

Introdução a linguagem Stan (*rstan*), um software para modelos bayesianos.



Universidade Federal do Oeste do Pará (UFOPA)
Campus de Monte Alegre – Engenharia de Aquicultura

Universidade Federal de Lavras (UFLA)
PPG – Estatística e Experimentação Agropecuária

AGRADECIMENTO E COLABORADORES:



UFOPA

Professor: Carlos Antônio Zarzar
E-mail: carloszarzar_@hotmail.com
carlos.zarzar@ufopa.edu.br

Data: 09/03/2022

➤ Sumário

- Estatística Bayesiana;
- Modelos Hierárquicos;
- Método computacional de integração MCMC;
- **Algoritmo de amostragem;**
 - Metropolis-Hasting;
 - Gibbs;
 - Hamiltoniano Monte Carlo;
 - NUTS;

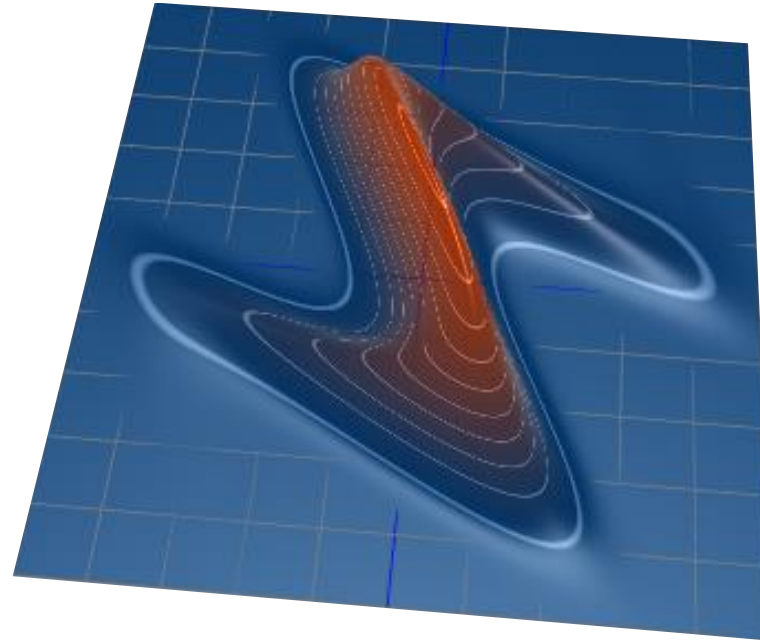


Algoritmo de amostragem



- Algoritmo de amostragem;
 - **Metropolis-Hasting;**
 - Gibbs;
 - Hamiltoniano Monte Carlo;
 - NUTS;

Espaço paramétrico da
distribuição *a posteriori*

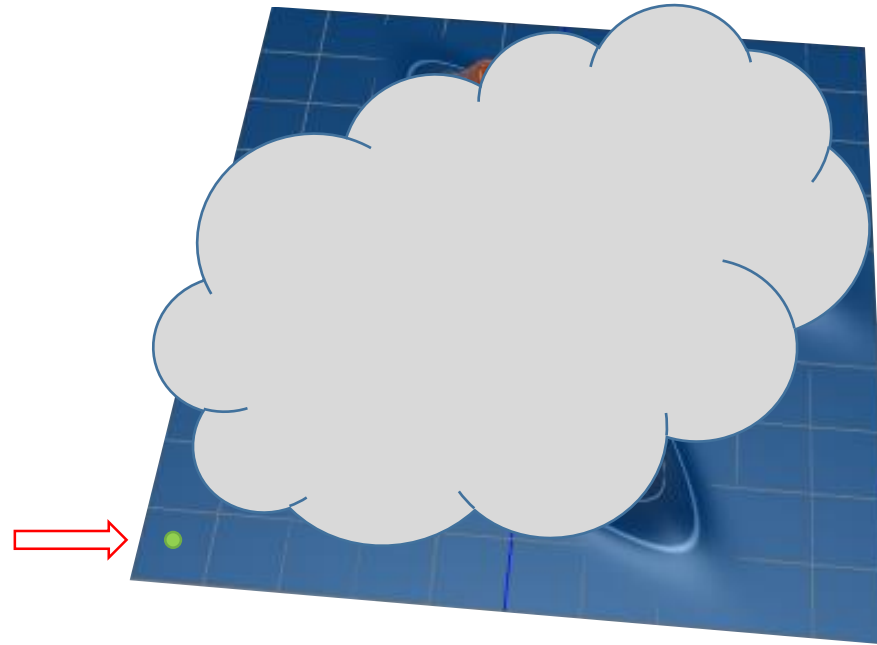


Espaço paramétrico da
distribuição *a posteriori* →

Desconhecido

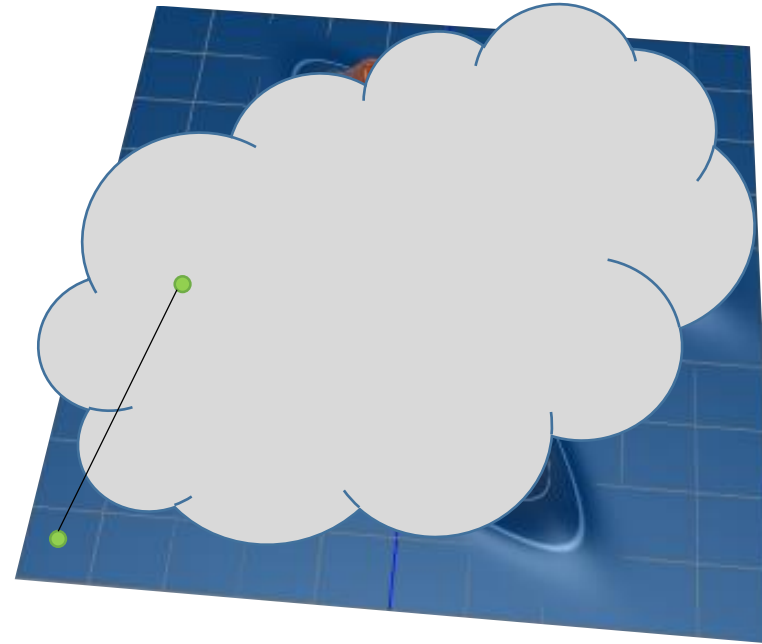


Espaço paramétrico da
distribuição *a posteriori*



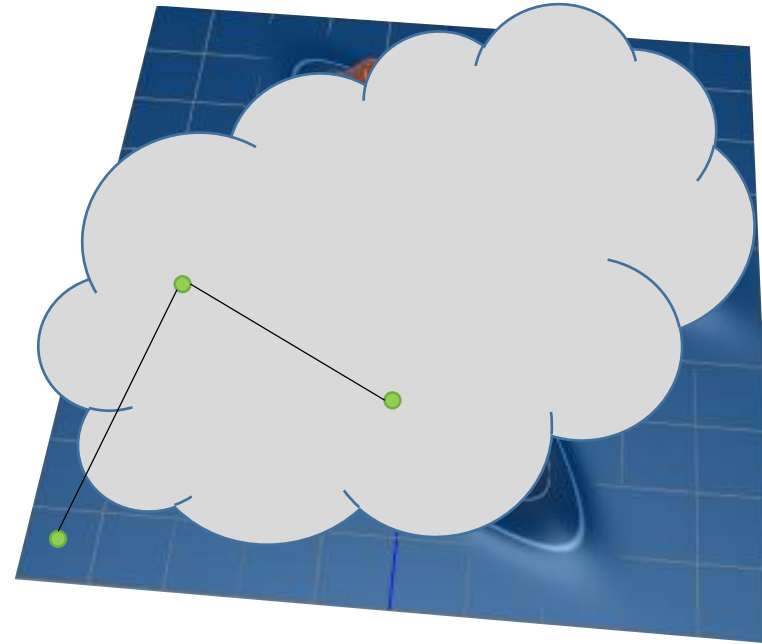
- *Random walk MCMC*

Espaço paramétrico da
distribuição *a posteriori*



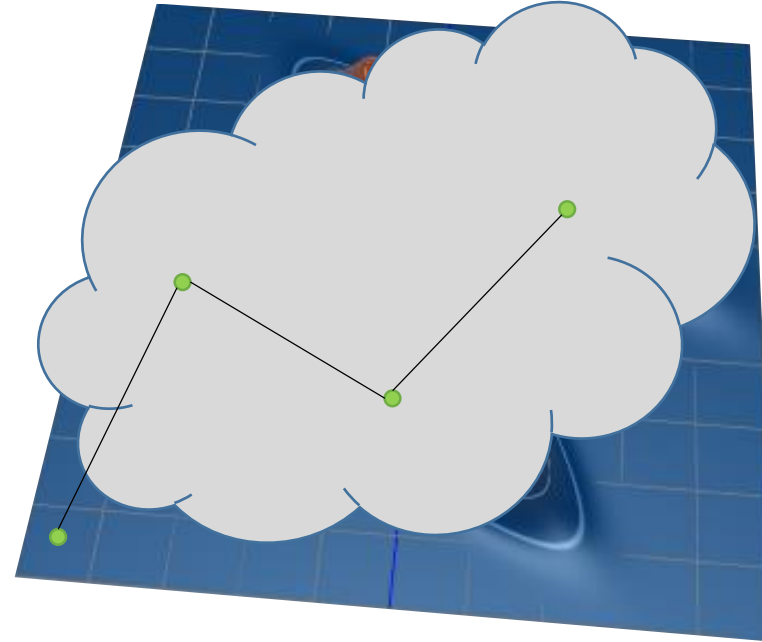
- *Random walk MCMC*

Espaço paramétrico da
distribuição *a posteriori*



- *Random walk MCMC*

Espaço paramétrico da
distribuição *a posteriori*

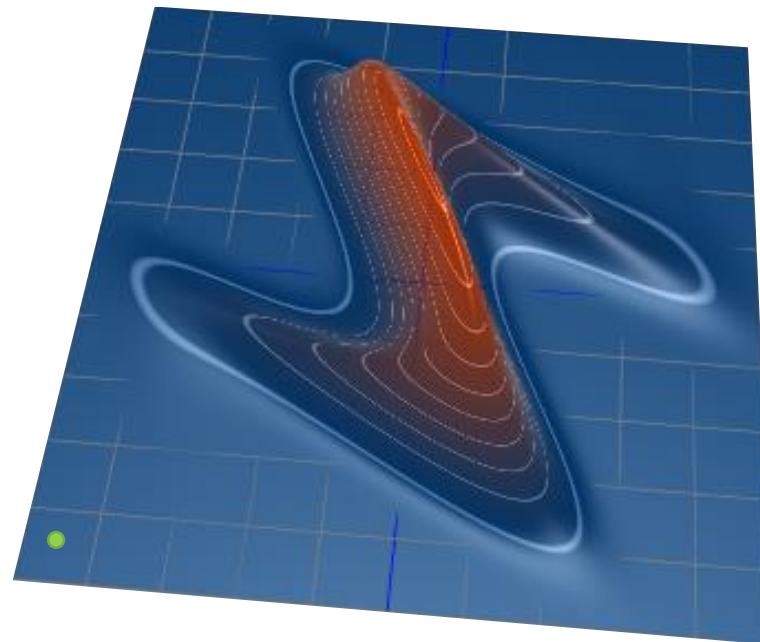


- *Random walk MCMC*

Custo computacional muito grande
para amostrar todo o espaço
paramétrico.



Espaço paramétrico da
distribuição *a posteriori*



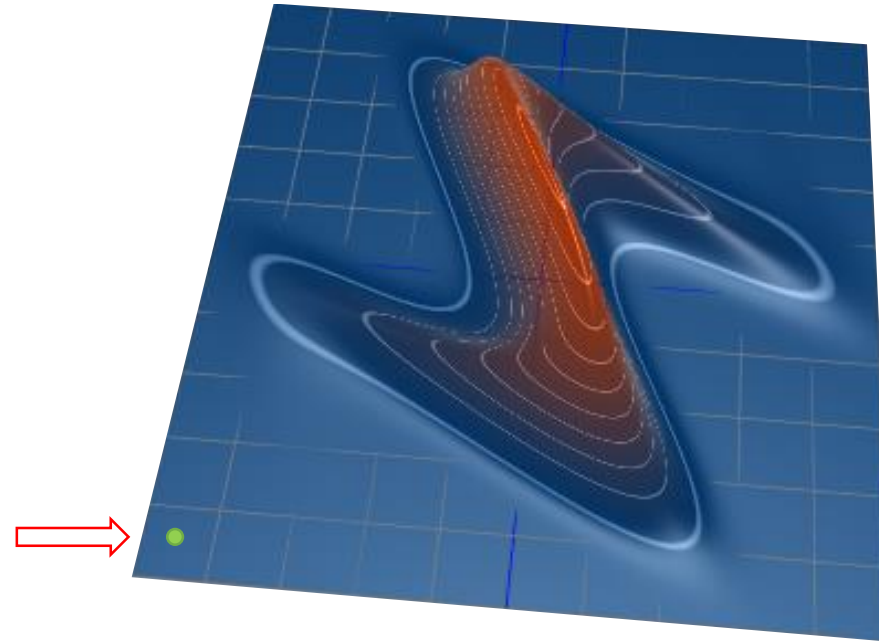
- *Random walk MCMC*

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



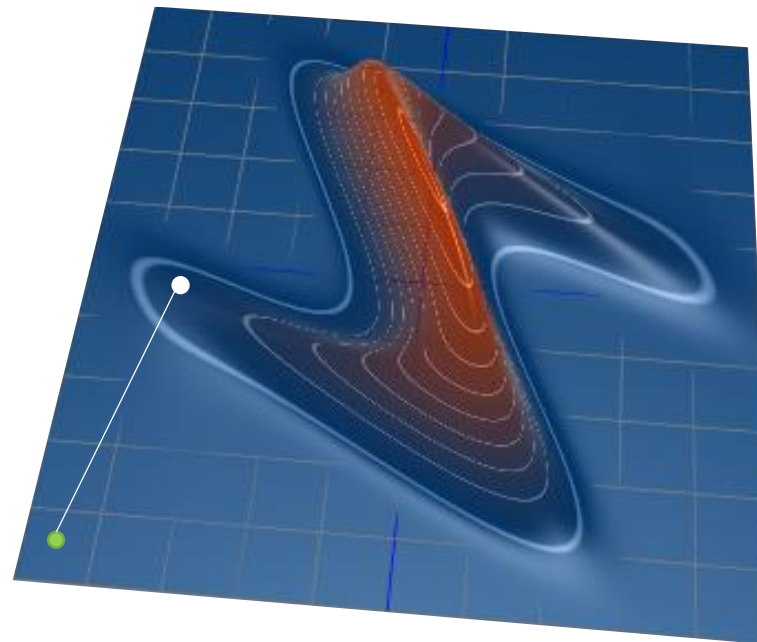
- *Random walk MCMC*

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



- *Random walk MCMC*

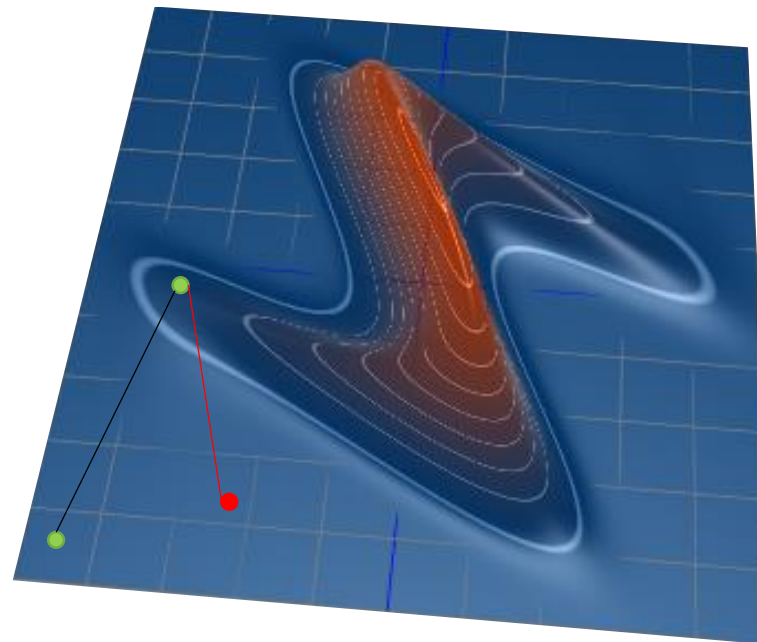
- Amostrador em avaliação;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

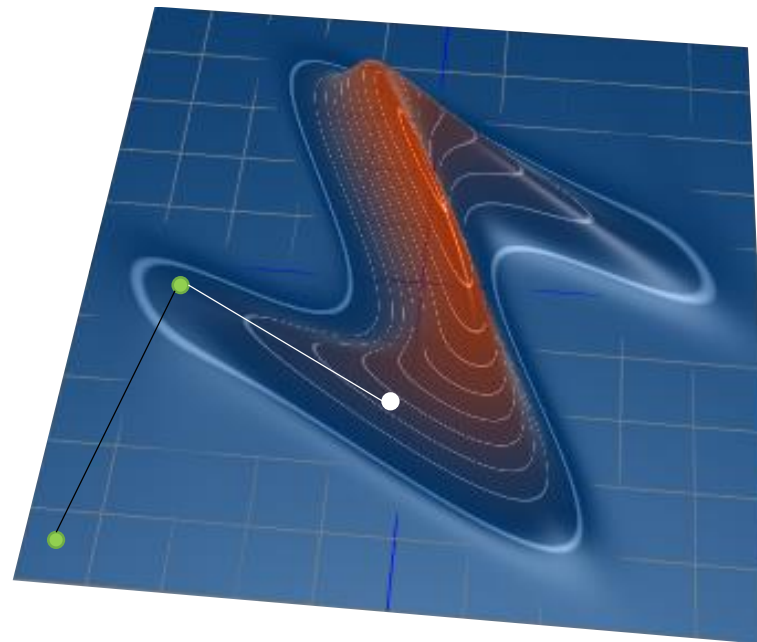
● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

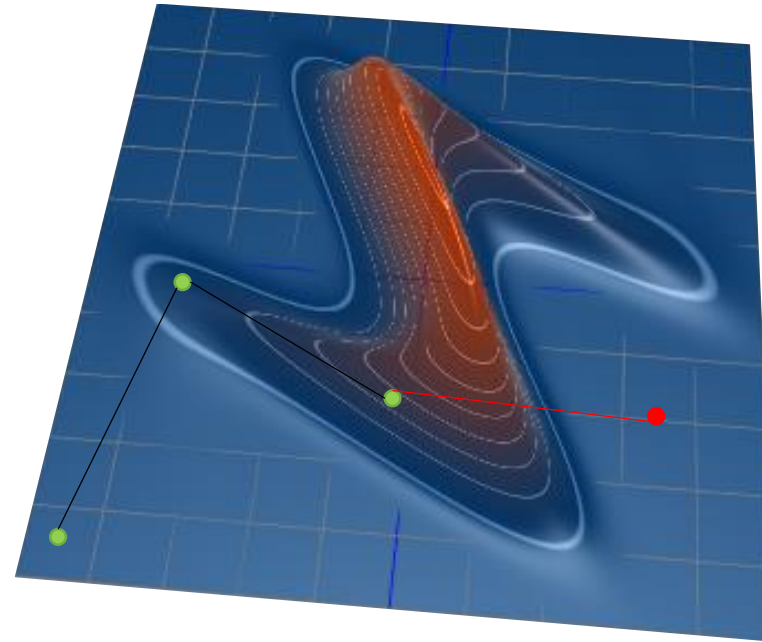
● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

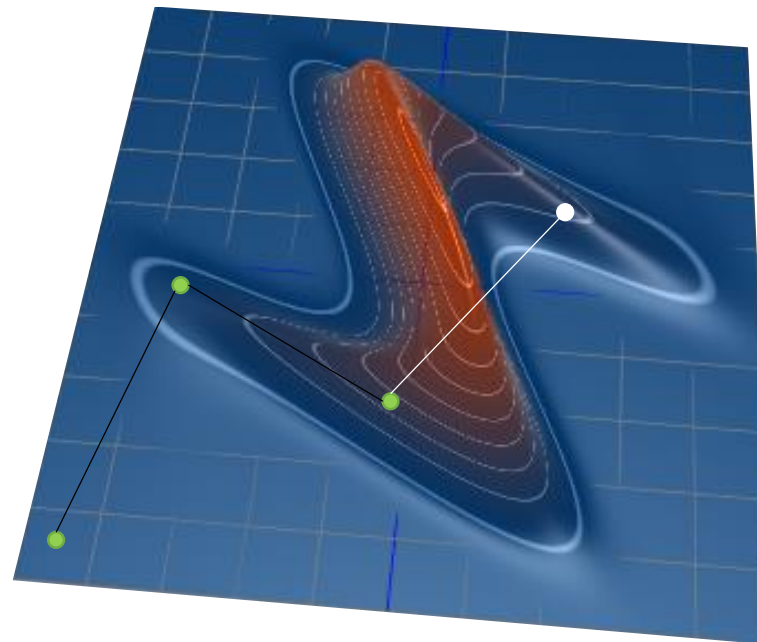
● Amostrador rejeitado;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

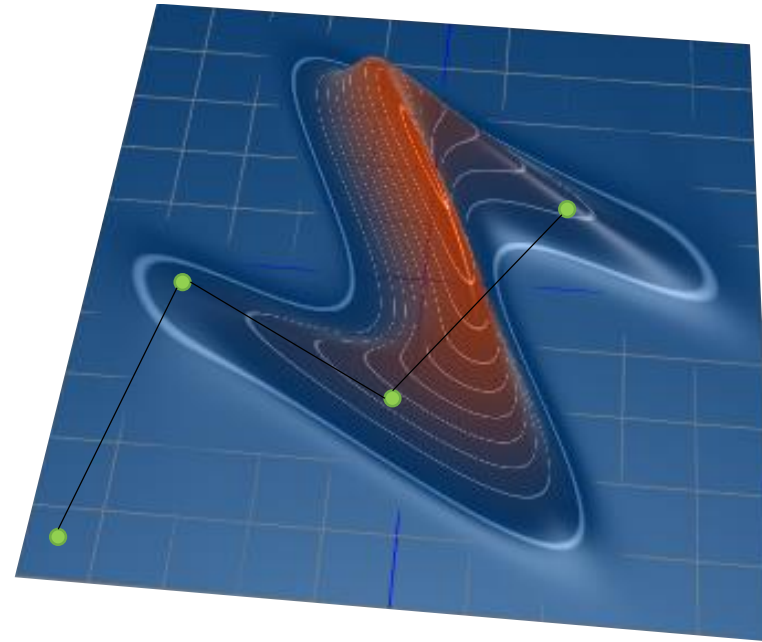
● Amostrador rejeitado;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

b)
$$r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

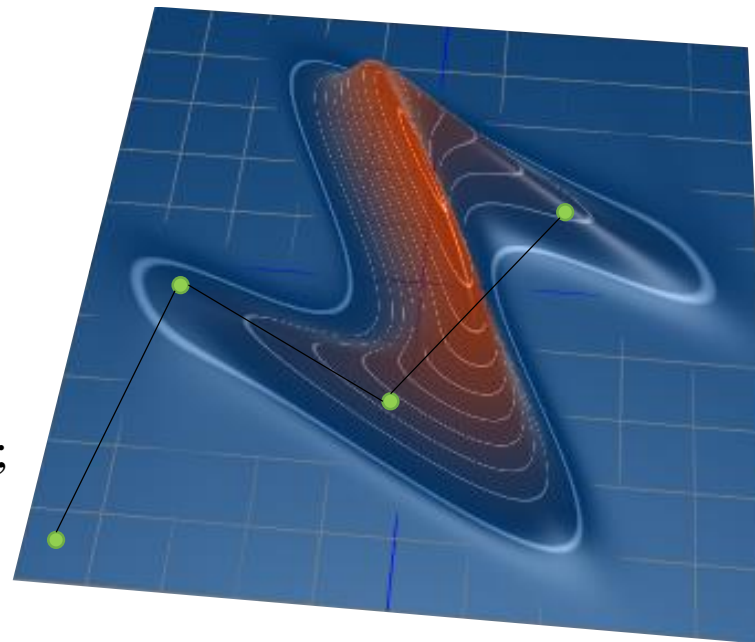
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

▪ $\alpha(\theta^*, \theta_{i-1}) = r$;

aceita a proposta θ^* e defina $\theta_i \leftarrow \theta^*$ com α Prob. Aceitabilidade;

rejeita a proposta θ^* e defina $\theta_i \leftarrow \theta_{i-1}$ com $1 - \alpha$ Prob. Aceitabilidade.

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

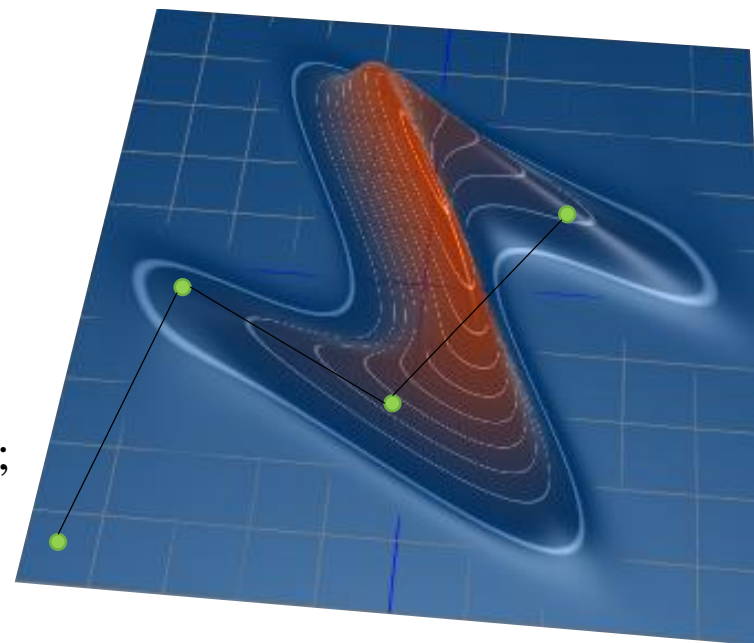
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, aceite a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

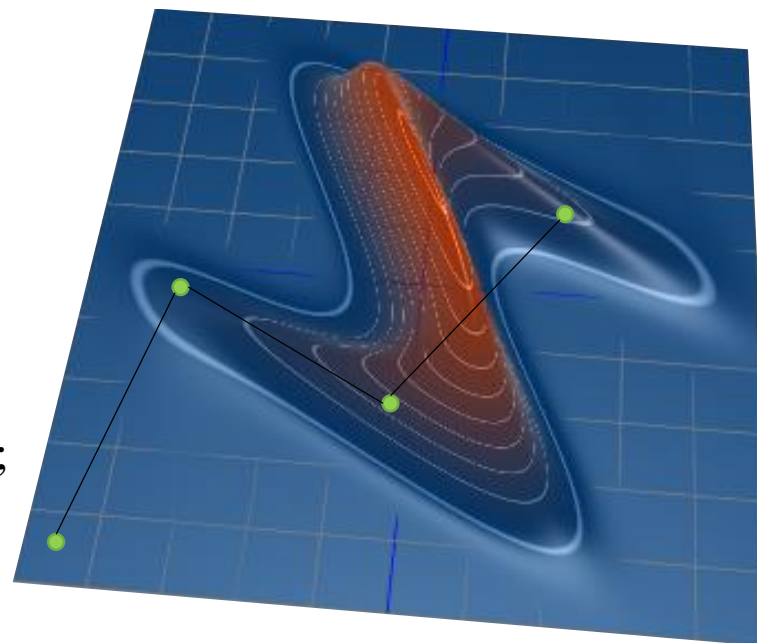
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite a proposta θ^*** e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

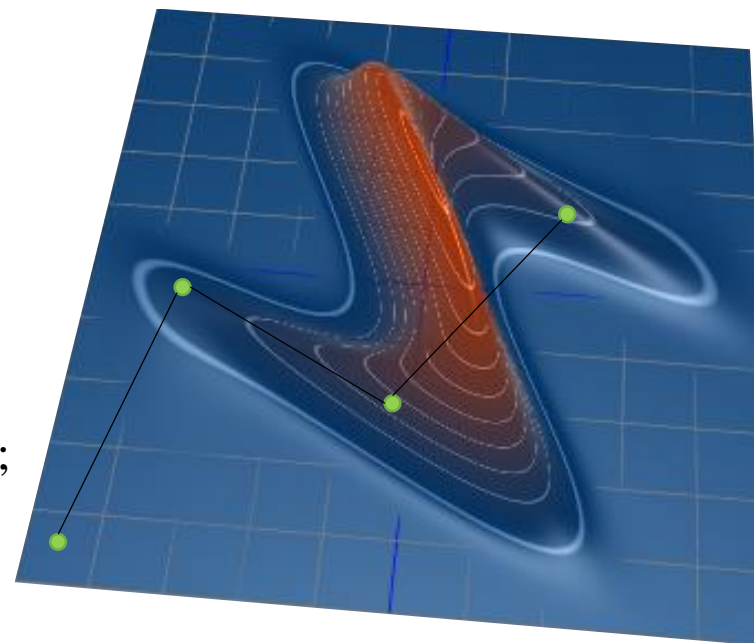
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, aceite a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

b) $r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

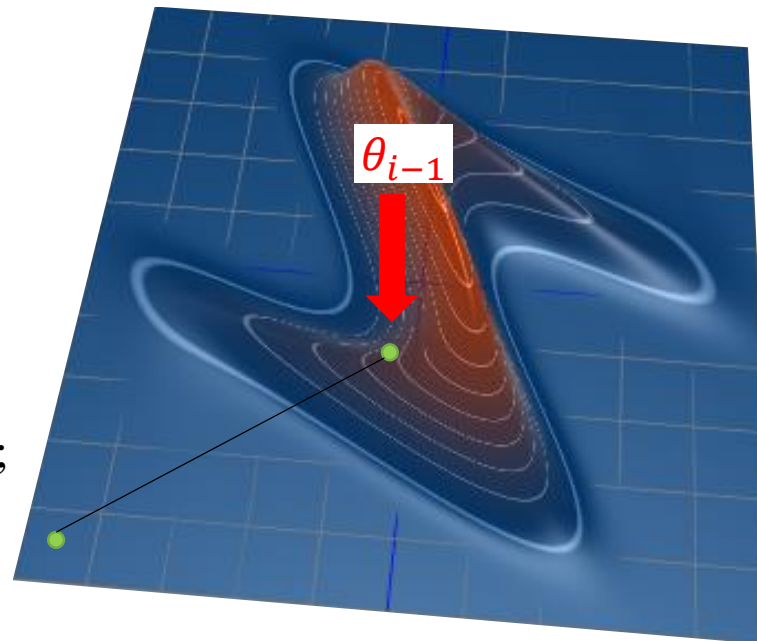
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, aceite a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

b) $r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

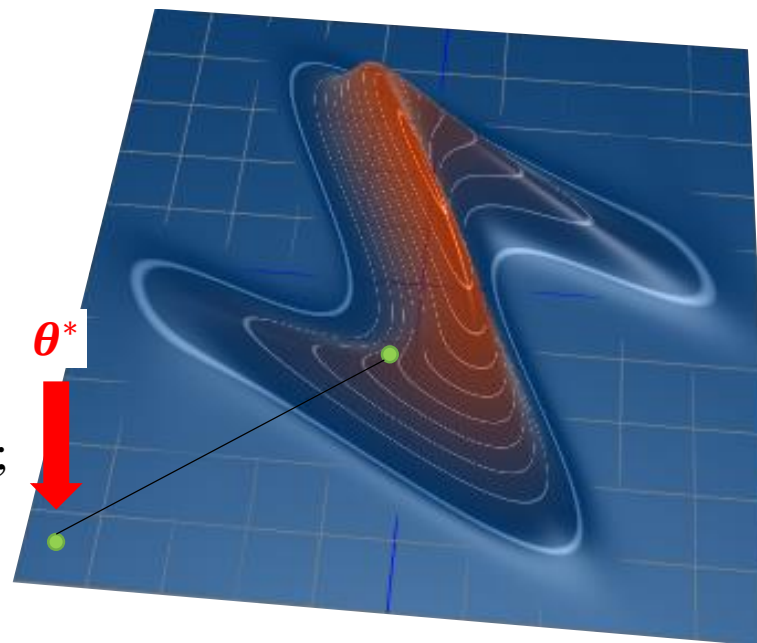
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, aceite a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

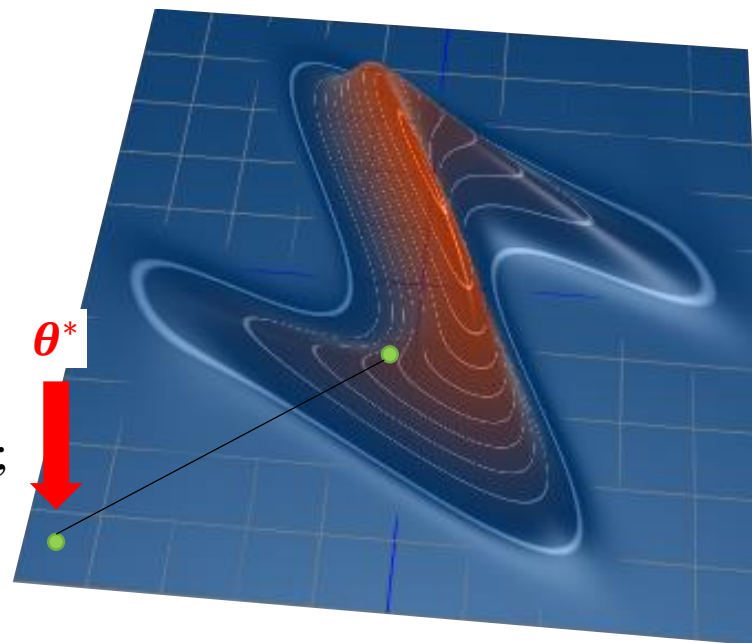
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

b) $r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

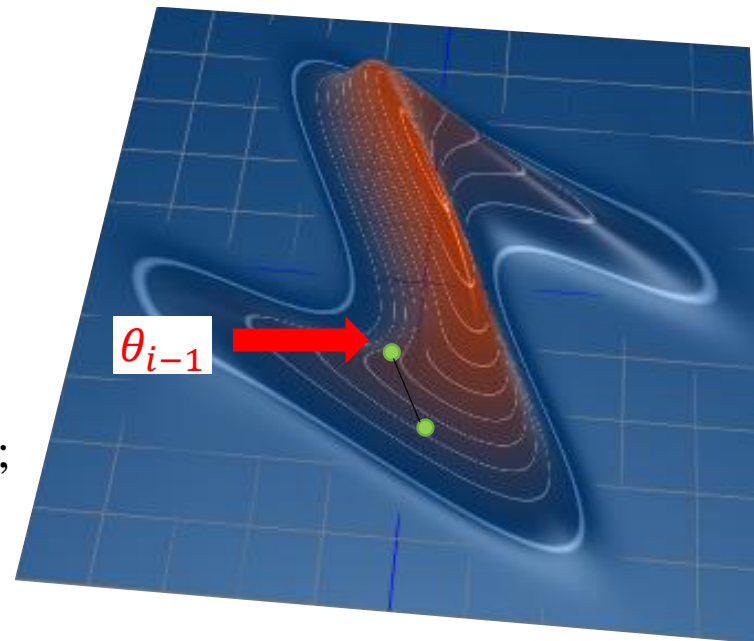
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

b) $r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

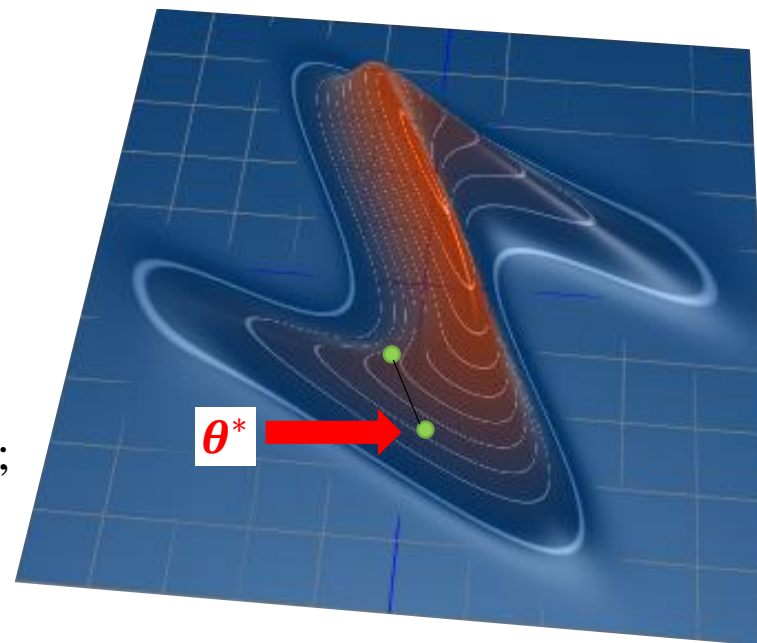
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim \text{Unif}(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

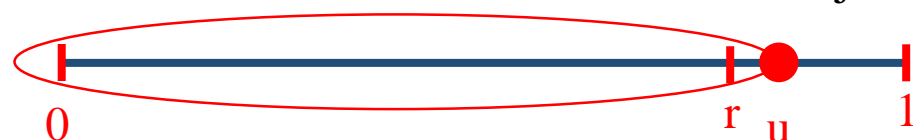
$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

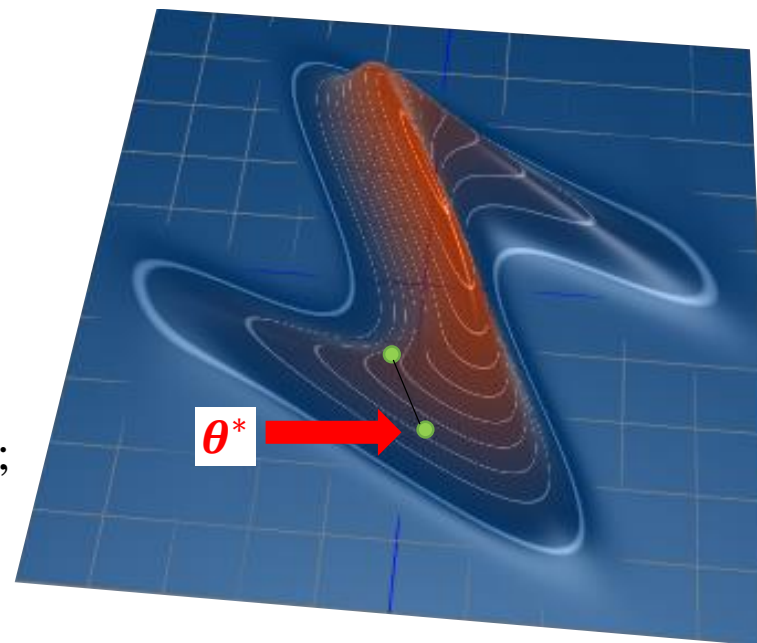
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim \text{Unif}(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

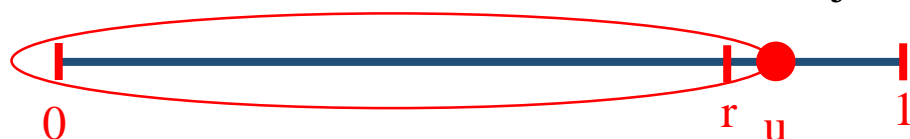
b) $r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

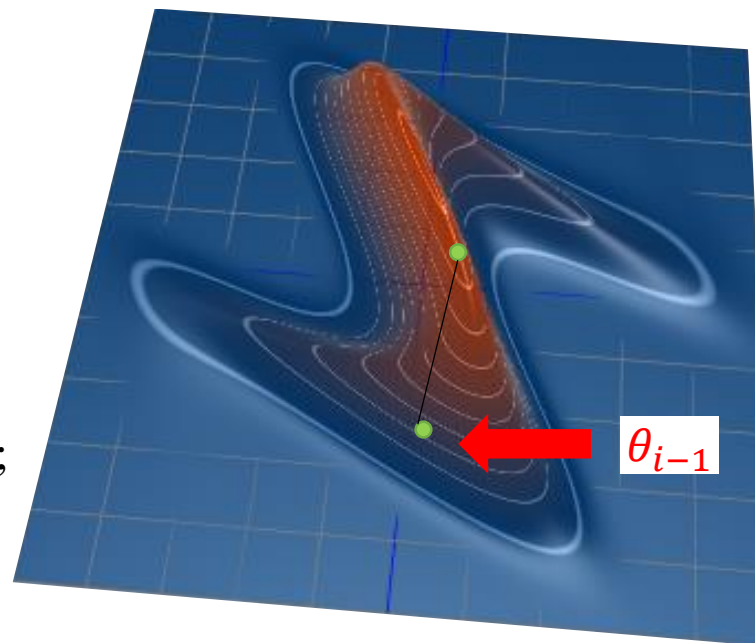
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim \text{Unif}(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

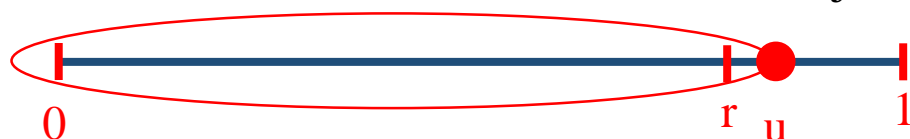
$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

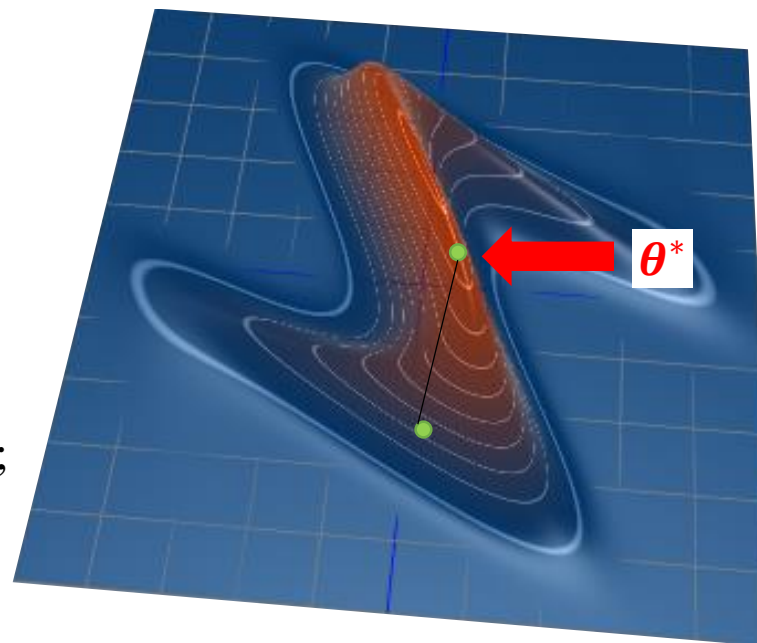
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

1. Selecione um valor inicial θ_0 ;

2. Para $i = 1, \dots, n$ repita:

a) Encontrar o candidato $\theta^* \sim q(\theta^* | \theta_{i-1})$;

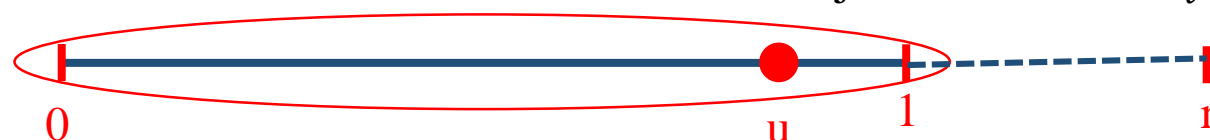
$$b) r = \frac{\pi(\theta^*)/q(\theta^*|\theta_{i-1})}{\pi(\theta_{i-1})/q(\theta_{i-1}|\theta^*)} = \frac{\pi(\theta^*)}{\pi(\theta_{i-1})} \cdot \frac{q(\theta_{i-1}|\theta^*)}{q(\theta^*|\theta_{i-1})}$$

c) Se $r \geq 1$ aceite o valor proposto θ^* e defina $\theta_i \leftarrow \theta^*$;

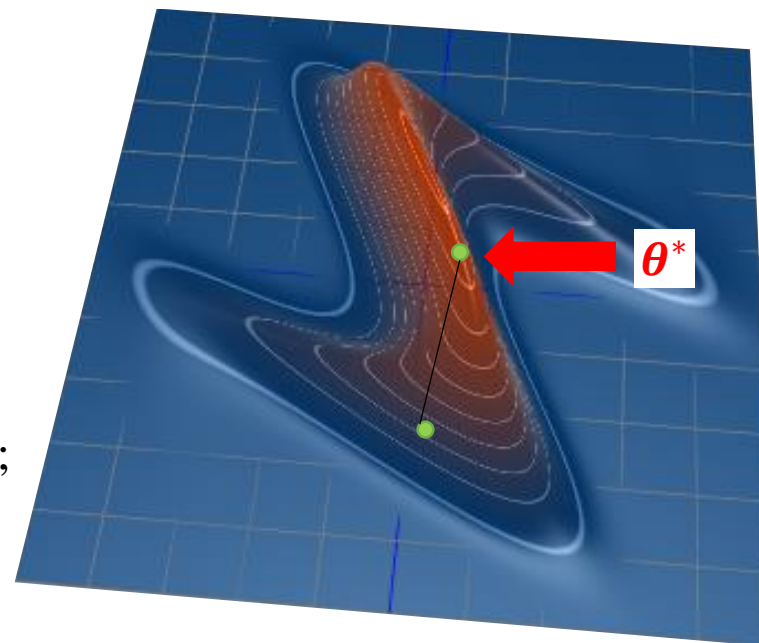
d) Se $0 \leq r < 1$, Calcule a Prob. Aceitabilidade:

- $\alpha(\theta^*, \theta_{i-1}) = \min\{r, 1\}$;
- Desenhe $u \sim Unif(0,1)$;
- Se $u < \alpha(\theta^*, \theta_{i-1})$, **aceite** a proposta θ^* e defina $\theta_i \leftarrow \theta^*$

caso contrário rejeita θ^* e defina $\theta_i \leftarrow \theta_{i-1}$.



Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

● Amostrador em avaliação;

● Amostrador rejeitado;

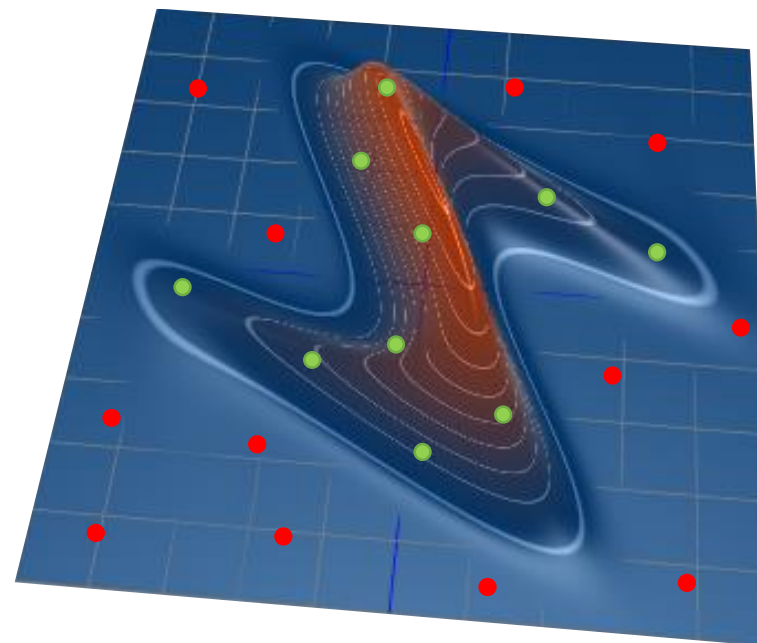
Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Se a maioria da iteração (amostras) for aceita e conseguirmos uma convergência para a distribuição alvo, teremos relativamente um processo mais rápida.

Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

1.000 iteração

600 iteração atingirmos
convergência

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

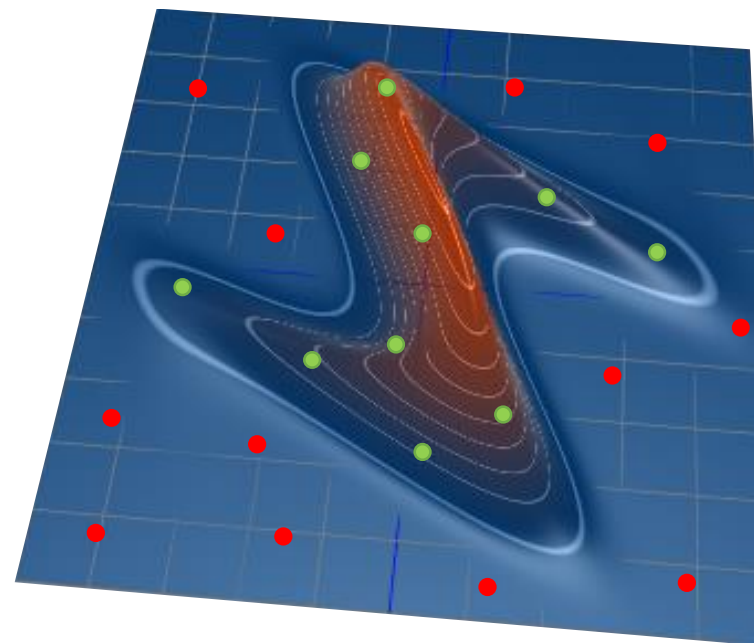
➤ Método a amostragem:

- Metropolis-Hasting;

Se a maioria da iteração (amostras) for aceita e conseguirmos uma convergência para a distribuição alvo, teremos relativamente um processo mais rápida.

Portanto o algoritmo Metropolis-Hastings tem essa informação (índice de aceitação e rejeição) como uma ferramenta de eficiência da algoritmo e consequentemente qualidade da convergência.

Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

1.000 iteração

600 iteração atingirmos convergência

Taxa de aceitabilidade

$$T_{\bullet} = \frac{\text{●}}{\text{●} + \text{●}}$$

Taxa de rejeitabilidade

$$T_{\bullet} = \frac{\text{●}}{\text{●} + \text{●}}$$

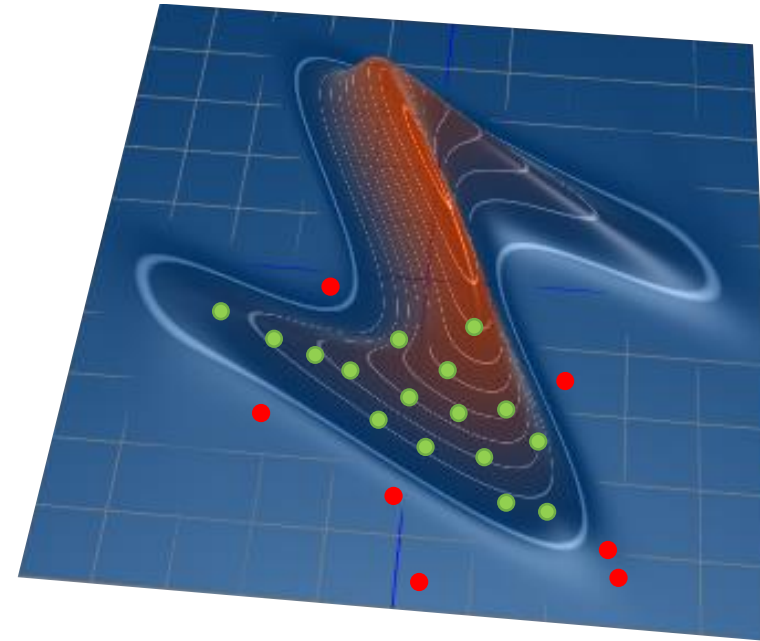
Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Porém não adianta ter uma maior proporção de amostras aceitas pois pode haver algum problema na qualidade da amostragem e uma equivocada convergência rápida.

Espaço paramétrico da distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Taxa de aceitabilidade

$$T_{\bullet} = \frac{\bullet}{\bullet + \bullet}$$

Taxa de rejeitabilidade

$$T_{\bullet} = \frac{\bullet}{\bullet + \bullet}$$

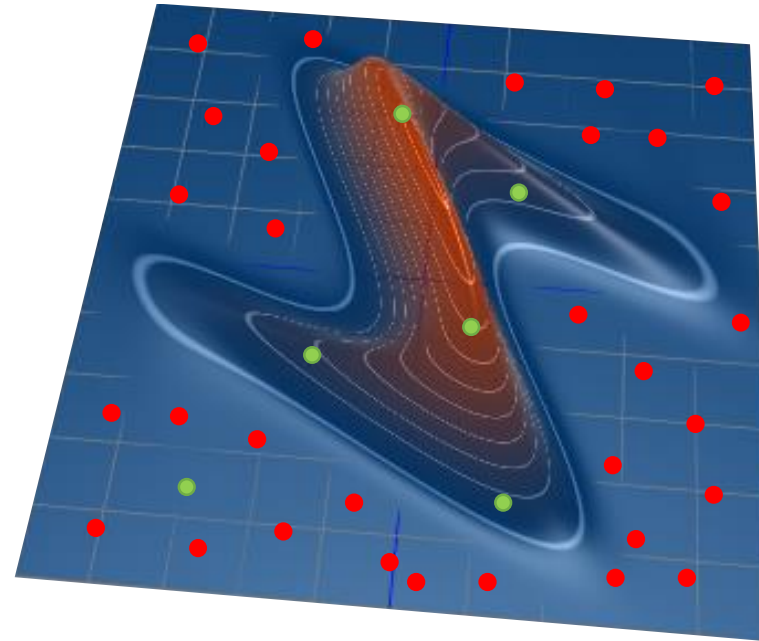
Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

➤ Método a amostragem:

- Metropolis-Hasting;

Espaço paramétrico da distribuição *a posteriori*

E caso tenhamos mais rejeição do que aceite, o custo computacional e o tempo da convergência serão muito altos.



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Taxa de aceitabilidade

$$T_{\bullet} = \frac{\text{●}}{\text{●} + \text{●}}$$

Taxa de rejeitabilidade

$$T_{\bullet} = \frac{\text{●}}{\text{●} + \text{●}}$$

Etapa de **aceitação** ou **rejeição** de cada iteração do passeio aleatório.

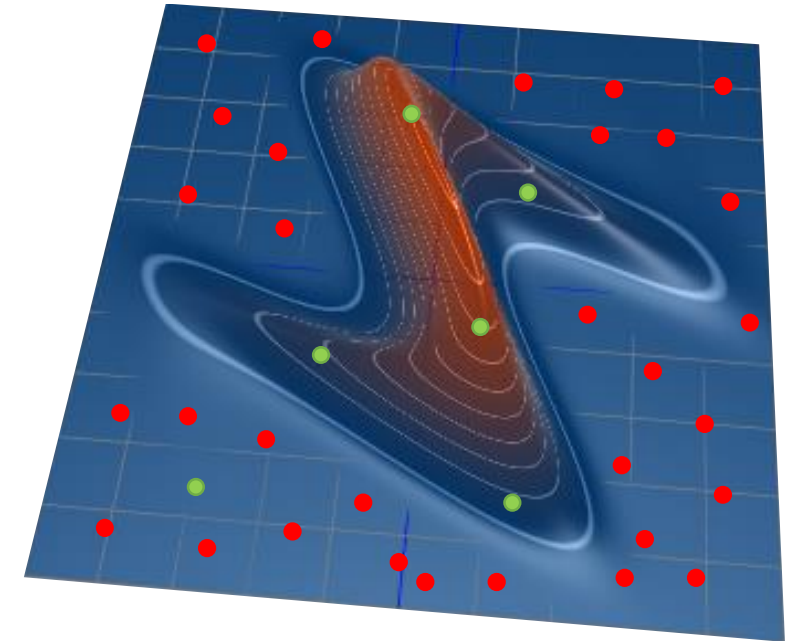
➤ Método a amostragem:

- Metropolis-Hasting;

➤ Resumo:

- É uma amostragem com etapas de **aceitação/rejeição** aplicada MC;

Espaço paramétrico da distribuição *a posteriori*



Taxa de aceitabilidade

$$T_{\bullet} = \frac{\text{green dot}}{\text{green dot} + \text{red dot}}$$

Taxa de rejeitabilidade

$$T_{\bullet} = \frac{\text{red dot}}{\text{green dot} + \text{red dot}}$$

➤ Método a amostragem:

- Metropolis-Hasting;

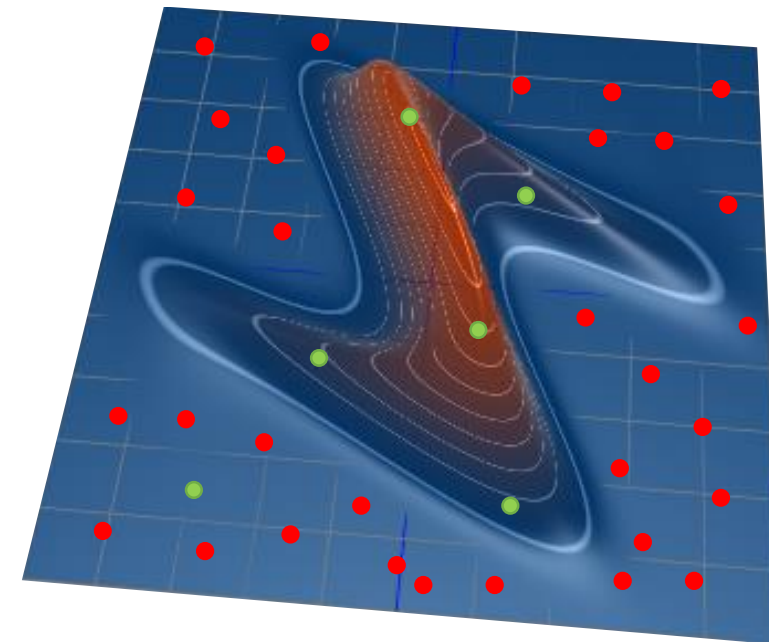
➤ Resumo:

- É uma amostragem com etapas de **aceitação/rejeição** aplicada MC;

➤ Prós:

- Dentre a família do método MC, pode escolher a mais conveniente;
- Funciona para distribuições não normalizadas;
- Fácil implementar;
- Amostragem da *a posteriori* já é uma distribuição conjunta de todos os parâmetros;

Espaço paramétrico da
distribuição *a posteriori*



Taxa de aceitabilidade

$$T_{\bullet} = \frac{\text{green dots}}{\text{green dots} + \text{red dots}}$$

Taxa de rejeitabilidade

$$T_{\bullet} = \frac{\text{red dots}}{\text{green dots} + \text{red dots}}$$

➤ Método a amostragem:

- Metropolis-Hasting;

➤ Resumo:

- É uma amostragem com etapas de **aceitação/rejeição** aplicada MC;

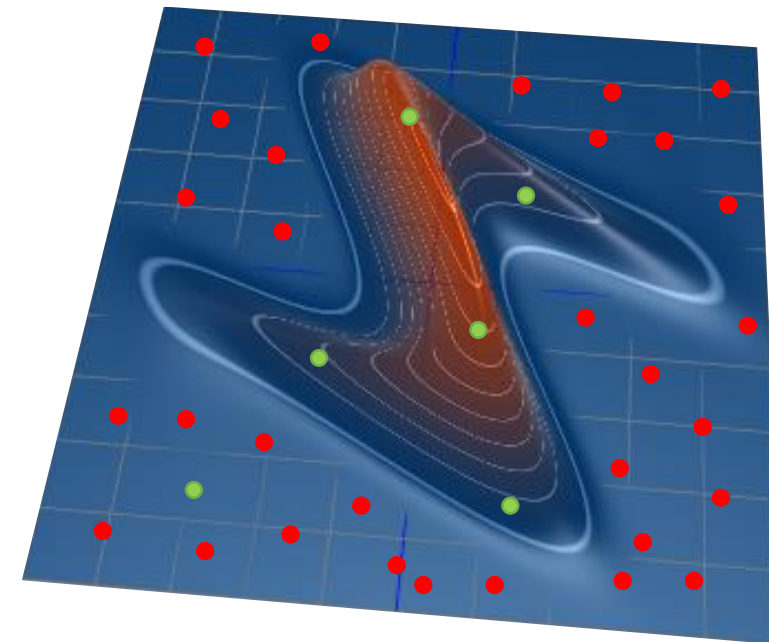
➤ Prós:

- Dentre a família do método MC, pode escolher a mais conveniente;
- Funciona para distribuições não normalizadas;
- Fácil implementar;
- Amostragem da *a posteriori* já é uma distribuição conjunta de todos os parâmetros;

➤ Contras:

- Amostras são correlacionadas;
- Lenta convergência.

Espaço paramétrico da
distribuição *a posteriori*



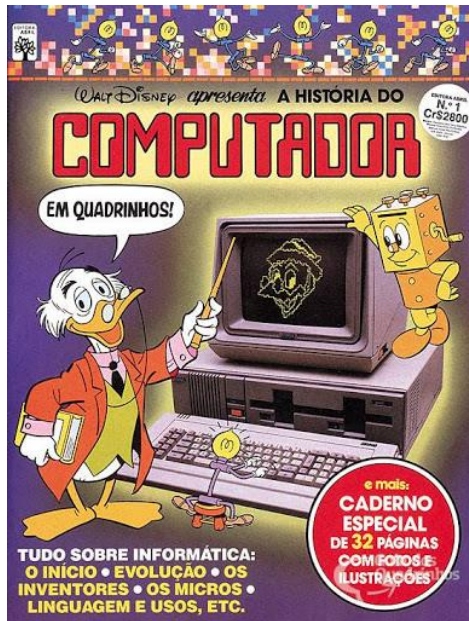
Taxa de aceitabilidade

$$T_{\bullet} = \frac{\text{aceitadas}}{\text{aceitadas} + \text{rejeitadas}}$$

Taxa de rejeitabilidade

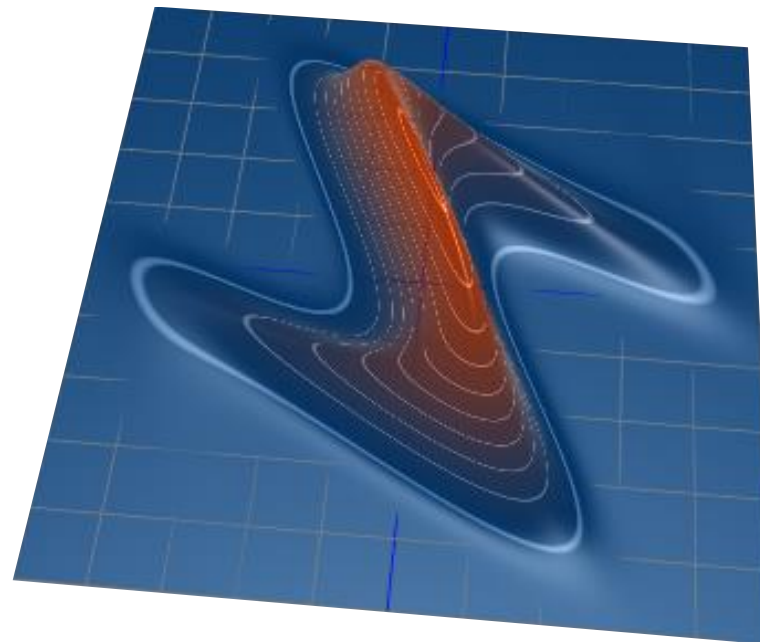
$$T_{\bullet} = \frac{\text{rejeitadas}}{\text{aceitadas} + \text{rejeitadas}}$$

Algoritmo de amostragem



- Algoritmo de amostragem;
 - Metropolis-Hasting;
 - **Gibbs**;
 - Hamiltoniano Monte Carlo;
 - NUTS;

Espaço paramétrico da
distribuição *a posteriori*



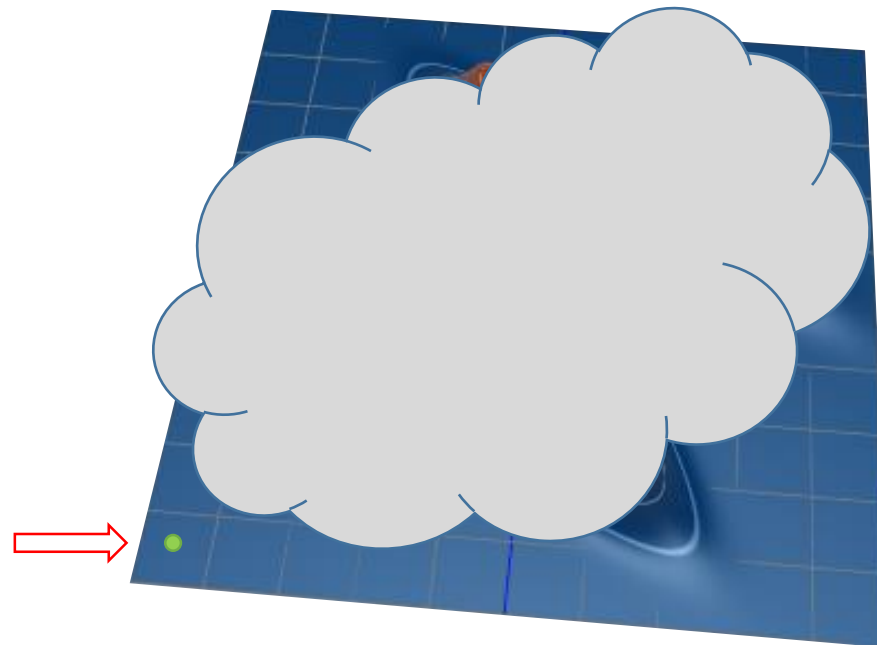
Espaço paramétrico da
distribuição *a posteriori*



Desconhecido

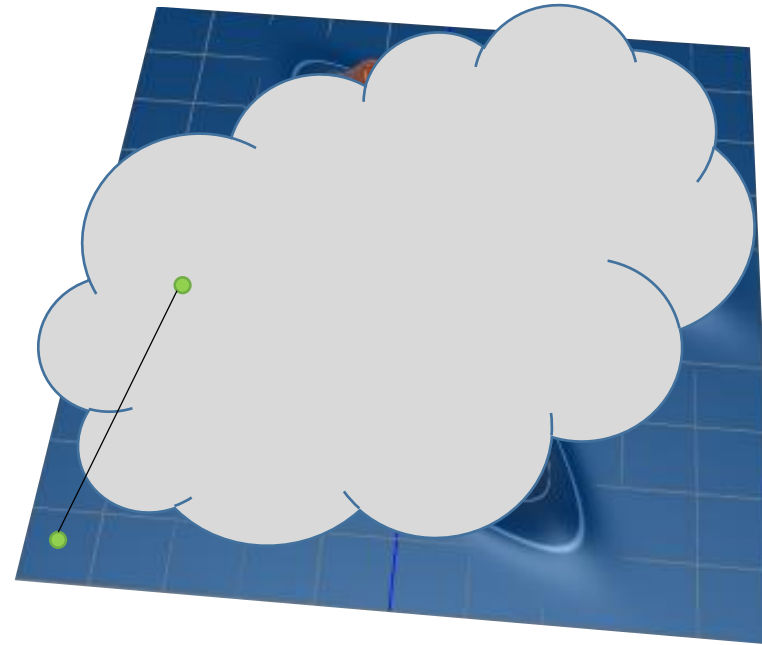


Espaço paramétrico da
distribuição *a posteriori*



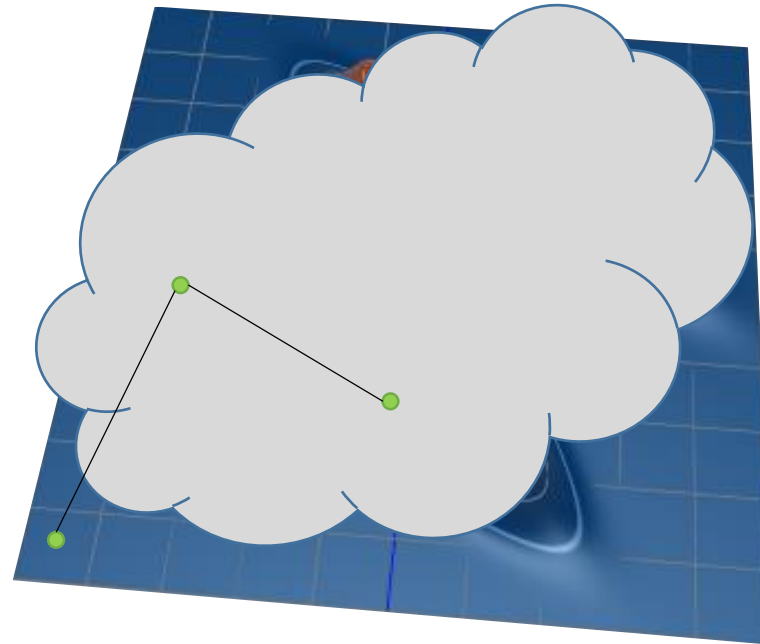
- *Random walk MCMC*

Espaço paramétrico da
distribuição *a posteriori*



- *Random walk MCMC*

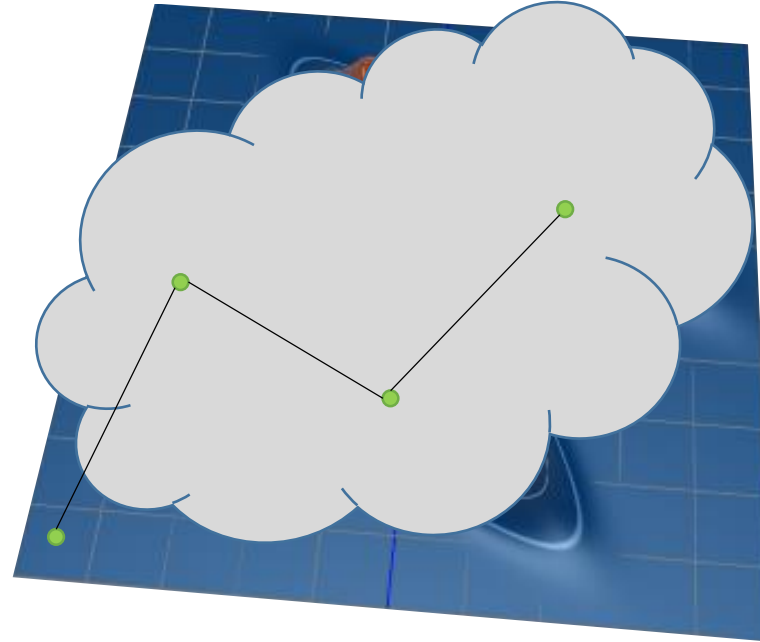
Espaço paramétrico da
distribuição *a posteriori*



- *Random walk MCMC*

Espaço paramétrico da
distribuição *a posteriori*

- *Random walk MCMC*

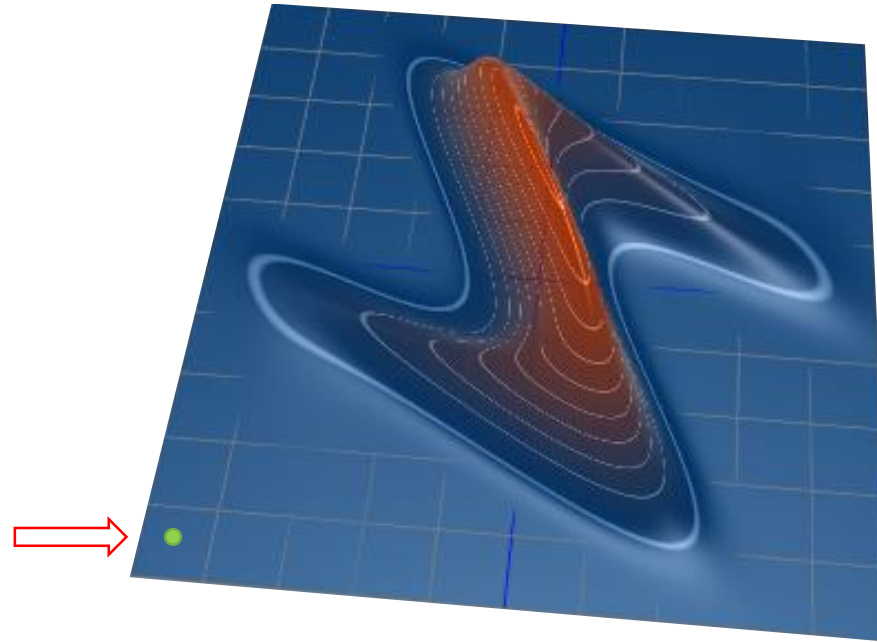


Custo computacional muito grande
para amostrar todo o espaço
paramétrico.



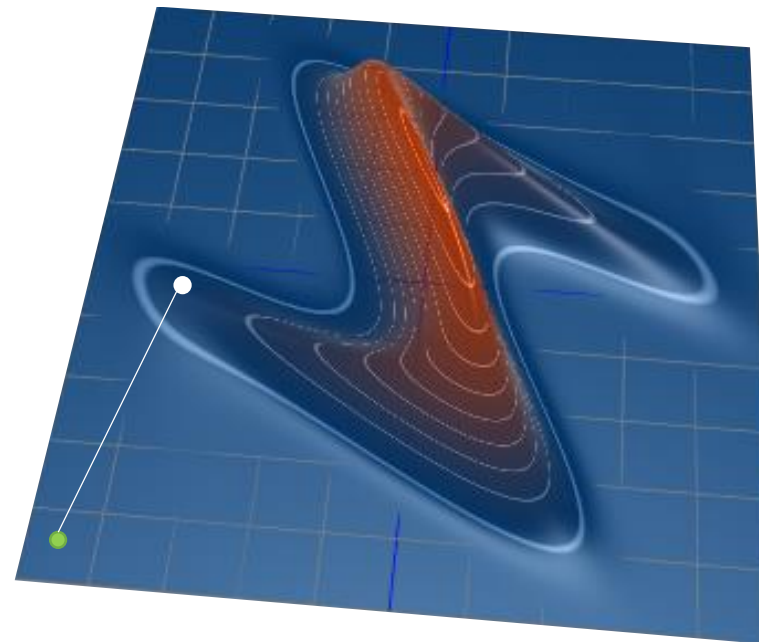
Espaço paramétrico da
distribuição *a posteriori*

- *Random walk MCMC*



Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*

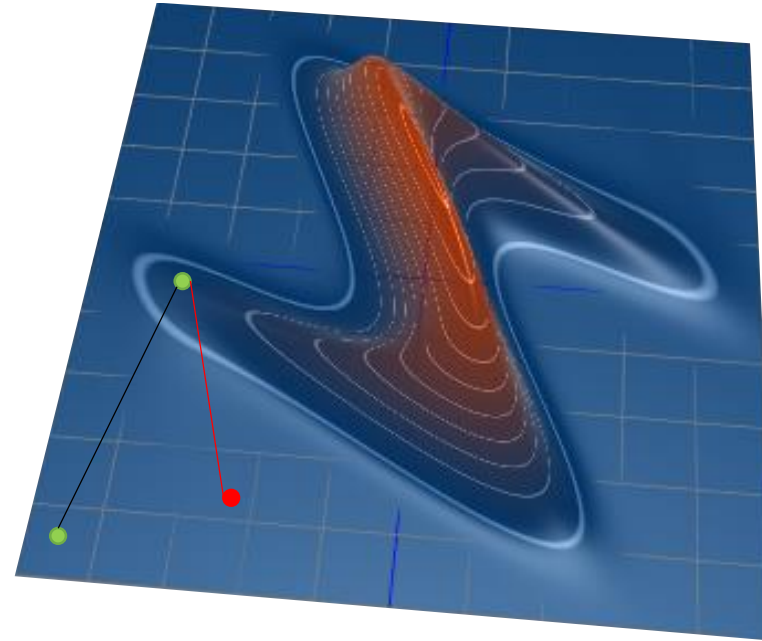


- *Random walk MCMC*

- Amostrador em avaliação;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*



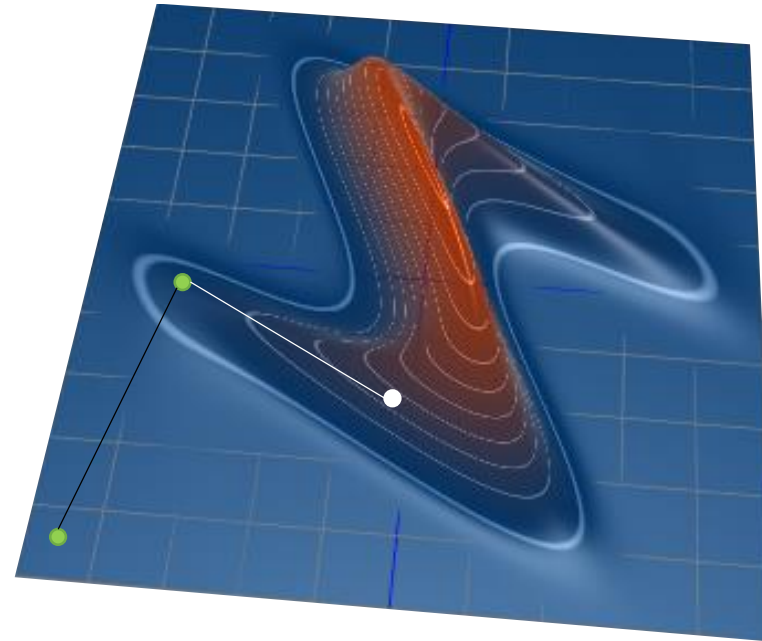
● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*



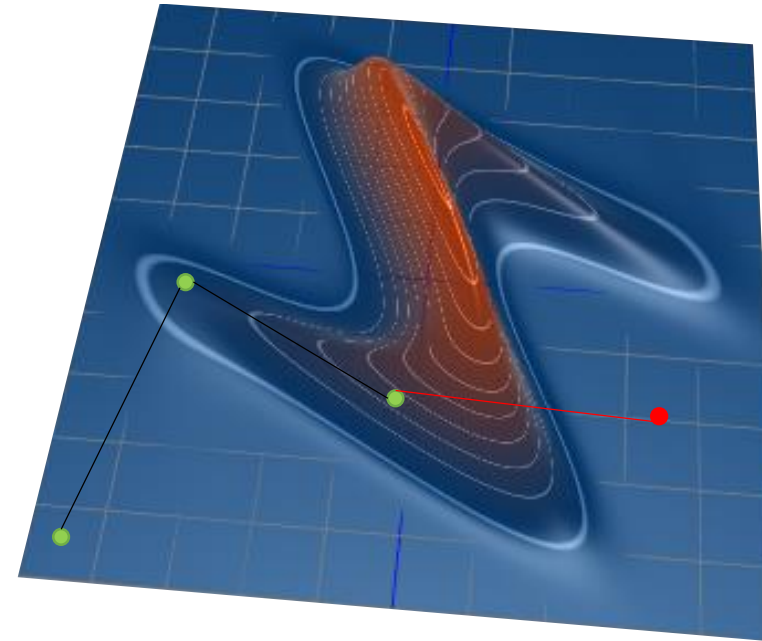
● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*



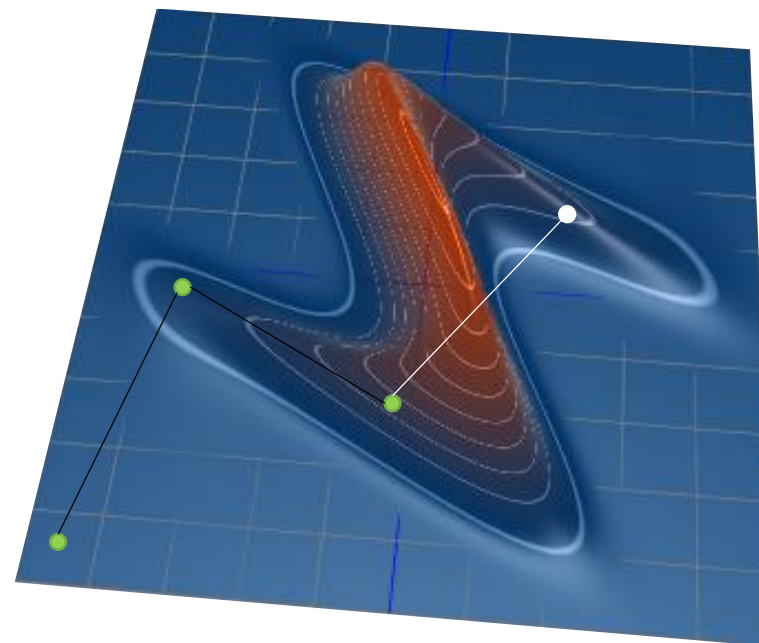
● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

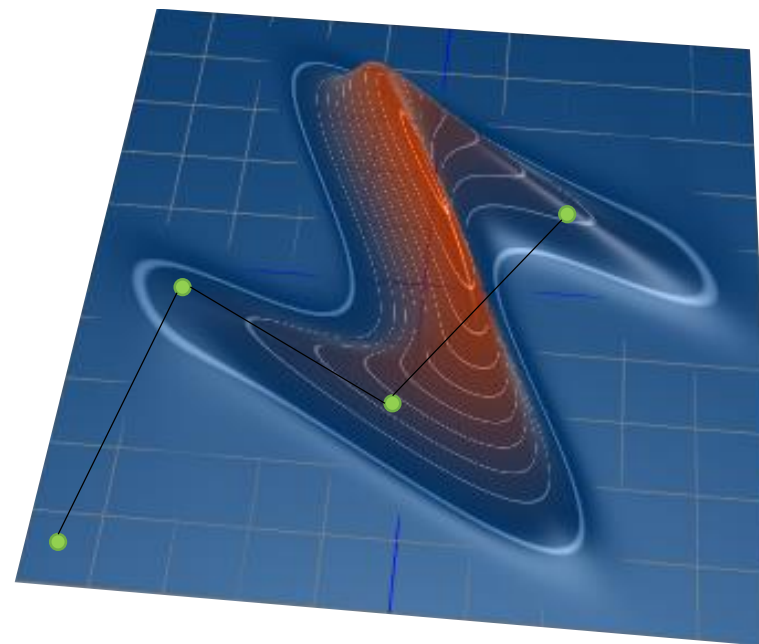
● Amostrador rejeitado;

• Método a amostragem:

- Gibbs
- Metropolis-Hasting
- Hamiltoniano MC
- NUTS

Etapa de **aceitação** ou
rejeição de cada iteração do
passeio aleatório.

Espaço paramétrico da
distribuição *a posteriori*



● *Random walk MCMC*

■ Amostrador em avaliação;

● Amostrador rejeitado;

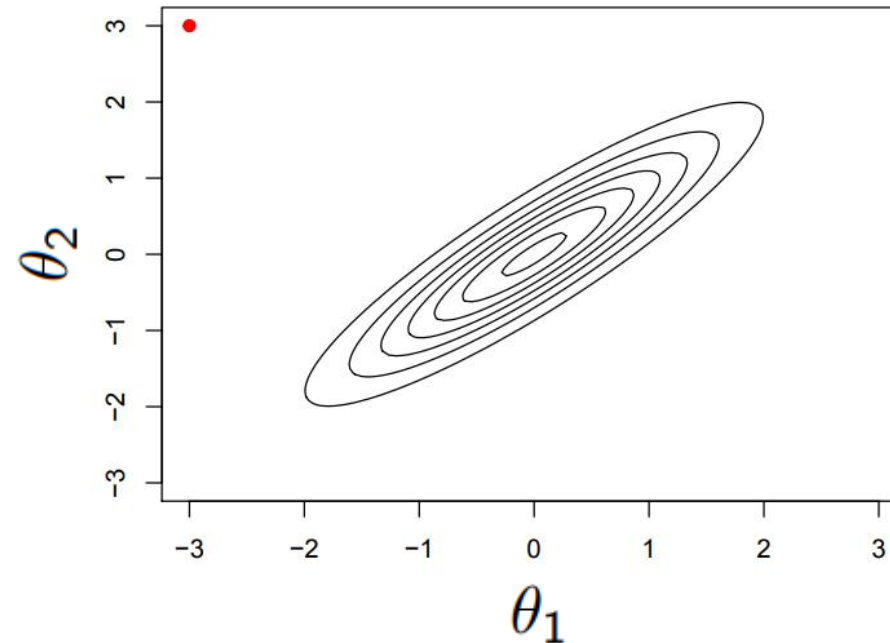
• Método a amostragem:

- Gibbs
- Metropolis-Hasting

atenção

- Amostragem *a posteriori* é marginal por parâmetro;
- Lenta convergência;
(Principalmente espaços
paramétricos de alta
dimensão).

Espaço paramétrico da
distribuição *a posteriori*



• Método a amostragem:

- Gibbs
- Metropolis-Hasting

atenção

- Amostragem *a posteriori* é marginal por parâmetro;

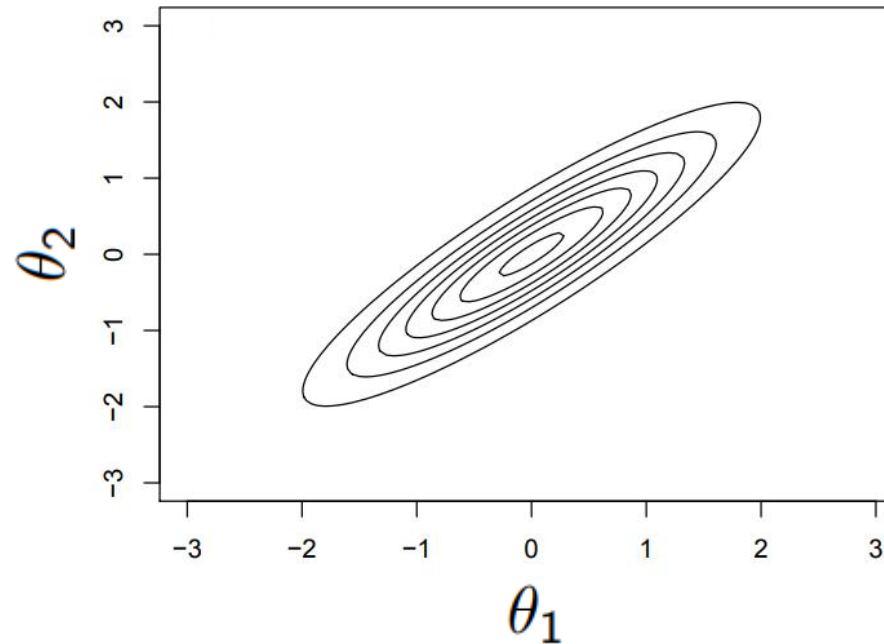
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



- Amostragem *a posteriori* é marginal por parâmetro;

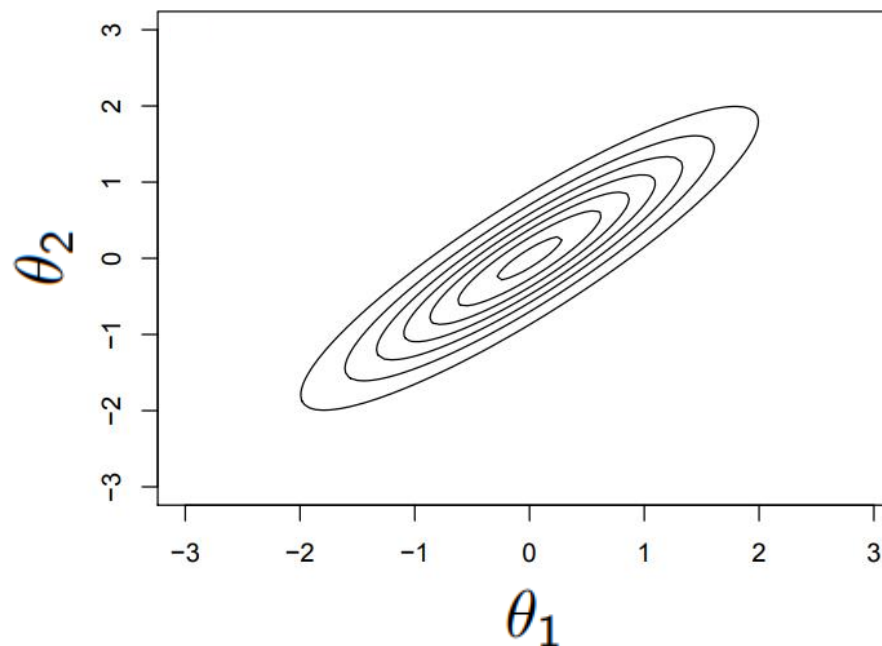
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



➤ Algoritmo:

I) $\left(\theta_1^{(t-1)}, \theta_2^{(t-1)} \right)$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

$\theta_1^{(t)} \sim p \left(\theta_1 \mid \theta_2^{(t-1)}, y \right)$ Passo 1.

$\theta_2^{(t)} \sim p \left(\theta_2 \mid \theta_1^{(t)}, y \right)$ Passo 2.

$\left(\theta_1^{(t)}, \theta_2^{(t)} \right)$ Valores t iteração;

- Amostragem *a posteriori* é marginal por parâmetro;

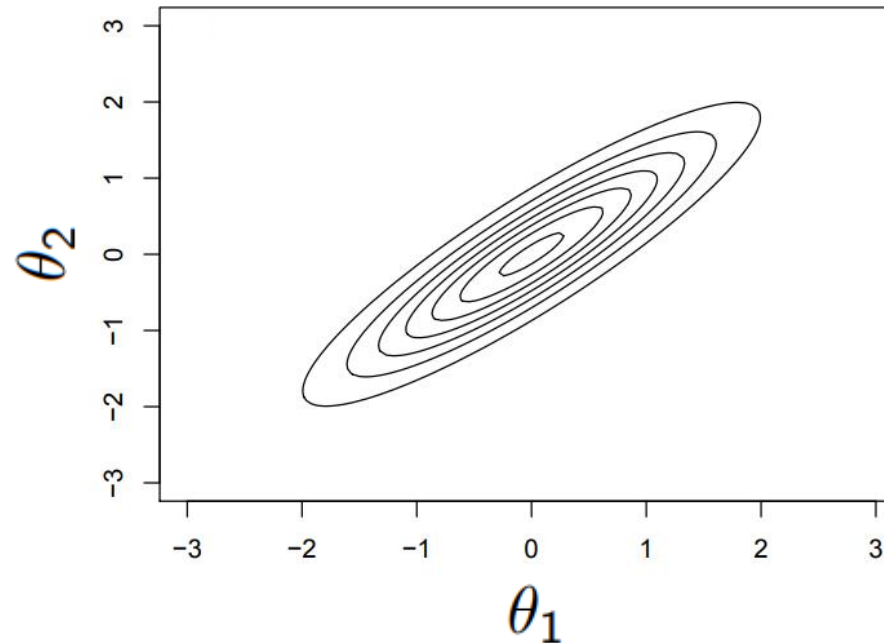
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



➤ Algoritmo:

I) $(\theta_1^{(t-1)}, \theta_2^{(t-1)})$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

⇒ $\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, y)$ Passo 1.

$\theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t)}, y)$ Passo 2.

$(\theta_1^{(t)}, \theta_2^{(t)})$ Valores t iteração;

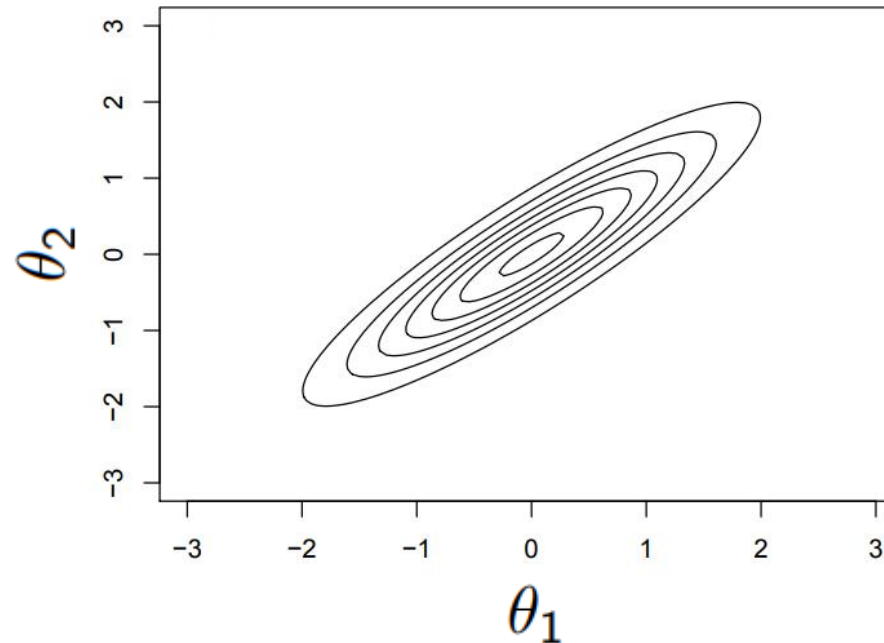
- Amostragem *a posteriori* é marginal por parâmetro;

Distribuição Alvo $p(\theta|y)$ $\xrightarrow{\text{com}}$ $\theta = (\theta_1, \theta_2)$ $\xrightarrow{\text{amostrar}}$ $\begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



➤ Algoritmo:

I) $(\theta_1^{(t-1)}, \theta_2^{(t-1)})$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

$\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, y)$ Passo 1.

$\Rightarrow \theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t)}, y)$ Passo 2.

$(\theta_1^{(t)}, \theta_2^{(t)})$ Valores t iteração;

- Amostragem *a posteriori* é marginal por parâmetro;

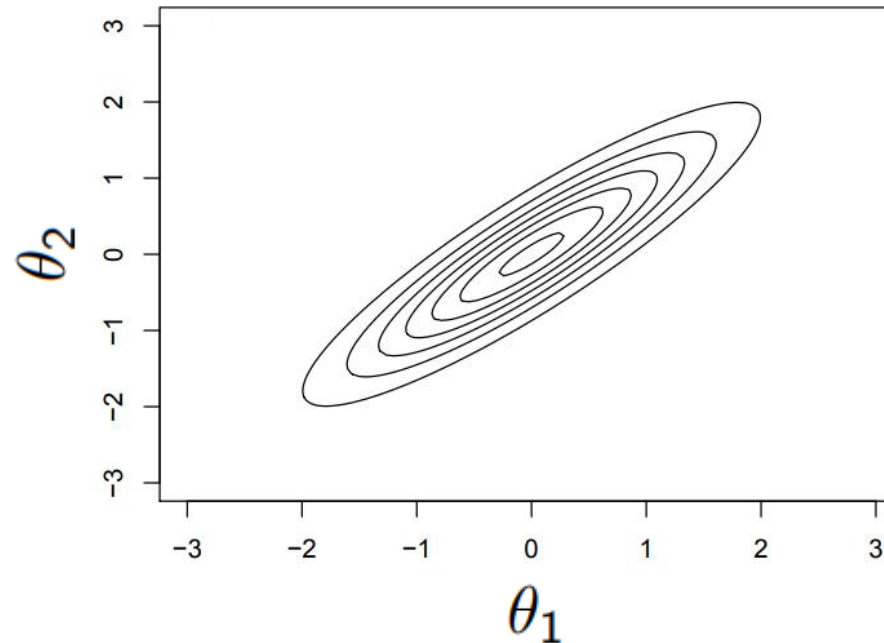
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*

➤ Algoritmo:

I) $(\theta_1^{(t-1)}, \theta_2^{(t-1)})$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

$\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, y)$ Passo 1.

$\theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t)}, y)$ Passo 2.

$(\theta_1^{(t)}, \theta_2^{(t)})$ Valores t iteração;

- Amostragem *a posteriori* é marginal por parâmetro;

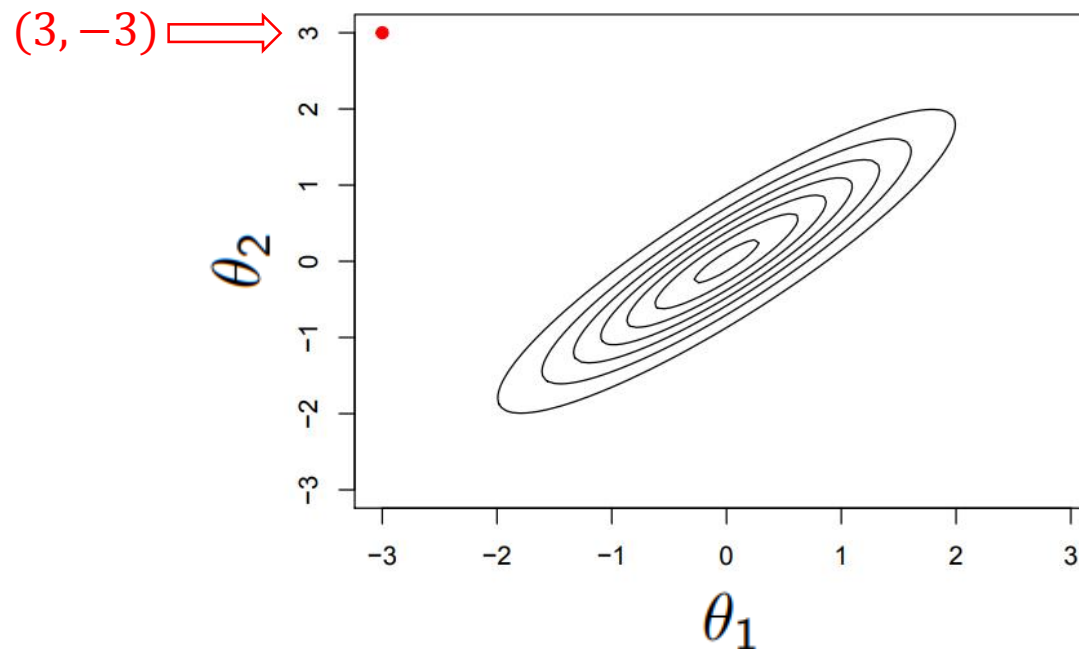
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

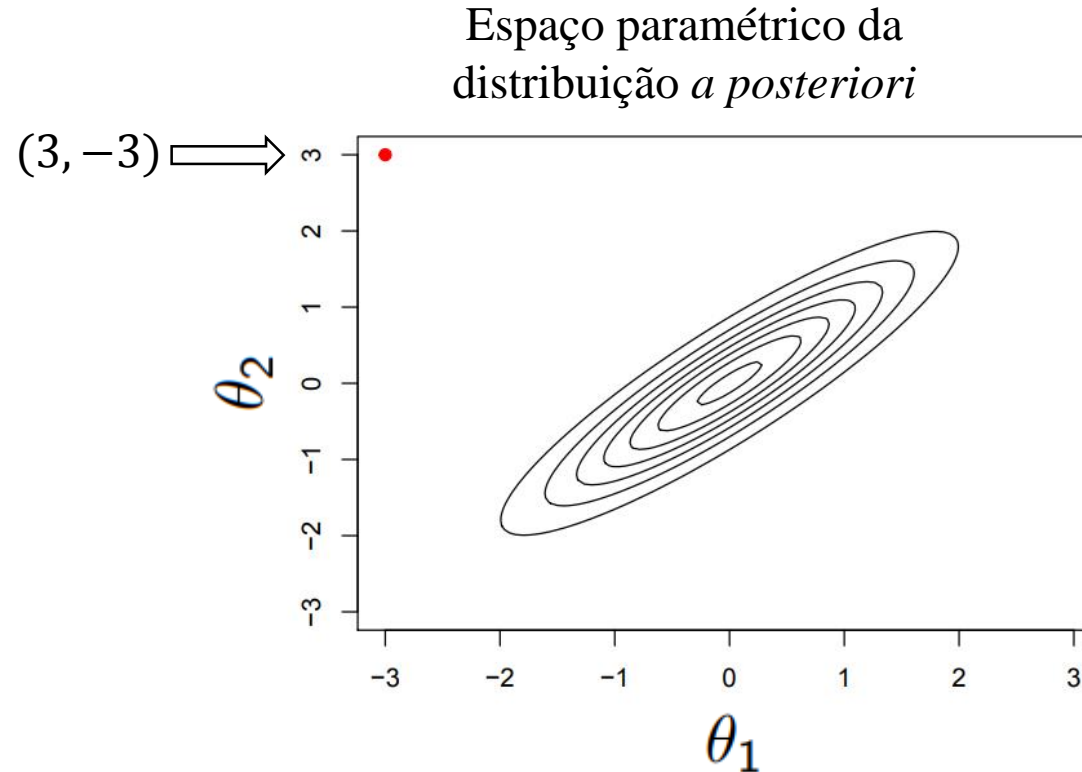
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y) \quad \text{Passo 1.}$$

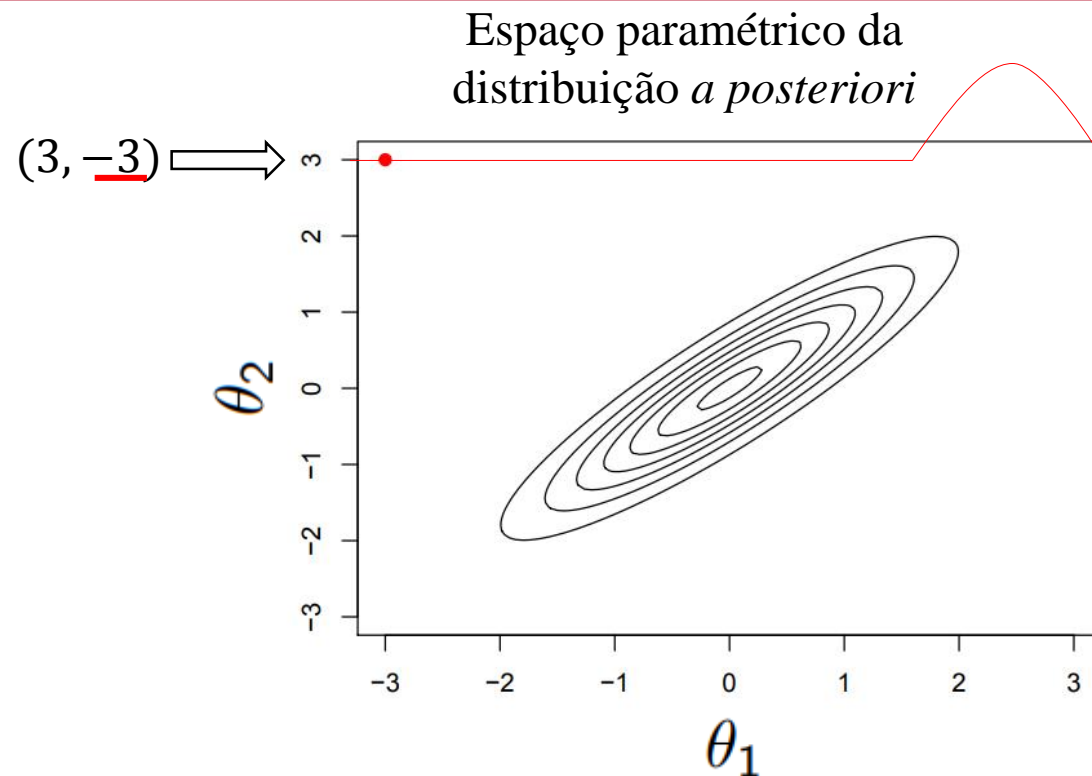
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração: **Fixar**
 $\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y)$ Passo 1.

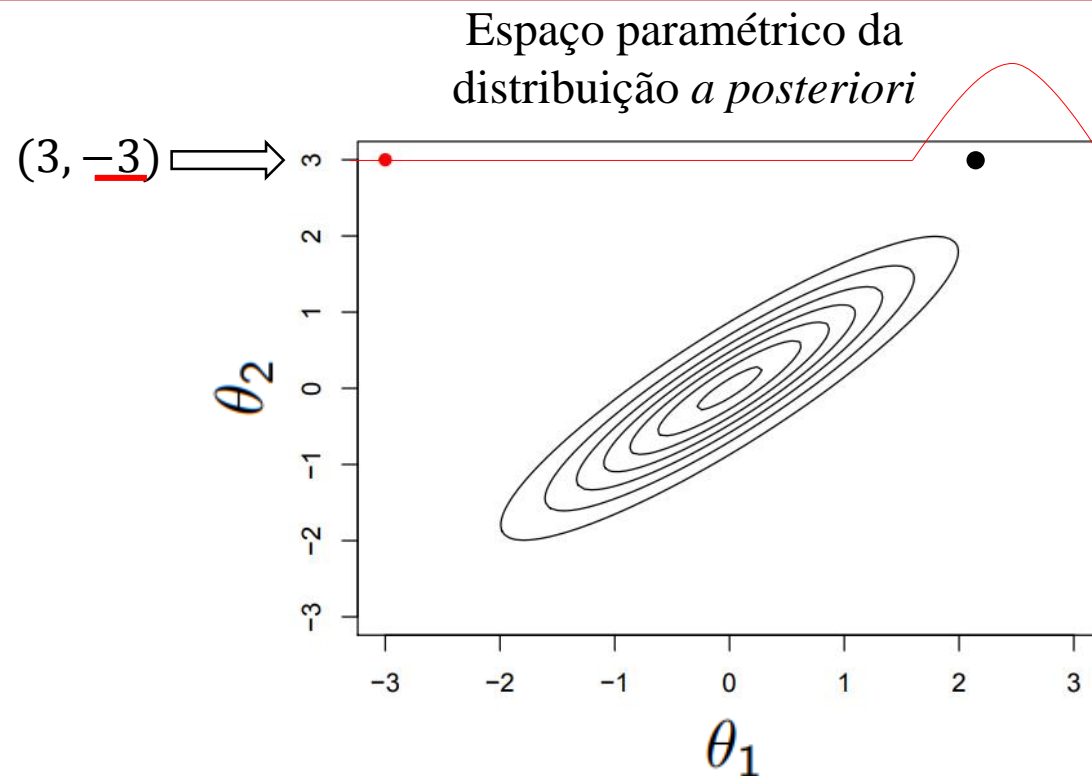
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$\theta_1^{(1)} \sim p(\theta_1 | \text{Fixar } \theta_2^{(0)}, y)$ Passo 1.

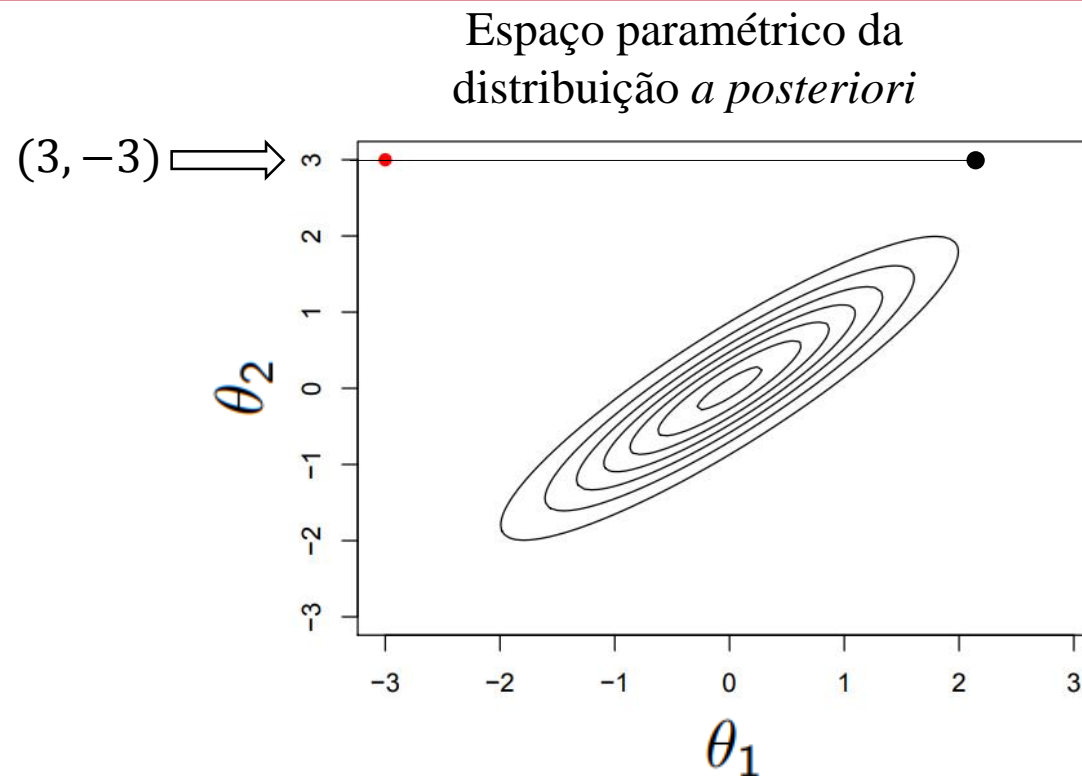
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y)$ Passo 1.

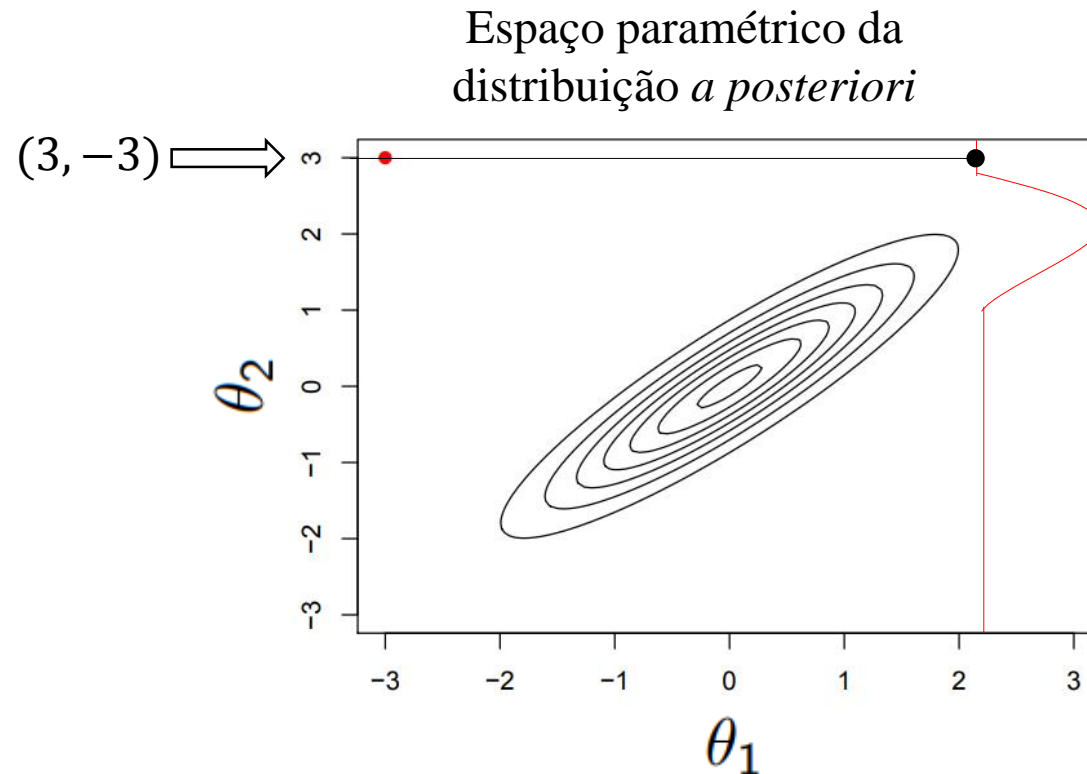
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y) \quad \text{Passo 1.}$$

$$\theta_2^{(1)} \sim p(\theta_2 | \theta_1^{(1)}, y) \quad \text{Passo 2.}$$

Fixar

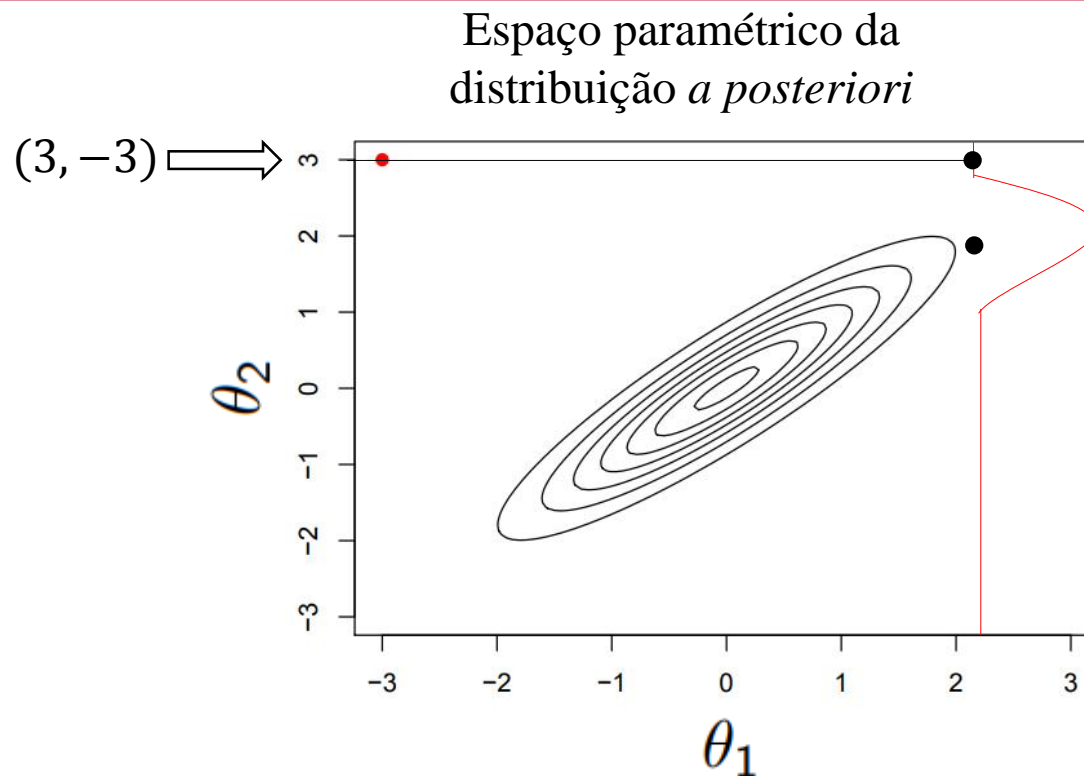
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y)$ Passo 1.

$\theta_2^{(1)} \sim p(\theta_2 | \theta_1^{(1)}, y)$ Passo 2.
Fixar

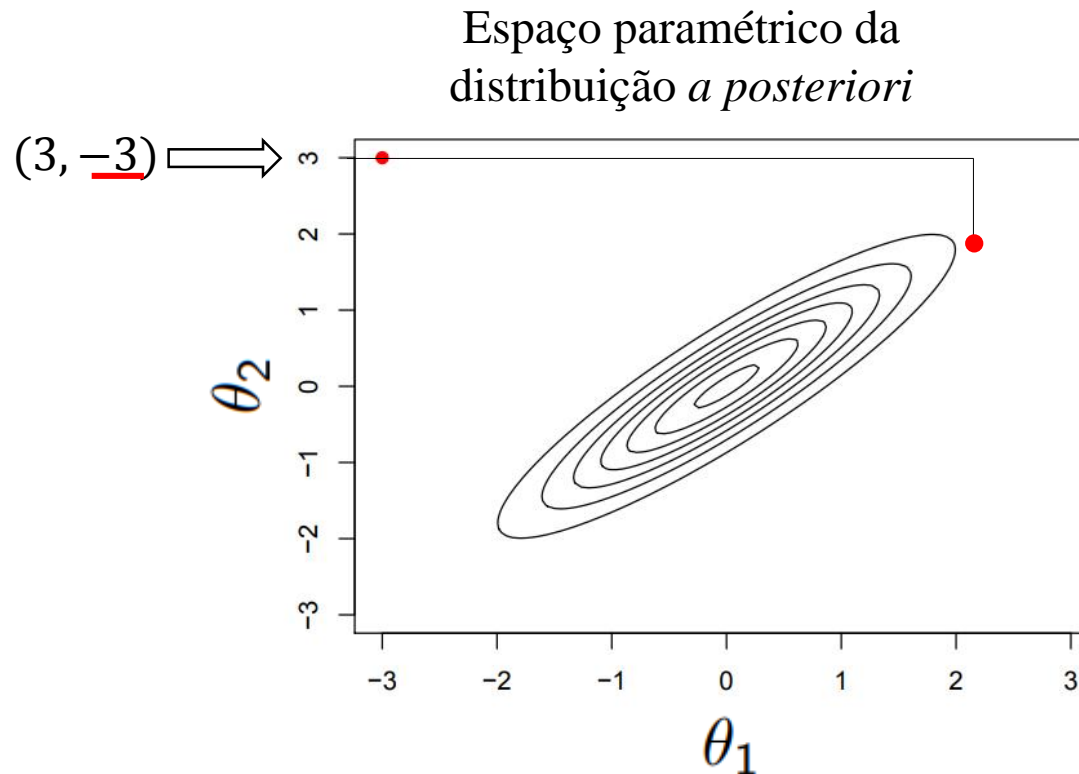
- Amostragem *a posteriori* é marginal por parâmetro;

$p(\theta|y)$ $\xrightarrow{\text{com}}$ $\theta = (\theta_1, \theta_2)$ $\xrightarrow{\text{amostrar}}$ $\begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y)$ Passo 1.

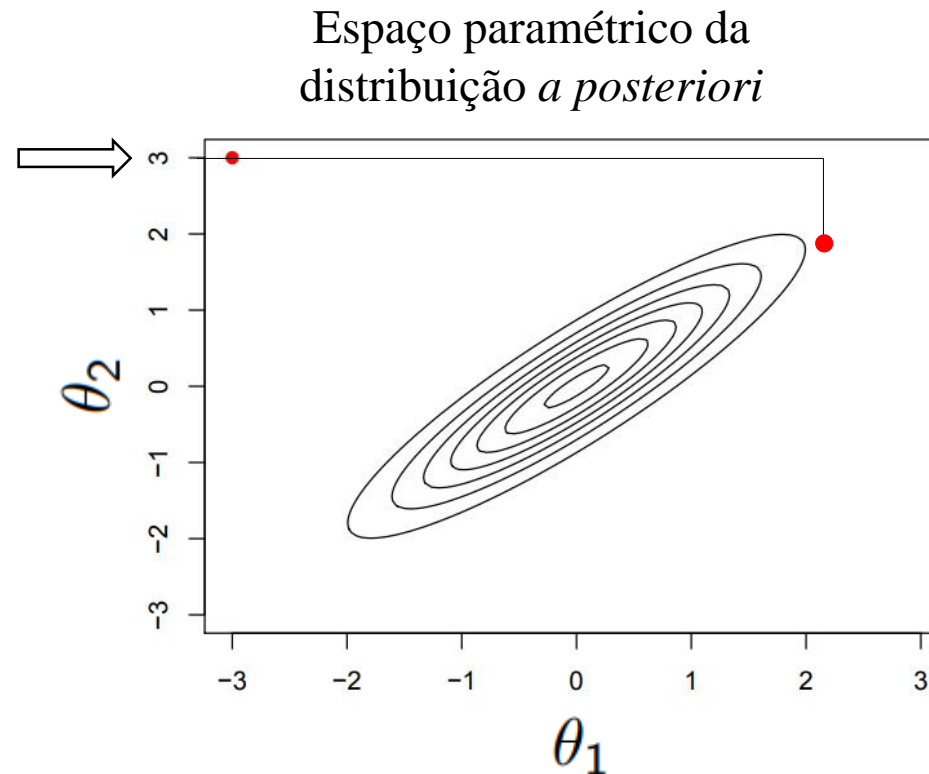
$\theta_2^{(1)} \sim p(\theta_2 | \theta_1^{(1)}, y)$ Passo 2.

- Amostragem *a posteriori* é marginal por parâmetro;

Distribuição Alvo $p(\theta|y)$ $\xrightarrow{\text{com}}$ $\theta = (\theta_1, \theta_2)$ $\xrightarrow{\text{amostrar}}$ $\begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(0)}, \theta_2^{(0)})$ Valor inicial (aleatório);

II) Iteração:

$\theta_1^{(1)} \sim p(\theta_1 | \theta_2^{(0)}, y)$ Passo 1.

$\theta_2^{(1)} \sim p(\theta_2 | \theta_1^{(1)}, y)$ Passo 2.

$(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

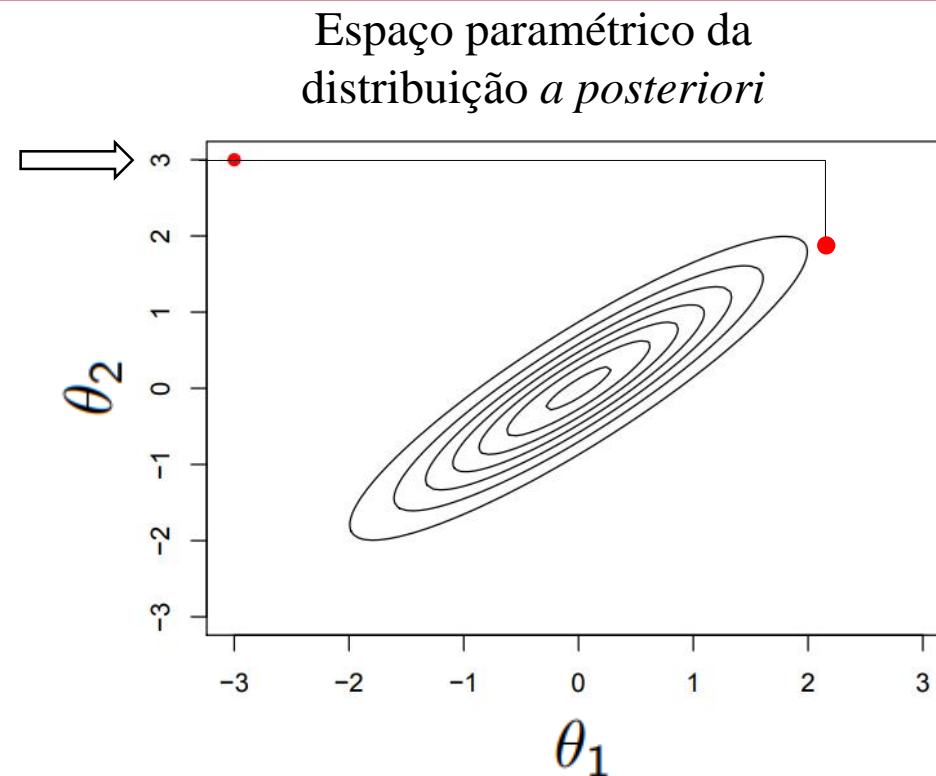
- Amostragem *a posteriori* é marginal por parâmetro;

Distribuição Alvo $p(\theta|y)$ $\xrightarrow{\text{com}}$ $\theta = (\theta_1, \theta_2)$ $\xrightarrow{\text{amostrar}}$ $\begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$

➤ Método a amostragem:

- Gibbs;

➤ Algoritmo:



I)

II) Iteração:

$(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

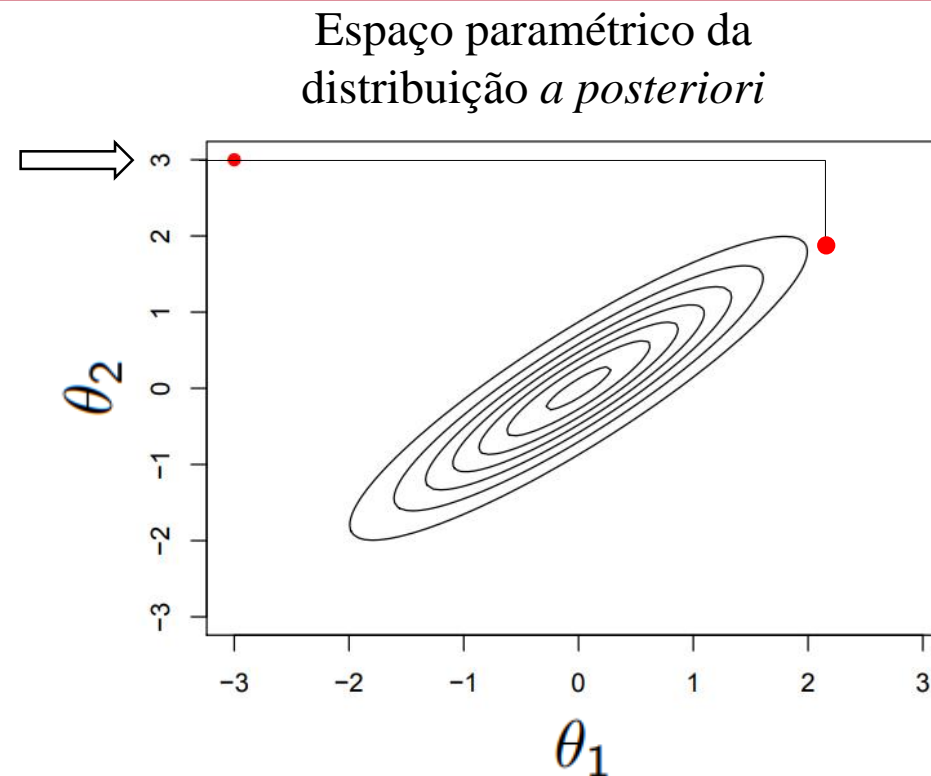
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

Fixar

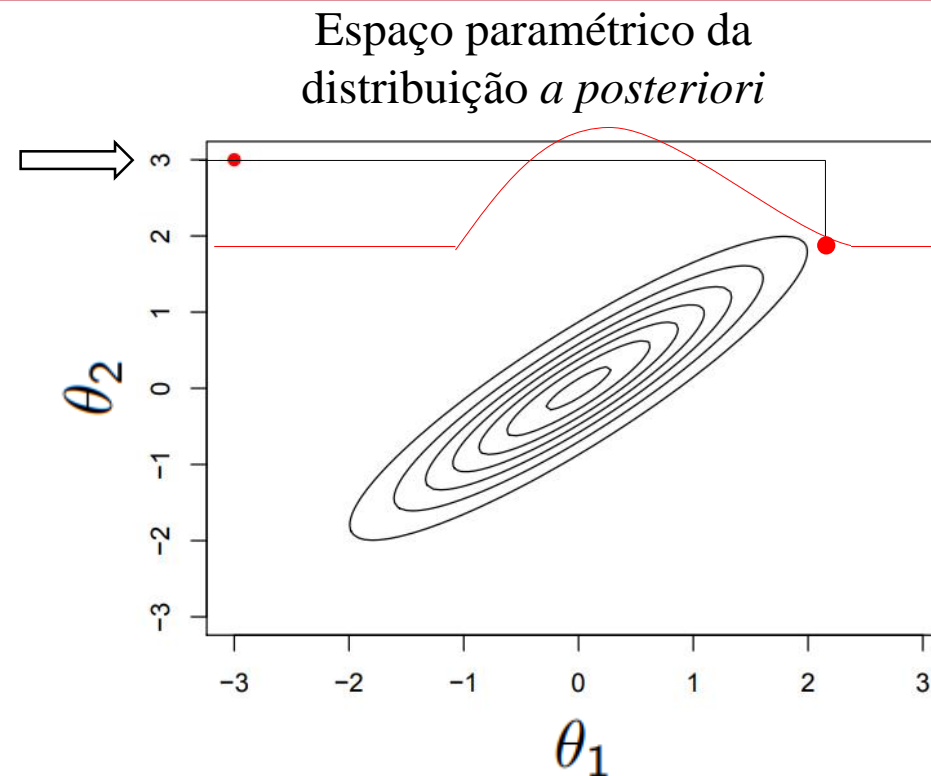
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

Fixar

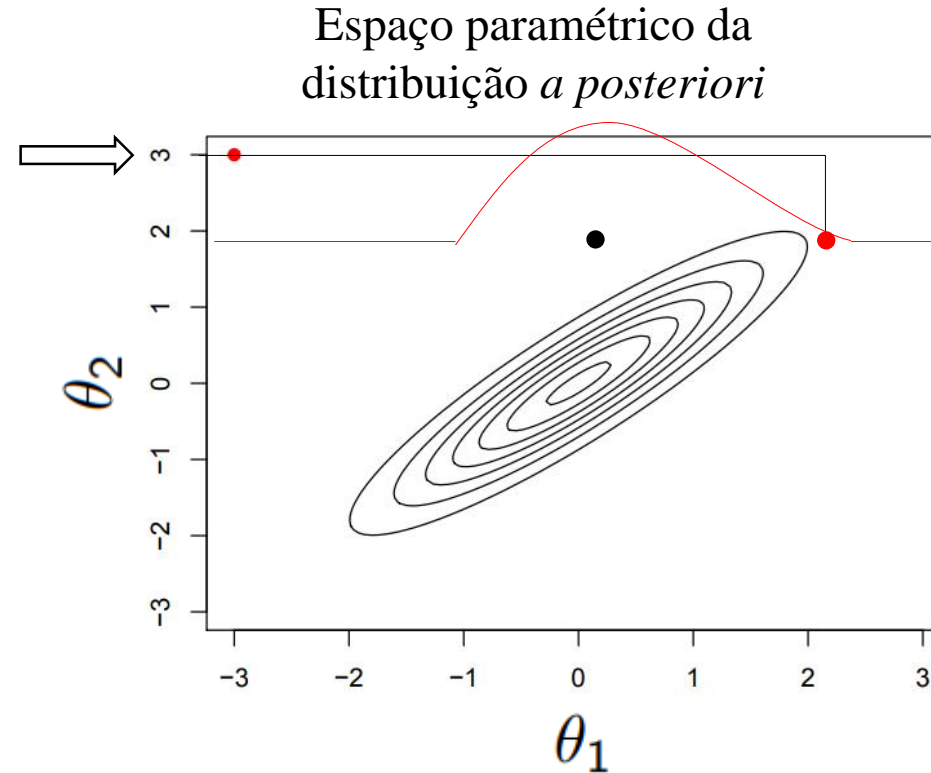
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y)$ Passo 1.

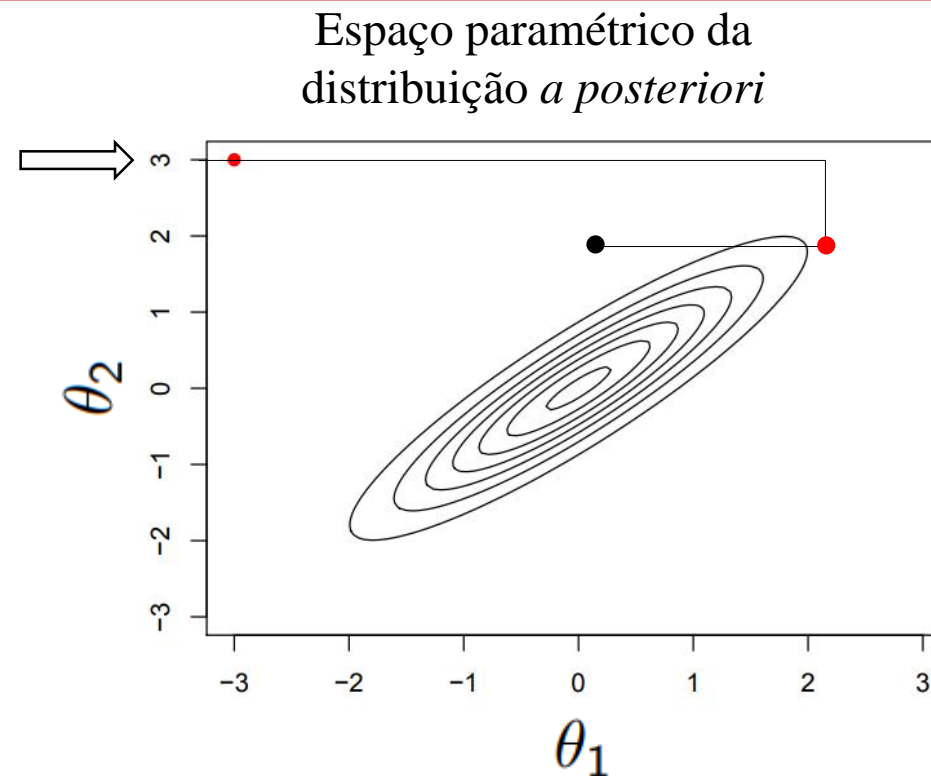
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y)$ Passo 1.

- Amostragem *a posteriori* é marginal por parâmetro;

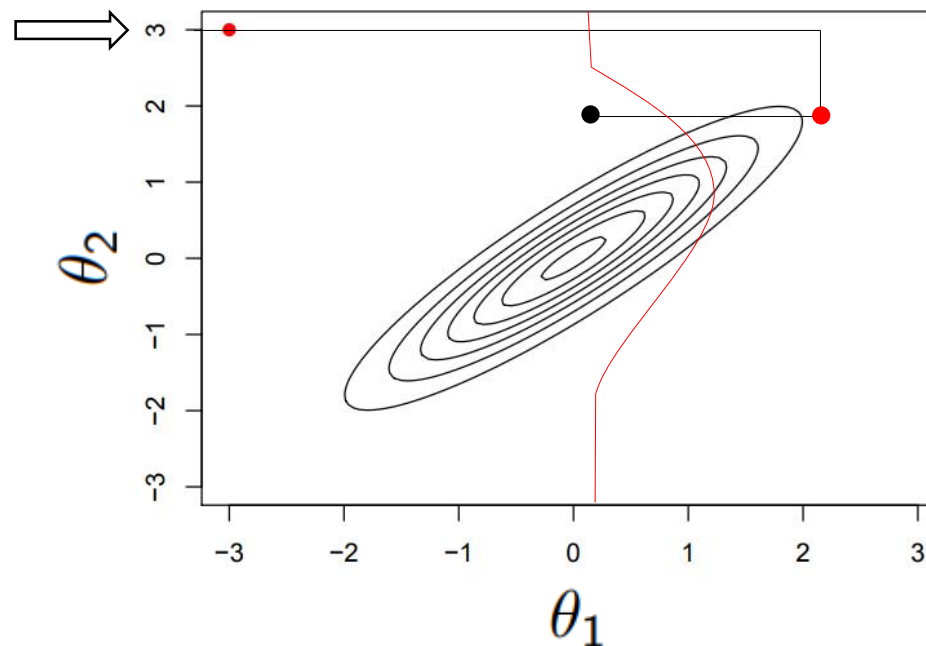
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

Espaço paramétrico da
distribuição *a posteriori*



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

$$\theta_2^{(2)} \sim p(\theta_2 | \theta_1^{(2)}, y) \quad \text{Passo 2.}$$

Fixar

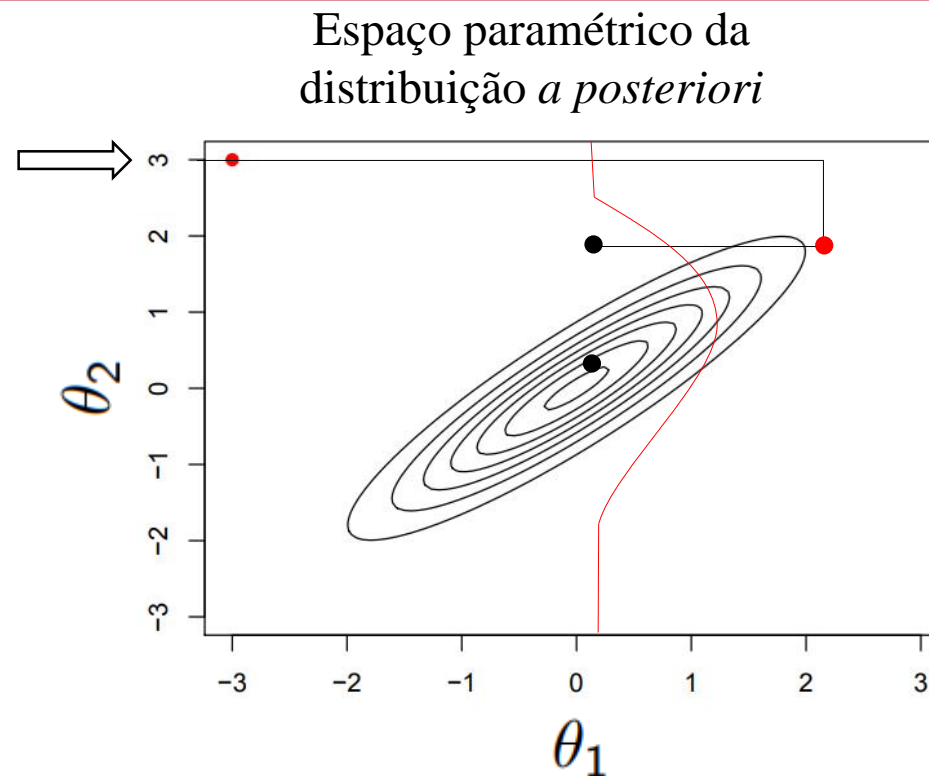
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

$$\theta_2^{(2)} \sim p(\theta_2 | \underline{\theta_1^{(2)}}, y) \quad \text{Passo 2.}$$

Fixar

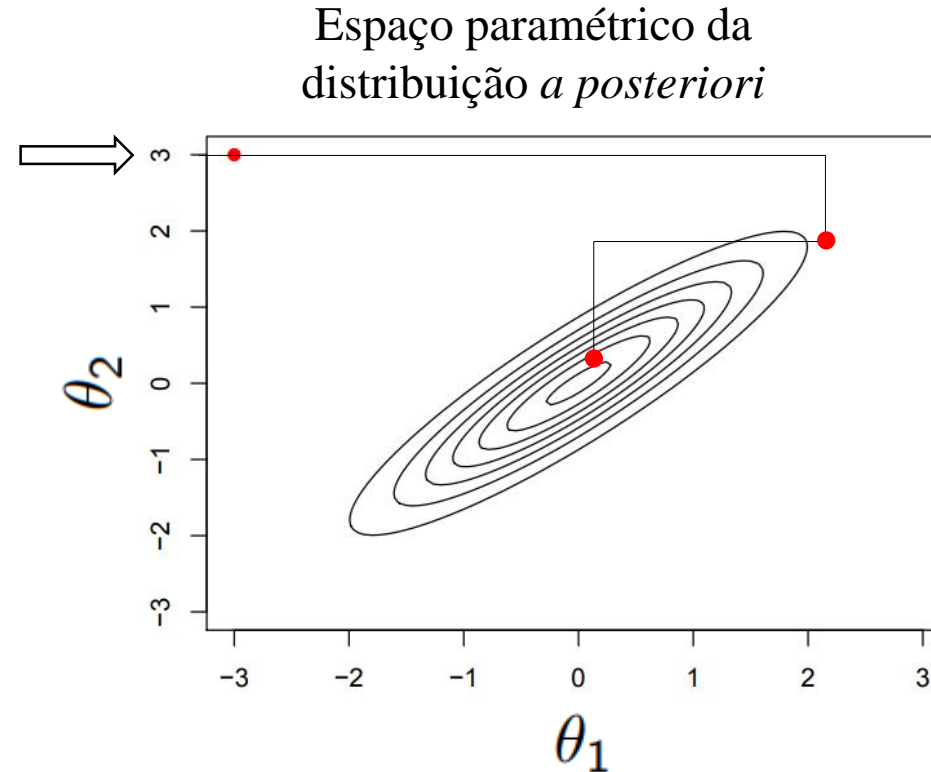
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

$$\theta_2^{(2)} \sim p(\theta_2 | \theta_1^{(2)}, y) \quad \text{Passo 2.}$$

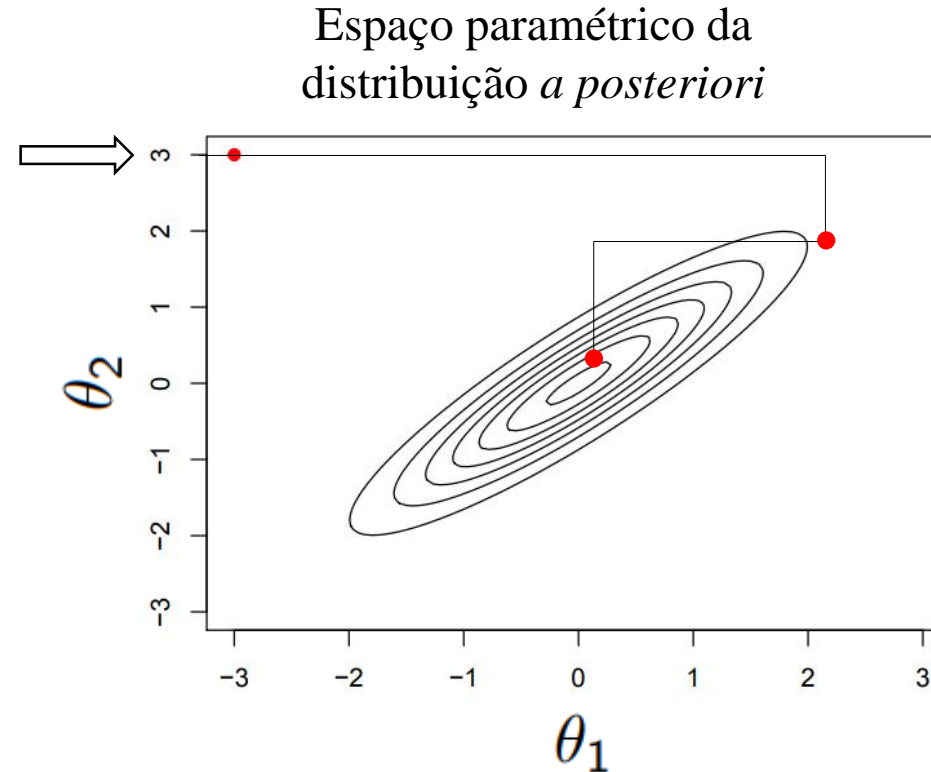
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(1)}, \theta_2^{(1)})$ Valores 1ª iteração;

II) Iteração:

$$\theta_1^{(2)} \sim p(\theta_1 | \theta_2^{(1)}, y) \quad \text{Passo 1.}$$

$$\theta_2^{(2)} \sim p(\theta_2 | \theta_1^{(2)}, y) \quad \text{Passo 2.}$$

$(\theta_1^{(2)}, \theta_2^{(2)})$ Valores 2ª iteração;

- Amostragem *a posteriori* é marginal por parâmetro;

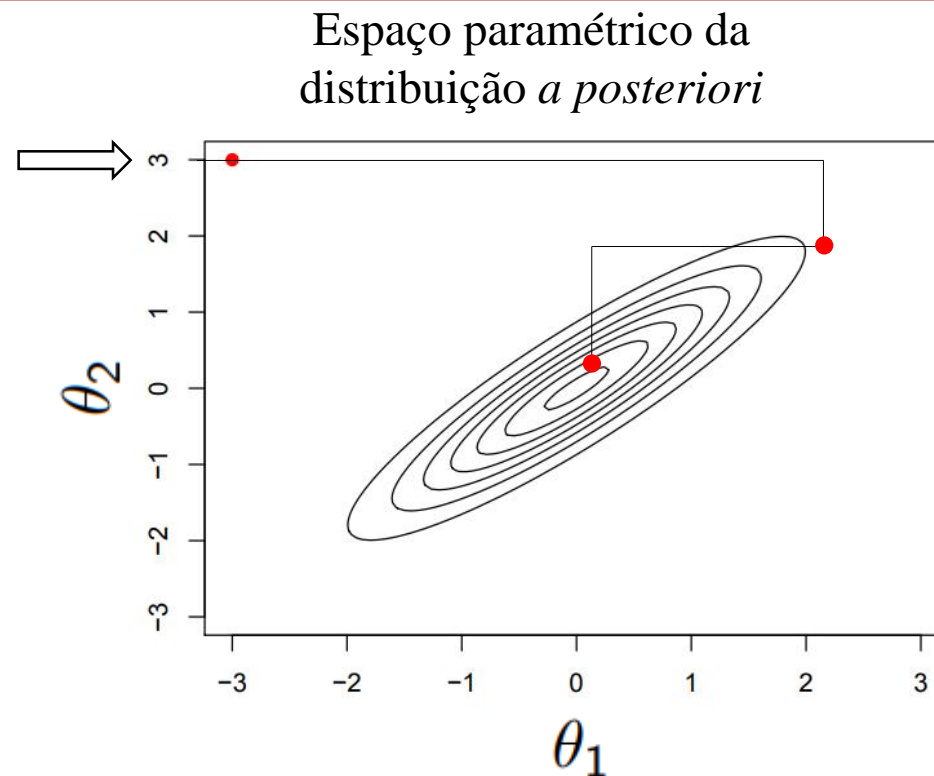
$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;

➤ Algoritmo:



I)

II) Iteração:

$(\theta_1^{(2)}, \theta_2^{(2)})$ Valores 2ª iteração;

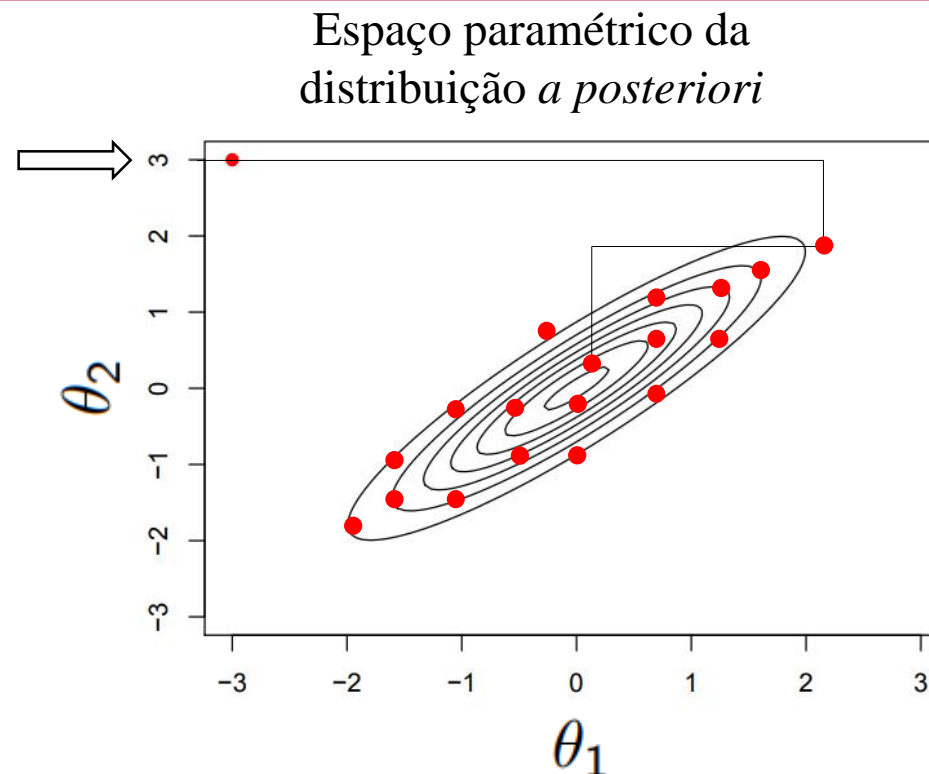
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(t-1)}, \theta_2^{(t-1)})$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

$\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, y)$ Passo 1.

$\theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t)}, y)$ Passo 2.

$(\theta_1^{(t)}, \theta_2^{(t)})$ Valores t iteração;

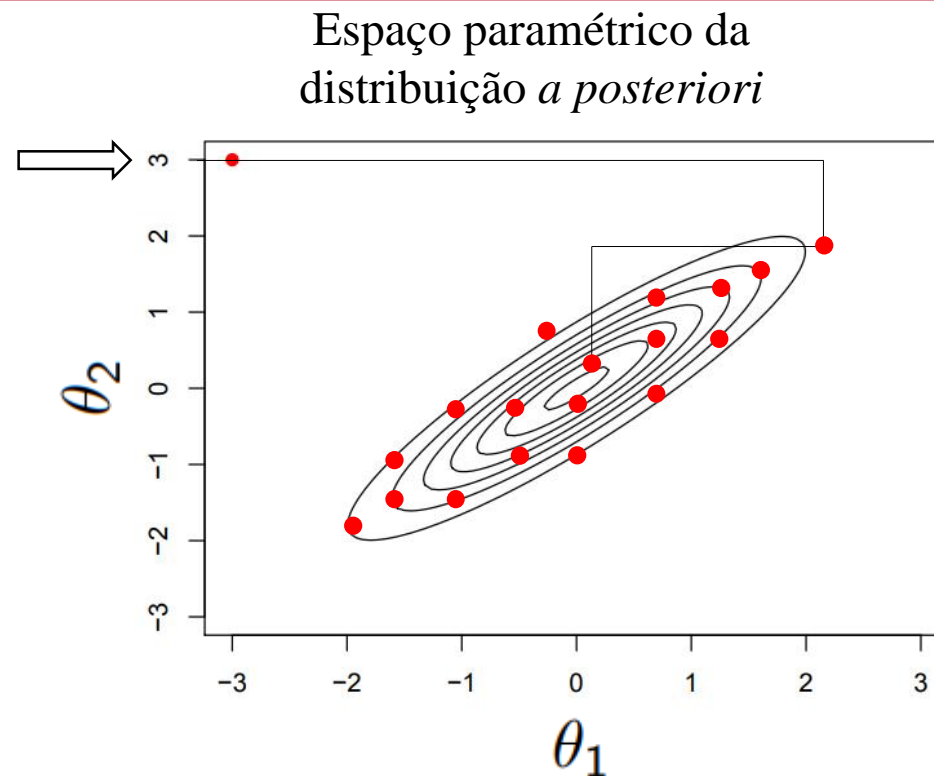
- Amostragem *a posteriori* é marginal por parâmetro;

$$p(\theta|y) \xrightarrow{\text{com}} \theta = (\theta_1, \theta_2) \xrightarrow{\text{amostrar}} \begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$$

Distribuição Alvo

➤ Método a amostragem:

- Gibbs;



➤ Algoritmo:

I) $(\theta_1^{(t-1)}, \theta_2^{(t-1)})$ $t - 1$ iteração
(Iteração anterior)

II) Iteração:

$\theta_1^{(t)} \sim p(\theta_1 | \theta_2^{(t-1)}, y)$ Passo 1.

$\theta_2^{(t)} \sim p(\theta_2 | \theta_1^{(t)}, y)$ Passo 2.

$(\theta_1^{(t)}, \theta_2^{(t)})$ Valores t iteração;

- Amostragem *a posteriori* é marginal **por parâmetro**;

Distribuição Alvo $p(\theta|y)$ $\xrightarrow{\text{com}}$ $\theta = (\theta_1, \theta_2)$ $\xrightarrow{\text{amostrar}}$ $\begin{cases} p(\theta_1|\theta_2, y) \\ p(\theta_2|\theta_1, y) \end{cases}$

- Lenta convergência;**

(Principalmente espaços paramétricos de alta dimensão).

➤ Método a amostragem:

- Gibbs;

➤ Resumo:

- Amostragem a posteriori é marginal por parâmetro;

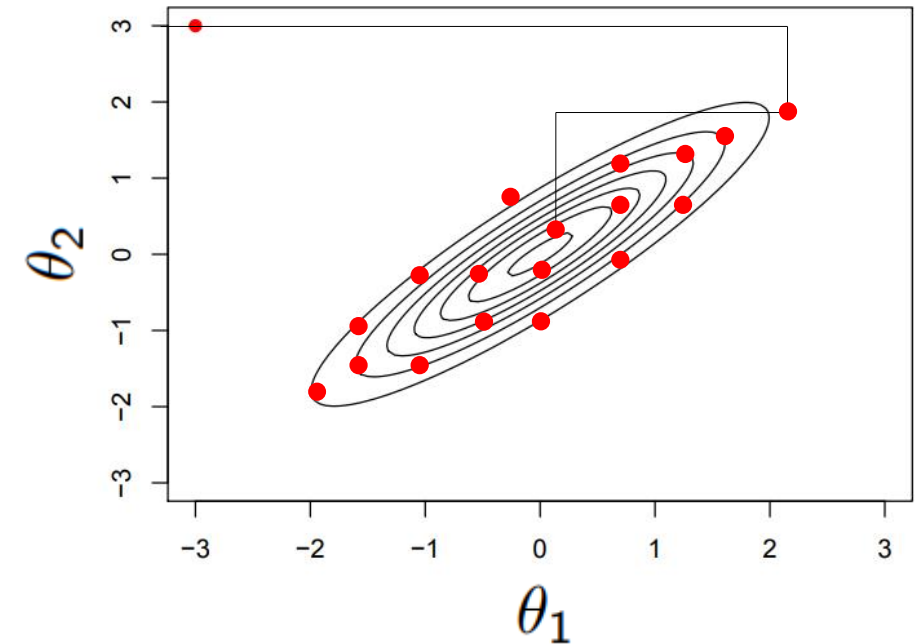
➤ Prós:

- Reduz complexas amostragem multidimensional para sequencias unidimensionais;
- Fácil implementação (linhas de código);

➤ Contras:

- Alta correlação nas amostras;
- Lenta convergência também (alta dimensão);
- Não permite programação paralela;

Espaço paramétrico da
distribuição *a posteriori*



Algoritmo de amostragem



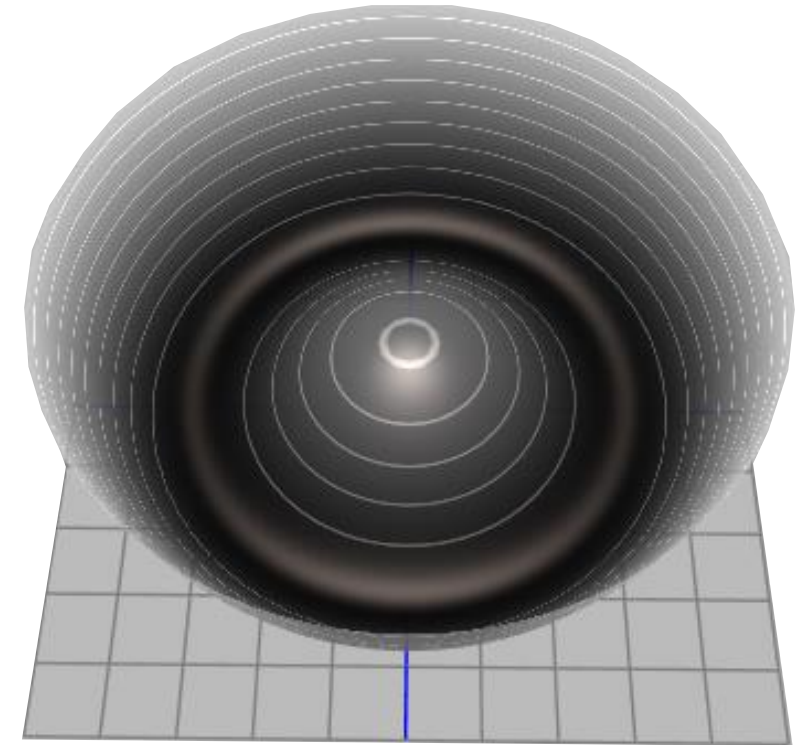
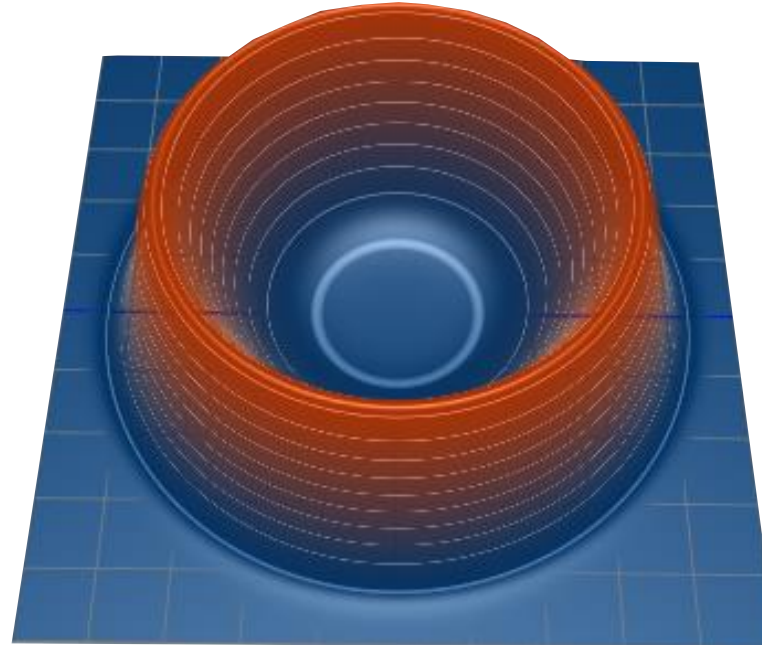
- Algoritmo de amostragem;
 - Metropolis-Hasting;
 - Gibbs;
 - **Hamiltoniano Monte Carlo;**
 - NUTS;

Algoritmo de amostragem



Radford Neal

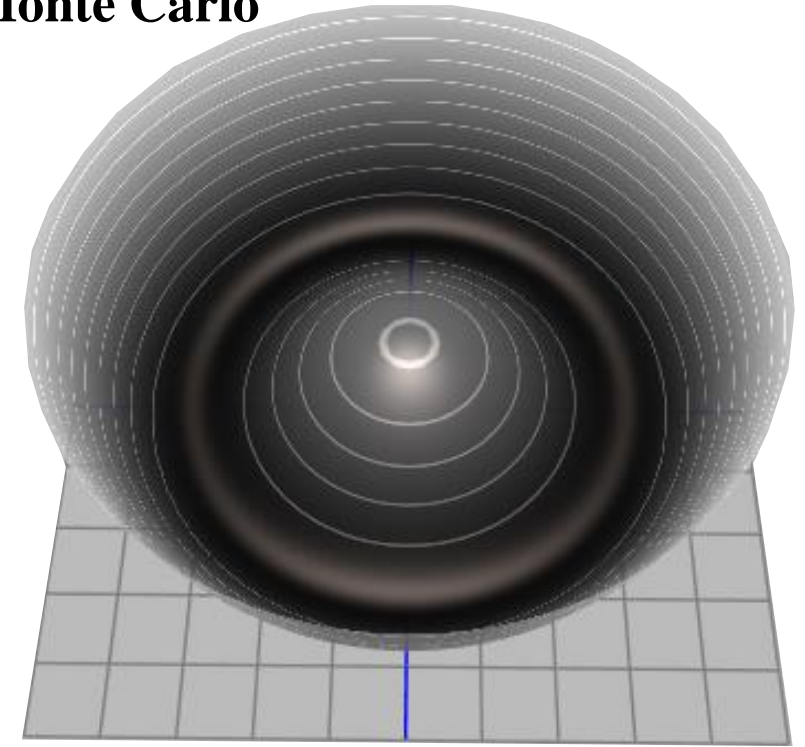
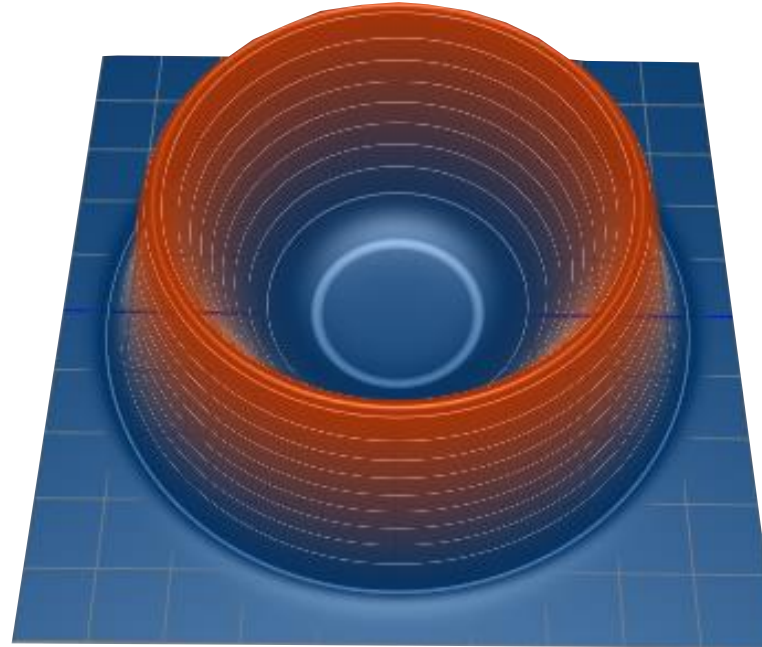
- Algoritmo de amostragem;
 - Metropolis-Hasting;
 - Gibbs;
 - **Hamiltoniano Monte Carlo;**
 - NUTS;



Inspiração na física → Estatística mecânica
Estatística física

- Amostrar por mais tempo regiões de interesse;
- Rápida convergência;

Método Hamiltoniano Monte Carlo



- Variável auxiliar *momento*;

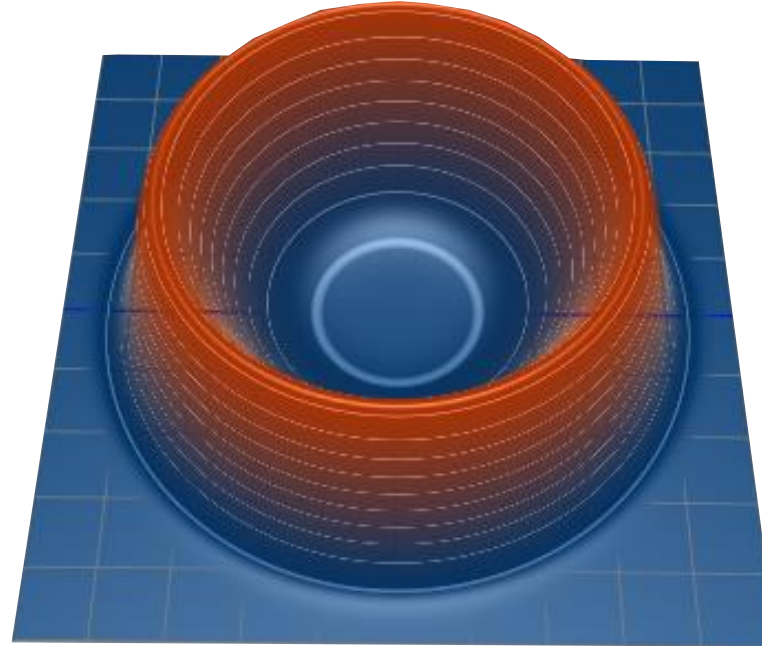
$$p(\rho, \theta) = p(\rho|\theta)p(\theta)$$

$$\rho \sim \text{MultiNormal}(0, M)$$

M = Matriz Euclidiana

M pode ser visto como uma **transformação** do **espaço paramétrico** que torna a **amostragem** mais **eficiente**.

Método Hamiltoniano Monte Carlo

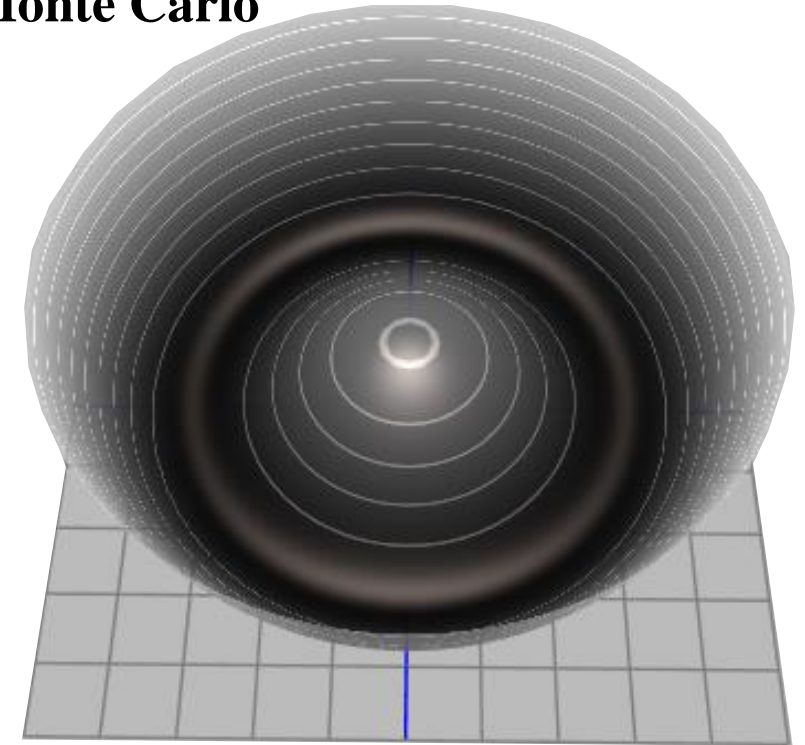


- Variável auxiliar *momento*;

$$p(\rho, \theta) = p(\rho|\theta)p(\theta)$$

$$\rho \sim \text{MultiNormal}(0, M)$$

M = Matriz Euclidiana



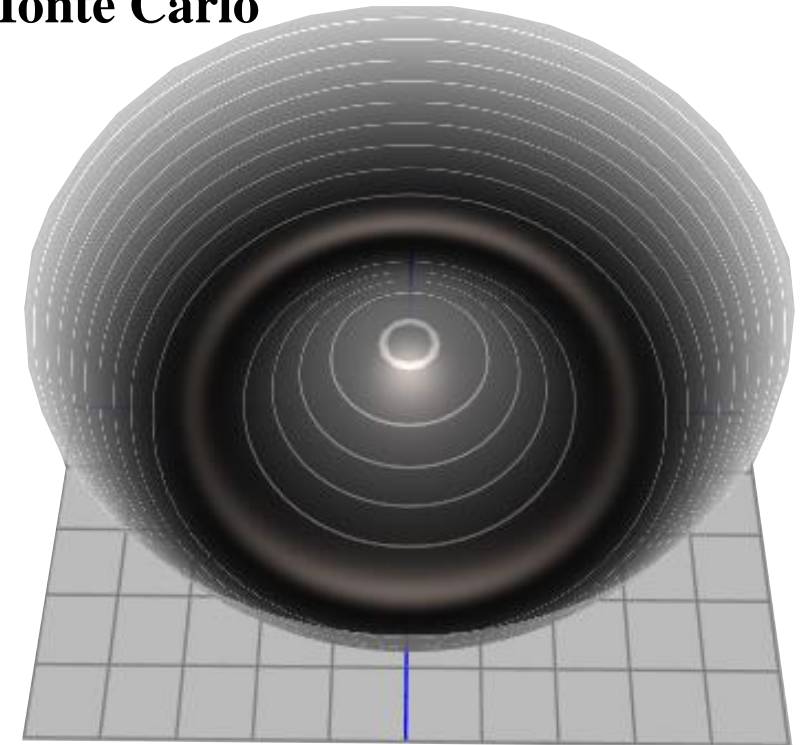
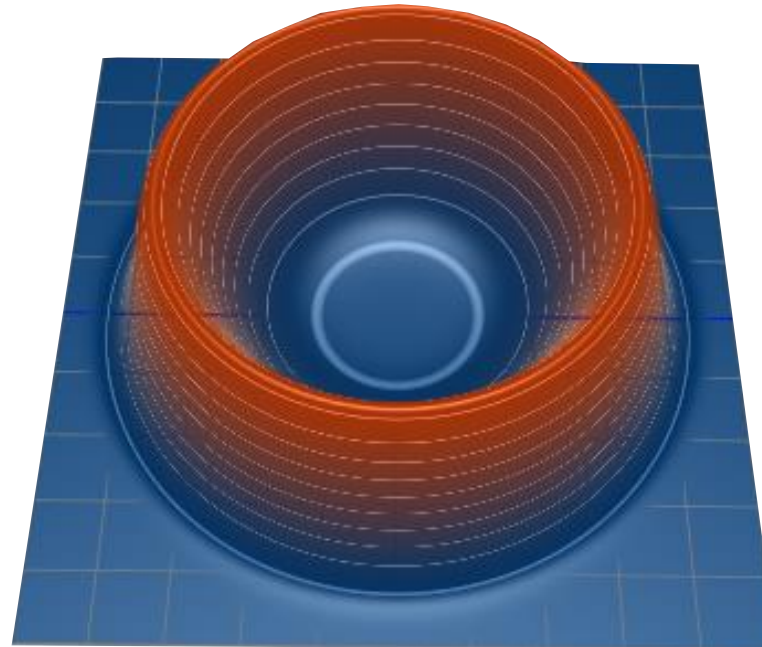
- Função Hamiltoniana;

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

Método Hamiltoniano Monte Carlo



- Variável auxiliar *momento*;

$$p(\rho, \theta) = p(\rho|\theta)p(\theta)$$

$$\rho \sim \text{MultiNormal}(0, M)$$

M = Matriz Euclidiana

- Função Hamiltoniana;

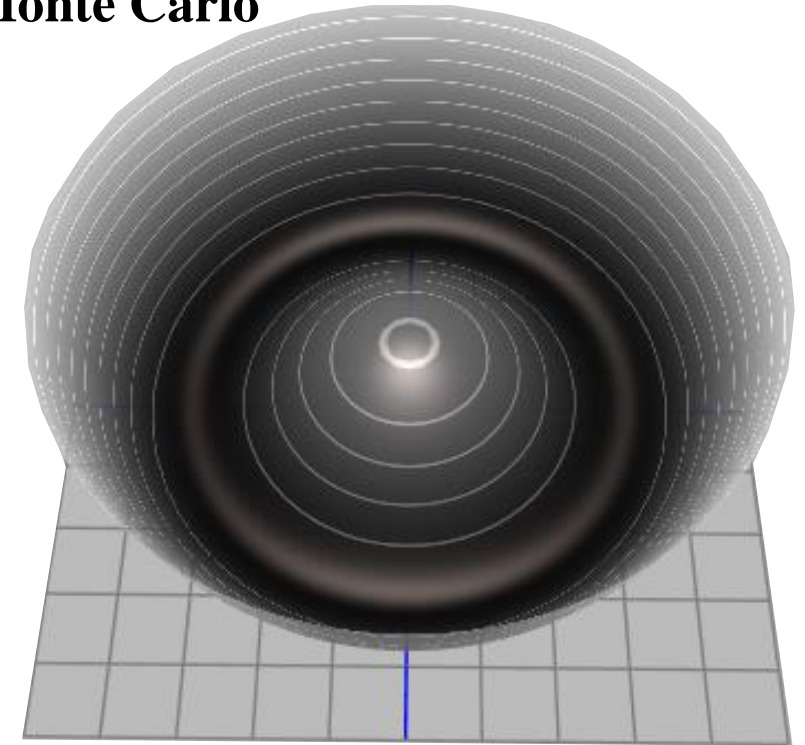
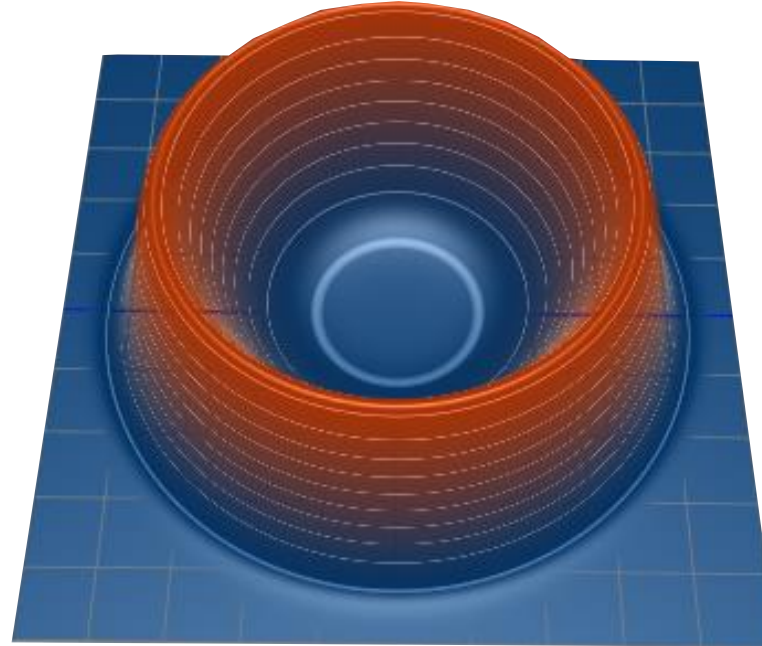
$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = \boxed{T(\rho|\theta)} + V(\theta)$$

Energia cinética ←

Método Hamiltoniano Monte Carlo



- Variável auxiliar *momento*;

$$p(\rho, \theta) = p(\rho|\theta)p(\theta)$$

$$\rho \sim \text{MultiNormal}(0, M)$$

M = Matriz Euclidiana

- Função Hamiltoniana;

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

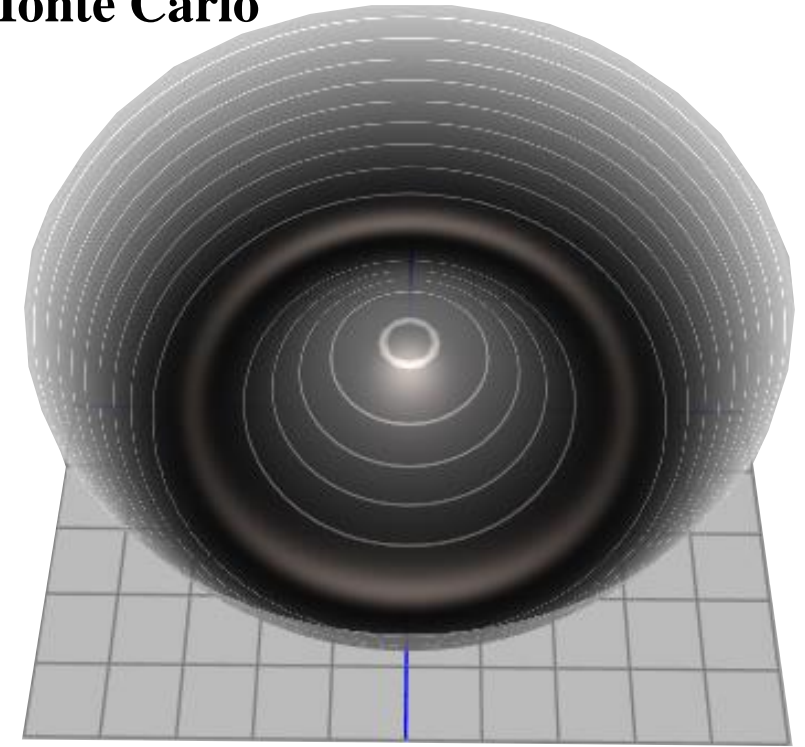
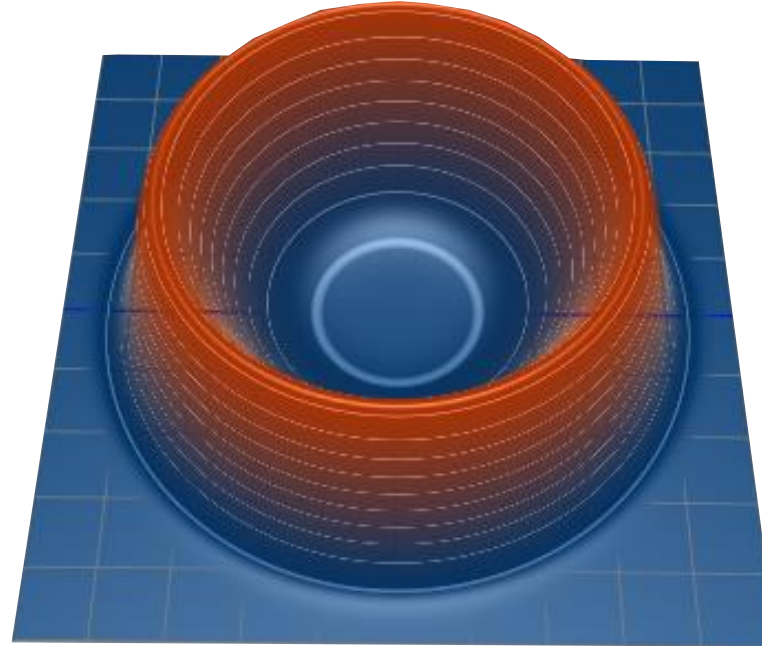
$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = \boxed{T(\rho|\theta)} + \boxed{V(\theta)}$$

Energia cinética ←

Energia Potencial ←

Método Hamiltoniano Monte Carlo



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

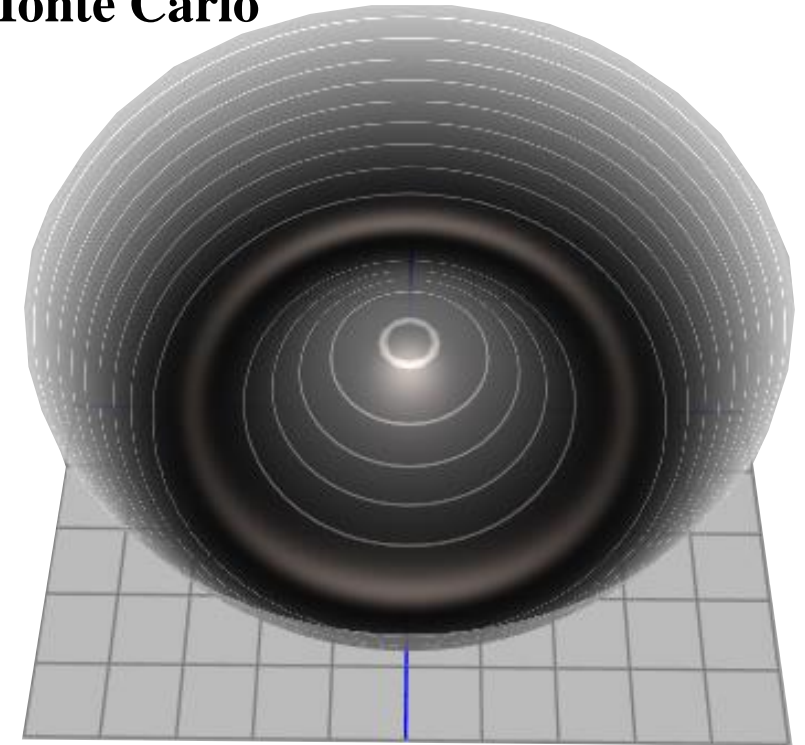
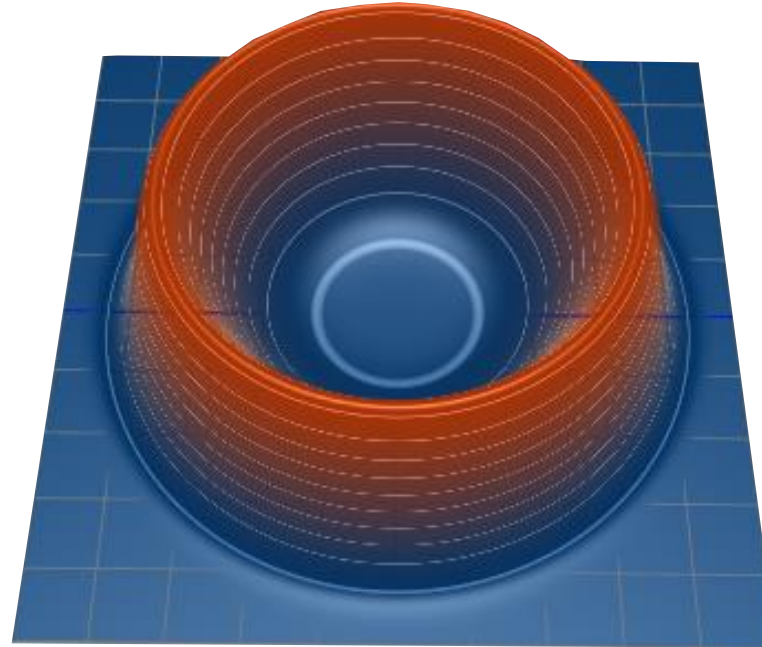
Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

Integrador *Leapfrog*

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

Método Hamiltoniano Monte Carlo



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = +\frac{\partial H}{\partial \rho} = +\frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = -\frac{\partial H}{\partial \theta} = -\frac{\partial T}{\partial \theta} - \frac{\partial V}{\partial \theta}$$

Integrador *Leapfrog*

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

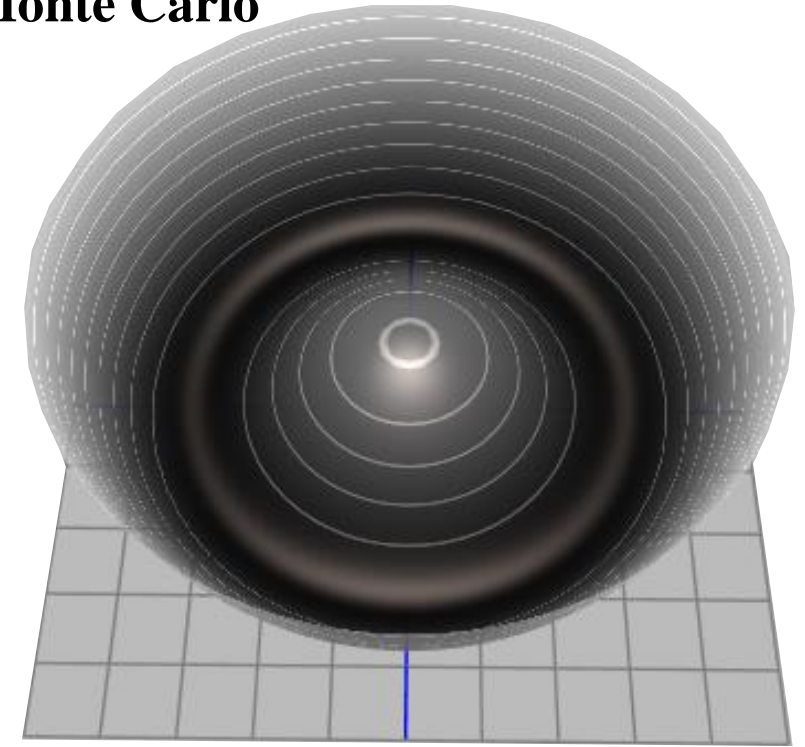
Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

Derivada com relação (θ)
Derivada constante = 0 (zero)



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = +\frac{\partial H}{\partial \rho} = +\frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = -\frac{\partial H}{\partial \theta} = -\frac{\partial T}{\partial \theta} - \frac{\partial V}{\partial \theta}$$

Momento
independente da
densidade alvo

$$p(\rho|\theta) = p(\rho)$$

Integrador Leapfrog

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

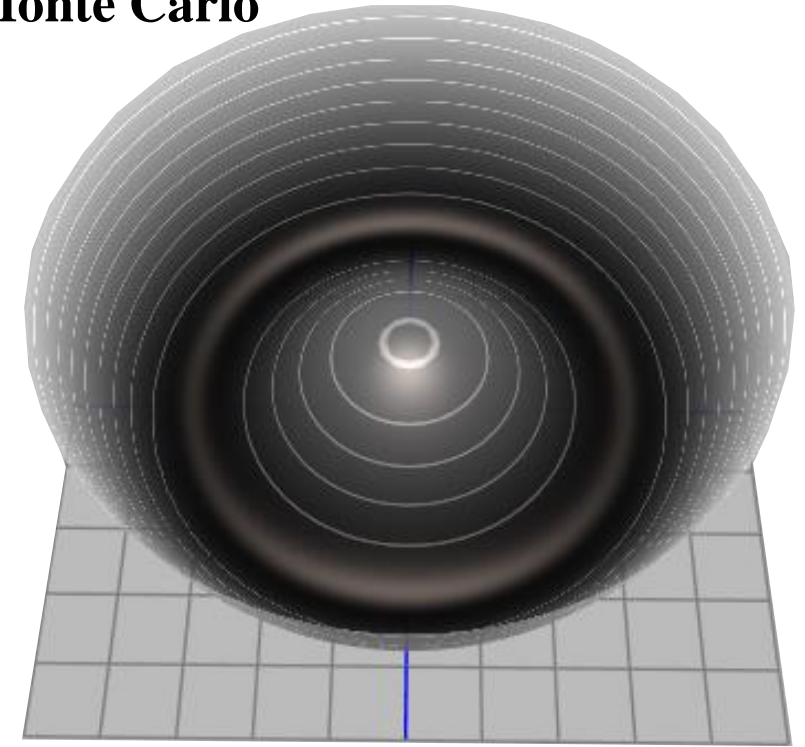
Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

Derivada com relação (θ)
Derivada constante = 0 (zero)



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = +\frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = -\frac{\partial V}{\partial \theta}$$

Momento
independente da
densidade alvo

$$p(\rho|\theta) = p(\rho)$$

Integrador Leapfrog

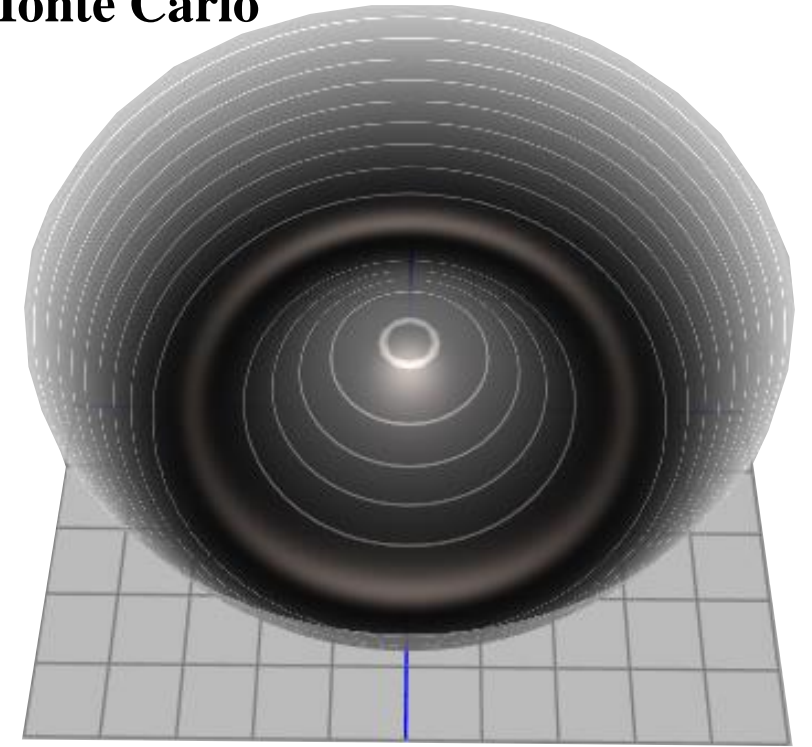
$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$

Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = + \frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = - \frac{\partial V}{\partial \theta}$$

3- Passo: Intervalo de tempo discretizado ϵ

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

$$\theta \leftarrow \theta + \epsilon M^{-1} \rho$$

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

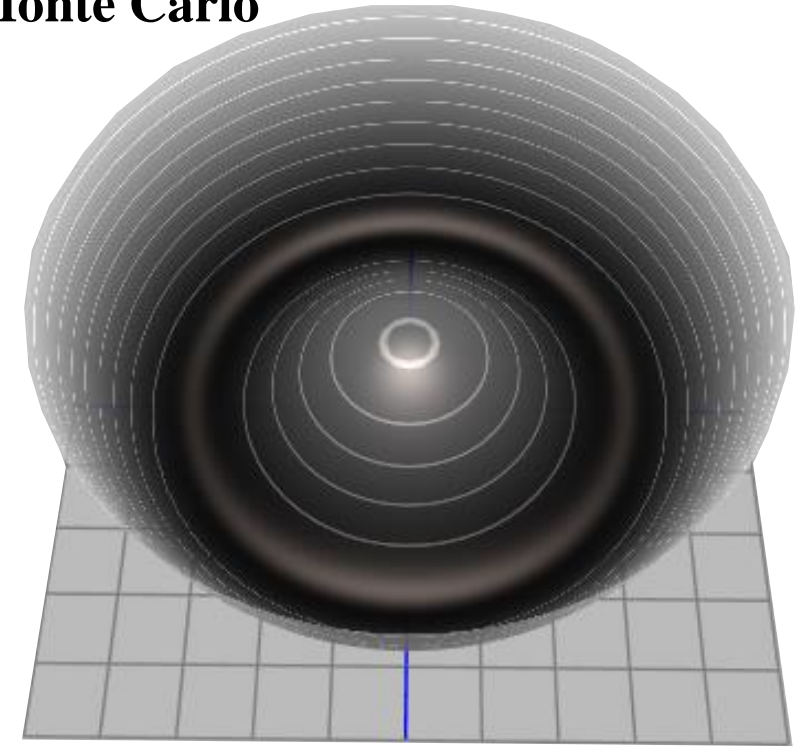
Integrador Leapfrog

Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = + \frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = - \frac{\partial V}{\partial \theta}$$

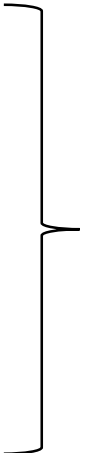
3- Passo: Intervalo de tempo discretizado ϵ

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

$$\theta \leftarrow \theta + \epsilon M^{-1} \rho$$

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

Atualiza meio passo
momento;

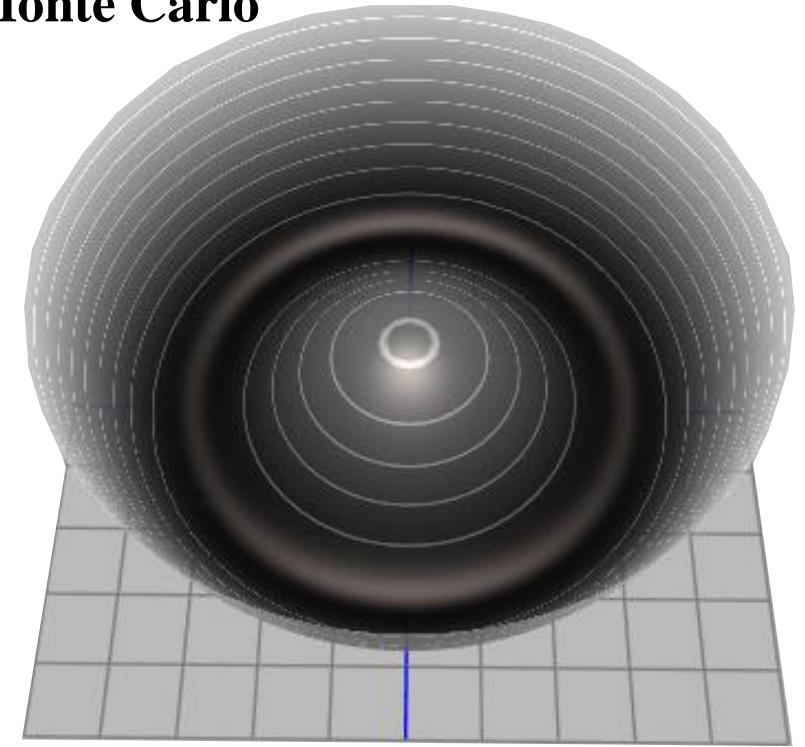


Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = + \frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = - \frac{\partial V}{\partial \theta}$$

3- Passo: Intervalo de tempo discretizado ϵ

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

Atualiza meio passo
momento;

$$\theta \leftarrow \theta + \epsilon M^{-1} \rho$$

Atualiza posição;

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

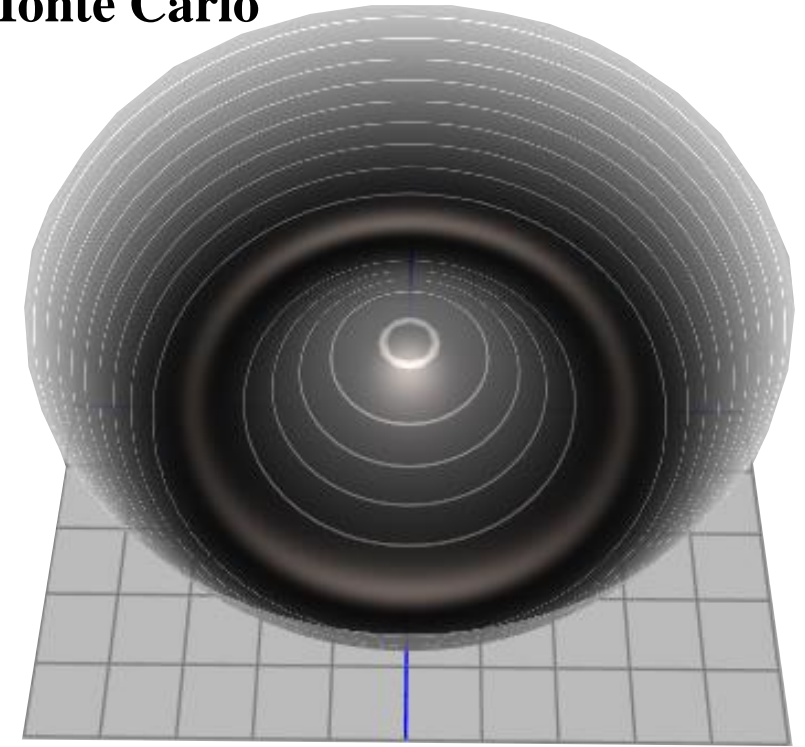
Integrador Leapfrog

Método Hamiltoniano Monte Carlo

$$H(\rho, \theta) = -\log p(\rho, \theta)$$

$$H(\rho, \theta) = -\log p(\rho|\theta) - \log p(\theta)$$

$$H(\rho, \theta) = T(\rho|\theta) + V(\theta)$$



➤ Gerando transição para as Equações Hamiltoniana

1- Passo:

Escolhe um valor aleatório para o momento

$$\rho \sim \text{MultiNormal}(0, M)$$

2- Passo:

$$\frac{d\theta}{dt} = + \frac{\partial T}{\partial \rho}$$

$$\frac{d\rho}{dt} = - \frac{\partial V}{\partial \theta}$$

3- Passo: Intervalo de tempo discretizado ϵ

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

Atualiza meio passo
momento;

$$\theta \leftarrow \theta + \epsilon M^{-1} \rho$$

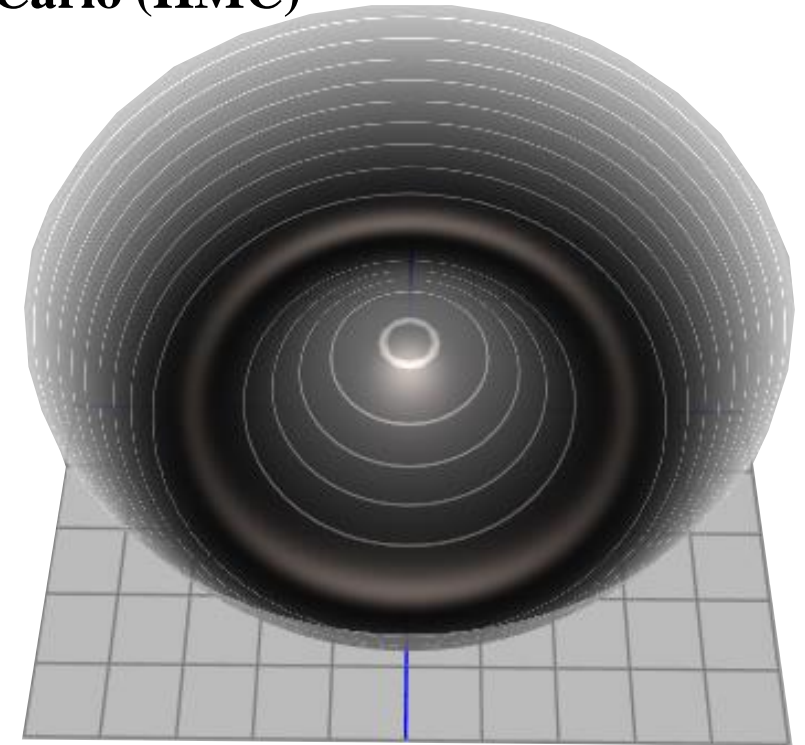
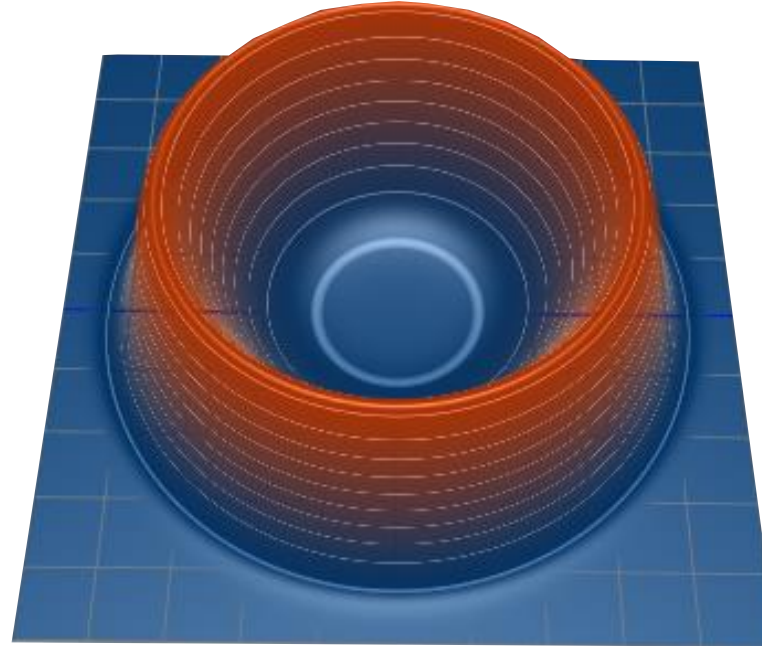
Atualiza posição;

$$\rho \leftarrow \rho - \frac{\epsilon}{2} \frac{\partial V}{\partial \theta}$$

Atualiza momento;
Repete = L vezes

Integrador Leapfrog

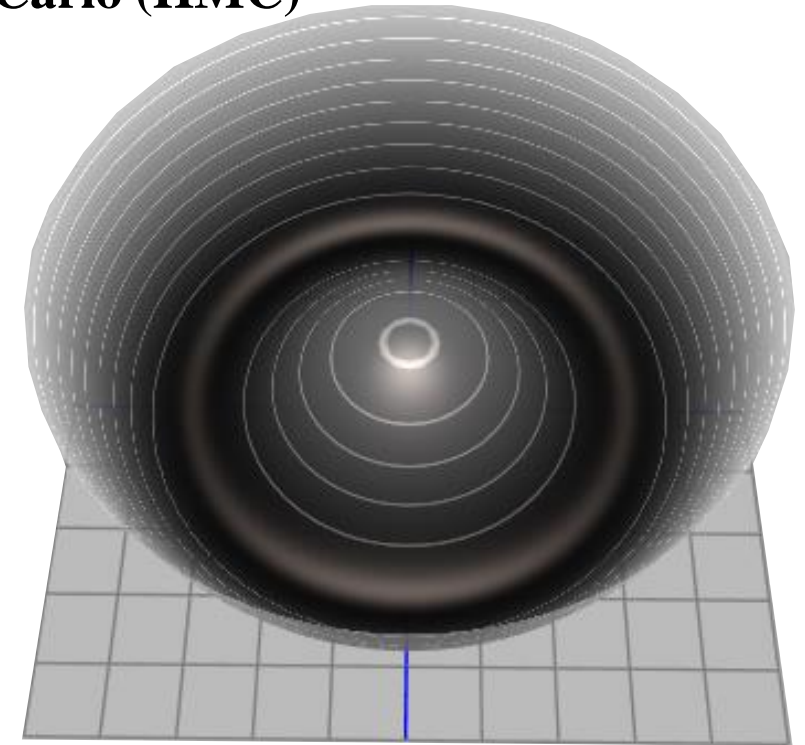
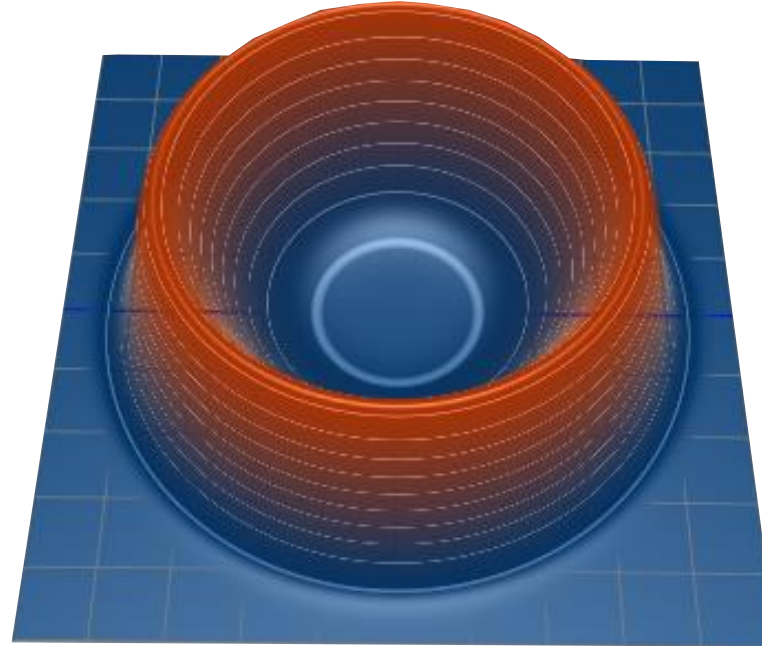
Método Hamiltoniano Monte Carlo (HMC)



O algoritmo HMC possui três parâmetros a serem definidos:

- discretização tempo ϵ ;
- métrica M ;
- número de passos dados L .

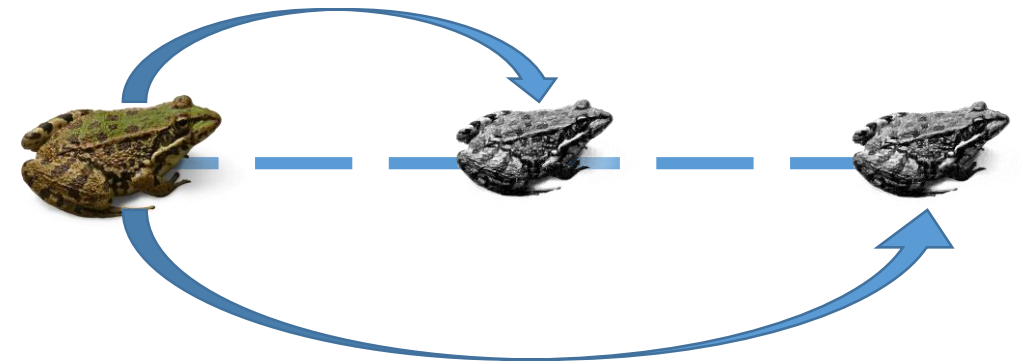
Método Hamiltoniano Monte Carlo (HMC)



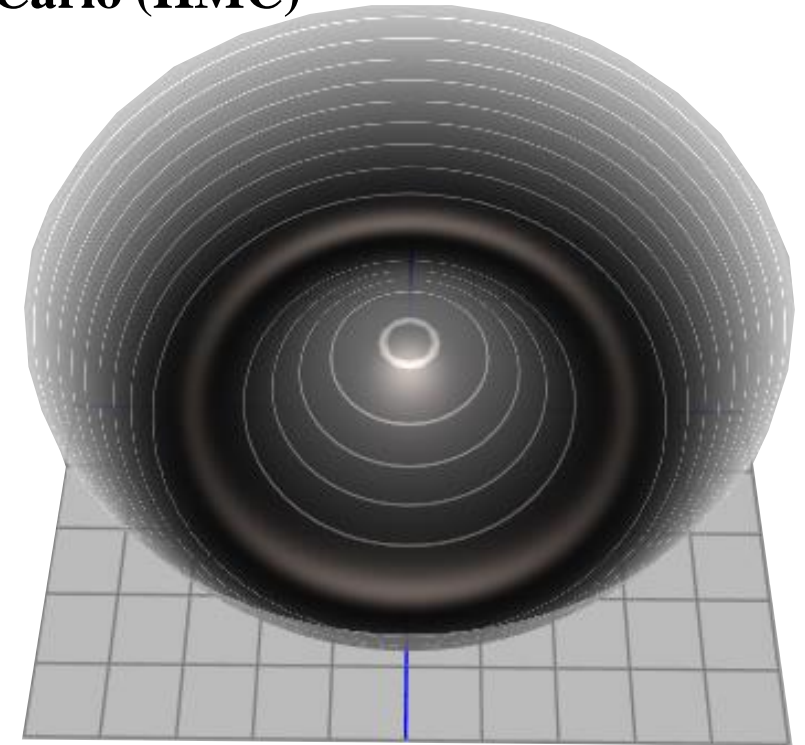
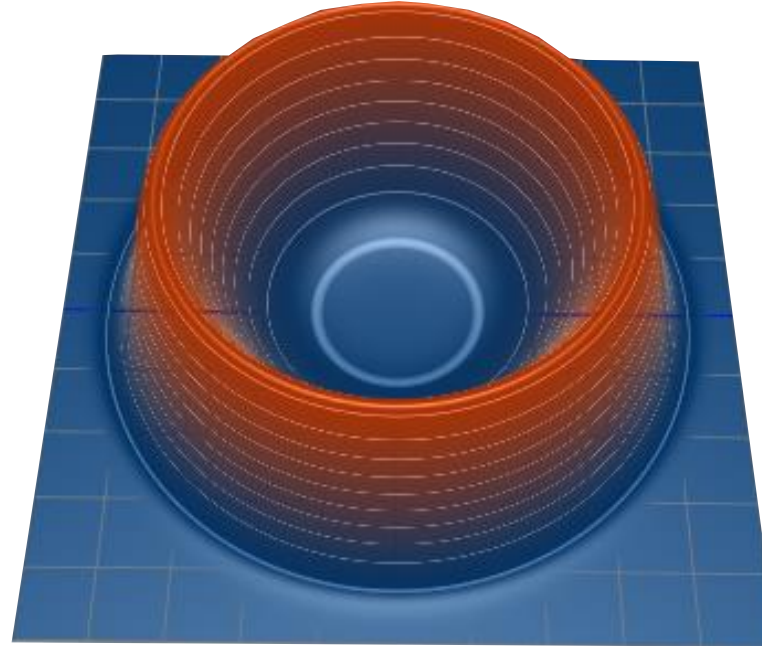
integrador *Leapfrog*

O algoritmo HMC possui três parâmetros a serem definidos:

- discretização tempo ϵ ; \longrightarrow **Tamanho do passo**
- métrica M ;
- número de passos dados L .



Método Hamiltoniano Monte Carlo (HMC)



O algoritmo HMC possui três parâmetros a serem definidos:

- discretização tempo ϵ ;

- métrica M ; \longrightarrow

- número de passos dados L .

- Variável auxiliar *momento*;

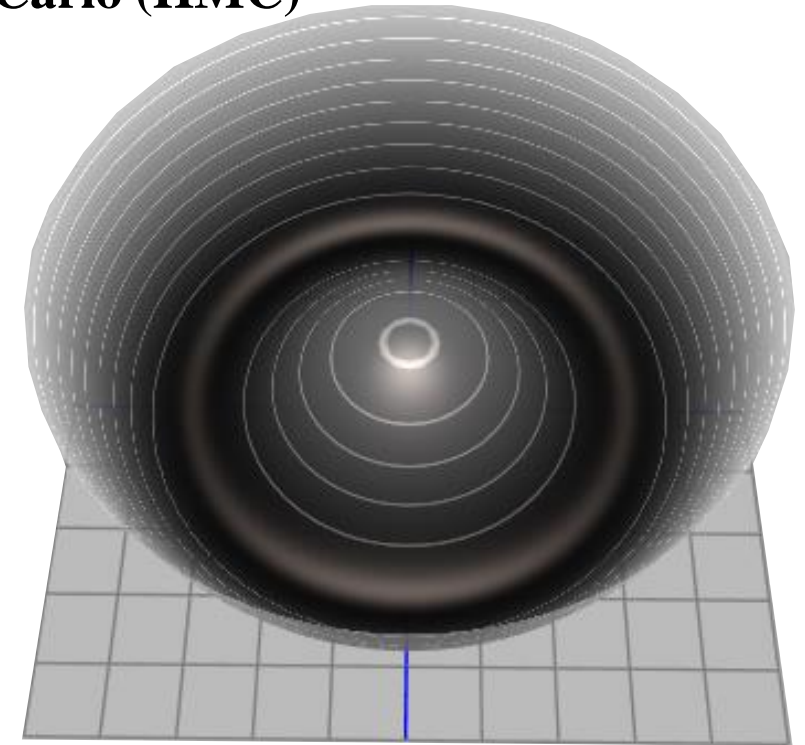
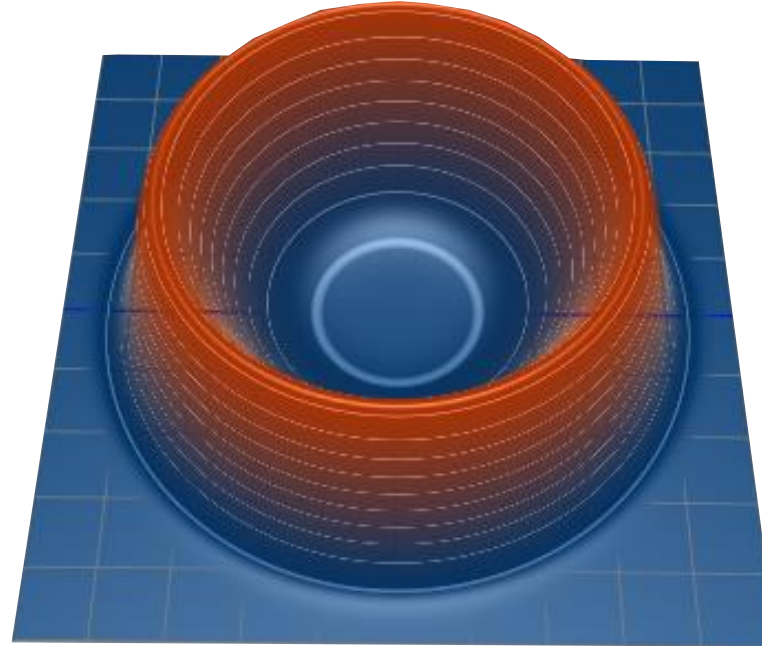
$$p(\rho, \theta) = p(\rho|\theta)p(\theta)$$

$$\rho \sim \text{MultiNormal}(0, M)$$

M = Matriz Euclidiana

M pode ser visto como uma transformação do espaço paramétrico que torna a amostragem mais eficiente.

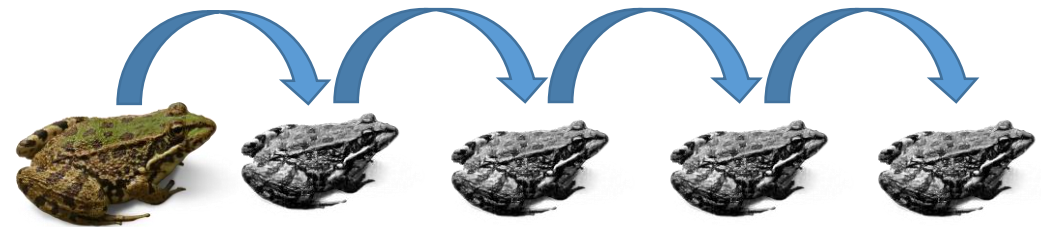
Método Hamiltoniano Monte Carlo (HMC)



O algoritmo HMC possui três parâmetros a serem definidos:

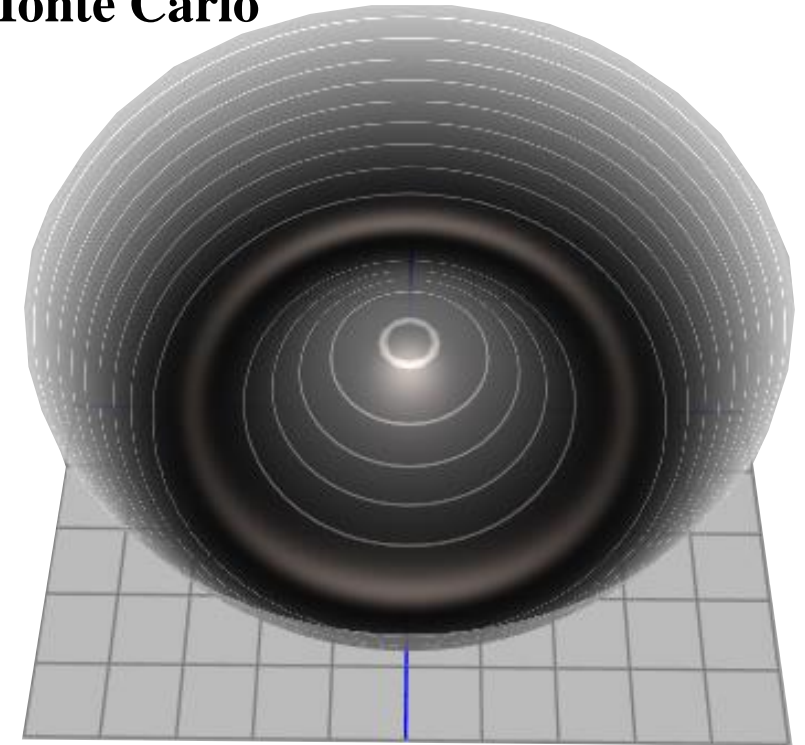
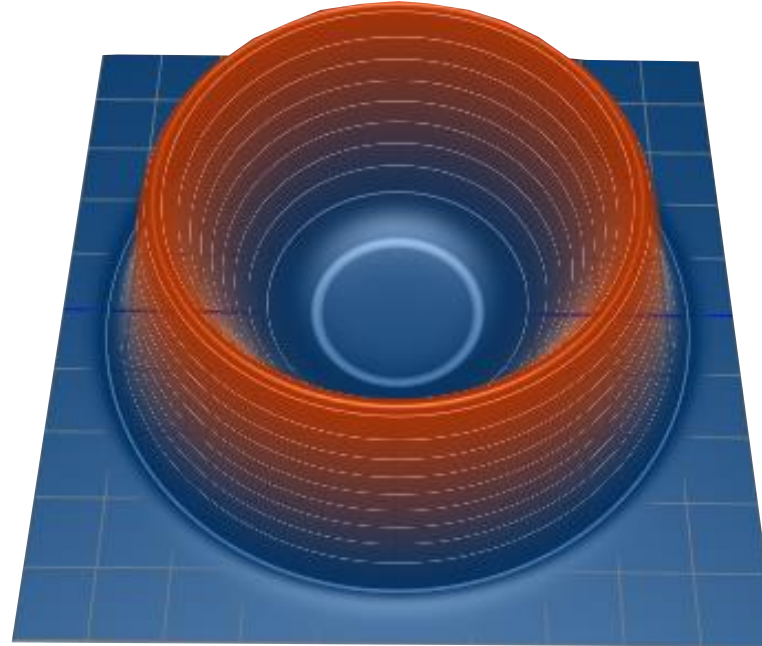
- discretização tempo ϵ ;
- métrica M ;
- número de passos dados L .

integrador *Leapfrog*



L número de passos

Método Hamiltoniano Monte Carlo

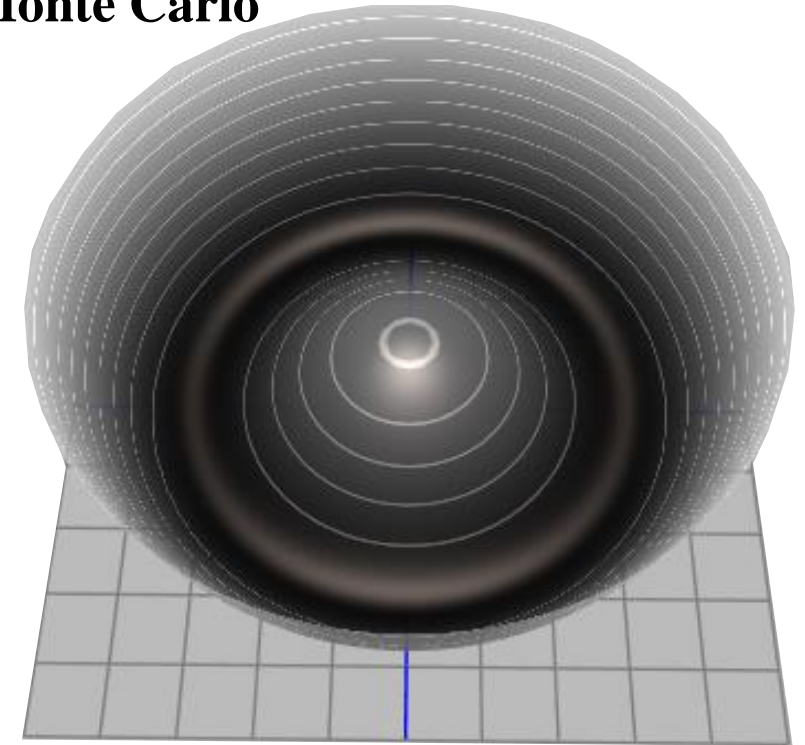
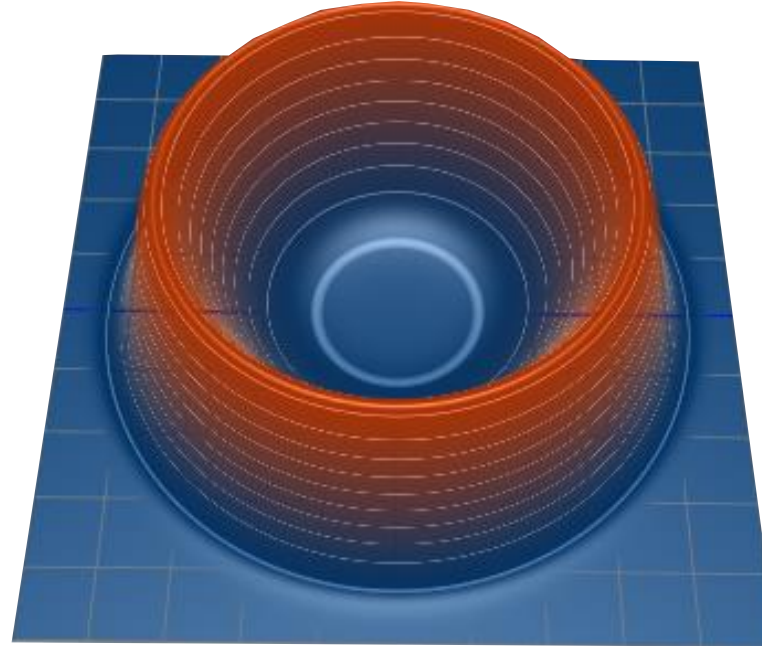


O algoritmo HMC possui três parâmetros a serem definidos:

- discretização tempo ϵ ; \longrightarrow
- métrica M ;
- número de passos dados L .

Intervalo de tempo (tamanho do passo) necessário para a maioria dos integradores numéricos.

Método Hamiltoniano Monte Carlo



O algoritmo HMC possui três parâmetros a serem definidos:

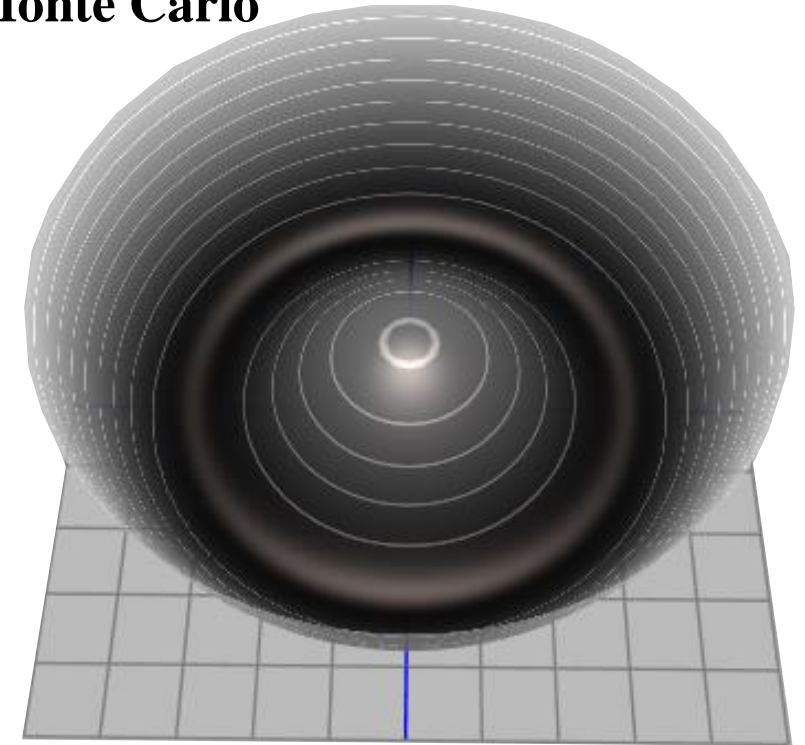
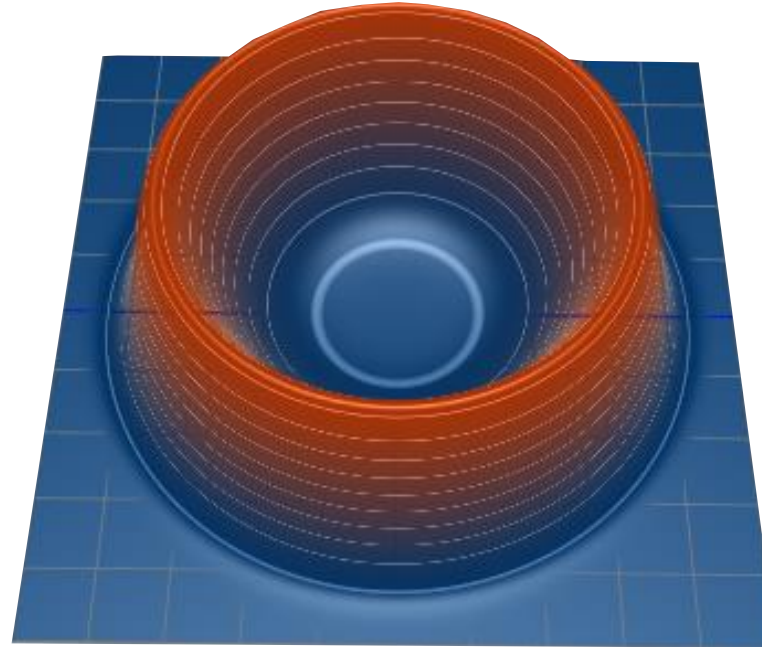
- discretização tempo ϵ ;

- métrica M ; \longrightarrow

M pode ser visto como uma **transformação** do **espaço paramétrico** que torna a **amostragem** mais **eficiente**.

- número de passos dados L .

Método Hamiltoniano Monte Carlo



O algoritmo HMC possui três parâmetros a serem definidos:

- discretização tempo ϵ ;
- métrica M ;
- número de passos dados L . \longrightarrow L número de passos dados no integrador *Leapfrog*.

Total de simulação:

$$L \times \epsilon.$$

Número de passos x intervalo tempo discretizado (saltos).

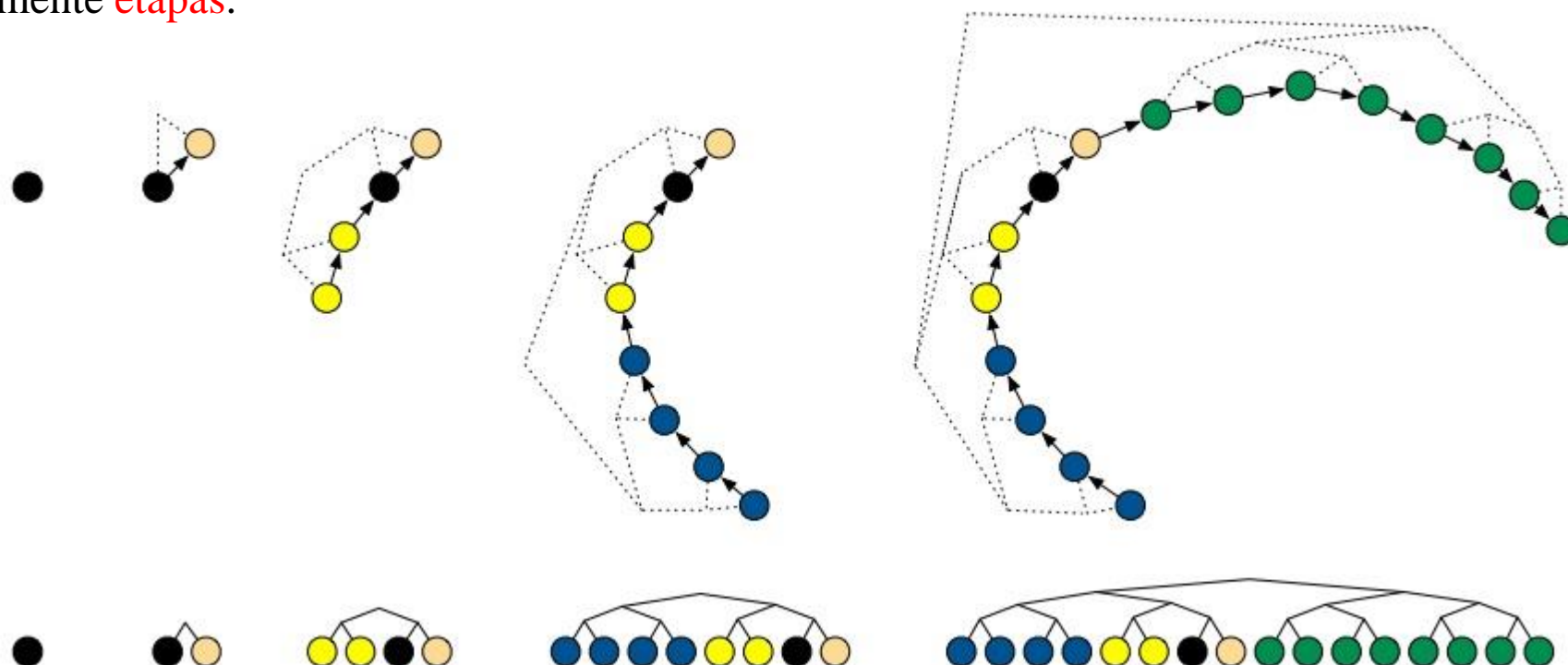
Algoritmo de amostragem



- Algoritmo de amostragem;
 - Metropolis-Hasting;
 - Gibbs;
 - Hamiltoniano Monte Carlo;
 - **NUTS**;

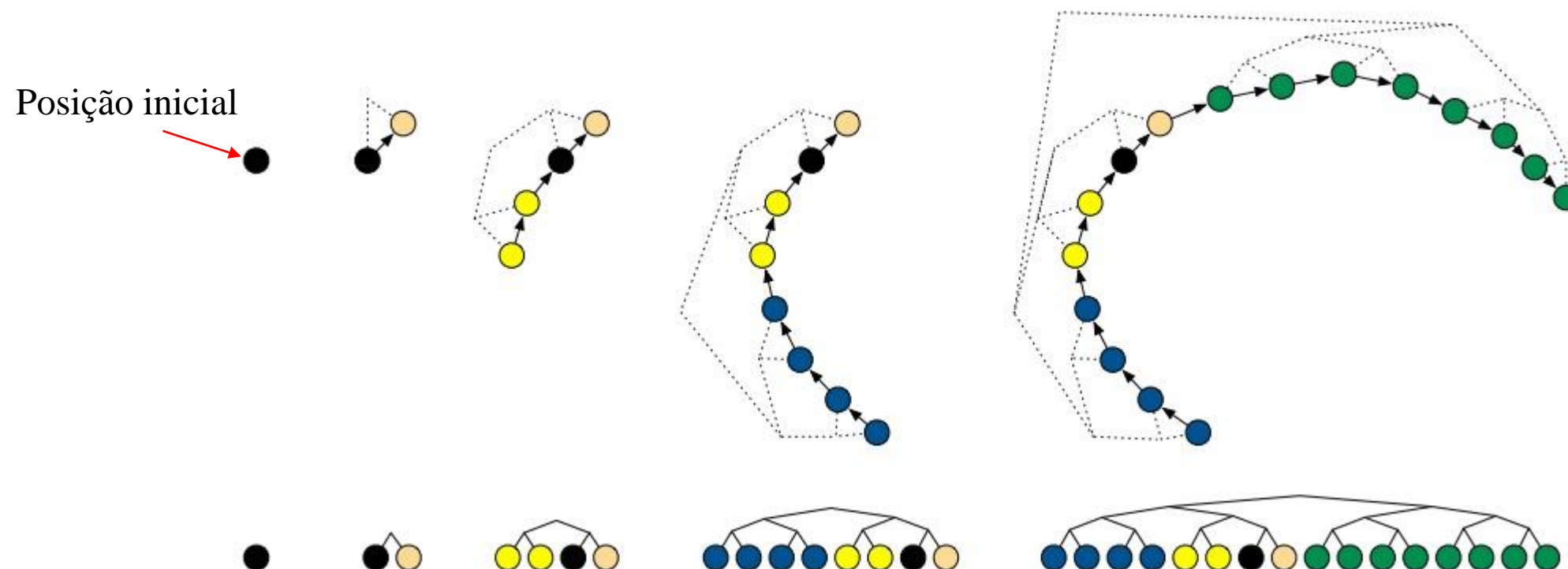
No-U-Turn (NUTS) HMC

Uma única **trajetória NUTS** é
construída **acumulando**
iterativamente **etapas**.



Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

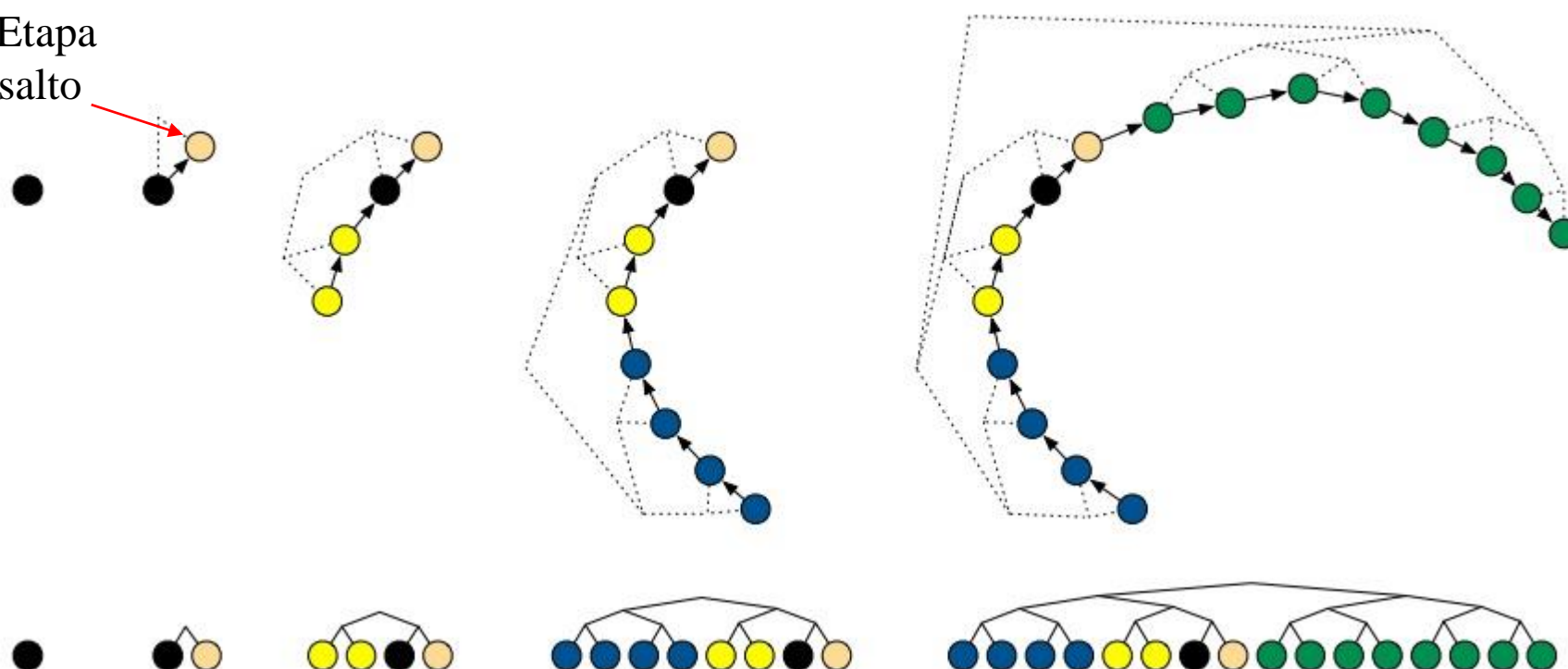
No-U-Turn (NUTS) HMC



Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.*, 15(1), 1593-1623.

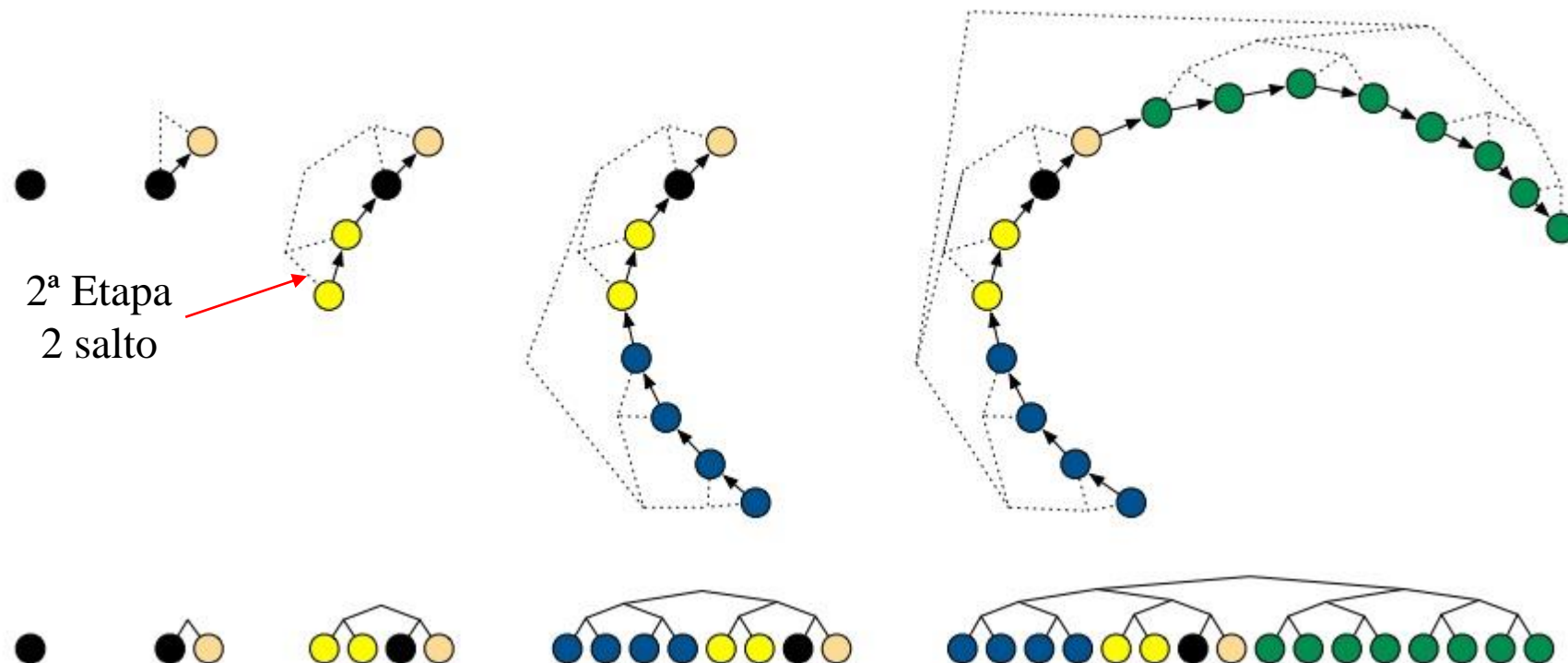
No-U-Turn (NUTS) HMC

1ª Etapa
1 salto



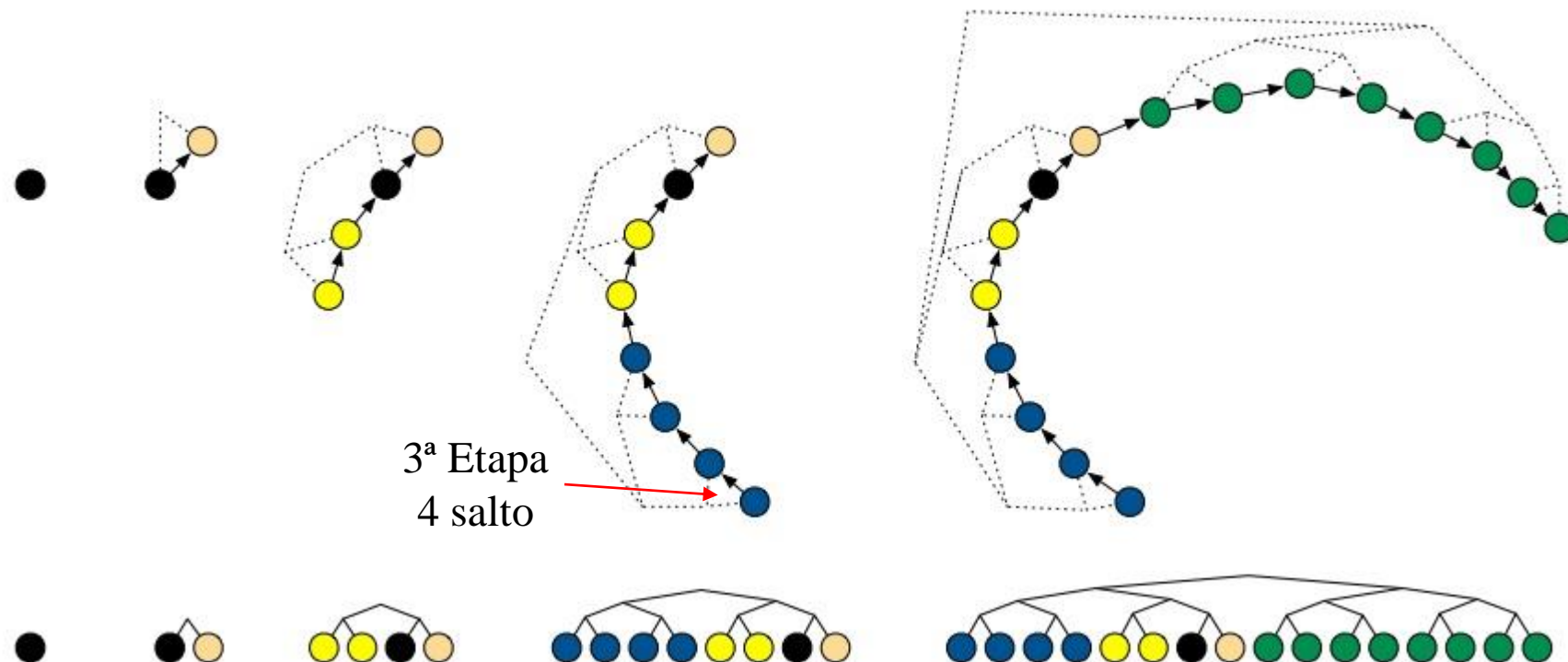
Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

No-U-Turn (NUTS) HMC



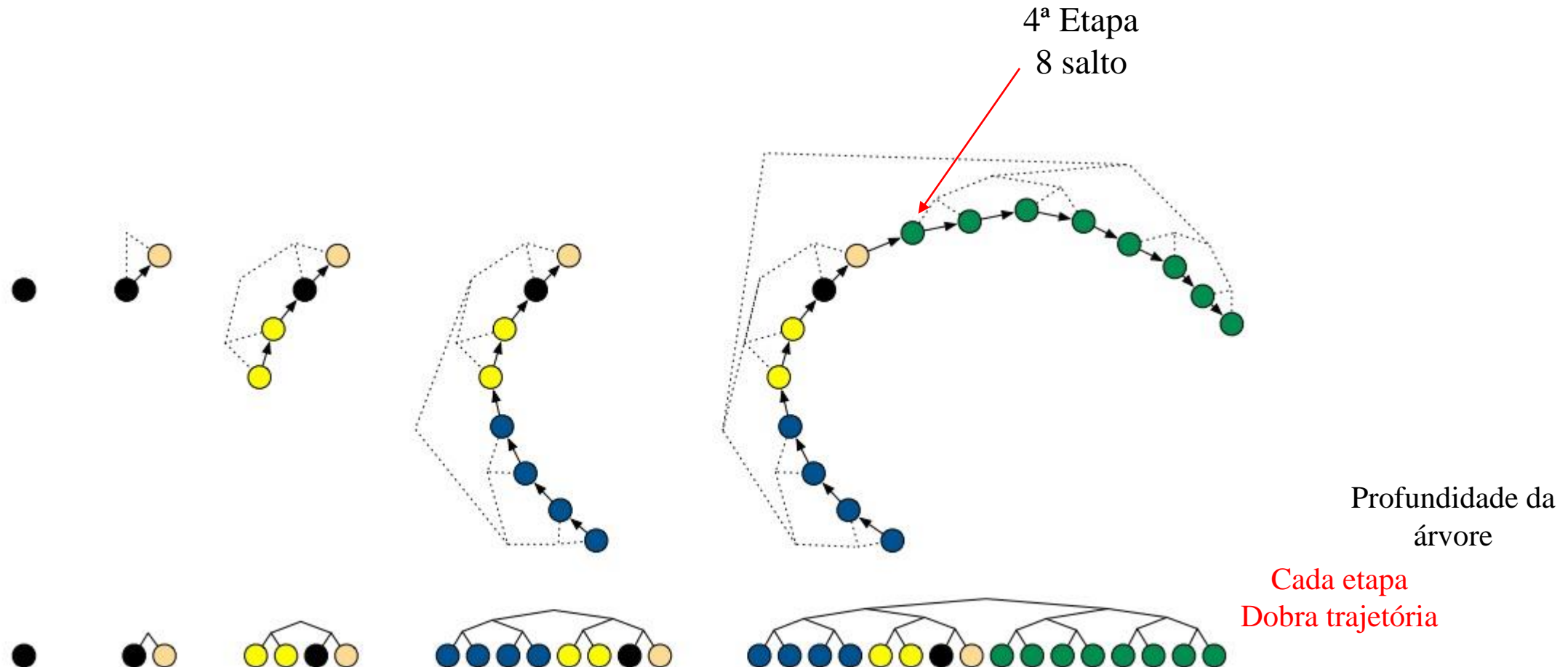
Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

No-U-Turn (NUTS) HMC



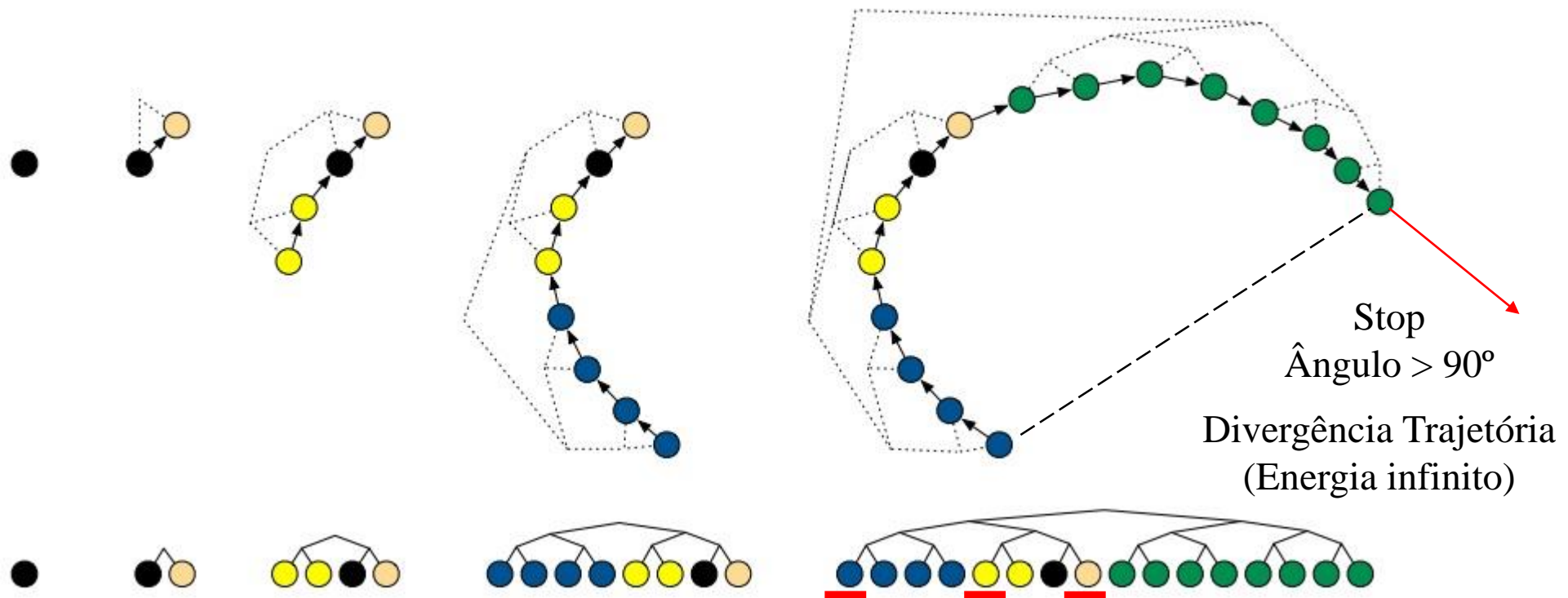
Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

No-U-Turn (NUTS) HMC



Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

No-U-Turn (NUTS) HMC



Desenha esquemático ilustrando o algoritmo NUTS HMC. Retirado do artigo Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Mach. Learn. Res., 15(1), 1593-1623.

Introdução a linguagem Stan (*rstan*), um software para modelos bayesianos.



Universidade Federal do Oeste do Pará (UFOPA)
Campus de Monte Alegre – Engenharia de Aquicultura

Obrigado !!!

Professor: Carlos Antônio Zarzar
E-mail: carloszarzar_@hotmail.com
carlos.zarzar@ufopa.edu.br

Data: 09/03/2022

AGRADECIMENTO E COLABORADORES:



UFOPA