Antecipação e adaptação: como incorporar o dinamismo do mundo financeiro

Igor Nascimento

Laboratório de Aprendizado de Máquina em Finanças e Organizações - LAMFO

28/02/2018

Sumário

Mundo financeiro

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

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

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:

$$r_1, r_2, ..., r_N$$

- retorno:

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

- variância:

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

- covariância (correlação):

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

Alocação

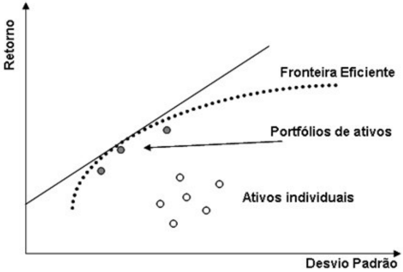
Determinar a locação, isto é, o percentual $p_1, p_2, ..., p_N$ que cada ativo representa da carteira:

$$E_{portfolio} = E(\mathbf{P}) = \sum_{i=1}^{N} p_i \times \mu_i \tag{1}$$

$$V_{portfolio} = V(\mathbf{P}) = \sum_{i=1}^{N} \sum_{j=1}^{N} p_i p_j \times \sigma_i \sigma_j \times \rho_{ij} \tag{2}$$

Fronteira Eficiente

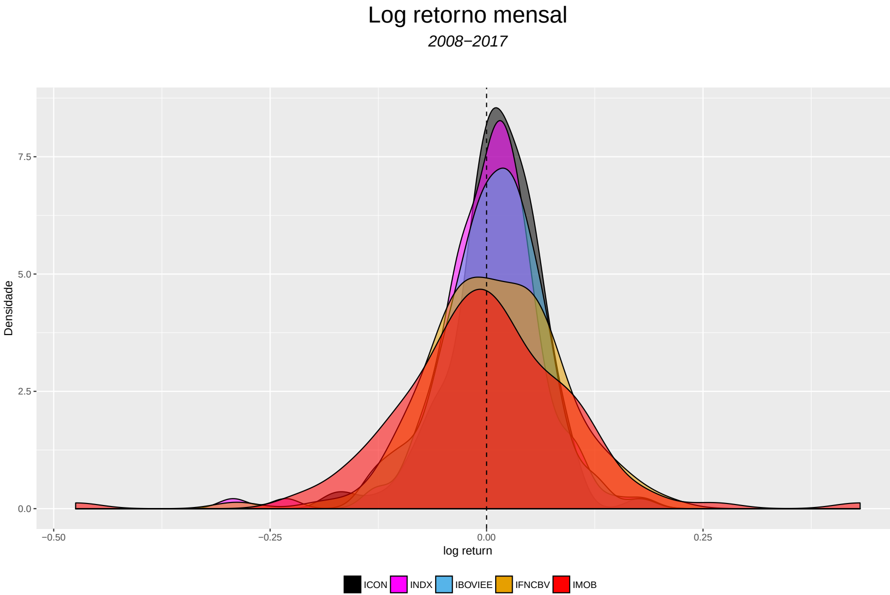
"Optimal weight of each asset, such that the overall portfolio provides the best return for a fixed level of risk, or conversely, the smallest risk for a given overall return?" Laloux et al. (1999)



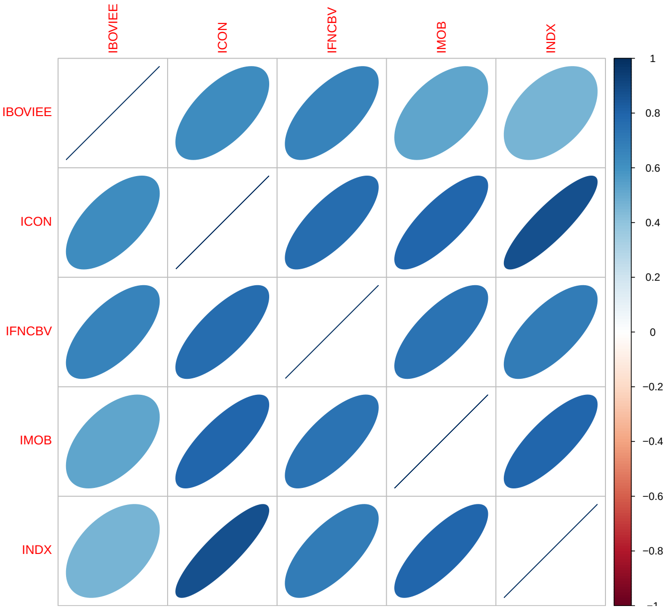
Minicaso

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

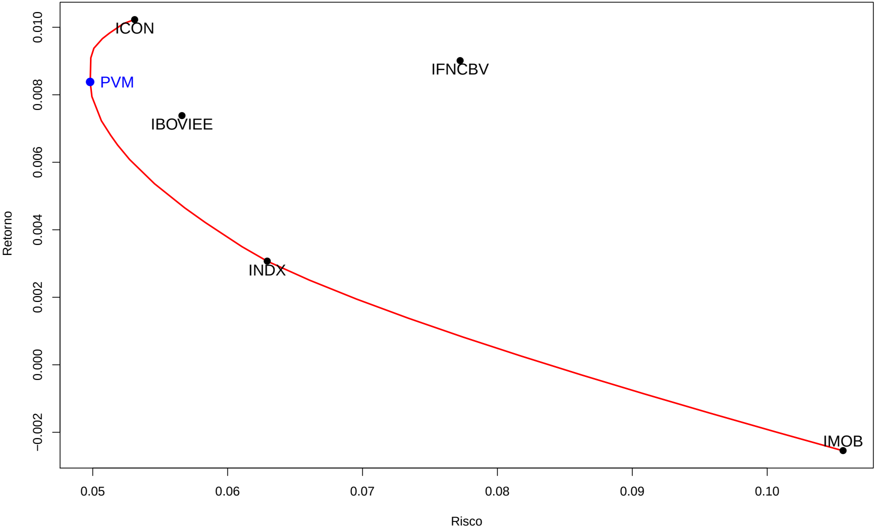
Índices IBOVESPA



Estatísticas



Fronteira



ICON
IFNCBV
PVM
IBOVIEE
INDX
IMOB
Retorno
Risco
0.010
0.008
0.006
0.004
0.002
0.000
−0.002
0.05
0.06
0.07
0.08
0.09
0.10

Amostragem aleatória

Stochastic Dynamic Programming

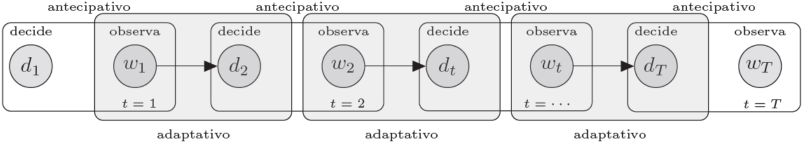
Um problema de programação dinâmica estocástica pode ser considerado, (Consigli and Dempster, 1998):

1. um problema em múltiplos estágios recursivo;
2. um processo de decisões d_t em $t = 1, \cdots, T$;
3. com restrições conhecidas e parâmetros w_t aleatórios.

Ao longo de todo o período $t = 1, \cdots, T$, (Kouwenberg and Zenios, 2008):

1. as decisões em d_t estão intercaladas pelo que foi observado no instante w_{t-1} e o que não se conhece sobre w_t
2. A relação entre a decisão tomada d_{t-1} e o parâmetro de incerteza w_{t-1} é **antecipativa** (condições de incerteza)
3. enquanto a relação entre d_t e w_{t-1} é **adaptativa** (ambiente de aprendizagem)

Estrutura de decisão



Estrutura de decisão

$$
\begin{aligned}
\text{Minimize:} \quad & Z = f(d_t) + E\left(q(d_t^*, w_t)\right) \\
\text{Sujeito} \quad & Ad_t = \text{b} \\
& W(w_t)d_t^* = h(w_t) - T(w_t)d_t
\end{aligned}
\tag{3}
$$

Otimizar, no instante t:

- função $f(d_t)$ de custo formada pelo estágio antecipativo
- e o valor esperado da função custo adaptativa $q(d_t^*, w_t)$.

Estrutura de decisão

$$
\begin{aligned}
\text{Minimize:} \quad & Z = f(d_t) + E\left(q(d_t^*, w_t)\right) \\
\text{Sujeito} \quad & Ad_t = \mathrm{b} \\
& W(w_t)d_t^* = h(w_t) - T(w_t)d_t
\end{aligned}
\tag{3}
$$

Otimizar, no instante t:

- função $f(d_t)$ de custo formada pelo estágio antecipativo
- e o valor esperado da função custo adaptativa $q(d_t^*, w_t)$.

1. A: matriz tecnológica para d_t
2. b: vetor de recursos para d_t
3. $T(w_t)$: matriz tecnológica que transforma d_t em recursos d_t^*
4. $h(w_t)$: de recursos para d_t^*
5. A:matriz tecnológica para d_t^*

Cenário discretizado

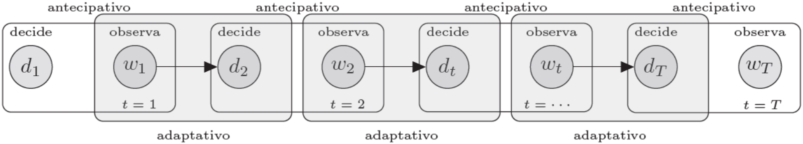
Discretizar o parâmetro de incerteza w_t em **cenários** $\Omega = \{w_t^1, w_t^2, \cdots, w_t^N\}$ (Kouwenberg and Zenios, 2008).

- $p^{(l)} = \pi(w_t = w_t^{(l)})$
- $\sum_{l=1}^{N} p^{(l)} = 1$

$$
\begin{aligned}
\text{Minimize:} \quad & Z = f(d_t) + \sum_{l=1}^{N} p^{(l)} q(d_t^{*(l)}, w_t^{(l)}) \\
\text{Sujeito} \quad & Ad_t = \mathrm{b} \\
& W(w_t^{(l)}) d_t^{*(l)} = h(w_t^{(l)}) - T(w_t^{(l)}) d_t
\end{aligned}
\tag{4}
$$

Estrutura de decisão

1. O processo de otimização se repete dinamicamente a medida em que se tenha acesso a novas informações sobre w_t.
2. A decisão d_t^* no instante t é a decisão d_{t+1} do instante $t+1$.



Função objetivo

A função objetivo pode ser adaptada para:

- minimizar custo de administração (Kouwenberg and Zenios, 2008)
- maximizar o valor final da carteira (Johannes et al., 2014)
- minimizar o risco (Ferstl and Weissensteiner, 2011) e (Quaranta and Zaffaroni, 2008).

Amostragem aleatória (Intuitivo)

Value at Risk (VaR)

Definida a carteira ótima:

- em um dia ruim, quanto eu posso perder?

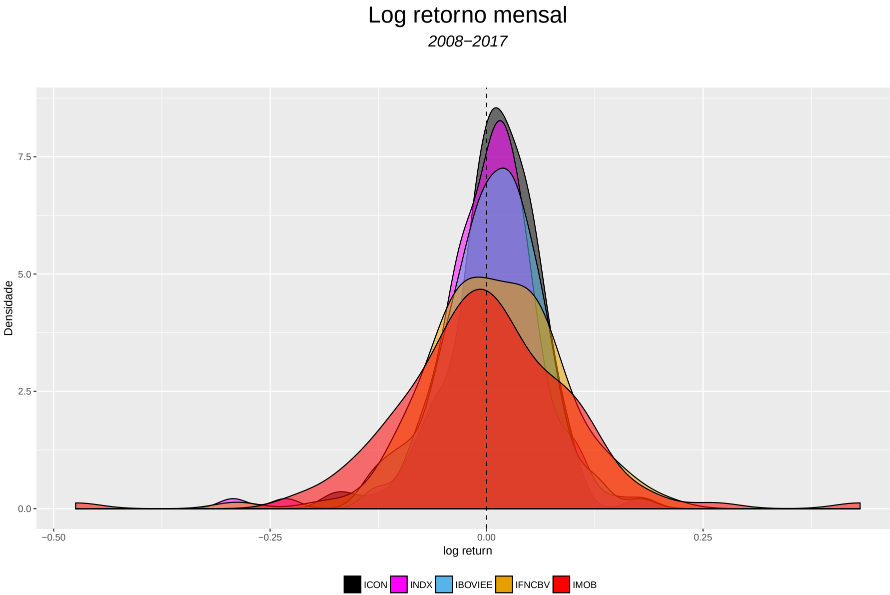
Value at Risk (VaR)

Definida a carteira ótima:

- em um dia ruim, quanto eu posso perder?
- quantos dias ruins eu suportaria para manter esse portfólio?

Monte Carlo

Monte Carlo



Log retorno mensal
2008−2017
Densidade
7.5
5.0
2.5
0.0
−0.50
−0.25
0.00
0.25
log return
ICON
INDX
IBOVIEE
IFNCBV
IMOB

Monte Carlo

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

Normal Multivariada

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

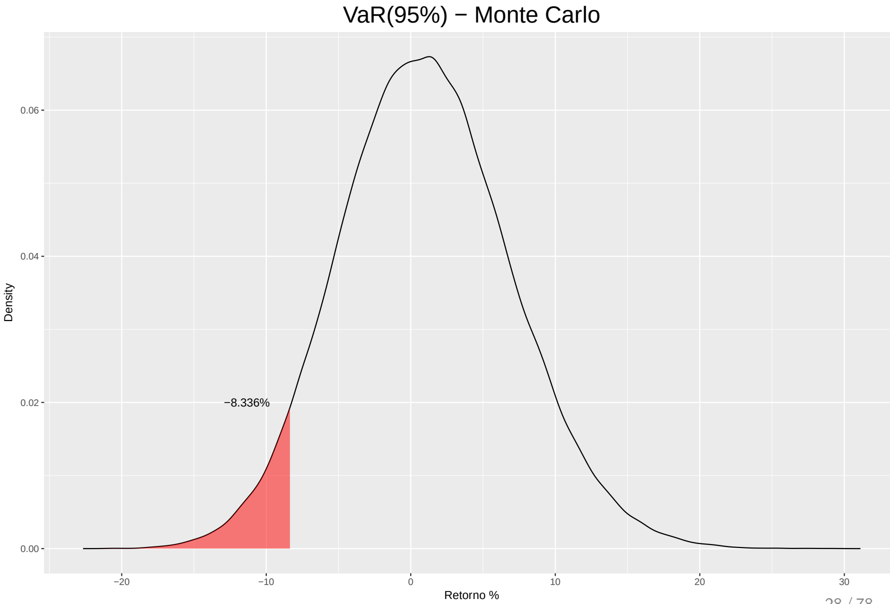
$$f(X) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{det\Sigma}} exp\left(-\frac{1}{2}(\tilde{X} - \tilde{\mu})\Sigma^{-1}(\tilde{X} - \tilde{\mu})\right)$$

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

Bootstrap

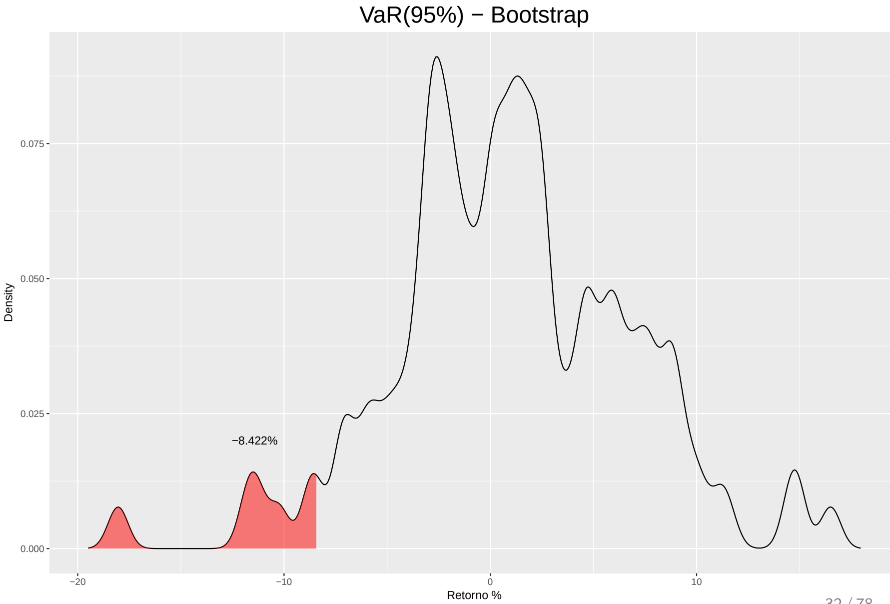
- descrição: (re)amostragem "não-paramétrica"
- vantagens: não requer conhecimento sobre a distribuição
- desvantagem: acessa espaço "realizado"

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 divida em janelas 5 meses

Value at Risk (Var) - Bootstrap



Modelos Dinâmicos

Modelo dinâmico

Parte não observável - sistema

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

Modelo dinâmico

Parte não observável - sistema

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

Parte observável

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

Modelo dinâmico

$$\begin{tikzcd}
\theta_0 \arrow[r] & \theta_1 \arrow[r] \arrow[d] & \theta_2 \arrow[r] \arrow[d] & \cdots \arrow[r] \arrow[d] & \theta_t \arrow[r] \arrow[d] & \theta_{t+1} \arrow[d] \\
 & Y_1 & Y_2 & \cdots & Y_t & Y_{t+1}
\end{tikzcd}$$

Modelo dinâmico

$$
\begin{tikzcd}
\theta_0 \arrow[r] & \theta_1 \arrow[r] \arrow[d] & \theta_2 \arrow[r] \arrow[d] & \cdots \arrow[r] \arrow[d] & \theta_t \arrow[r] \arrow[d] & \theta_{t+1} \arrow[d] \\
 & Y_1 & Y_2 & \cdots & Y_t & Y_{t+1}
\end{tikzcd}
$$

Equação do Sistema

$$\theta_t = G_t \theta_{t-1} + w_t$$

Modelo dinâmico

$$
\begin{tikzcd}
\theta_0 \arrow[r] & \theta_1 \arrow[r] \arrow[d] & \theta_2 \arrow[r] \arrow[d] & \cdots \arrow[r] \arrow[d] & \theta_t \arrow[r] \arrow[d] & \theta_{t+1} \arrow[d] \\
 & Y_1 & Y_2 & \cdots & Y_t & Y_{t+1}
\end{tikzcd}
$$

Equação do Sistema

$$\theta_t = G_t\theta_{t-1} + w_t$$

Equação das Observações

$$y_t = F_t\theta_t + v_t$$

Propriedades

A.1: $\theta_t, t = 1, ..., n$ 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: $\theta_t, t = 1, ..., n$ 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 $t = 1, ..., n$.

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

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:

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

Seja v_t e w_t perturbações com distribuição Gaussiana e independentes, West and Harrison (1997) apresenta um esquema sequencial para o processo de filtragem, conhecido como **Filtro de Kalman**.

- **(a) Posteriori de θ em $t-1$:** $(\theta_{t-1}|D_{t-1}) \sim N[m_{t-1}, C_{t-1}]$;
- **(b) Atualização da priori em $t-1$:** $(\theta_t|D_{t-1}) \sim N[m_{t-1}, R_t]$;
- **(c) Predição para y em t:** $(y_t|D_{t-1}) \sim N[f_t, Q_t]$;
- **(d) Posteriori de θ em t:** $(\theta_t|D_t) \sim N[m_t, C_t]$;

Sendo: $R_t = C_{t-1} + W_t$, $f_t = m_{t-1}$, $Q_t = R_t + V_t$, $\theta_t = m_{t-1} + A_t e_t$, $C_t = A_t V_t$, $A_t = R_t/Q_t$ e $e_t = Y_t - f_t$.

Monte Carlo (formal)

Definição

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

Definição

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

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

Definição

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

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

Aproximação e dada por:

$$\hat{\pi}_n(\theta_1, ..., \theta_n) = \frac{1}{N} \sum_i^N \delta_{\Theta^i_{1:n}}(\theta_{1:n}),$$

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

Definição

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

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

Definição

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

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

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

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

Definição

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

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

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

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

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

Desvantagens

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

Desvantagens

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

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

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** $q(.)$ para obter o mesmo suporte da distribuição $\pi(.)$.

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

Problema 1

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

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

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

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

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,

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

Sendo:

$$W_n(\theta_{1:n}) = \frac{w_n(\Theta^i_{1:n})}{\sum_{j=1}^{N} w_n(\Theta^j_{1:n})}$$

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:

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

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

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 $\phi_n(.)$ 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, $q(.)$, 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, $q(.)$, possa ser escrita por:

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

Tem-se:

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

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

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

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 reponsável 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:

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

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**.

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 $g_{t|t-1}(.)$ 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:

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

Petris et al. (2009) indica utilizar de como critério para regeneração das partículas $N/2$

Reamostragem

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

Reamostragem

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

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

Reamostragem

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

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

partículas tem peso $w_t^i = 1/N$.

Reamostragem

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

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

Filtro de partículas

Partículas

Definição

Parte de um todo

Definição

Parte de um todo

1. Parte: Partículas.

Definição
Parte de um **todo**

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

Definição

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.

Filtro de Partículas

Filtro Bootstrap

Sistema Dinâmico

Proposto por Gordon et al. (1993).

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t),$$

Sistema Dinâmico

Proposto por Gordon et al. (1993).

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t),$$

Equação das observações

$$y_t = h(\theta_t, v_t),$$

Sistema Dinâmico

Proposto por Gordon et al. (1993).

Equação do sistema

$$\theta_t = f(\theta_{t-1}, w_t),$$

Equação das observações

$$y_t = h(\theta_t, v_t),$$

Considere:

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

Objetivo

Considere a *posteriori*:

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

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:

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

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** $h(.)$.

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** $h(.)$.

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**.

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

Algoritmo

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

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

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

Stochastic Volatility Models (SVM)

SVM

- inicialmente com distribuição gaussiana (Hull and White, 1987).
- atualmente distribuições de caudas pesadas e misturas de normais (Virbickaite et al., 2016).

Uma possível formulação para o modelo SVM é:

$$\begin{aligned} y_t &= exp\left(\frac{x_t}{2}\right)\varepsilon_t \\ x_t &= \alpha + \beta x_{t-1} + w_t \end{aligned} \tag{5}$$

Linearlização

Linearização $r_t = log(y_t^2)$ (Kim et al., 1998):

$$\begin{aligned} r_t &= x_t + v_t \\ x_t &= \beta x_{t-1} + w_t \end{aligned} \tag{6}$$

Caso $\varepsilon_t \sim N(0,1)$ a distribuição exata é $log(\chi^2)$, sendo χ^2 distribuição Qui-Quadrado. Aproximação com mistura de normais com os seguintes parâmetros:

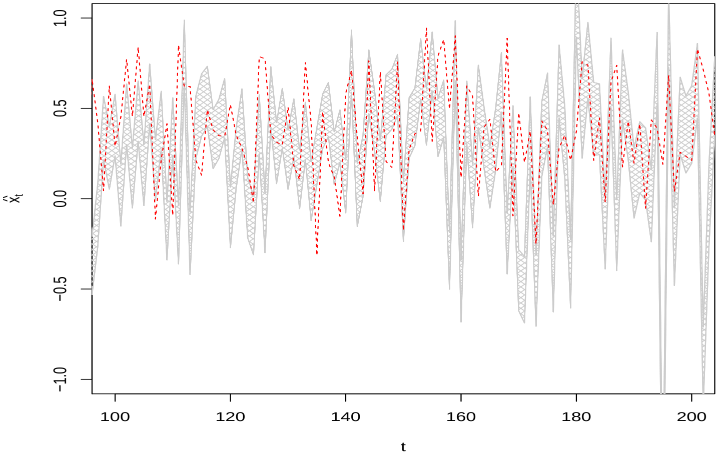
$$v_t \sim log(\chi^2) \approx \sum_{i=1}^{7} \pi_i f_N(\mu_i, \sigma^2) \tag{7}$$

Simulação

Considere o modelo SV básico com parâmetros

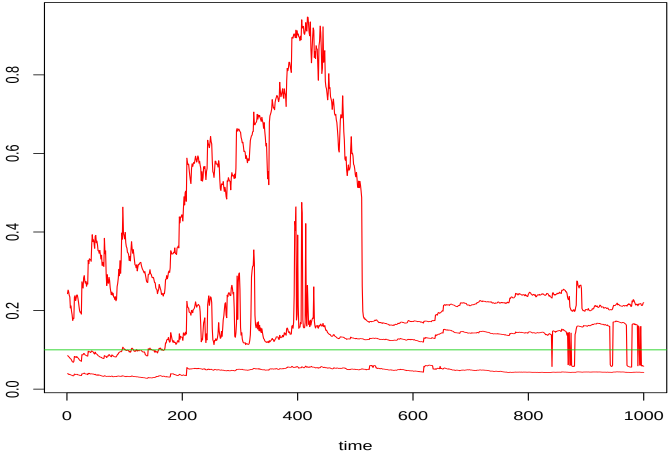
1. $\tau^2 = 0.1$
2. $\beta = -0.2$
3. $\alpha = 0.5$
4. simulado para um período de $T = 1000$
5. $M = 5000$ partículas

Estados latentes (variâncias)



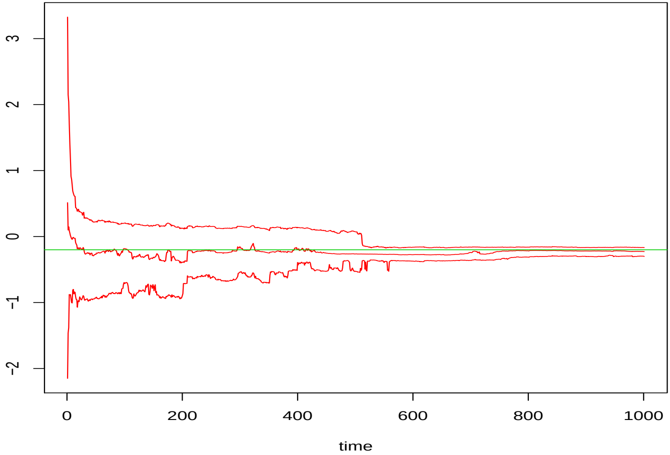
Volatilidade estimada pelo modelo SV simulado entre $t \in [100; 200]$. A região em cinza representa o intervalo de credibilidade ao nível de 95% e em vermelho os valores verdadeiros.

SVM - parâmetros



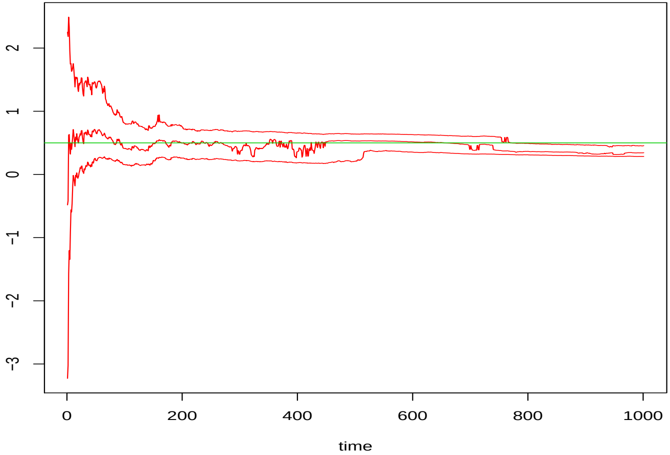
Estimativas dinâmicas para o parâmetro τ^2.

SVM - parâmetros



Estimativas dinâmicas para o parâmetro β

SVM - parâmetros



Estimativas dinâmicas para o parâmetro α

SVM - parâmetros

	$\tau = \sqrt{0.1} = 0.31$				$\beta = -0.2$				$\alpha = 0.5$			
	$\bar{\tau}$	$\tilde{\tau}$	$q_{5\%}$	$q_{95\%}$	$\bar{\beta}$	$\tilde{\beta}$	$q_{5\%}$	$q_{95\%}$	$\bar{\alpha}$	$\tilde{\alpha}$	$q_{5\%}$	$q_{95\%}$
Simulação	0.34	0.24	0.21	0.45	-0.23	-0.23	-0.29	-0.17	0.34	0.34	0.29	0.44

Distribuição a posteriori dos hiperparâmetros nos modelos SV simulado com $M = 5000$.

Aplicação

Simulação

Aplicação de R$ 1000 baseado nos dados 2008 – 2017 dos ativos:

- IFN: índice setor financeiro
- IMOB: índice do setor imobiliário
- ICON: índice de consumo
- IEE: índice de energia
- INDX: índice da indústria
- IBOV: índice geral da bolsa de São Paulo
- LTN: Letra do Tesouro Nacional

Ajustar um modelo SV para cada série individualmente.

IBOVESPA

1. simulação de 10000 carteiras aleatórias
2. 10000 cenários
3. utilizando os parâmetros estimados para o modelo SV
4. 252 períodos

Alocação no ativo (%)							Valor final da carteira			
ICON	IEE	IFN	IMOB	INDX	IBOV	LTN	$VaR_{99\%}$	$VaR_{95\%}$	$\bar{R}$	$\sigma(W)$
11.21	2.67	11.81	1.27	35.79	9.16	28.09	624.84	711.95	7.50	241.79
34.27	2.89	11.98	16.57	2.97	1.78	29.54	597.63	733.85	8.34	242.06
34.35	0.66	7.85	2.53	0.79	25.12	28.71	633.07	728.82	8.08	249.42
17.78	0.86	10.60	15.77	23.91	5.93	25.14	607.13	711.50	6.27	251.81
20.13	2.77	5.07	13.93	6.28	20.94	30.89	611.37	701.44	7.00	254.07

Informações das 5 melhores carteiras no modelo SV independente. Análise de 10000 carteiras, simuladas em 10000 cenários de 252 períodos.

IBOVESPA

- retorno médio esperado de 7.5%
- *VaR* ao nível 99% de R$ 624.84
- *VaR* ao nível 95% de R$ 711.95
- pelo menos 25% dos recursos alocados no ativo livre de risco.

Obrigado

Perguntas?

igor.ferreira.n@gmail.com

lamfo.unb.br

lamfo-unb.github.io

Consigli, G. & Dempster, M. a. H. (1998). Annals of Operations Research, 81(October):131 – 161.

Ferstl, R. & Weissensteiner, A. (2011). Asset-liability management under time-varying investment opportunities. Journal of Banking and Finance, 35(1):182–192.

Gordon, N. J., Salmond, D. J., & Smith, A. F. M. (1993). Novel approach to nonlinear/non-gaussian bayesian state estimation. IEEE Proceedings F on Radar and Signal Processing, (140):107–113.

Hull, J. & White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. The Journal of Finance.

Johannes, M., Korteweg, A., & Polson, N. (2014). Sequential learning, predictability, and optimal portfolio returns. Journal of Finance, 69(2):611–644.

Kim, S., Shepherd, N., & Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. Review of Economic Studies, 65(3):361–393.

Kouwenberg, R. & Zenios, S. A. (2008). Stochastic Programming Models for Asset Liability Management. In: Handbook of Asset and Liability Management - Set, volume 1, pages 253–303. incollection.

Laloux, L., Cizeau, P., Bouchaud, J. P., & 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.

Petris, G., Petrone, S., & Campagnoli, P. (2009). Dynamic Linear Models with R. Springer.

Quaranta, A. G. & Zaffaroni, A. (2008). Robust optimization of conditional value at risk and portfolio selection. Journal of Banking and Finance, 32(10):2046–2056.

Virbickaite, A., Lopes, H. F., Ausín, M. C., & Galeano, P. (2016). Particle Leaning for Bayesian Non-Parametric Markov Switching Stochastic Volatility Model. R&R Bayesian Analysis, 2(1):1–28.

West, M. & Harrison, J. (1997). Bayesian Forecast and Dynamic Models. Springer-Verlag.

I