

Projeto e Análise de Algoritmos (PAA)

Algoritmo de Contração Probabilística de Karger

Probabilidade Condicional

- Espaço Amostral U com todos os resultados possíveis das escolhas aleatórias
 $\forall x \in U (P_n(x) \geq 0)$ e $\sum_{x \in U} P_n(x) = 1$

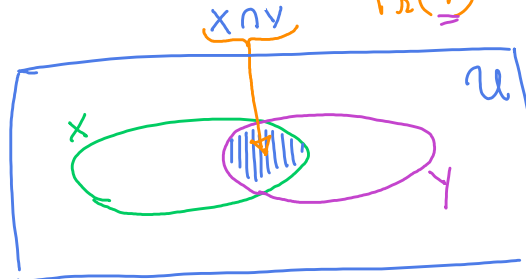
- Evento subconjunto $S \subseteq U$
 $P_n(S) = \sum_{x \in S} P_n(x)$

- Variável aleatória X é uma função
 $X: U \rightarrow \mathbb{R}$

- Esperança de X
 $E[X] = \sum_{x \in U} X(x) \cdot P_n(x)$

→ - Dados eventos X e Y Temos

$$P_n(\underline{X|Y}) = \frac{P_n(\underline{X \cap Y})}{P_n(\underline{Y})}$$



$\underbrace{x \text{ dado } Y}$

- $P_n(X|Y)$ pode ser interpretada como a restrição/renapeamento/normalização do universo p/ os resultados que ocorrem em Y

- Quiz: Suponha que rolem dois dados e a soma dos resultados foi 7. Qual a probabilidade de um dos dados ter sido 1?

$$R: P_n(\underline{X|Y}) = \frac{P_n(\underline{X \cap Y})}{P_n(\underline{Y})} = \frac{2 \cdot 1/36}{6 \cdot 1/36} = 1/3$$

$Y = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

⇒ (Im) Dependência entre ^{Eventos e} Variáveis

- Def.: eventos X e Y são independentes se, e somente se, $P_r(X \cap Y) = P_r(X) \cdot P_r(Y)$

- Equivalências:

$$\boxed{P_r(X|Y) = P_r(X \cap Y)} = \frac{P_r(X \cap Y)}{P_r(Y)} = \frac{P_r(X) \cdot P_r(Y)}{P_r(Y)} = P_r(X)$$

$$\boxed{P_r(Y|X) = P_r(Y)}$$

Prova: $E[A \cdot B] = \sum_{a,b} (a \cdot b) \cdot P_r(A=a \wedge B=b) \rightarrow A \text{ e } B \text{ são variáveis independentes}$

Exemplo: X_1 e X_2 e $\{0, 1\}$ e/por. uniforme e $\underline{X_3} = X_1 \oplus X_2$

$$U = \{(0, 0, \underline{0}), (0, 1, \underline{1}), (1, 0, \underline{1}), (1, 1, \underline{0})\}$$

Quiz 1: X_1 e X_3 são independentes?

R: Sim, $P_r(X_1=1) = 1/2$ e $P_r(X_3=1|X_1=1) = 1/2$
 $P_r(X_3=1|X_1=0) = 1/2$

- Def.: variáveis A e B (definidas sobre U) são independentes se $\forall (a, b) \in U \quad P_r(A=a \wedge B=b) = P_r(A=a) \cdot P_r(B=b)$

Produto das esperanças:

se A e B são variáveis independentes, então $E[A \cdot B] = E[A] \cdot E[B]$

$$= \sum_a a P_r(A=a) \left(\sum_b b \cdot P_r(B=b) \right) = E[A] E[B]$$

Quiz 2: X_1, X_3 e X_2 são independentes? R: Não

$$E[(X_1, X_3) \cdot X_2] = \sum (n_1, n_2, n_3) \cdot P_r(X_1=n_1 \wedge X_2=n_2 \wedge X_3=n_3) = 0$$

$$E[X_1, X_3] \cdot E[X_2] = E[X_1] E[X_3] E[X_2] = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8} \neq 0$$

prod. esp. var. ind

Problema do Corte Mínimo

lembrando que um corte é um subconjunto de V : $\emptyset \neq S \neq V$

Entrada: um grafo não-orientado $G = (V, E)$ # de arestas atravessando a fronteira de S

Solução: um corte $S \subset V$: $|S(S)|$ é mínima, i.e., o # de arestas atravessando a fronteira de S é $\leq |S(S')|$ p/ qualquer $S' \subset V$ e $\emptyset \neq S' \neq V$

Algoritmo de Contração Probabilística de Karger

enquanto # de vértices > 2 :

- escolha uma aresta $e = \{u, v\}$ c/ prob. uniforme nas arestas?
 - contraia os vértices u e v , mantendo arestas múltiplas, mas removendo auto-laços que surjam
- devolva o corte S correspondente a um dos vértices restantes

Obs: Sempre existe corte mínimo que induz subgrafo conexo. Qui: Porque?

R.: Por absurdo, suponha um corte mínimo desconexo c/ partes S e S' : $\exists (u, v) \in E$:
 $u \in S$ e $v \in S'$

Assim $|S(S \cup S')| = |S(S)| + |S(S')| \Rightarrow |S(S)| < |S(S \cup S')|$ e $|S(S')| < |S(S \cup S')|$ pois $|S(S)| > 0 \forall S \subset V$

Análise probabilística de uma execução/iteração do algoritmo de Karger

- Qual a probabilidade de sucesso?

ou $V \setminus S$

- Seja $S \subset V$ um corte mínimo. Vamos calcular a prob. do alg. devolver S
 - $\Pr(\text{Sucesso}) \geq \Pr(\text{devolver } S) = 1 - \Pr(\text{não devolver } S)$
 - O alg. não devolve S apenas se em alg. iter. contrain uma aresta de $\bar{S}(S)$
 - $X_j \in \{0, 1\}$ é a variável aleatória que indica se uma aresta de $\bar{S}(S)$ foi contraída na iter. j , $j = 1$ até $n-2$ (quando sobram apenas 2 supernúos)
 - $\Pr(\text{devolver } S) = \Pr(\sim X_1 \wedge \sim X_2 \wedge \sim X_3 \wedge \dots \wedge \sim X_{n-2})$ c/ $\Pr(\sim X_j) = 1 - \Pr(X_j)$
 - $\Pr(X_1) = |\bar{S}(S)|/m$, c/ $\left[2m = \sum_{v \in V} |\bar{S}(v)| \right]$ e $|\bar{S}(S)| \leq |\bar{S}(v)| \quad \forall v \in V \Rightarrow 2m \geq n \cdot |\bar{S}(S)|$

$\leq |\bar{S}(S)|/m \cdot |\bar{S}(S)|/2 = 2/m \Rightarrow \Pr(\sim X_1) \geq 1 - 2/m = \frac{n-2}{n}$

o grau de cada supernó é $\geq |\bar{S}(S)|$
- ⊗ Fazn p/ X_2 antes de generalizar p/ X_j
- $\Pr(X_j | \sim X_1 \wedge \dots \wedge \sim X_{j-1}) = |\bar{S}(S)| / \# \text{arestas} \leq 2|\bar{S}(S)| / ((n-j+1)|\bar{S}(S)|) = 2/(n-j+1)$
 - $\Pr(\sim X_j | \sim X_1 \wedge \dots \wedge \sim X_{j-1}) = 1 - \Pr(X_j | \sim X_1 \wedge \dots \wedge \sim X_{j-1}) \geq 1 - \frac{2}{n-j+1} = \frac{n-j-1}{n-j+1}$

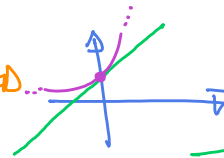
⊗ cada aresta acrescenta em 1 o grau de 2 nódos

Uso de repetições com escolhas aleatórias independentes
 - para obter alta probabilidade de sucesso

$$P_n(\text{devolver } S) = P_n(\sim X_1) \cdot P_n(\sim X_2 | \sim X_1) \cdot P_n(\sim X_3 | \sim X_1, \sim X_2) \dots P_n(\sim X_{m-2} | \sim X_1, \dots, \sim X_{m-3})$$

$$\geq \frac{n-2}{n} \cdot \frac{n-3}{n-1} \cdot \frac{n-4}{n-2} \cdot \frac{n-5}{n-3} \dots \frac{2}{4} \cdot \frac{1}{3} = \frac{2}{m(m-1)} \geq \frac{1}{n^2}$$

prob. pequena p/ n grande, mas muito maior que $1/2^{m-2}$

De cálculo: $(1+x) \leq e^x \Rightarrow$  $\Rightarrow P_n(\text{falha}) \leq 1 - P_n(\text{devolver } S)$

Fazendo N repetições independentes e guardando o melhor corte $\leq (1 - \frac{1}{n^2}) \leq e^{-1/n^2}$

$$P_n(\text{falha } N \text{ vezes}) = P_n(\text{falha})^N \leq e^{-N/n^2} = \frac{1}{e^{N/n^2}}$$

p/ $N = n^2$ Temos $\leq 1/e^{n^2/n^2} = 1/e$

p/ $N = n^2 \ln n \leq 1/e^{\frac{n^2 \ln n}{n^2}} = 1/e^{\ln n} = 1/n$

p/ $N = n^2 \cdot 2 \ln n \leq 1/e^{\frac{n^2 \cdot 2 \ln n}{n^2}} = 1/e^{2 \ln n} = 1/n^2$

c/ $> n^2 \ln n$ repetições
 a prob. de falha tende a 0
 muito rápido

Extra: qual o número máximo de cortes mínimos que um grafo pode ter?

Seja t o # de cortes mínimos de um grafo G c/ n vértices

chamemos de $S_i, i=1, \dots, t$, cada corte mínimo

Nossa análise mostrou que $P_n(\text{alg. deriv. } S_i) \geq \frac{2}{n(n-1)}$

$\forall i$ temos S_i são eventos disjuntos $\Rightarrow P_n(\bigcup_{i=1}^t \text{alg. der. } S_i) = \sum_{i=1}^t P_n(\text{alg. der. } S_i)$

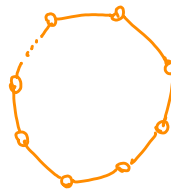
$\sum_{i=1}^t P_n(\text{alg. deriv. } S_i) \leq \text{Prob. de alg. deriv. qualquer corte} = 1$

(alg. der. S_i)

p # de cortes mínimos

$$1 \geq \sum_{i=1}^t P_n(\text{alg. deriv. } S_i) \geq \sum_{i=1}^t \frac{2}{n(n-1)} = t \cdot \frac{2}{n(n-1)} \Leftrightarrow t \leq \frac{n(n-1)}{2} = \binom{n}{2}$$

Esse caso máximo ocorre em ciclos,
pois todo par de arestas corresponde
à fronteira de um corte mínimo



em um ciclo c/ n vértices temos n arestas