Método de Monte Carlo
Resolução de Integrais

Ref: H.Gould e J. Tobochnik

Para integrais em uma dimensão as regras do trapezóide e de Simpson são melhores, mais rápidas. A técnica de resolução de integrais por Monte Carlo é superior para integrais multidimensionais.

Vamos estudar dois métodos:

- *Hit or Miss*
- Método da média - simples e com *importance sampling*

Hit or Miss

Como usar um monte de pequenas pedras para medir a área de um lago?

Suponha que o lago esteja dentro de um campo de área A conhecida.

Se jogarmos as pedras uma a uma aleatoriamente de forma a caírem sempre dentro do campo, podemos contar quantas caem dentro do lago. Assim temos a relação:

$$\frac{A_{lago}}{A} = \frac{n_{splash}}{N}$$

Temos uma função $f(x)$ e queremos calcular a integral definida $=$ área entre a e b.

Escolhemos um retângulo de altura H e largura $(b-a)$.

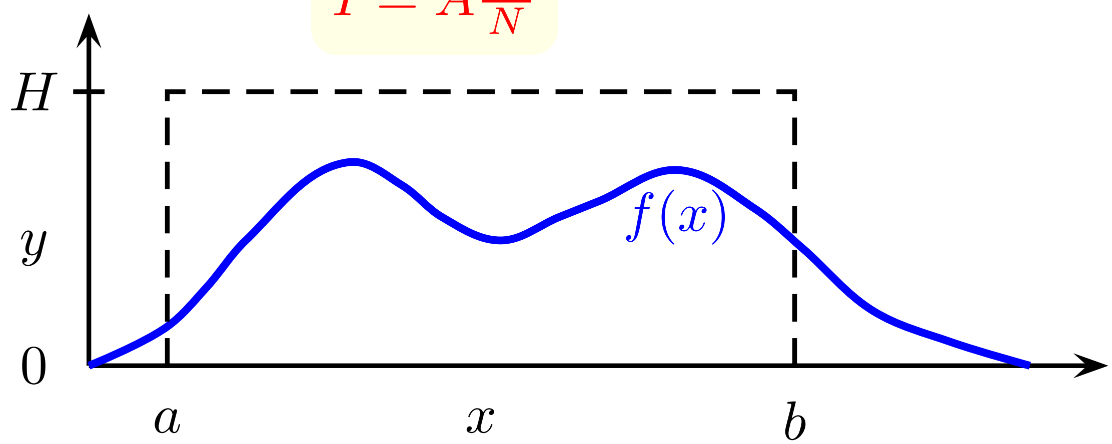
Sorteamos N vezes dois números aleatórios uniformemente distribuídos:

$$a \leq x_i \leq b \qquad 0 \leq y_i \leq H$$

Contamos quantas vezes $y_i \leq f(x_i) \equiv n_s$

e temos então que

$$I = A\frac{n_s}{N}$$



Primeira Tarefa de Hoje

- Escreva um programa que calcule a integral $I = \int_0^1 \sqrt{1-x^2}\mathrm{d}x$ usando o método *Hit or Miss*
- Utilize um retângulo de altura $H = 1$
- O número de pontos sorteados deve ser lido do teclado
- o valor exato da integral é $\pi/4$.
- imprima:
  - o número de sorteios
  - a diferença entre o resultado obtido com o método e o valor exato
- Rode o programa algumas vezes modificando o número de sorteios e observe como muda o valor da diferença

Método da Média

Uma maneira de se calcular uma integral é usar um teorema do cálculo que diz:

$$\int_a^b f(x)\mathrm{d}x = (b-a)\langle f\rangle$$

Como calcular a média? Se tivermos uma lista de números aleatórios uniformemente distribuídos entre a e b podemos obter a média através de:

$$\langle f\rangle = \tfrac{1}{N}\sum_{i=1}^{N} f(x_i)$$

e então:

$$\int_a^b f(x)\mathrm{d}x \approx (b-a)\frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

Quanto maior N melhor a aproximação.

Erro no Método de Monte Carlo

No exemplo dado, calculei $\boxed{I = \int_0^1 4\sqrt{1-x^2}\ \mathrm{d}x}$ para $N = 10.000$ e obtive $I_N = 3.1489$. O resultado exato é $I = \pi = 3.1416$. Portanto para este N, $\epsilon = 0.0073$.

A pergunta é: Como saberemos se $n = 10.000$ atingirá a precisão desejada?

Veremos que o melhor que podemos fazer é calcular a probabilidade de que o valor verdadeiro I esteja num certo intervalo centrado em I_N

Queremos achar σ_m tal que

I_N tem 68% de chance de estar entre $I - \sigma_m$ e $I + \sigma_m$

Um possível chute para a estimativa do erro é o desvio padrão σ :

$$\sigma^2 = \langle f^2 \rangle - \langle f \rangle^2$$

onde $\langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$ e $\langle f^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} f(x_i)^2$

Repare que se f não depende de x, isto é, se a distribuição for uniforme, σ seria 0.

Para o nosso exemplo obtemos $\sigma = 0.8850$ que é muito diferente do valor 0.0073 → o desvio padrão não é uma boa estimativa.

Já era de se esperar pois o erro deve diminuir com N e o desvio padrão não diminui.

Outra tentativa

Uma maneira de se obter uma estimativa do erro é rodar algumas m vezes (experimento), cada vez com o mesmo número N de pontos. Para cada experimento obtemos o valor da média M e calculamos o desvio padrão da média

$$\sigma_m^2 = \langle M^2 \rangle - \langle M \rangle^2$$

com $\quad \langle M \rangle = \frac{1}{m} \sum_{j=1}^{m} M_j \quad$ e $\quad \langle M^2 \rangle = \frac{1}{m} \sum_{j=1}^{m} M_j^2$

Como a sequência de números aleatórios é diferente, M varia para cada experimento.

Vamos fazer 10 experimentos com $N = 10.000$ cada:

j	M	σ
1	3.14892	0.88501
2	3.13255	0.89865
3	3.14042	0.88924
4	3.14600	0.88525
5	3.15257	0.88757
6	3.13972	0.89698
7	3.13107	0.89700
8	3.13585	0.89406
9	3.13442	0.89746
10	3.14047	0.89213

Com estes valores obtemos $\sigma_m = 0.0068$ que é consistente com 0.0073

Portanto σ_m é a medida do erro

O resultado fica escrito como:

$I_n = 3.149 \pm 0.007$ para $n = 10000$

Este método de gerar vários experimentos para determinar σ_m não é nada útil, mas pode-se mostrar que

$$\boldsymbol{\sigma_m \approx \frac{\sigma}{\sqrt{N}}}$$

que se torna exato quando $N \to \infty$

Note que σ_m decresce com $\sqrt{N}$.

Para o primeiro valor que obtivemos

$$\sigma = 0.8850 \Rightarrow \sigma_m = 0.8850/\sqrt{10000} = 0.009$$

que é da mesma ordem de 0.007

Segunda Tarefa de Hoje

- Escreva um programa que calcule a integral $I = 4\int_0^1 \sqrt{1-x^2}\mathrm{d}x$ usando o método da média.
- Rode duas vezes o programa: 50 experimentos com $N = 100$ e 50 experimentos com $N = 10000$.
- o valor exato da integral é π.
- imprima em um arquivo (tudo na mesma linha):
  - o í ndice do experimento
  - o valor da integral calculado com o método
  - o desvio padrão da função $\sigma = \sqrt{\langle f^2 \rangle - \langle f \rangle^2}$
  - a diferenç a entre o resultado obtido com o método e o valor exato
  - $\sigma_m = \sigma/\sqrt{N}$
- Observe através do gnuplot se as afi rmativas que fi zemos sobre o erro do método se comprovam

```
Cria uma funcao que dado x retorna 4*sqrt(1-x*x)

loop no numero de experimentos
   inicializa a semente
   loop no numero de pontos
      gera um numero aleatorio x entre a e b
      calcula f(x)
      calcula soma f(x)
      calcula soma f(x)^2
   fim loop
   calcula integral
   calcula sigma
   calcula erro da media
   imprime
fim do loop
```

Importance Sampling

Vimos que o erro da integral pelo método de Monte Carlo é proporcional à variância do integrando, isto é, ao quanto $f(x)$ varia.

Vamor ver uma forma de diminuir esta variância, ou em outras palavras, transformar o integrando em algo mais uniforme.

Queremos calcular

$$I = \int_a^b f(x)\mathrm{d}x$$

Vamos multiplicar e dividir o integrando por uma função $p(x)$ tal que

$$\int_a^b p(x)\mathrm{d}x = 1$$

$$I = \int_a^b \frac{f(x)}{p(x)} p(x) \mathrm{d}x$$

Fazemos agora uma mudança de variável de x para y de forma que

$$I = \int_a^b \frac{f(x)}{p(x)} \mathrm{d}y$$

com $p(x)dx = dy \Longrightarrow y(x) = \int p(x)dx$

Se $p(x)$ for escolhida como uma que se comporta como $f(x)$, o integrando vai ser $\approx$ cte. O que estamos fazendo aqui é gerar uma distribuição de acordo com $p(x)$.

Se $p(x) \approx f(x)$ vamos gerar mais pontos onde $f(x)$ é grande e pouco recurso é gasto gerando pontos onde $f(x)$ é pequeno.

Exemplo

Queremos resolver $I = \int_0^1 e^x \mathrm{d}x$

Uma função parecida é a sua própria expansão em série de Taylor em torno de 0: $\quad p(x) = 1 + x$

$$I = \int_0^1 \frac{e^x}{1+x}(1+x)\ \mathrm{d}x \qquad \text{com} \quad \mathrm{d}y = (1+x)\mathrm{d}x$$

$$y = \int_0^x (1+x')\mathrm{d}x' = x + \frac{x^2}{2}$$

$$\frac{x^2}{2} + x - y = 0 \Longrightarrow x = \frac{-1\pm\sqrt{1+4\frac{1}{2}y}}{2\ 1/2} = -1 \pm \sqrt{1+2y}$$

para $x = 0 \implies y = 0$ para $x = 1 \implies y = 1 + 1/2 = 3/2$

para que $0 \leq y \leq 3/2$ o sinal positivo deve ser escolhido.

Então geramos um número aleatório y entre 0 e 3/2 e calculamos

$$I = \int_0^{3/2} \frac{e^{\sqrt{1+2y}-1}}{\sqrt{1+2y}}\ \mathrm{d}y$$

Terceira Tarefa de Hoje

- Escreva um programa que calcule a integral $I = \int_0^1 e^{-x^2} \mathrm{d}x$ usando o método da média.
- Use $N = 1000$ imprima o valor da integral e o erro.
- Use a função peso $p(x) = Ae^{-x}$ onde A é escolhido de forma a satisfazer $\int_0^1 p(x)\mathrm{d}x = 1$. Quando você fi zer as contas verá que a integral a resolver será

$$\frac{1}{A} \int_{-A}^{-A/e} e^{-x^2+x} dy$$

com $x = -ln(-y/A)$
- Use $N = 1000$ imprima o valor da integral e o erro. Compare com o resultado anterior.