

Antecipação e adaptação: como incorporar os dinamismos do mundo financeiro

Igor Nascimento

Laboratório de Aprendizado de Máquina em Finanças e Organizações - LAMFO

28/02/2018

Visão geral

1. Contexto
2. Amostragem aleatória
3. Modelos Dinâmicos
4. Filtro de Partículas

O investidor (banco, pessoa física, fundo de investimento, fundo de pensão) possui um capital e deseja utilizá-lo para atingir um objetivo:

- ▶ Rendimento superior a taxa de captação
- ▶ Segurança financeira
- ▶ Lucro ao investidor
- ▶ Aposentadoria

- ▶ Ações (PETR3, VALE3, IBOVESPA)

Ativos

- ▶ Ações (PETR3, VALE3, IBOVESPA)
- ▶ Títulos de dívida pública (NTN-B, LTN)

Ativos

- ▶ Ações (PETR3, VALE3, IBOVESPA)
- ▶ Títulos de dívida pública (NTN-B, LTN)
- ▶ Empresa de terceiros (Debêntures)]

Ativos

- ▶ Ações (PETR3, VALE3, IBOVESPA)
- ▶ Títulos de dívida pública (NTN-B, LTN)
- ▶ Empresa de terceiros (Debêntures)]
- ▶ Empresas própria (Empresário)

Ativos

- ▶ Ações (PETR3, VALE3, IBOVESPA)
- ▶ Títulos de dívida pública (NTN-B, LTN)
- ▶ Empresa de terceiros (Debêntures)]
- ▶ Empresas própria (Empresário)
- ▶ Outros (criptomoeda, Avestrus Master, Hinode)

Alocação de ativos ou portfólio é **escolher** um ou mais ativos.

Alocação

- ▶ Retorno: qual o valor esperado ao final do investimento
- ▶ Risco: quais são os valores possíveis para o retorno

O trabalho seminal de [Markowitz, 1952] sobre alocação de portfólio e fronteira eficiente.

[Markowitz, 1952]

- ▶ ativos:

$$r_1, r_2, \dots, r_N$$

- ▶ retorno:

$$E(r_1) = \mu_1, E(r_2) = \mu_2, \dots, E(r_N) = \mu_N$$

- ▶ variância:

$$V(r_1) = \sigma_1^2, V(r_2) = \sigma_2^2, \dots, V(r_N) = \sigma_N^2$$

- ▶ covariância (correlação):

$$COR(r_i, r_j) = \rho_{ij}$$

Alocação

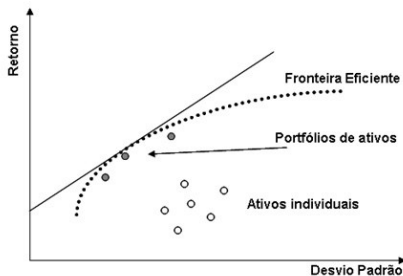
Determinar a locação, isto é, o percentual w_1, w_2, \dots, w_N que cada ativo representa da carteira:

$$E_{portfolio} = E(\mathbf{w}) = \sum_{i=1}^N w_i \times \mu_i \quad (1)$$

$$V_{portfolio} = V(\mathbf{w}) = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \times \sigma_i \sigma_j \times \rho_{ij} \quad (2)$$

Fronteira Eficiente

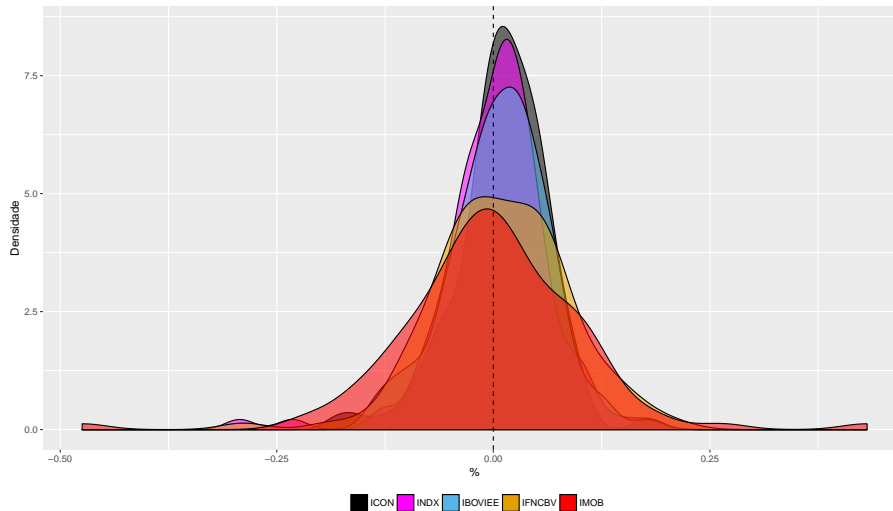
"Optimal weight of each asset, such that the overall portfolio provides the best return for a fixed level of risk, or conversely, the smallest risk for a given overall return?" [Laloux et al., 1999]

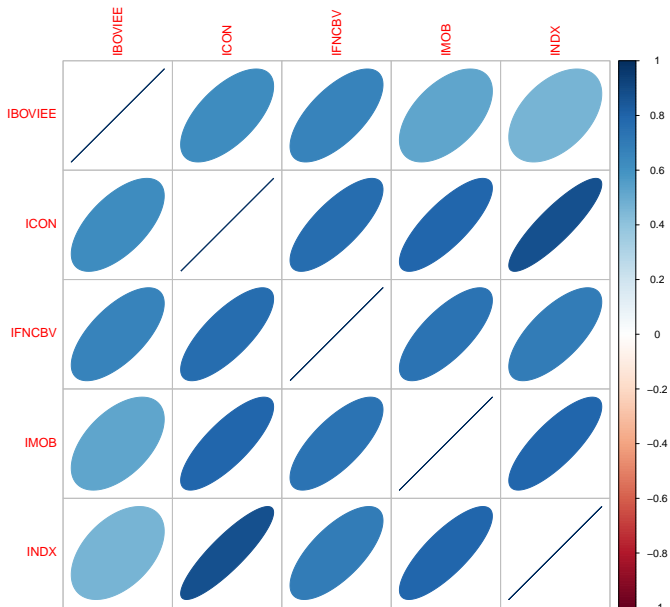


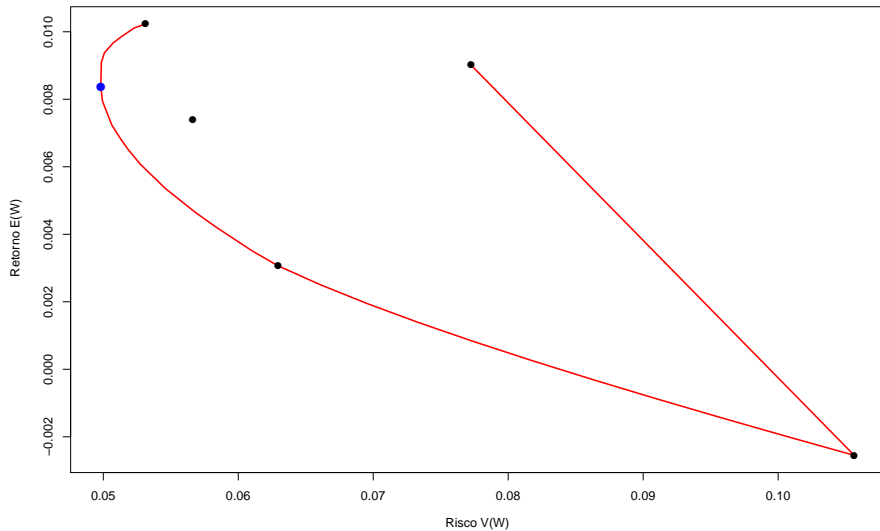
Minicaso

- ▶ IFN: índice setor financeiro
- ▶ IMOB: índice do setor imobiliário
- ▶ ICON: índice de consumo
- ▶ IEE: índice de energia
- ▶ INDX: índice da indústria

Log retorno mensal 2008-2017







Objetivo

Apresentar:

- ▶ Métodos numéricos para séries temporais
- ▶ Desenvolvimentos recentes
- ▶ Principais aplicações

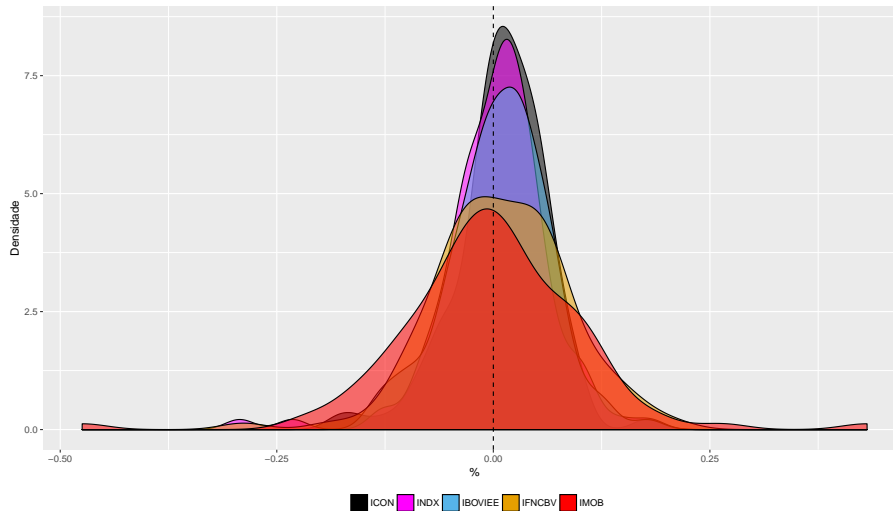
Expectativa ao final

- ▶ Identificar diferenças entre os principais métodos
- ▶ Relacionar a aplicações em Finanças
- ▶ Conhecer referências clássicas e recentes na área

Value at Risk (Var)

- ▶ Em um dia ruim, quanto eu posso perder em um determinado portfólio?
- ▶ Quantos dias ruins eu suportaria nesse portfólio?

Log retorno mensal 2008-2017



Monte Carlo

- ▶ descrição: (re)amostragem “paramétrica”
- ▶ vantagens: acessa “todo” espaço
- ▶ desvantagem: conhecimento prévio da distribuição

Normal Multivariada

$$X \sim N_k(\tilde{\mu}, \Sigma)$$

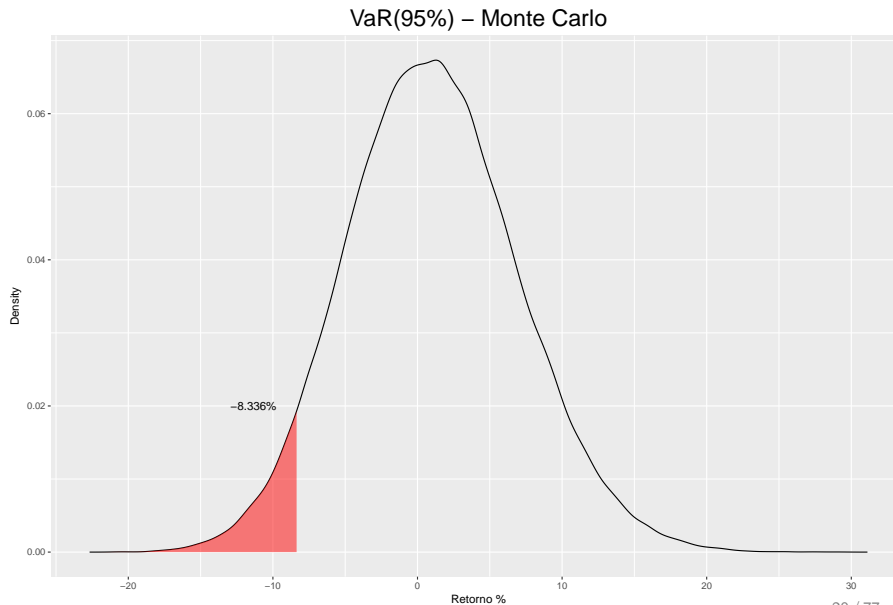
$$f(X) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\det \Sigma}} \exp \left(-\frac{1}{2} (\tilde{X} - \tilde{\mu}) \Sigma^{-1} (\tilde{X} - \tilde{\mu}) \right)$$

1. modelo
2. parâmetros
3. amostragem

Code - MC

```
set.seed(1052210218)
library(MASS)
n_mc <- 100000
mc <- mvrnorm(n=n_mc,mu,Sigma)
mc <- matriz_mc %*% w_optm
qlim <- quantile(mc,p=0.05)
```

Value at Risk (Var)



Pontos importantes

- ▶ Suposição da distribuição:
- ▶ Estrutura de dependência longitudinal?

Bootstrap

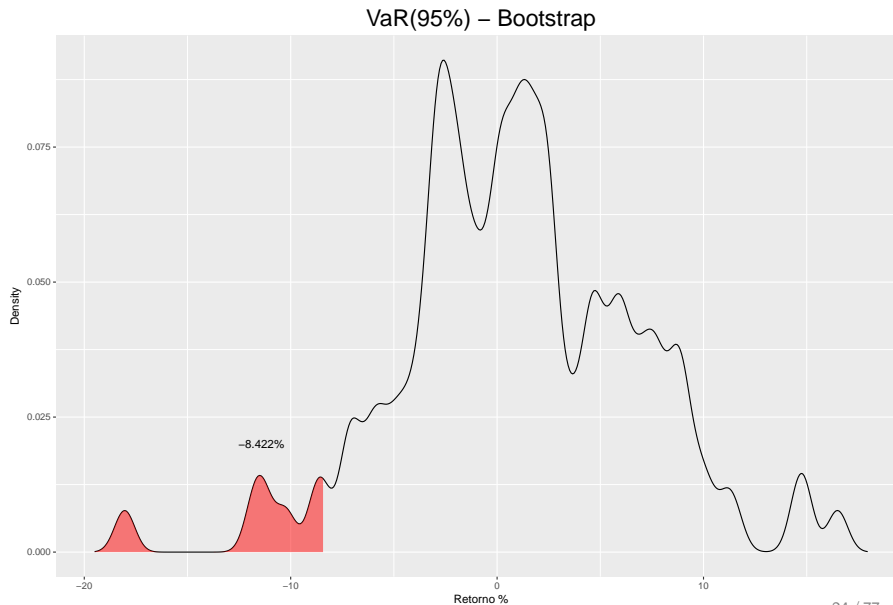
- ▶ descrição: (re)amostragem “não-paramétrica”
- ▶ vantagens: não requer conhecimento sobre a distribuição
- ▶ desvantagem: acessa espaço “realizado”

Bootstrap

- ▶ descrição: (re)amostragem “não-paramétrica”
- ▶ vantagens: não requer conhecimento sobre a distribuição
- ▶ desvantagem: acessa espaço “realizado”

- ▶ Suposição da distribuição: amostragem aleatória nos próprios dados
- ▶ Estrutura de dependência longitudinal: Base dividida em janelas 5 meses

Value at Risk (Var) - Bootstrap



Método

1. Monte Carlo
2. Bootstrap
3. Markov Chain Monte Carlo

Monte Carlo

História do Método Monte Carlo.

Monte Carlo

Por que isso é interessante ?

O que é Bootstrap

Sistemas Complexos

Por que isso é interessante ?

Markov Chain Monte Carlo

O que é Markov Chain Monte Carlo?

Markov Chain Monte Carlo

Por que isso é interessante ?

Objetivos da dissertação

- ▶ Revisão bibliográfica do filtro de partículas

Objetivos da dissertação

- ▶ Revisão bibliográfica do filtro de partículas
- ▶ Estimação dos estados latentes nos modelos dinâmicos utilizando a técnica

Objetivos da dissertação

- ▶ Revisão bibliográfica do filtro de partículas
- ▶ Estimação dos estados latentes nos modelos dinâmicos utilizando a técnica
- ▶ Estimação de parâmetros estáticos e desconhecidos presentes nas equações dos modelos

Objetivos da dissertação

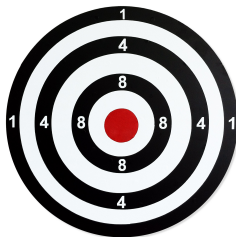
- ▶ Revisão bibliográfica do filtro de partículas
- ▶ Estimação dos estados latentes nos modelos dinâmicos utilizando a técnica
- ▶ Estimação de parâmetros estáticos e desconhecidos presentes nas equações dos modelos
- ▶ Apresentar uma dissertação didática para o estudo da técnica

Problema

Estimar um parâmetro desconhecido.

Problema

Estimar um parâmetro desconhecido.



Estimação

Filtro de partículas

Partículas

Definição Naive

Parte de um todo

Partículas

Definição Naive

Parte de um todo

1. Parte: Partículas.

Partículas

Definição Naive

Parte de um **todo**

1. Parte: Partículas.
2. Todo: Espaço paramétrico.

Partículas

Definição Naive

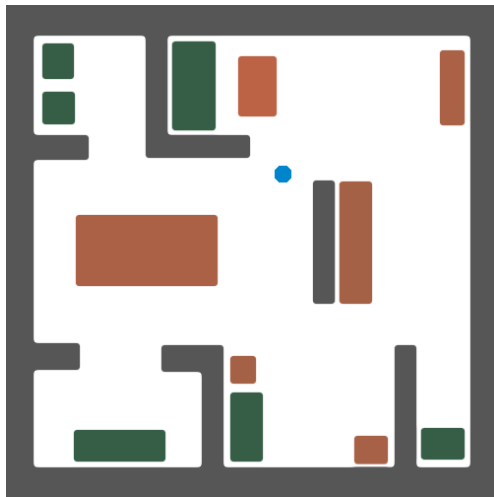
Parte de um todo.

1. Parte: Partículas.
2. Todo: Espaço paramétrico.

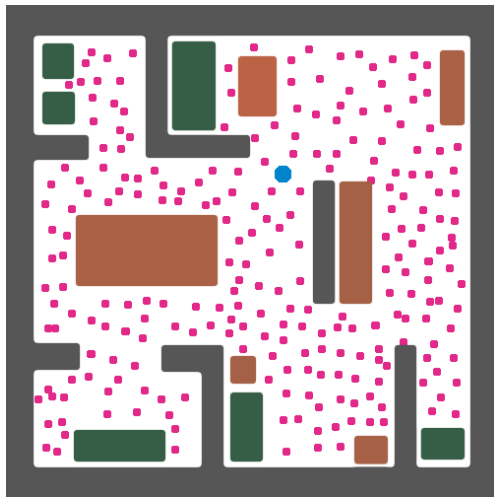
Definição

Partículas são realizações de um experimento cujos valores possíveis estão definidos no espaço paramétrico.

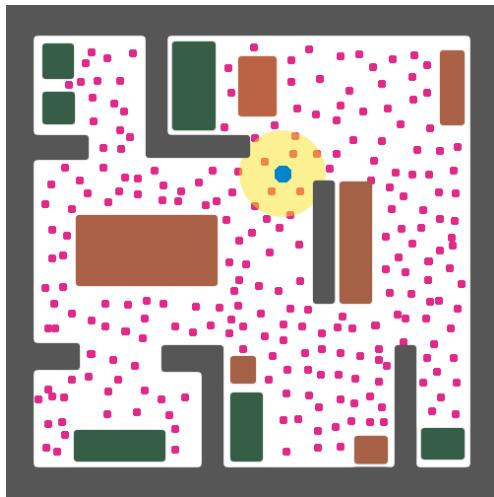
Exemplo do uso de Filtro de partículas na Robótica



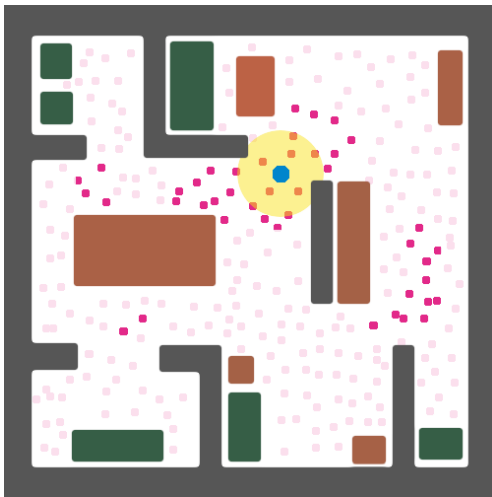
Partículas



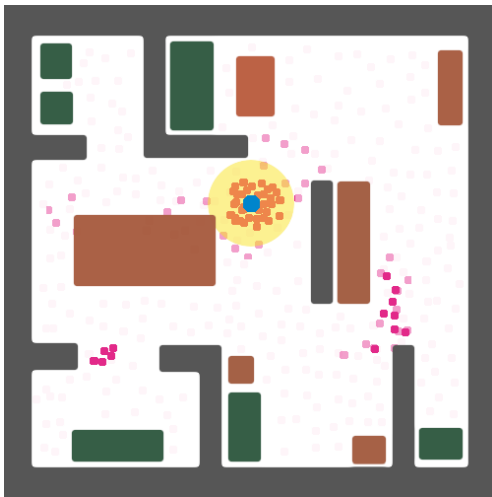
Medida de proximidade



Aproximação



Propagação



Modelo dinâmico

Parte não observável - sistema

Sequência θ_t com estrutura de dependência de um Processo Markoviano, $\{\theta_t, t = 1, \dots, n\}$.

Modelo dinâmico

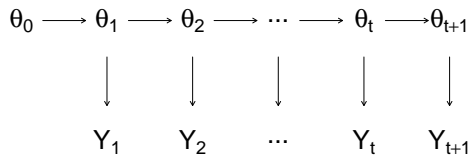
Parte não observável - sistema

Sequência θ_t com estrutura de dependência de um Processo Markoviano, $\{\theta_t, t = 1, \dots, n\}$.

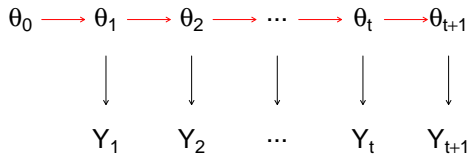
Parte observável

Sequência y_t $\{y_t, t = 1, \dots, n\}$ que é dependente, exclusivamente, de θ_t .

Modelo dinâmico



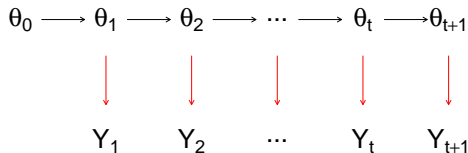
Modelo dinâmico



Equação do Sistema

$$\theta_t \sim \pi(\cdot | \theta_{t-1})$$

Modelo dinâmico



Equação do Sistema

$$\theta_t \sim \pi(\cdot | \theta_{t-1})$$

Equação das Observações

$$y_t \sim \pi(\cdot | \theta_t)$$

Propriedades

A.1: $\theta_t, t = 1, \dots, n$ é uma sequência de estados de um Processo Markoviano;

$$\pi(\theta_{1:n}) = \pi(\theta_1) \prod_{k=2}^n \pi(\theta_k | \theta_{k-1})$$

Propriedades

A.1: $\theta_t, t = 1, \dots, n$ é uma sequência de estados de um Processo Markoviano;

$$\pi(\theta_{1:n}) = \pi(\theta_1) \prod_{k=2}^n \pi(\theta_k | \theta_{k-1})$$

A.2: $y_{1:t}$ é um vetor de observações que são condicionalmente independentes dado $\theta_{1:t}$, para cada $t = 1, \dots, n$.

$$\pi(y_t | \theta_{1:t}, y_{1:t-1}) = \pi(y_t | \theta_t)$$

Distribuição conjunta

Essas propriedades permitem descrever a **distribuição conjunta** das observações e dos estados como o produto das seguintes **distribuições condicionais**:

$$\pi(y_{1:n}, \theta_{1:n}) = \pi(\theta_0) \prod_{t=1}^n \pi(\theta_t | \theta_{t-1}) \pi(y_t | \theta_t)$$

Interesse:

$$\pi(\theta_{1:n} | y_{1:n}) = \frac{\pi(\theta_{1:n}, y_{1:n})}{\pi(y_{1:n})}$$

Estimação

Métodos numéricos

Estimação

Monte Carlo

Definição

Distribuição alvo : $\pi_n(\theta_1, \dots, \theta_n)$ para n fixo.

Definição

Distribuição alvo : $\pi_n(\theta_1, \dots, \theta_n)$ para n fixo.

Gera-se N amostras independentes da variável aleatória

$$\Theta_{1:n}^i \sim \pi_n(\theta_1, \dots, \theta_n), \quad i = 1, \dots, N.$$

Definição

Distribuição alvo : $\pi_n(\theta_1, \dots, \theta_n)$ para n fixo.

Gera-se N amostras independentes da variável aleatória $\Theta_{1:n}^i \sim \pi_n(\theta_1, \dots, \theta_n)$, $i = 1, \dots, N$.

Aproximação e dada por:

$$\hat{\pi}_n(\theta_1, \dots, \theta_n) = \frac{1}{N} \sum_i^N \delta_{\Theta_{1:n}^i}(\theta_{1:n}),$$

sendo $\delta_{\Theta_{1:n}^i}(\theta_{1:n})$ uma função indicadora de massa no ponto $\theta_{1:n}$.

Definição

Considerar mensurar uma função ϕ_n em $\pi(\theta_1, \dots, \theta_n)$:

$$E(\phi_n) = \int \phi_n(\theta_1, \dots, \theta_n) \pi_n(\theta_1, \dots, \theta_n) d\theta_{1:n}$$

Definição

Considerer mensurar uma função ϕ_n em $\pi(\theta_1, \dots, \theta_n)$:

$$E(\phi_n) = \int \phi_n(\theta_1, \dots, \theta_n) \pi_n(\theta_1, \dots, \theta_n) d\theta_{1:n}$$

Pelo método Monte Carlo, **avaliamos** ϕ_n no suporte simulado de $\pi(\theta_1, \dots, \theta_n)$ por meio das **amostras**:

$$\hat{E}(\phi_n) = \int \phi_n(\theta_1, \dots, \theta_n) \hat{\pi}_n(\theta_1, \dots, \theta_n) d\theta_{1:n}$$

Definição

Considerer mensurar uma função ϕ_n em $\pi(\theta_1, \dots, \theta_n)$:

$$E(\phi_n) = \int \phi_n(\theta_1, \dots, \theta_n) \pi_n(\theta_1, \dots, \theta_n) d\theta_{1:n}$$

Pelo método Monte Carlo, **avaliamos** ϕ_n no suporte simulado de $\pi(\theta_1, \dots, \theta_n)$ por meio das **amostras**:

$$\hat{E}(\phi_n) = \int \phi_n(\theta_1, \dots, \theta_n) \hat{\pi}_n(\theta_1, \dots, \theta_n) d\theta_{1:n}$$

$$\hat{E}(\phi_n) = \frac{1}{N} \sum_{i=1}^N \phi_n(\Theta_{1:n}^i)$$

Desvantagens

Problema 1: Não é fácil se gerar $\pi_n(\theta_1, \dots, \theta_n)$.

Desvantagens

Problema 1: Não é fácil se gerar $\pi_n(\theta_1, \dots, \theta_n)$.

Problema 2: Ainda que fosse possível se gerar, a dimensão de $\pi_n(\theta_1, \dots, \theta_n)$ pode ser muito grande para se obter amostras multivariadas. Problema encontrado no método **MCMC**.

Estimação

Amostragem de importância

Problema 1

A amostragem de importância trata o **problema 1**: não se sabe gerar $\pi(\theta_{1:n})$.

Problema 1

A amostragem de importância trata o **problema 1**: não se sabe gerar $\pi(\theta_{1:n})$.

Utiliza-se uma **distribuição de importância** $q(\cdot)$ para obter o mesmo suporte da distribuição $\pi(\cdot)$.

$$E_{\pi}(\phi) = \int \phi(\theta)\pi(\theta)d\theta = \int \frac{\phi(\theta)\pi(\theta)}{q(\theta)}q(\theta)d\theta$$

Problema 1

A amostragem de importância trata o **problema 1**: não se sabe gerar $\pi(\theta_{1:n})$.

Utiliza-se uma **distribuição de importância** $q(\cdot)$ para obter o mesmo suporte da distribuição $\pi(\cdot)$.

$$E_{\pi}(\phi) = \int \phi(\theta)\pi(\theta)d\theta = \int \frac{\phi(\theta)\pi(\theta)}{q(\theta)}q(\theta)d\theta$$

A quantidade $\frac{\pi(\theta)}{q(\theta)}$ fornece um ajuste da discrepância entre **gerar um suporte** de $\pi(\theta)$ por meio da distribuição geradora $q(\theta)$.

Considere a seguinte representação, [?]:

$$\pi(\theta_{1:n}) = \frac{\gamma_n(\theta_{1:n})}{Z_n(\theta_{1:n})}, \quad Z_n = \int \gamma_n(\theta_{1:n}) d\theta_{1:n}$$

Considere a seguinte representação, [?]:

$$\pi(\theta_{1:n}) = \frac{\gamma_n(\theta_{1:n})}{Z_n(\theta_{1:n})}, \quad Z_n = \int \gamma_n(\theta_{1:n}) d\theta_{1:n}$$

Considere uma distribuição auxiliar, tal que:

$$\pi(\theta_{1:n}) > 0 \rightarrow q_n(\theta_{1:n}) > 0$$

A distribuição de interesse:

$$\pi(\theta_{1:n}) = \frac{\gamma_n(\theta_{1:n})}{Z_n(\theta_{1:n})} = \frac{\overbrace{\gamma_n(\theta_{1:n})}^{W_n}}{q_n(\theta_{1:n})} \frac{q_n(\theta_{1:n})}{Z_n(\theta_{1:n})}$$

$$\pi(\theta_{1:n}) = \frac{w_n(\theta_{1:n})q_n(\theta_{1:n})}{Z_n(\theta_{1:n})}$$

A distribuição de interesse:

$$\pi(\theta_{1:n}) = \frac{\gamma_n(\theta_{1:n})}{Z_n(\theta_{1:n})} = \frac{\overbrace{\gamma_n(\theta_{1:n})}^{W_n}}{q_n(\theta_{1:n})} \frac{q_n(\theta_{1:n})}{Z_n(\theta_{1:n})}$$

$$\pi(\theta_{1:n}) = \frac{w_n(\theta_{1:n})q_n(\theta_{1:n})}{Z_n(\theta_{1:n})}$$

O denominador é:

$$Z(\theta_{1:n}) = \int w_n(\theta_{1:n})q_n(\theta_{1:n})d\theta_{1:n}$$

Aproximação

Considere N amostras independentes $\Theta_{1:n}^i \sim q_n(\theta_{1:n})$.

Aproximação

Considere N amostras independentes $\Theta_{1:n}^i \sim q_n(\theta_{1:n})$.

Aproximação Monte Carlo é:

$$\hat{q}_n(\theta_{1:n}) = \frac{1}{N} \sum_i^N \delta_{\Theta^i}(\theta_{1:n})$$

Dessa forma, a aproximação para a distribuição de interesse é:

Aproximação

Considere N amostras independentes $\Theta_{1:n}^i \sim q_n(\theta_{1:n})$.

Aproximação Monte Carlo é:

$$\hat{q}_n(\theta_{1:n}) = \frac{1}{N} \sum_i^N \delta_{\Theta^i}(\theta_{1:n})$$

Dessa forma, a aproximação para a distribuição de interesse é:

$$\hat{\pi}_n(\theta_{1:n}) = \frac{w_n(\theta_{1:n}) \hat{q}_n(\theta_{1:n})}{\int w_n(\theta_{1:n}) \hat{q}_n(\theta_{1:n}) d\theta_{1:n}}$$

Aproximação

A distribuição de interesse pode ser expressa por:

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\frac{1}{N} \sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\int w_n(\theta_{1:n}) \left[\frac{1}{N} \sum_{i=1}^N \delta_{\Theta_{1:n}^i}(\theta_{1:n}) \right] d\theta_{1:n}}$$

Aproximação

A distribuição de interesse pode ser expressa por:

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\frac{1}{N} \sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\int w_n(\theta_{1:n}) \left[\frac{1}{N} \sum_{i=1}^N \delta_{\Theta_{1:n}^i}(\theta_{1:n}) \right] d\theta_{1:n}}$$
$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \left[\int \delta_{\Theta_{1:n}^i}(\theta_{1:n}) d\theta_{1:n} \right]},$$

Aproximação

A distribuição de interesse pode ser expressa por:

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\frac{1}{N} \sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\int w_n(\theta_{1:n}) \left[\frac{1}{N} \sum_{i=1}^N \delta_{\Theta_{1:n}^i}(\theta_{1:n}) \right] d\theta_{1:n}}$$

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \left[\int \delta_{\Theta_{1:n}^i}(\theta_{1:n}) d\theta_{1:n} \right]},$$

sendo,

$$\int \delta_{\Theta_{1:n}^i}(\theta_{1:n}) d\theta_{1:n} = 1$$

$$\hat{Z}_n(\theta_{1:n}) = \sum_{i=1}^N w_n(\Theta_{1:n}^i)$$

Aproximação

Dessa forma,

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i)}$$

Aproximação

Dessa forma,

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i)}$$

Sendo:

$$W_n(\theta_{1:n}) = \frac{w_n(\Theta_{1:n}^i)}{\sum_{j=1}^N w_n(\Theta_{1:n}^j)}$$

Aproximação

Dessa forma,

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i)}$$

Sendo:

$$W_n(\theta_{1:n}) = \frac{w_n(\Theta_{1:n}^i)}{\sum_{j=1}^N w_n(\Theta_{1:n}^j)}$$

$$\hat{\pi}_n(\theta_{1:n}) = \sum_{i=1}^N W_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})$$

Aproximação

Dessa forma,

$$\hat{\pi}_n(\theta_{1:n}) = \frac{\sum_{i=1}^N w_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})}{\sum_{i=1}^N w_n(\Theta_{1:n}^i)}$$

Sendo:

$$W_n(\theta_{1:n}) = \frac{w_n(\Theta_{1:n}^i)}{\sum_{j=1}^N w_n(\Theta_{1:n}^j)}$$

$$\hat{\pi}_n(\theta_{1:n}) = \sum_{i=1}^N W_n(\Theta_{1:n}^i) \delta_{\Theta_{1:n}^i}(\theta_{1:n})$$

A aproximação Monte Carlo para a função $\phi_n(\cdot)$ é:

$$\hat{E}(\phi_n) = \frac{1}{N} \sum_{i=1}^N W_n(\Theta_{1:n}^i) \phi_n(\Theta_{1:n}^i).$$

Amostragem de importância sequencial (AIS)

Problema 2

Apesar do uso da amostragem de importância, amostrar $q_n(\theta_{1:n})$ pode ser **inviável devido a dimensão n** .

Problema 2

Apesar do uso da amostragem de importância, amostrar $q_n(\theta_{1:n})$ pode ser **inviável devido a dimensão n** .

Considere que a distribuição auxiliar escolhida, $q(\cdot)$, possa ser escrita por:

$$q(\theta_{1:n}) = q(\theta_{1:n-1})q(\theta_n|\theta_{1:n-1})$$

Problema 2

Apesar do uso da amostragem de importância, amostrar $q_n(\theta_{1:n})$ pode ser **inviável devido a dimensão n** .

Considere que a distribuição auxiliar escolhida, $q(\cdot)$, possa ser escrita por:

$$q(\theta_{1:n}) = q(\theta_{1:n-1})q(\theta_n|\theta_{1:n-1})$$

Tem-se:

$$q(\theta_{1:n}) = q(\theta_1) \prod_{k=2}^n q(\theta_k|\theta_{1:k-1})$$

Decomposição sequencial

Assim, os pesos no tempo $t = 2$:

$$w(\theta_{1:2}) = \frac{\gamma(\theta_{1:2})}{q(\theta_{1:2})} = \frac{\gamma(\theta_{1:2})}{q(\theta_1)q(\theta_2|\theta_1)}$$

Decomposição sequecial

Assim, os pesos no tempo $t = 2$:

$$w(\theta_{1:2}) = \frac{\gamma(\theta_{1:2})}{q(\theta_{1:2})} = \frac{\gamma(\theta_{1:2})}{q(\theta_1)q(\theta_2|\theta_1)}$$

Algebricamente,

$$\begin{aligned} w(\theta_{1:2}) &= \frac{\gamma(\theta_1)}{q(\theta_1)} \frac{\gamma(\theta_{1:2})}{\gamma(\theta_1)q(\theta_2|\theta_1)} \\ w(\theta_{1:2}) &= w_1(\theta_1) \frac{\gamma(\theta_{1:2})}{\gamma(\theta_1)q(\theta_2|\theta_1)} \end{aligned}$$

Generalização

Por indução, tem-se, $\forall t > 1$:

$$w_n(\theta_{1:n}) = w_{n-1}(\theta_{1:n-1}) \frac{\gamma(\theta_{1:n})}{\gamma(\theta_{1:n-1})q(\theta_n|\theta_{1:n-1})}$$

Generalização

Por indução, tem-se, $\forall t > 1$:

$$w_n(\theta_{1:n}) = w_{n-1}(\theta_{1:n-1}) \frac{\gamma(\theta_{1:n})}{\gamma(\theta_{1:n-1})q(\theta_n|\theta_{1:n-1})}$$

Pesos são atualizados **iterativamente**. Isso reduz a dimensão da amostragem para a dimensão de θ_t , em cada tempo t .

$$\alpha_n(\theta_{1:n}) = \frac{\gamma(\theta_{1:n})}{\gamma(\theta_{1:n-1})q(\theta_n|\theta_{1:n-1})}$$

Na contexto da filtragem, tem-se que $\gamma(\theta_{1:n}) = p(\theta_{1:n}, y_{1:n})$.

Assim:

$$\alpha_n(\theta_{1:n}) = \frac{p(\theta_{1:n}, y_{1:n})}{p(\theta_{1:n-1}, y_{1:n-1})q_n(\theta_n|\theta_{1:n-1})}$$

Na contexto da filtragem, tem-se que $\gamma(\theta_{1:n}) = p(\theta_{1:n}, y_{1:n})$.

Assim:

$$\alpha_n(\theta_{1:n}) = \frac{p(\theta_{1:n}, y_{1:n})}{p(\theta_{1:n-1}, y_{1:n-1})q_n(\theta_n|\theta_{1:n-1})}$$

$$\alpha_n(\theta_{1:n}) = \frac{p(y_n|\theta_{1:n-1}, \theta_n, y_{1:n-1})p(\theta_n|\theta_{1:n-1}, y_{1:n-1})p(\theta_{1:n-1}, y_{1:n-1})}{p(\theta_{1:n-1}, y_{1:n-1})q_n(\theta_n|\theta_{1:n-1})}$$

Na contexto da filtragem, tem-se que $\gamma(\theta_{1:n}) = p(\theta_{1:n}, y_{1:n})$.

Assim:

$$\alpha_n(\theta_{1:n}) = \frac{p(\theta_{1:n}, y_{1:n})}{p(\theta_{1:n-1}, y_{1:n-1})q_n(\theta_n|\theta_{1:n-1})}$$

$$\alpha_n(\theta_{1:n}) = \frac{p(y_n|\theta_{1:n-1}, \theta_n, y_{1:n-1})p(\theta_n|\theta_{1:n-1}, y_{1:n-1})p(\theta_{1:n-1}, y_{1:n-1})}{p(\theta_{1:n-1}, y_{1:n-1})q_n(\theta_n|\theta_{1:n-1})}$$

$$\alpha_n(\theta_{1:n}) = \frac{p(y_n|\theta_{1:n-1}, \theta_n, y_{1:n-1})p(\theta_n|\theta_{1:n-1}, y_{1:n-1})}{q_n(\theta_n|\theta_{1:n-1})}$$

Amostragem de importância sequencial

As propriedades Markovianas A.1 e A.2 garantem:

$$\alpha_n(\theta_{1:n}) = \frac{p(y_n|\theta_n)p(\theta_n|\theta_{n-1})}{q_n(\theta_n|\theta_{1:n-1})}$$

Amostragem de importância sequencial

As propriedades Markovianas A.1 e A.2 garantem:

$$\alpha_n(\theta_{1:n}) = \frac{p(y_n|\theta_n)p(\theta_n|\theta_{n-1})}{q_n(\theta_n|\theta_{1:n-1})}$$

Chamada de **peso de incremento**, a parte responsável pelo processo **sequencial** de estimação.

$$w_n(\theta_{1:n}) = w_{n-1}(\theta_{1:n-1})\alpha_n(\theta_{1:n})$$

Amostragem de Importância sequencial

Considere o processo sequencial de reponderação do processo de Amostragem sequencial de Importância:

$$w_t \propto \frac{\pi(y_t|\theta_t)\pi(\theta_t|\theta_{t-1})}{g_{t|t-1}(\theta_t|\theta_{0:t-1}, y_{1:t})} w_{t-1}$$

Amostragem de Importância sequencial

Considere o processo sequencial de reponderação do processo de Amostragem sequencial de Importância:

$$w_t \propto \frac{\pi(y_t|\theta_t)\pi(\theta_t|\theta_{t-1})}{g_{t|t-1}(\theta_t|\theta_{0:t-1}, y_{1:t})} w_{t-1}$$

A função $g_{t|t-1}(\theta_t|\theta_{0:t-1}, y_{1:t})$ é a responsável por gerar as propostas de partículas, conhecida como **função de transição de importância**.

Amostragem de Importância sequencial

Considere o processo sequencial de reponderação do processo de Amostragem sequencial de Importância:

$$w_t \propto \frac{\pi(y_t|\theta_t)\pi(\theta_t|\theta_{t-1})}{g_{t|t-1}(\theta_t|\theta_{0:t-1}, y_{1:t})} w_{t-1}$$

A função $g_{t|t-1}(\theta_t|\theta_{0:t-1}, y_{1:t})$ é a responsável por gerar as propostas de partículas, conhecida como **função de transição de importância**.

Os tipos de filtro de partículas são definidos pelo tipo de equação $g_{t|t-1}(\cdot)$ escolhida.

Reamostragem

Degeneração

É possível que a distribuição dos pesos w_t^i se degenerem. Isso ocorre quando tem-se pesos com valores próximos de 1.

Degeneração

É possível que a distribuição dos pesos w_t^i se degenerem. Isso ocorre quando tem-se pesos com valores próximos de 1.

Adota-se então um critério de degeneração da distribuição dos pesos calculando o seguinte valor em cada vetor de partículas:

$$N_{eff}^t = \frac{1}{\sum_i^N (w_t^i)^2}$$

Degeneração

É possível que a distribuição dos pesos w_t^i se degenerem. Isso ocorre quando tem-se pesos com valores próximos de 1.

Adota-se então um critério de degeneração da distribuição dos pesos calculando o seguinte valor em cada vetor de partículas:

$$N_{eff}^t = \frac{1}{\sum_i^N (w_t^i)^2}$$

[?] indica utilizar de $N/2$ como critério para regeneração das partículas, substituindo os pesos das partículas por $1/N$.

O processo de reamostragem está presente no processo de estimação para reduzir a discrepância, o que em casos extremos causa a chamada **degeneração dos pesos de importância**.

O processo de reamostragem está presente no processo de estimação para reduzir a discrepância, o que em casos extremos causa a chamada **degeneração dos pesos de importância**.

- ▶ Reamostragem Multinomial;
- ▶ Reamostragem Residual;
- ▶ Reamostragem Estratificada;
- ▶ Reamostragem Sistemática.

Reamostragem

O processo de reamostragem está presente no processo de estimação para reduzir a discrepância, o que em casos extremos causa a chamada **degeneração dos pesos de importância**.

- ▶ Reamostragem Multinomial;
- ▶ Reamostragem Residual;
- ▶ Reamostragem Estratificada;
- ▶ Reamostragem Sistemática.

Reamostrar com probabilidade proporcional ao peso de importância.

Reamostragem multinomial

$$\{\theta_t^i, w_t^i\}, i = 1, \dots, N$$

Reamostragem multinomial

$$\{\theta_t^i, w_t^i\}, i = 1, \dots, N$$

As partículas θ_t^i são reamostradas com w_t^i como peso

Reamostragem multinomial

$$\{\theta_t^i, w_t^i\}, i = 1, \dots, N$$

As partículas θ_t^i são reamostradas com w_t^i como peso As novas partículas tem peso $w_t^i = 1/N$.

Reamostragem multinomial

$$\{\theta_t^i, w_t^i\}, i = 1, \dots, N$$

As partículas θ_t^i são reamostradas com w_t^i como peso As novas partículas tem peso $w_t^i = 1/N$.

Partículas com baixa probabilidade são descartadas

Estimação

Filtro de Partículas

Método

Filtro Bootstrap

Sistema Dinâmico

Proposto por [?].

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t), \quad f : \mathbb{R}^n \times \mathbb{R}^m \Rightarrow \mathbb{R}^n$$

Sistema Dinâmico

Proposto por [?].

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t), \quad f : \mathbb{R}^n \times \mathbb{R}^m \Rightarrow \mathbb{R}^n$$

Equação das observações

$$y_t = h(\theta_t, v_t), \quad h : \mathbb{R}^n \times \mathbb{R}^r \Rightarrow \mathbb{R}^p$$

Sistema Dinâmico

Proposto por [?].

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t), \quad f : \mathbb{R}^n \times \mathbb{R}^m \Rightarrow \mathbb{R}^n$$

Equação das observações

$$y_t = h(\theta_t, v_t), \quad h : \mathbb{R}^n \times \mathbb{R}^r \Rightarrow \mathbb{R}^p$$

Considere:

- ▶ $w_t \sim p_1(\cdot)$ e $v_t \sim p_2(\cdot)$
- ▶ $p(\theta_1 | D_0) = p(\theta_1)$

Objetivo

Considerare a *posteriori*:

$$p(\theta_t|D_t) = \frac{h(y_t|\theta_t)p(\theta_t|D_{t-1})}{\int h(y_t|\theta_t)p(\theta_t|D_{t-1})d\theta_t}$$

Objetivo

Considere a *posteriori*:

$$p(\theta_t|D_t) = \frac{h(y_t|\theta_t)p(\theta_t|D_{t-1})}{\int h(y_t|\theta_t)p(\theta_t|D_{t-1})d\theta_t}$$

Três grandes tarefas:

3: *Posteriori*

$$\overbrace{p(\theta_t|D_t)}$$

=

2: Atualização

$$\overbrace{\left[\frac{h(y_t|\theta_t)}{\int h(y_t|\theta_t)p(\theta_t|D_{t-1})d\theta_t} \right]}$$

1: Propagação

$$\overbrace{p(\theta_t|D_{t-1})}$$

1: Propagação

Por meio do suporte de $p(\theta_{t-1}|D_{t-1})$,

$$p(\theta_t|D_{t-1}) = \int f(\theta_t|\theta_{t-1})p(\theta_{t-1}|D_{t-1})d\theta_{t-1}$$

1: Propagação

Por meio do suporte de $p(\theta_{t-1}|D_{t-1})$,

$$p(\theta_t|D_{t-1}) = \int f(\theta_t|\theta_{t-1})p(\theta_{t-1}|D_{t-1})d\theta_{t-1}$$

e “atravessando” a equação dos estados latente do sistema pelo suporte de w_t ,

$$p(\theta_t|\theta_{t-1}) = \int f(\theta_t|\theta_{t-1}, w_t)p(w_t|\theta_{t-1})dw_t$$

1: Propagação

Por meio do suporte de $p(\theta_{t-1}|D_{t-1})$,

$$p(\theta_t|D_{t-1}) = \int f(\theta_t|\theta_{t-1})p(\theta_{t-1}|D_{t-1})d\theta_{t-1}$$

e “atravessando” a equação dos estados latente do sistema pelo suporte de w_t ,

$$p(\theta_t|\theta_{t-1}) = \int f(\theta_t|\theta_{t-1}, w_t)p(w_t|\theta_{t-1})dw_t$$

obtém-se de forma determinística,

$$p(\theta_t|D_{t-1}) = \int \int f(\theta_t|\theta_{t-1}, w_t)p(w_t)p(\theta_{t-1}|D_{t-1})dw_t d\theta_{t-1}$$

1: Propagação

Com isso, pode-se obter $p(\theta_t | D_{t-1})$, faz-se:

1: Propagação

Com isso, pode-se obter $p(\theta_t|D_{t-1})$, faz-se:

- 1- gerar suporte de θ_{t-1} a partir de $p(\theta_{t-1}|D_{t-1})$. **Todo suporte da distribuição θ_{t-1} .**

1: Propagação

Com isso, pode-se obter $p(\theta_t|D_{t-1})$, faz-se:

- 1- gerar suporte de θ_{t-1} a partir de $p(\theta_{t-1}|D_{t-1})$. **Todo suporte da distribuição θ_{t-1} .**
- 2- gerar suporte de w_t a partir de $p(w_t)$. **Todo suporte da distribuição w_t .**

1: Propagação

Com isso, pode-se obter $p(\theta_t|D_{t-1})$, faz-se:

- 1- gerar suporte de θ_{t-1} a partir de $p(\theta_{t-1}|D_{t-1})$. **Todo suporte da distribuição θ_{t-1} .**
- 2- gerar suporte de w_t a partir de $p(w_t)$. **Todo suporte da distribuição w_t .**
- 3- Obter de forma determinística, o suporte da distribuição de interesse por meio de $\{\theta_{t-1}, w_t\}$, **obtidos pela função $h(\cdot)$.**

1: Propagação

Com isso, pode-se obter $p(\theta_t|D_{t-1})$, faz-se:

- 1- gerar suporte de θ_{t-1} a partir de $p(\theta_{t-1}|D_{t-1})$. **Todo suporte da distribuição** θ_{t-1} .
- 2- gerar suporte de w_t a partir de $p(w_t)$. **Todo suporte da distribuição** w_t .
- 3- Obter de forma determinística, o suporte da distribuição de interesse por meio de $\{\theta_{t-1}, w_t\}$, **obtidos pela função** $h(\cdot)$.

A equação de interesse é representada por:

$$p(\theta_t|D_{t-1}) = \int \int \overbrace{h(\theta_t|\theta_{t-1}, w_{t-1})}^3 \overbrace{p(w_{t-1})}^2 \overbrace{p(\theta_{t-1}|D_{t-1})}^1 dw_{t-1} d\theta_{t-1}$$

2: Atualização

A distribuição gerada, “atravessa” a equação observável do sistema pelo suporte de v_t ,

$$p(y_t|\theta_t) = \int h(y_t|\theta_t, v_t)p(v_t)dv_t,$$

2: Atualização

A distribuição gerada, “atravessa” a equação observável do sistema pelo suporte de v_t ,

$$p(y_t|\theta_t) = \int h(y_t|\theta_t, v_t)p(v_t)dv_t,$$

e é avaliada por:

$$\pi_t \propto h(y_t|\theta_t, v_t)$$

.

3: Posteriori

Obtém-se a distribuição de $p(\theta_t|D_t)$ por meio da combinação entre **1:Propagação** e **2:Atualização**.

3: Posteriori

Obtém-se a distribuição de $p(\theta_t|D_t)$ por meio da combinação entre **1:Propagação** e **2:Atualização**.

$$\overbrace{p(\theta_t|D_t)}^{3:Posteriori} = \overbrace{\left[\frac{h(y_t|\theta_t)}{\int h(y_t|\theta_t)p(\theta_t|D_{t-1})d\theta_t} \right]}^{2:Atualização} \underbrace{p(\theta_t|D_{t-1})}_{1: Propagação}$$

Algoritmo

1.1 Para $t = 1$: gerar N amostras $\{\theta_0^i, i = 1, \dots, N\} \sim p(\theta_0)$;

1.2 Para $t > 1$: gerar N amostras
 $\{\theta_{t-1}^i, i = 1, \dots, N\} \sim p(\theta_{t-1}|D_{t-1})$

2 Gerar N amostra para $w_t^i \sim p_1(w)$

3 Obter valores de θ_t^{i*} , de forma determinística,
 $\theta_{t*}^i = f(\theta_{t-1}^i, w_t^i)$

4 Sendo v_t uma estatística conhecida, atualiza-se o peso de θ_t^{i*} usando:

$$\pi_t^i = \frac{p(y_t|\theta_t^{i*}, v_t)}{\sum_j^N p(y_t|\theta_t^{j*}, v_t)}$$

5 Reamostrar N vezes $\{\theta_t^{i*}, i = 1, \dots, N\}$ com probabilidade igual a π_t^i .

Laloux, L., Cizeau, P., Bouchaud, J. P., and Potters, M. (1999). Noise dressing of financial correlation matrices. *Physical Review Letters*, 83(7):1467–1470.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance*.

I