Métodos Probabilísticos e Algébricos Aplicados em Combinatória

Domingos Dellamonica Junior
ddj@ime.usp.br

Orientador: Yoshiharu Kohayakawa
yoshi@ime.usp.br

Departamento de Ciência da Computação

1 Métodos Probabilísticos

Muitas vezes queremos provar a existência de um objeto matemático com propriedades desejáveis. Para não ficar tão vago, poderíamos pensar no seguinte exemplo: dado um grafo $G = (V, E)$, existe algum conjunto de vértices $S \subset V$ tal que o corte $\delta(S) \doteq \{\{x, y\} \in E \mid x \in S, y \in V \setminus S\}$ tenha pelo menos metade das arestas de E?

A resposta é sim, e podemos usar uma técnica para obter este resultado muito facilmente. Suponha que vértices de V são escolhidos de forma independente e com probabilidade $1/2$ para entrarem ou não no conjunto S. Qual o número esperado de arestas de $\delta(S)$? Seja X_e a variável aleatória indicadora do evento $e \in \delta(S)$. Seja $X = \sum_{e \in E} X_e$. Temos

$$\begin{aligned}\mathbf{E}[X] &= \mathbf{E}\Big[\sum_{e \in E} X_e\Big] = \sum_{e \in E} \mathbf{E}[X_e] \\ &= \sum_{\{u,v\} \in E} (\mathbf{P}[u \in S, v \in V \setminus S] + \mathbf{P}[u \in V \setminus S, v \in S]) \\ &= \sum_{e \in E} \Big(\frac{1}{4} + \frac{1}{4}\Big) = \frac{1}{2} \sum_{e \in E} 1 = \frac{1}{2}|E|.\end{aligned}$$

Na segunda igualdade estamos utilizando a linearidade da esperança (uma propriedade importantíssima), depois estamos utilizando a definição da esperança de uma variável indicadora e aplicando a hipótese de independência na escolha dos vértices.

Mas como isso prova o que foi pedido? Isso é muito simples. Se o valor esperado para o tamanho do corte de S para uma escolha aleatória de S é $|E|/2$ então, se para todo conjunto $S' \subset V$, tivermos $|\delta(S')| < |E|/2$, a esperança (que pode ser encarada como média) deve ser menor que $|E|/2$.

Paul Erdős, o pioneiro do Método Probabilístico.

Note que essa demonstração não é construtiva. Com efeito, não temos uma "receita de bolo" para obter um conjunto S cujo corte tenha pelo menos $|E|/2$ arestas. De fato, os métodos probabilísticos não são construtivos. No entanto, eles fornecem algoritmos probabilísticos para a construção desses objetos e, para alguns algoritmos, são conhecidas técnicas de desaleatorização que fornecem algoritmos determinísticos. Vamos mostrar uma técnica desse tipo a seguir.

Cara ou coroa?

A técnica que empregaremos é chamada de *Método das Esperanças Condicionais*. Vamos decidir se um vértice de V deve ir para S considerando um vértice por vez e nunca mudando a escolha. Para aqueles familiarizados com o jargão da computação, este é um algoritmo guloso. No i-ésimo passo do algoritmo, temos conjuntos S_i e T_i, onde todo vértice analisado até o i-ésimo passo que foi escolhido para entrar em S está em S_i e os que foram rejeitados estão em T_i. Como nunca mudamos nossas escolhas, $S_i \subset S$ e $T_i \subset V \setminus S$. Para o $(i+1)$-ésimo passo, precisamos decidir se $v \in V$ deve ir ou não para S. Calculamos as esperanças condicionais:

$$\mathbf{E}[X \mid S_{i+1} = S_i \cup \{v\}, T_{i+1} = T_i] \text{ e}$$
$$\mathbf{E}[X \mid S_{i+1} = S_i, T_{i+1} = T_i \cup \{v\}].$$

A decisão é baseada em qual o maior valor obtido. Se for o primeiro, então v é escolhido para S, caso contrário, v fica de fora. Fica a cargo do leitor interessado (!) demonstrar que o máximo das esperanças condicionais nunca é menor do que o valor das esperanças condicionais do passo anterior. Em particular, a cada passo, o máximo das esperanças é pelo menos $\mathbf{E}[X] = |E|/2$. Ao definirmos para onde deve ir o último elemento de V não há mais nada aleatório e temos um conjunto S com $|\delta(S)| \geq |E|/2$.

2 Álgebra e Combinatória

Um dos assuntos estudados nesta iniciação científica foi o emprego de técnicas algébricas em problemas combinatórios. Em particular, a álgebra linear é muito útil na demonstração de resultados que parecem inatingíveis a partir de argumentos puramente combinatórios.

2.1 Comunidades do Orkut

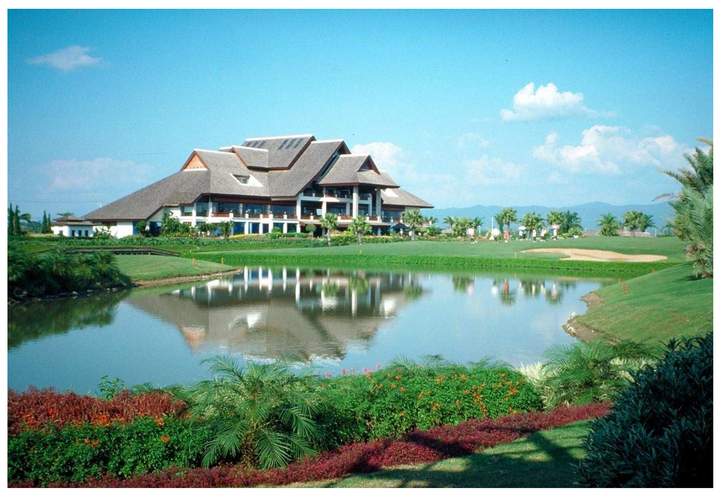
Os patrocinadores do Orkut estão preocupados com a imensa quantidade de comunidades criadas todos os dias por seus usuários. O intenso movimento dessas comunidades faz com que os servidores do site estejam sempre congestionados e, enquanto o departamento administrativo não libera verbas para a compra de mais hardware, você voi contratado para impor regras para as comunidades de forma a evitar a explosão do sistema!

Um projetista do sistema teve a seguinte idéia.
— Não vamos permitir que duas comunidades distintas possuam exatamente os mesmos membros. Além disso, todas as comunidades deverão ter um número par de membros e os membros em comum entre duas comunidades deverão ser uma quantidade par!

A idéia parecia boa mas... e se todo casal de namorados no Orkut resolvesse participar das mesmas comunidades? Com certeza as comunidades teriam tamanho par e os membros em comum seriam sempre casais e, portanto, pares! Se o número de casais no Orkut é K, podemos ter até $2^K - 1$ comunidades diferentes seguindo essa regra. Este número certamente vai derrubar todos os servidores disponíveis.

Você sugere uma leve modificação na regra.
— Vamos forçar as comunidades a possuirem número ímpar de membros e os membros em comum de duas comunidades serão sempre um número par.

Quantas comunidades podem ser feitas dessa forma? Parece difícil até mesmo rabiscar qualquer idéia de solução. Tente... ... desistiu? Tudo bem, mas depois não vá dizer que você pensaria nisso, hein? Vamos usar fatos bem básicos de álgebra linear e mostrar que se há n pessoas no Orkut, o número de comunidades não pode passar de n. Isso é bem razoável e a turma da engenharia já avisou que os servidores não vão ficar caindo com esse número de comunidades.



Todo mundo quer participar dessa comunidade!

Bem, vamos a prova então. Podemos construir um vetor característico para todas as comunidades formadas. Seja P o conjunto de todos os usuários do Orkut e $\mathbb{F}_2$ o corpo finito com 2 elementos (calma, não vamos usar nada de Álgebra que você odeie, basta lembrar que $1 + 1 = 0$ nesse corpo). O vetor característico de uma comunidade, $\mathbf{v} : P \to \mathbb{F}_2$ é tal que $v_i = 1$ se $i \in P$ e $v_i = 0$ se $i \notin P$.

Agora só precisamos lembrar de algo básico sobre independência linear. Dizemos que um conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ de vetores num espaço vetorial sobre um corpo $\mathbb{F}$ é linearmente dependente se e somente se existem $\lambda_1, \ldots, \lambda_m \in \mathbb{F}$, não todos nulos, tais que

$$\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}.$$

Seja $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ o conjunto dos vetores característicos de todas as comunidades do Orkut. Sejam $\lambda_1, \ldots, \lambda_m \in \mathbb{F}_2$ tais que $\lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m = \mathbf{0}$. Denotamos por $\langle \mathbf{u}, \mathbf{v} \rangle$ o produto interno dos vetores $\mathbf{u}$ e $\mathbf{v}$. Note que, como somar um número par de 1's em $\mathbb{F}_2$ resulta em 0, pelas regras impostas sobre as comunidades, temos

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1, & \text{se } i = j, \\ 0, & \text{se } i \neq j. \end{cases}$$

Lembre-se que o produto interno de um vetor por $\mathbf{0}$ resulta em 0, que o produto é distributivo, i.e, $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$ e que, para um escalar μ, temos $\langle \mathbf{u}, \mu \mathbf{v} \rangle = \mu \langle \mathbf{u}, \mathbf{v} \rangle$. Tendo tudo isso em mente, veja que, para todo i, temos

$$\begin{aligned} 0 = \langle \mathbf{v}_i, \mathbf{0} \rangle &= \langle \mathbf{v}_i, \lambda_1 \mathbf{v}_1 + \cdots + \lambda_m \mathbf{v}_m \rangle \\ &= \lambda_1 \langle \mathbf{v}_i, \mathbf{v}_1 \rangle + \cdots + \lambda_m \langle \mathbf{v}_i, \mathbf{v}_m \rangle = \lambda_i. \end{aligned}$$

Mas então, $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0$. Isso prova que os vetores de V são linearmente independentes. Como a dimensão do espaço desses vetores é $|P|$, temos $m = |V| \leq |P| = n$. Fica demonstrado que o número de comunidades não ultrapassa o número de usuários do Orkut.

A propósito, utilizando outras técnicas de álgebra linear é possível mostrar que se as primeiras regras propostas pelo projetista fossem aplicadas então o número máximo de comunidades seria realmente $2^K - 1$.

Referências

[1] N. Alon and J. Spencer, *The Probabilistic Method*. New York: John Wiley and Sons, 2nd ed., 2000.

[2] L. Babai and P. Frankl, *Linear Algebra Methods in Combinatorics*. preliminary version 2 ed., 1992.