Algoritmos de Monte Carlo e Cadeias de Markov – CPS 767 2021/1

Prof. Daniel R. Figueiredo

Quarta Lista de Exercícios

Dica: Para ajudar no processo de aprendizado responda às perguntas integralmente, mostrando o desenvolvimento das respostas.

Questão 1: Sequências binárias restritas. Considere uma sequência de dígitos binários (0s e 1s) de comprimento s. Uma sequencia é dita válida se ela não possui 1s adjacentes. Considerando a distribuição uniforme, queremos determinar o valor esperado do número de 1s de uma sequência válida, denotado por μ_s.

1. Considerando $s = 4$, determine todas as sequências válidas e calcule μ_4.
2. Construa uma cadeia de Markov sobre o conjunto de sequências válidas, deixando claro como funcionam as transições de estado. Argumente que a cadeia é irredutível.
3. Desenhe a cadeia de Markov para o caso de $s = 4$, mostrando todas as transições.
4. Mostre como aplicar Metropolis-Hastings para resolver o problema de estimar μ_s. Deixe claro as probabilidades de aceite, e o funcionamento do estimador.

Questão 2: Amostras de Modelos de Mistura (*Mixture Models*). Considere a seguinte função de probabilidade:

$$p(x) = \alpha p_B(x; n, p_1) + (1 - \alpha)p_B(x; n, p_2) \ ,$$

onde $p_B(x; n, p)$ é a probabilidade associada ao valor x da binomial com parâmetros n e p, e $\alpha \in [0, 1]$ é um peso. Repare que $p(x)$ é um modelo de mistura de duas binomiais com diferentes valores de p com pesos dado por α e $1 - \alpha$. Considere duas variáveis aleatórias X e K, representando o valor de $X \in [0, n]$ e a binomial utilizada $K \in [1, 2]$. Queremos gerar amostras de acordo com $p(x)$.

1. Determine a distribuições de probabilidade condicionais $P(X|K)$ e $P(K|X)$. Dica: utilize regra de Bayes no segundo caso.
2. Determine a distribuição de probabilidade conjunta $P(X, K)$.
3. Utilize a técnica de Gibbs Sampling para gerar amostras de X. Mostre com construir a cadeia de Markov e determine a transição entre os estados.
4. Para $n = 2$, $p_1 = 0.1$, $p_2 = 0.8$, $\alpha = 0.3$, desenhe a cadeia de Markov com todas as transições.
5. Descreva como usar a cadeia de Markov para gerar amostras.

Questão 3: Amostrando triângulos. Considere um grafo conexo qualquer. Desejamos gerar amostras de triângulos deste grafo (cliques de tamanho 3), tal que todo triângulo tem igual probabilidade de ser amostrado. Ou seja, uma distribuição uniforme sobre o conjunto de triângulos do grafo.

1. Mostre como gerar amostras de forma direta, utilizando a distribuição uniforme (dica: pense em amostragem por rejeição). Determine a eficiência desse método.

2. Mostre como gerar amostras utilizando Metropolis-Hastings. Determine os estados da CM, as transições da cadeia base (que deve ser irredutível), e a probabilidade de aceite na cadeia modificada pelo método Metropolis-Hastings.

3. Intuitivamente, discuta quando a abordagem via Metropolis-Hastings é mais eficiente (do ponto de vista computacional) do que a abordagem via amostragem por rejeição.

Questão 4: Quebrando o código. Você encontrou uma mensagem que foi cifrada com a código da substituição (neste código, cada letra é mapeada em outra letra, de forma bijetora). Você deseja encontrar a chave do código para ler a mensagem. Repare que a chave é um mapeamento σ entre as letras, por exemplo $\sigma(a) = x$, $\sigma(b) = h$, $\sigma(c) = e$, $\ldots$. Considere uma função $f : \Omega \to [0, 1]$ que avalia a capacidade de uma pessoa entender a mensagem cifrada dado um mapeamento $\sigma \in \Omega$. Repare que $f(\sigma) = 1$ significa que é possível entender por completo a mensagem decifrada com o mapeamento σ, e $f(\sigma) = 0$ se o mapeamento σ não revela nenhuma informação sobre a mensagem. Utilize o método de *Simulated Annealing* para resolver este problema! Mostre todos os passos necessários para que o método possa ser aplicado (não é necessário fazer a implementação).