | \mathbf{x}_{\setminus \{i,j\}}) f(\mathbf{x}_j | \mathbf{x}_{\setminus \{i,j\}}),$$

em que $\mathbf{x}_{\setminus \{i,j\}}$ representa o conjunto de todas as variáveis exceto $\mathbf{x}_i$ e $\mathbf{x}_j$.

Definição

Um **clique** é definido como um subconjunto de vértices em um grafo tal que existe uma aresta entre todos os pares de vértices do subconjunto.

Um conjunto de vértices em um **clique** são completamente conectados.

Um **clique máximo** é um **clique** na qual não é possível incluir nenhum outro vértice do grafo no conjunto sem perder a propriedade de ser um clique.

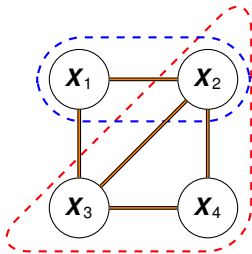


Figura: Um grafo não direcionado de 4 vértices com um clique (em azul) e um clique máximo (em vermelho).

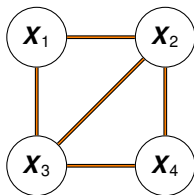


Figura: Um grafo não direcionado de 4 vértices.

O grafo não direcionado na figura acima tem a seguinte propriedade:

- 5 cliques de 2 vértices cada,

$\{X_1, X_2\}$, $\{X_1, X_3\}$, $\{X_2, X_3\}$, $\{X_2, X_4\}$ e $\{X_3, X_4\}$;

- 2 cliques máximos de 3 vértices cada,

$\{X_1, X_2, X_3\}$ e $\{X_2, X_3, X_4\}$.

Observe que $\{X_1, X_2, X_3, X_4\}$ não é um clique porque não existe uma aresta entre X_1 e X_4 .

Agora, podemos definir os fatores da decomposição da distribuição conjunta como funções das variáveis em um clique.

Na verdade, utilizaremos os cliques máximos sem perda de generalidade.

Denotaremos um clique por $\mathcal{C}$ e um conjunto de variáveis em um clique por $\mathbf{x}_{\mathcal{C}}$.

Assim, a distribuição conjunta pode ser escrita como um produto de funções potenciais $\psi_{\mathcal{C}}(\mathbf{x}_{\mathcal{C}})$ sobre os cliques máximos de um grafo:

$$f(\mathbf{x}) = \frac{1}{\kappa} \prod_{\mathcal{C}} \psi_{\mathcal{C}}(\mathbf{x}_{\mathcal{C}}),$$

em que κ é uma constante de normalização, conhecida como a função partição, dada por

$$\kappa = \sum_{\mathbf{x}} \prod_{\mathcal{C}} \psi_{\mathcal{C}}(\mathbf{x}_{\mathcal{C}}).$$

Devemos restringir as funções potenciais $\psi_{\mathcal{C}}(\mathbf{x}_{\mathcal{C}})$ a serem estritamente positivas para qualquer escolha de $\mathbf{x}_{\mathcal{C}}$ para que existam uma relação entre independência condicional e a fatorização em modelos gráficos não direcionados.

Exemplo

O objetivo é remover ruídos de uma imagem binária.

Podemos considerar uma matriz de pixels $\{X_{ij}\}$, $i = 1, 2, \dots, q$ e $j = 1, 2, \dots, p$, tal que $X_{ij} \in \{-1, +1\}$.

Poderíamos organizar tudo em um vetor $\mathbf{Y}$ tal que $Y_i \in \{-1, +1\}$, com $i = 1, 2, \dots, N$ (com $N = qp$).

A imagem com ruídos é obtida via uma imagem “limpa” denotada por $Z_{ij} \in \{-1, +1\}$. Os pixels são invertidos com uma probabilidade pequena.

Nesta aplicação, foi utilizada uma probabilidade de 0,1 (10%).

O nível de ruído é considerado baixo. Assim, há uma forte dependência (correlação) entre Z_{ij} e X_{ij} .

Observe que os pixels em uma vizinhança X_{ij} também são altamente dependentes.

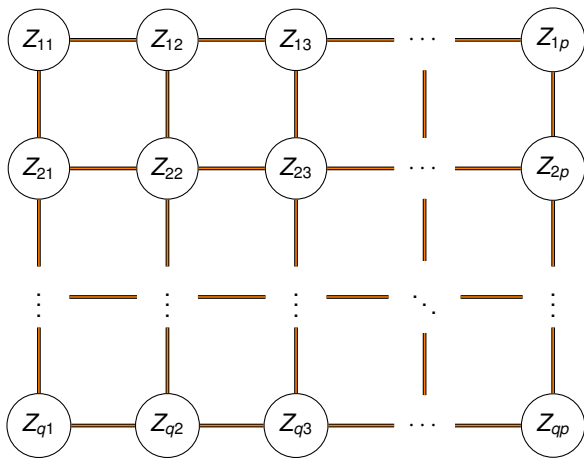


Figura: Representação parcial da estrutura espacial dos pixels de uma figura em um modelo gráfico não direcionado.

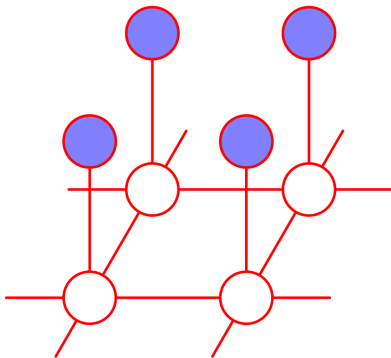


Figura: Representação da estrutura espacial dos pixels de uma figura em um modelo gráfico não direcionado. Os círculos não preenchidos representam os vértices Z_{ij} e os círculos preenchidos em azul representam os vértices X_{ij} .

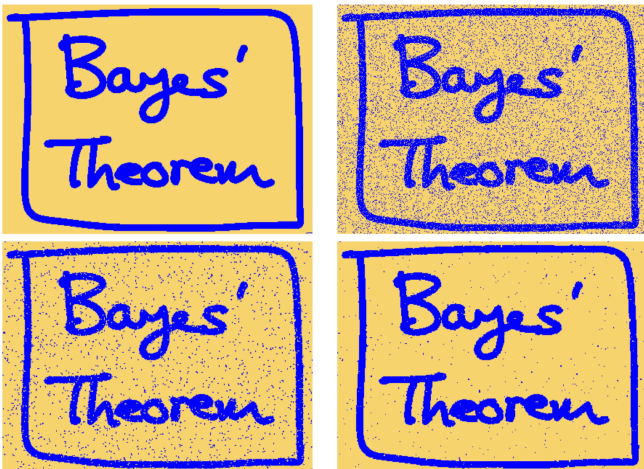


Figura: Removendo ruído de uma imagem via campos aleatórios de Markov. Na sequência temos a imagem original, imagem corrompida em 10% dos pixels, imagem recuperada via modelo condicional iterado (96%) e imagem recuperada via algoritmo *grafo-corte* (99%).