Algoritmos Probabilísticos

Gilson Evandro Fortunato Dias

Orientador: José Coelho de Pina

1 de Dezembro de 2008

Resumo

A proposta para este trabalho é estudar o livro *Probability and Computing, M. Mitzenmacher e E. Upfal*, e produzir um texto sobre algoritmos probabilísticos. Esta monografia faz parte do meu trabalho de formatura, da disciplina MAC0499 - Trabalho de Formatura Supervisionado. Fui orientado pelo professor José Coelho de Pina, tendo estudado também com o Israel Lacerra. No trabalho, apresento algumas técnicas para desenvolver algoritmos probabilísticos e métodos para poder analisá-los.

Conteúdo

1 Introdução

Algoritmos probabilísticos são caraterizados por fazerem escolhas aleatórias durante a sua execução. É como se durante a sua execução fosse jogada uma moeda e dependendo do resultado, cara ou coroa, o algoritmo decide entre uma ou outra execução. Algoritmos determinísticos são o oposto do que aqui iremos estudar.

Existem muitas áreas onde estes algoritmos são bastante usados, e tentaremos sempre que possível, apresentar aplicações para os algoritmos que aqui estudarmos. Mas podemos destacar a Criptografia (primalidade, dada a importância de números primos grandes), Protocolos de Rede (a Ethernet por exemplo), Estruturas de Dados, Otimização Combinatória e outras.

As grandes motivações para estudarmos este tipo de algoritmos estão relacionadas ao fato deles geralmente serem mais simples e/ou mais eficazes que algoritmos determinísticos. Mas estas caraterísticas acabam por ter as suas desvantagens que estão diretamente relacionadas com a corretude e o seu tempo de execução. Quanto a corretude, estes algoritmos podem apresentar uma certa probabilidade de dar a resposta errada, como veremos em exemplos mais para frente. E o tempo de execução dos mesmos agora já não depende única e exclusivamente da entrada do problema, mas principalmente das escolhas aleatórias que forem feitas durante a sua execução, o que pode alterar significativamente o desempenho do algoritmo.

Vamos estudar alguns métodos para analisar estes algoritmos e aprender a aleatorizar algoritmos determinísticos, alguns bem conhecidos já. Assim, na segunda seção vamos rever alguns conceitos de probabilidade, variáveis aleatórias e seus momentos, esperança, variância e algumas técnicas para limitar variáveis aleatórias.

Na terceira seção, veremos a distribuição de Poisson para apresentarmos o modelo de bolas e sacos, grafos aleatórios e cadeias de Markov. Como bônus desta seção é apresentado o conceito de Monte Carlo e Las Vegas, bem como suas aplicações.

E na última seção da monografia, farei uma pequena descrição sobre as atividades realizadas, métodos de estudo, reuniões com o Israel e o professor Coelho, conclusões sobre o trabalho, vou falar também sobre as experiências adquiridas na realização do trabalho, dificuldades, desafios e outros.

2 Conceitos de Probabilidade

Nesta seção vamos aproveitar a rever alguns conceitos de probabilidade de forma superficial, pois vamos precisar saber calcular algumas probabilidades mais para frente. A medida que novos conceitos forem introduzidos, vamos aproveitar para definir a notação usada ao longo do texto e utilizar alguns exemplos também.

2.1 Axiomas

O *espaço amostral* tem uma importância fundamental quando estamos a falar de probabilidades, ele é composto pelas seguintes caraterísticas:

1. um *espaço amostral* Ω é o conjunto de todos os resultados de um experimento aleatório.

2. Quaisquer subconjuntos de Ω, são chamados de *eventos*. Os subjconjuntos unitários de Ω definem os eventos um por um do experimento.

3. **Pr**: $F \rightarrow \Re$, é a função de probabilidade, onde F é o conjunto de eventos de Ω.

Por sua vez, uma *função de probabilidade* **Pr** precisa respeitar as seguintes condições:

1. Para todo evento $E \subseteq \Omega$, $0 \leq \Pr[E] \leq 1$.

2. $\Pr[\Omega] = 1$.

3. Para eventos $E_1, E_2, \ldots$ disjuntos dois-a-dois, então $\Pr[\cup_i E_i] = \sum_i \Pr[E_i]$.

Por uma questão de simplicidade, e sem perda de generalidade, vamos trabalhar apenas com espaços de probabilidade *discreta*, onde o espaço amostral é finito ou enumerável, e portanto, os eventos correspondem a todos os subconjuntos de Ω.

Vamos considerar, como exemplo o lançamento de um dado não viciado de seis faces, com o espaço probabilidade definido por $\Omega = 1, 2, 3, 4, 5, 6$ conjunto de resultados possíveis e a função de probabilidade Pr que associa cada evento i como: E_i = resultado obtido foi i, onde $i = 1, \ldots, 6$ e temos então,

$$\Pr[E_i] = \frac{1}{6}.$$

Independência e Probabilidades Condicionais

Vamos começar com duas definições que serão muito úteis mais a frente.

Definição 1.1: Dois eventos A e B são independentes se e somente se

$$\Pr(A \cap B) = \Pr(A).\Pr(B).$$

Definição 1.2: A probabilidade condicional de o evento A ocorrer, dado que o evento B aconteceu é

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(F)}.$$

Outra forma de verificar se os eventos A e B são independentes é se $\Pr(A|B) = \Pr(A)$, pois neste caso teríamos que:

$$\Pr(A|B) = \frac{\Pr(A \cap B}{\Pr(B)} = \frac{\Pr(A).\Pr(B)}{\Pr(B)} = \Pr(A).$$

2.2 Variáveis Aleatórias e Esperança

Definição 1.3: Uma variável aleatória é uma função. Mais concretamente, uma variável aleatória X em Ω é uma função tal que $X : \Omega \to \Re$. Uma variável aleatória *discreta* é uma variável aleatória que pega somente um número finito ou enumerável de valores.
Vamos denotar a $Pr(X = a)$, onde X é uma variável aleatória e a um valor real, como a probabilidade do evento acontecer, onde o evento é definido pelo conjunto $s \in \Omega | X(s) = a$.

A *esperança* ou o *valor esperado* de uma variável aleatória X, denotada por $E[X]$ é definida por

$$E[X] = \textstyle\sum_i i\Pr(X = i).$$

Por exemplo, a esperança de uma variável aleatória X que representa um dado de seis faces é

$$E[X] = \frac{1}{6}.1 + \frac{1}{6}.2 + \ldots + \frac{1}{6}.6 = \frac{21}{6}.$$

2.2.1 Linearidade da Esperança

Uma propriedade muito importante da esperança que simplifica muito a nossa tarefa na hora de calcular a esperança da soma de variáveis aleatórias ou a soma das esperanças de variáveis aleatórias é definida pelo teorema abaixo.

Teorema 2.1 [Linearidade da Esperança]: Para qualquer conjunto de variáveis aleatórias discretas $X_1, X2, \ldots, X_n$, com esperanças finitas, temos que

$$E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i].$$

Um resultado diretamente relacionado a linearidade da esperança é dado pelo seguinte lema.

Lema 2.2: Para qualquer constante c e uma variável aleatória X temos que,

$$E[cX] = cE[X].$$

2.3 Algumas Variáveis Aleatórias

Vamos apresentar raidamente as variáveis aleatórias que mais usaremos ao longo do texto, e no final desta seção apresentamos um exemplo para alguma delas.

Bernoulli

É uma variável aleatória que pode assumir dois valores 0 (normalmente chamada de fracasso, quando não ocorre determinado evento) ou 1 (sucesso, quando ocorre determinado evento). Seja X uma variável aleatória tal que, $X = 1$ com probabilidade p e $X = 0$ com probabilidade $1 - p$. Então X é chamada de *Bernoulli*. Calculando a esperança de uma variável aleatória de Bernoulli, temos que,

$$E[X] = p.1 + (1 - p).0 = p = \Pr(X = 1).$$

Binomial

Se realizarmos um experimento de Bernoulli n vezes, obtemos uma variável aleatória *Binomial*, $B(n, p)$, onde n é o número de experimentos e p a probabilidade de sucesso para cada experimento e tem distribuição para $j = 1, \ldots, n$:

$$\Pr(X = j) = p^j (1 - p)^{n-j}.$$

Sabendo que $X = \sum_{i=1}^{n} X_i$, e lembrando que cada X_i é uma Bernoulli com probabilidade de sucesso igual a p. Então, claramente $E[X_i] = p$ e pela linearidade da esperança temos que,

$$E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = np.$$

Geométrica

Aqui, estamos preocupados com o número de tentivas necessárias até obtermos o primeiro sucesso de um experimento aleatório. Podemos considerar o caso em que lançamos uma moeda e queremos saber qual é o número de lançamentos que precisamos fazer até sair cara. Este número é dado pela variável aleatória *geométrica*.
Então, a $\Pr(X = n) = p(1-p)^{n-1}$, para $n = 1, 2, 3, \ldots$, ou seja, é a probabilidade de precisarmos realizar n tentivas até obtermos sucesso. No exemplo da moeda, seria o mesmo que lançarmos n vezes a moeda até sair *cara*.

Mais para frente vamos estudar outras distribuições como a de *Poisson* e a *Estacionária*.

2.4 Variância e Momentos de uma Variável Aleatória

Definição 2.1: O *k-ésimo momento* de uma variável aleatória X é $E[X^k]$. E a esperança de x, $E[X]$, é chamada de primeiro momento de X.

Para introduzirmos um técnica que vai ajudar-nos a criar limitações para variáveis aleatórias, chamada *desigualdade de Chebyshev*, precisamos definir o que é a variância de uma variável aleatória. E para isso, vamos usar o primeiro e o segundo momento de X.

Definição 2.2: A *variância* de uma variável aleatória X é defina como

$$Var[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

O *desvio padrão* de uma variável aleatória X é dado por

$$\sigma[X]^2 = Var[X].$$

2.5 Desigualdades de Markov e Chebyshev

Nesta subseção vamos apresentar duas ferramentas que iremos usar para limitar variáveis aleatórias e começamos pela desigualdade de *Markov*, definida pelo teorema a seguir.

Teorema 2.1 [Desigualdade de Markov]: Seja X uma variável aleatória que assume apenas valores positivos. Então, para todo $a > 0$,

$$\Pr(X \geq a) \leq \frac{E[X]}{a}.$$

Teorema 2.1 [Desigualdade de Chebyshev]: Para qualquer $a > 0$,

$$\Pr(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}.$$

2.6 Observações

Nesta seção não foi provado nenhum dos teoremas enunciados, assim como algumas propriedades ficaram por demonstrar. A maior parte delas estudamos e provamos durante as reuniões que mantive com o professor Coelho, estando nos meus apontamentos. Mas devido o atraso em que se encontra a minha monografia, não consegui acrescentá-las aqui. Pretendo revisar e acrescentar todas as demonstrações, mais exemplos sobre as definições apresentadas e alguns detalhes que não foram mencionados aqui.

3 Grafos e Passeios Aleatórios

Neste seção, vou apresentar os assuntos do capítulo 5 e 7 do livro texto, assuntos esses que ainda estou a estudar com o meu orientador. Mas deixo aqui a estrutura esperada para esta seção, onde vamos introduzir alguns conceitos como distribuição de Poisson e Cadeias de Markov, dando exemplos de algortimos aleatórios onde são aplicados esses conceitos, nomeadamente em Grafos Aleatórios e Passeios Aleatórios.

Por enquanto fica aqui a estrutura esperada para estar seção da monografia, esta é a principal seção do meu trabalho, que irei concluir após terminar de estudar os conceitos que serão aqui abordados.

3.1 Modelo Bolas-e-Sacos

Por concluir...

3.2 Distribuição de Poisson

Por concluir...

3.3 Grafos Aleatórios

Por concluir...

3.4 Cadeias de Markov

Por concluir...

3.5 Passeios Aleatórios

Por concluir...

3.6 Monte Carlo e Las Vegas

Tópico de bônus, também por concluir...

4 Conclusões

Nesta seção final da monografia quero falar um pouco sobre o que significou este trabalho para mim, quais foram as coisas que mais aprendi, que dificuldades tive, o que pretendo realizar após esta etapa e outras. Mas como não terminei de escrever a minha monografia, deixo as considerações finais para quando terminar o meu trabalho de formatura.

4.1 Atividades Realizadas

Após a escolha do orientador, mantive e venho mantendo reuniões semanais com o professor Coelho e o Israel Lacerra. Nestas reuniões estudamos alguns assuntos do livro, fazemos demonstrações dos teoremas, definições, lemas e outros que encontramos no livro. Geralmente um de nós, eu e o Israel preparamos as reuniões, apresentamos os assuntos estudados para o nosso orientador e vamos esclarecendo assunto por assunto, dúvidas e provas dos resultados a que chegamos.

Pretendo continuar a reunir-me com eles agora nos meses de Dezembro e Janeiro, para concluir o estudo do livro a que me propus estudar e consequentemente ir escrevendo o restante da monografia a medida que formos avançando nas nossas reuniões.

4.2 Experiência Obtida

Foram muitas, mas pretendo enunciá-las quando terminar a monografia. Pois ainda venho aprendendo muita coisa, tenho tido dificuldades em alguns aspetos, conceitos novos surgem a cada reunião, portanto, seria desajustado falar apenas das experiências que tive até agora apenas. Mas posso mencionar a mais importante de todas elas até então, o trabalho em grupo. A filosofia de trabalhar em grupo.

4.3 Desafios

Por concluir..

4.4 Conclusões

Por concluir...

Referências

[1] M. Mitzenmacher e E. Upfal, Probability and Computing, *Cambridge*, 2005.

[2] L. Breiman, Probability and Stochastic Processes, *Mifflin*, 1969.

[3] W. Bussab e P. Morettin, Estatística Básica, 2004.

[4] www.ime.usp.br/$\sim$ pf.