

Métodos Computacionais B

Agenor Hentz¹ Leonardo Brunnet¹ Heitor Fernandes¹

¹Instituto de Física, Universidade Federal do Rio Grande do Sul, Brasil

Semestre 2016-1

1 Resolução de Integrais por Monte Carlo

- Método 1: Amostragem Uniforme
- Método 2: Amostragem Seletiva
- Método 2: Amostragem Seletiva

2 Caminhante Aleatório

3 Caminhante Aleatório

- Modelo Discreto Unidimensional
- Modelo Discreto Unidimensional
- Modelo Generalizado
- Caminhante Aleatório e o Teorema do Limite Central
- Distribuições Sem Momento Definido

Resolução de Integrais por Monte Carlo

O método de Monte Carlo para a resolução numérica de integrais consiste em se realizar a seguinte substituição:

$$\int_{x_0}^{x_f} f(x') dx' \equiv I = (x_f - x_0) E[f(X)], \quad (1)$$

onde $E[f(x)]$ é o valor esperado da função $f(x)$ no intervalo de integração $[x_0, x_f]$. A determinação do valor esperado da função pode ser realizada de acordo com diferentes estratégias, descritas a seguir.

Método 1: Amostragem Uniforme

Este método é bastante intuitivo. Considere uma função $f(x)$ conhecida na qual se deseja calcular sua integral dentro do intervalo $[x_0; x_f]$. Considere agora que definimos dois valores de y : y_{min} e y_{max} menor e maior, respectivamente, do que os valor mínimo e máximo de $f(x)$ dentro do intervalo $[x_0; x_f]$. Desta maneira os valores de x_0, x_f, y_{min} e y_{max} formam uma região retangular na qual $f(x)$ está contida. Considere agora que são gerados aleatoriamente N pontos dentro desta região retangular, cada ponto tendo coordenadas $\{x_i; y_i\}$. Para cada ponto gerado, verificamos se o mesmo se encontra na região entre $f(x)$ e o eixo das coordenadas de tal forma que dos N pontos gerados, um total de S se encontram nesta região. Se os pontos gerados foram sorteados de maneira que se possa considerar aleatória, é de se esperar que a integral de $f(x)$ na região analisada seja proporcional à S , na mesma medida em que a área da região retangular seja proporcional à N . Matematicamente:

$$\int_{x_0}^{x_f} f(x') dx' = \lim_{N \rightarrow \infty} \frac{S}{N} [(x_f - x_0)(y_{max} - y_{min})],$$

onde a quantidade entre colchetes é a área da região retangular.

Método 2: Amostragem Seletiva

Considerando ainda a integral I (eq. 1), o valor esperado de $f(x)$ pode ser **inferido** a partir do valor médio $\tilde{\mu}_N$ calculado como:

$$E[f(x)] \sim \tilde{\mu}_N = \frac{1}{N} \sum_{i=1}^N f(X_i), \quad (2)$$

onde X_i são números aleatórios sorteados dentro do intervalo $[x_0; x_f]$. Podemos facilmente imaginar que dependendo da forma de f e do intervalo a ser integrado, uma amostragem uniforme sobre todo o domínio poderá ser ineficiente para calcularmos o valor médio $\tilde{\mu}_N$. Este fato fica evidente se considerarmos que o valor esperado da função dentro do intervalo $E[f(x)]$ dificilmente será igual ao valor médio $\tilde{\mu}_N$.

Método 2: Amostragem Seletiva

Por este motivo, define-se o cálculo da integral por Monte Carlo como sendo:

$$I = (x_f - x_0) \tilde{\mu}_N \pm \sigma_N,$$

onde σ_N é definido como:

$$\sigma_N = \left[\frac{1}{N-1} \sum_{i=1}^N (f(X_i) - \tilde{\mu}_N)^2 \right]^{\frac{1}{2}}.$$

Uma forma de diminuirmos o erro no cálculo de I é sortermos valor de f que estejam próximos ao valor médio da função no intervalo de integração. Assim, fica claro que sortear números segundo uma distribuição uniforme, por exemplo, pode ser não ser a melhor estratégia para o cálculo de I para qualquer função f .

Método 2: Amostragem Seletiva

Podemos transformar a integral I de uma função $f(x)$ da seguinte maneira:

$$I = \int_{x_0}^{x_f} f(x)dx = \int_{x_0}^{x_f} \left[\frac{f(x)}{p(x)} \right] p(x)dx.$$

O ideal é escolhermos a função $p(x)$ o mais próximo possível da função $f(x)$, de forma que a razão $f(x)/p(x)$ seja o mais próxima possível da unidade. A idéia central do método consiste em se escolher os valores aleatórios de X_i a serem utilizados na equação (2) à partir da equação peso $p(x)$, e não à partir da equação $f(x)$: se $p(x)$ for escolhida simples o suficiente, é fácil utilizarmos algum método simples de amostragem, por exemplo o método da transformada inversa, a fim de sortearmos X_i , tarefa esta que pode ser difícil, se não impossível, de ser feita se $f(x)$ for analiticamente complicada.

Método 2: Amostragem Seletiva

Assim temos:

$$I = (x_f - x_0) \left\{ \frac{1}{N} \sum_{i=1}^N \left[\frac{f(X_i)}{p(X_i)} \right] \right\} \pm \sigma_N,$$

onde os valores de X_i são escolhidos de acordo com a distribuição $q(x)$ e o erro é dado por:

$$\sigma_N = \left\{ \frac{1}{N-1} \sum_{i=1}^N \left[\frac{f(X_i)}{p(X_i)} - \left\langle \frac{f(X_i)}{p(X_i)} \right\rangle \right]^2 \right\}^{\frac{1}{2}}$$

Método 2: Amostragem Seletiva

Em resumo, uma boa candidata para função $p(x)$ deve ter as seguintes características:

- deve ter maior densidade ao redor da região onde o integrando de I é maior;
- $p(x) > 0$ para qualquer ponto do domínio a ser integrado;
- deve ser suficientemente simples para que uma distribuição aleatória possa ser facilmente calculada a partir dela.

Até este momento consideramos somente integrais unidimensionais, mas este método pode ser facilmente estendido para casos multi-dimensionais.

O caminhante aleatório é um sistema onde um objeto pontual realiza passos aleatórios sucessivos que podem ter comprimento Δx fixo ou variável, em passos de tempo Δt que são, em geral, discretos. Este tipo de sistema visa simular tanto fenômenos físicos quanto outros observados em sistemas que possuem características estocásticas. Entre os fenômenos físicos que podem ser simulados pelo caminhante aleatório, podemos citar o movimento de determinada molécula em um meio gasoso ou líquido. Neste caso, se soubéssemos a posição e velocidade de todas as moléculas que estão no meio em questão, além da forma precisa do potencial de interação das mesmas entre si, poderíamos caso tivéssemos o tempo e a capacidade computacional para tanto, calcular a trajetória de uma determinada molécula neste meio.

Caminhante Aleatório

Neste caso, o movimento desta molécula é caracterizado por momentos em que esta viaja com velocidade e direções constantes, quando está relativamente longe das demais moléculas a ponto de não sofrer influência do potencial das mesmas; estes momentos são intercalados com colisões com as demais moléculas onde tanto a velocidade quanto a direção provavelmente mudam devido à esta interação. Do ponto de vista de um observador externo, o comportamento desta molécula é de um caminhante que anda em linha reta por determinada distância, mudando abruptamente de direção repetidas vezes, de forma periódica mas aleatória. Outras áreas onde o estudo do caminhante aleatório pode ser aplicado são economia (flutuação dos preços de ações e de *commodities*), ecologia (movimento de cardumes de peixes e de aves quando em grupo) e física (movimento browniano), entre outros.

Modelo Discreto Unidimensional

Na presente publicação, estudaremos o caminhante aleatório mais simples: o caminhante aleatório unidimensional, com passo fixo $\Delta t = \pm l$, que tem igual probabilidade, de 50% ou 0.5, de dar um passo para a direção positiva do eixo x (o passo será, neste caso $+l$, e a nova posição $x_{n+1} = x_n + l$, onde x_n é a posição do caminhante no tempo t_n) quanto de dar um passo na direção negativa do eixo x (o passo será, neste caso $-l$, e a nova posição será $x_{n+1} = x_n - l$). Em geral, o caminhante aleatório inicia seu movimento na origem do eixo x , ou seja, $x_0 = 0$. Assim, no primeiro passo ($t_1 = \Delta t$), segundo as regras acima, o caminhante pode pular tanto para a posição $x_1 = +l$ (caso o passo aleatório tenha sido dado no sentido positivo de x) quanto na posição $x_1 = -l$ (caso o passo aleatório tenha sido dado no sentido oposto).

Modelo Discreto Unidimensional

No passo seguinte as possíveis configurações do caminhante aleatório aumentam: caso ele tenha caminhado no sentido positivo de x no primeiro passo, ele pode dar mais um passo no sentido positivo ficando em $x_2 = 2l$ ou dar um passo no sentido oposto, ficando em $x_2 = 0$; outras possibilidades são relacionadas com a situação onde o primeiro passo do caminhante tenha sido no sentido negativo de x , já que neste caso ele pode ficar tanto em $x_2 = -2l$ caso o segundo passo também seja no mesmo sentido, ou ficar em $x_2 = 0$ caso o segundo passo seja no sentido oposto. Fica claro que depois de dois passos os únicos lugares poderão estar ocupados são as posições $x_2 = -2l, 0, 0, +2l$. A posição $x_2 = 0$ está repetida na lista anterior porque dois caminhos diferentes resultam nesta mesma posição final: o caminhante dar um passo adiante e outro para trás ou o caminhante dar um passo inicial para trás e depois um passo para frente. Isto resulta no fato de que a probabilidade do caminhante retornar à posição original depois de dois passos é duas vezes maior do que ele dar dois passos adiante, por exemplo. Todos os passo seguintes seguem a mesma lógica.

Modelo Discreto Unidimensional

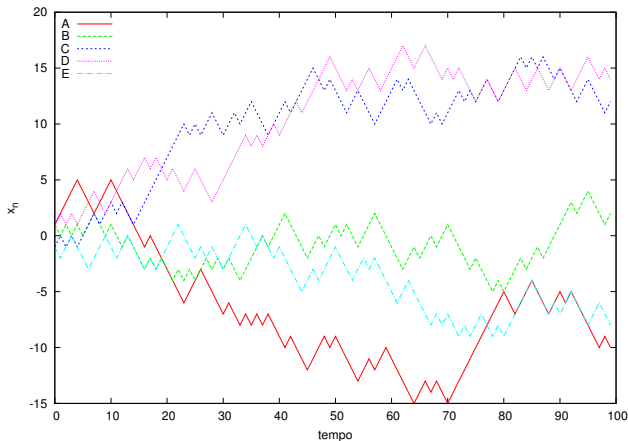


Figura: Exemplo de caminhos seguidos por 5 diferentes caminhantes aleatórios, inicialmente na origem e com igual probabilidade de darem um salto com $l = 1$ para frente ou para trás. Foram contabilizados os primeiros 100 passos para cada caminhante

Outro ponto de vista do caminhante aleatório é através do uso do conceito de *ensemble*, cuja definição é um conjunto virtual formado por, idealmente, um número infinitamente grande de indivíduos inicialmente idênticos que visitam todos os possíveis estados que este sistema pode assumir através de sua dinâmica intrínseca. Em outras palavras, um *ensemble* representa a distribuição de probabilidade de um dado sistema.

Suponha agora que o caminhante aleatório tenha passo contínuo, segundo uma distribuição contínua $f(x)$, ou seja, cada passo do caminhante aleatório é escolhido de acordo com $f(x)$, e o tamanho do passo é determinado segundo a expressão (??).

Podemos fazer a análise, assim como no caso discreto, da posição média e desvio padrão calculado para um *ensemble* de caminhantes. Vamos estudar o caso em que o passo é determinado segundo uma distribuição gaussiana na seção abaixo.

Caminhante Aleatório e o Teorema do Limite Central

Suponha um *ensemble* de N caminhantes aleatórios cuja dinâmica individual seja dada por uma série de passos aleatórios independentes que seguem uma distribuição $f(x)$ com média μ e variância σ^2 . O teorema do limite central diz que para valores grandes de N :

$$\sum_{t'=0}^t \mathbb{E}[X(t')] \sim N\mu,$$

onde:

$$\mathbb{E}[X(t)] = \frac{X_1(t) + X_2(t) + \dots + X_N(t)}{N},$$

e onde $X_i(t)$ é o caminho total percorrido por um caminhante aleatório no instante t . Além disso, o teorema também diz que os valores de X_i formarão uma distribuição normal (Gaussiana) ao redor de $N\mu$, com variância dada por $\sim N\sigma^2$.

Distribuições Sem Momento Definido

Considere o exemplo de um caminhante aleatório que dá passos segundo uma distribuição de Cauchy:

$$f(x) = \frac{1}{\pi(1+x^2)}. \quad (3)$$

Podemos ver facilmente, conforme explicado na seção ?? que para gerar números aleatórios X que seguem esta distribuição, à partir de um gerador de valores R com distribuição uniforme, podemos usar a seguinte expressão:

$$X = \tan \left[\pi \left(R + \frac{1}{2} \right) \right].$$

Os gráficos resultantes mostram que apesar da média temporal da posição dos caminhantes permanecer próxima de zero, a largura da distribuição difere do que seria esperado para o caso de difusão normal. Isto pode ser explicado porque a distribuição (3) não possui segundo momento definido.

Atividades Sugeridas

- 1 Utilize o gerador aleatório congruente linear para calcular os primeiros 200 números pseudo-aleatórios gerados para os seguintes valores: a) $(x_0 = 27; a = 17; b = 43; m = 100)$ b) $(x_0 = 1, 2, 3, 4; a = 13; b = 0; m = 2^6)$. Para as sequências geradas verifique qual o período da mesma.
- 2 Faça um programa que simule a difusão de M caminhantes aleatórios em um espaço unidimensional. a) Faça histogramas da distribuição espacial de 10^5 caminhantes após 10000 passos quando esses se deslocarem respectivamente de 1, 2 e 3 unidades de distância a cada passo de tempo. b) Faça gráficos no gnuplot da distribuição espacial dos caminhantes. c) Ajuste o logaritmo dos histogramas encontrados a parábolas do tipo: $f(x) = b(x^2) + \log(a)$. d) Determine o desvio quadrático médio por passo em cada caso. e) Relacione o valor do parâmetro b ajustado com a distância percorrida por passo e com o tempo.

Atividades Sugeridas

- 3 Use o método da transformada inversa para encontrar um gerador de números aleatórios cuja distribuição, $p(x)$, seja dada por $p(x) = \frac{A^2}{x^2 + b^2}, x \in \{-\infty, \infty\}$. Encontre a relação entre b e A que normaliza a distribuição. Determine os dois primeiros cumulantes dessa distribuição.
- 4 Faça um programa que simule a difusão em uma dimensão de M caminhantes de passos contínuos. A probabilidade de um passo de tamanho x é dada por $p(x) = \frac{1}{2\lambda} e^{-|x|/\lambda}, x \in \{-\infty, \infty\}$. Faça um histograma de distribuição de caminhantes, use $\lambda = 2$.
- 5 Refaça o problema 4 usando o gerador encontrado no problema 3. Use $b = 5$.

Atividades Sugeridas

- 6 Use o método de Monte Carlo para calcular $\int_{-\infty}^{\infty} p(x) dx$ onde $p(x)$ é definido na questão 3. Use $b = 5$ e a distribuição de números aleatórios é uma exponencial como definida na questão 4.
- 7 Mostre que o k -ésimo momento de uma distribuição uniforme definida em um intervalo L é dada por $\langle x^k \rangle = L^k / (k + 1)$. Faça um programa que calcule numericamente o resultado obtido no item anterior.
- 8 Calcule $\int_0^3 x^{3/2} \exp(-x) dx$ através do método de Monte Carlo utilizando a função $p(x) = A \exp(-x)$ como função de amostragem.
- 9 Utilize a função $p(x) = A \exp(-ax)$ como função de amostragem para calcular a integral $\int_0^{\pi} (x^2 + \cos^2(x))^{-1} dx$. Determine o valor de a que minimiza a variância da integral.