Применение Фильтра Калмана

Павел Мозгунов

31 января 2014 г.

Модели пространства состояний (SSM)

X_k - ненаблюдаемая переменная (*состояние*), k=1,2,...,N.

Y_k - наблюдаемая переменная, k=1,2,...,N.

(1) $X_{k+1} = A_k X_k + V_{k+1} \quad \in R^m, \quad k=0,1,\ldots$ *(уравнение состояний)*

(2) $Y_k = C_k X_k + W_k \quad \in R^d, \quad k=0,1,\ldots$ *(уравнения наблюдений)*

A_k - $(m \times m)$, C_k - $(d \times m)$

V_k - модельный шум. V_k *iid* $N(0, Q_k)$,

W_k - шум наблюдений, W_k *iid* $N(0, R_k)$.

$X_0 \sim N(\mu, \Sigma)$

$E[V_k W_l^*] = 0.$

V_k, W_l и X_0 - независимы.

Модели представимые в форме пространства состояний

- Регресионная модель с меняющимися коэффициентами (TVP)
- Процесс скользящего среднего - MA(q)
- Авторегрессионный процесс - AR(p)
- Модель авторегрессии — скользящего среднего - ARMA (p,q)
- Сезонная модель с шумом.
- Динамическая факторная модель.
- Динамическая факторная модель с общим стохастическим трендом.
- ...

Пример

Модель с ненаблюдаемыми компонентами.(Clark, 1987)

$y_k = n_k + z_k$ (логарифм реального ВВП)

$n_{k+1} = \rho n_k + g_k + \varepsilon_{k+1}$ (стохастический тренд)

$g_k = \gamma g_{k-1} + w_k$ (дрифт в тренде)

$z_k = \phi_1 z_{k-1} + \phi_2 z_{k-2} + e_k$ (стационарная циклическая компонента)

ε_k *iid* $\mathcal{N}(0, \sigma_\varepsilon^2)$; *k=1,2,...*;

w_k *iid* $\mathcal{N}(0, \sigma_w^2)$; *k=1,2,...*;

e_k *iid* $\mathcal{N}(0, \sigma_e^2)$; *k=1,2,...*;

$E[e_k \varepsilon_l] = E[e_k w_l] = E[\varepsilon_k w_l] = 0$ для любых k и l.

Пример

Модель с ненаблюдаемыми компонентами. (Clark, 1987)

$$X_k = \begin{bmatrix} n_k \\ z_k \\ z_{k-1} \\ g_k \end{bmatrix}; \quad X_{k+1} = \begin{bmatrix} \rho & 0 & 0 & 1 \\ 0 & \phi_1 & \phi_2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma \end{bmatrix} X_k + \begin{bmatrix} \varepsilon_{k+1} \\ e_{k+1} \\ 0 \\ w_{k+1} \end{bmatrix};$$

$$Y_k = [1, 1, 0, 0] X_k$$

Обозначения.

$\mathcal{Y}_k = \sigma\{Y_0, ..., Y_k\}$,$k$=0,1,2,...,$N$.

$\hat{X}_{k|k}$=$E[X_k|\mathcal{Y}_k]$

$\Sigma_{k|k}$= $E[(X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^*|\mathcal{Y}_k]$

ν_k - инновации (innovations)

H_k - дисперсия инноваций

G_k - усиление Калмана (Kalman Gain)

Фильтр Калмана (1961)

θ_k= $[\mu, \Sigma, A_k, C_k, Q_k, R_k, S_k]$ - известны. Необходимо найти:
$\hat{X}_{k|k}$=$E[X_k|\mathcal{Y}_k]$; $\Sigma_{k|k}$= $E[(X_k - \hat{X}_{k|k})(X_k - \hat{X}_{k|k})^*]$, k=0,1,2,. . . ,
$X_0 \sim N(\mu, \Sigma)$; $\quad Y_0 = C_0 X_0 + W_0$;
Используя свойство многомерного нормального распределения:
$\hat{X}_{0|0}$=$\mu + \Sigma {C_0}^* (C_0 \Sigma {C_0}^* + R_0)^{-1}(Y_0 - C_0\mu)$
$\Sigma_{0|0} = \Sigma - \Sigma {C_0}^* (C_0 \Sigma {C_0}^* + R_0)^{-1} C_0 \Sigma$
Затем, предположим, что на моменте k, нам известны,
$\hat{X}_{k|k}$
$\Sigma_{k|k}$
Прогнозируем ,используя уравнение (1):
$\hat{X}_{k+1|k}$=$A_k \hat{X}_{k|k}$
$\Sigma_{k+1|k}$=$A_k \Sigma_{k|k} {A_k}^* + Q_{k+1}$

Фильтр Калмана (1961)

Поступает новое наблюдение Y_{k+1} и начинается *"корректировка"*

$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + G_{k+1}\nu_{k+1}$

$cov(\theta, \theta|\xi) = \Sigma_{k+1|k+1}$

Фильтр Калмана (1961)

Уравнения прогнозирования:

$\hat{X}_{k+1|k} = A_k \hat{X}_{k|k}$

$\Sigma_{k+1|k} = A_k \Sigma_{k|k} {A_k}^* + Q_{k+1}$

Уравнения инноваций:

$\nu_{k+1} = Y_{k+1} - C_{k+1} \hat{X}_{k+1|k}$

$H_{k+1|k} = C_{k+1} \Sigma_{k+1|k} {C_{k+1}}^* + R_{k+1}$

Усиление Калмана:

$G_{k+1} = \Sigma_{k+1|k} {C_{k+1}}^* {H_{k+1|k}}^{-1}$

Уравнения корректировки:

$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k} + G_{k+1} \nu_{k+1}$

$\Sigma_{k+1|k+1} = (I - G_{k+1} C_{k+1}) \Sigma_{k+1|k}$

Сглаживание Калмана

Необходимо найти:

$\hat{X}_{k|N} = E[X_k | \mathcal{Y}_N]$

$\Sigma_{k|N} = E[(X_k - \hat{X}_{k|N})(X_k - \hat{X}_{k|N})^* | \mathcal{Y}_N]$.

Сглаживание Калмана:

$J_k = \Sigma_{k|k} A_k^* \Sigma_{k+1|k}^{-1}$

$\hat{X}_{k|N} = \hat{X}_{k|k} + J_k(\hat{X}_{k+1|N} - \hat{X}_{k+1|k})$

$\Sigma_{k|N} = \Sigma_{k|k} + J_k[\Sigma_{k+1|N} - \Sigma_{k+1|k}] J_k^*$

Пример. ФК и СК при известных матрицах

$$x_{k+1} = x_k + \varepsilon_{k+1} \; ; \quad \varepsilon_k \sim \mathcal{N}(0, 20)$$

$$\begin{bmatrix} y_{1k} \\ y_{2k} \end{bmatrix} = \begin{bmatrix} 1.2 \\ 0.65 \end{bmatrix} x_k + \begin{bmatrix} \epsilon_k \\ \eta_k \end{bmatrix};$$

$$\begin{pmatrix} \epsilon_k \\ \eta_k \end{pmatrix} \sim \mathcal{N}(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 80 & -20 \\ -20 & 100 \end{pmatrix})$$

$$x_0 \sim \mathcal{N}(100, 30)$$

Пример. ФК при известных матрицах

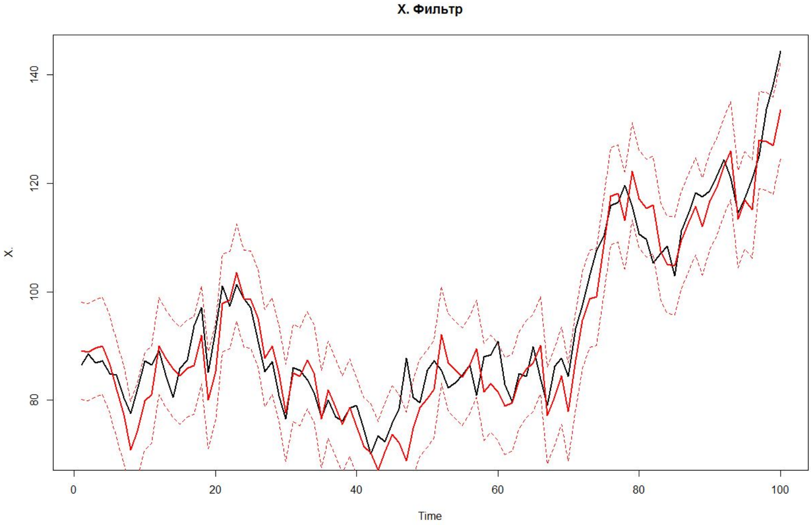
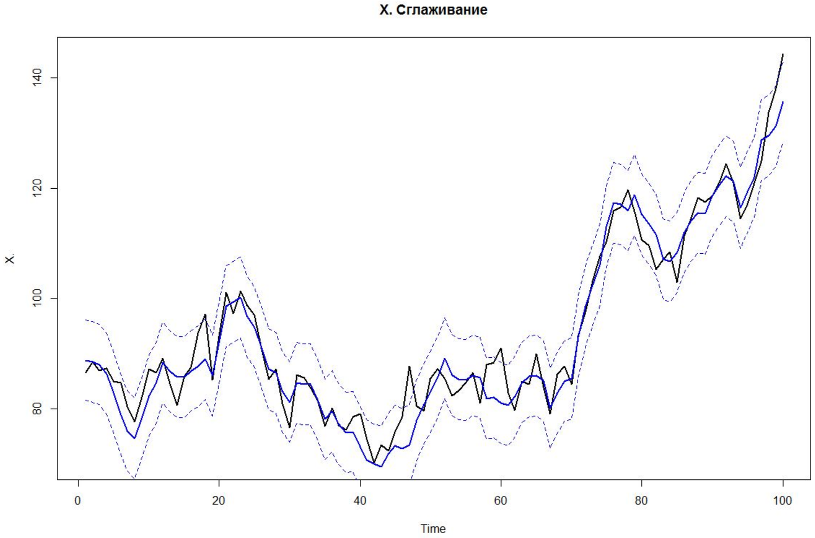


Рис. : Фильтр и Сглаживание при известных матрицах

Пример. СК при известных матрицах



Пример. ФК и СК при известных матрицах

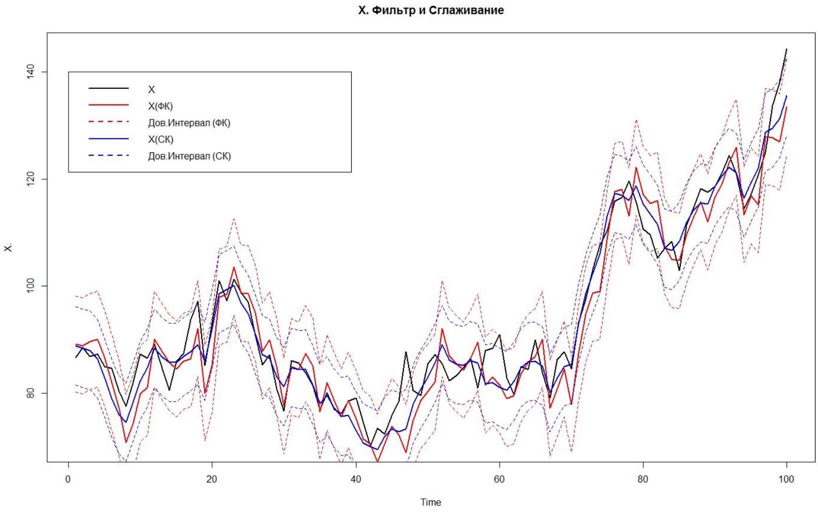


Рис. : Фильтр и Сглаживание при известных матрицах.

Методы оценивания параметров SSM

$\theta = [\mu, \Sigma, A, C, Q, R]$

Метод Максимального Правдоподобия

EM - алгоритм (Expectation - Maximization Algorithm)

Метод Максимального Правдоподобия

Зная оценку вектора параметров θ на предыдущем шаге:

$\nu_k(\theta) = y_k - C\hat{X}_{k|k-1}, k = 1, 2, ..., N$

$H_{k|k-1}(\theta) = C\Sigma_{k|k-1}C^* + R$

Далее строится функция правдоподобия для инноваций:

$$L_Y(\theta) = \prod_{k=1}^{N} \frac{1}{(2\pi)^{n/2}} |H_{k|k-1}(\theta)|^{-1/2} exp(-\frac{1}{2}\nu_k(\theta)^* H_{k|k-1}^{-1}\nu_k(\theta))$$

Максимизируя данное выражение получаем новую оценку $\hat{\theta}$, используя процедуры численной оптимизации.

Метод Максимального Правдоподобия

Алгоритм оценивания.

1. $\theta^{(0)}$
2. С помощью ФК вычисляются $\nu_k(\theta^{(0)})$ и $H_{k|k-1}(\theta^{(0)})$ для k=1,2,...,N. Вычисляется функция правдоподобия.
3. Полученная функция максимизируется по вектору параметров (θ).
4. Получаем новый вектор параметров $\theta^{(1)}$.
5. Имея, новый вектор параметров, повторяются шаги 2-4 до сходимости.

EM-алгоритм.
Необходимые понятия.

$\bar{P} : (\Omega, \mathcal{F}, (\mathcal{F}_t), \bar{P})$ ($\bar{P}$ reference probability).
$X_k \sim \mathcal{N}(0, I_n)$; $Y_l \sim \mathcal{N}(0, I_m)$
$P : (\Omega, \mathcal{F}, (\mathcal{F}_t), P)$, такую что справедливо (1) и (2):
$\frac{dP}{d\bar{P}} = \bar{\Lambda}_t$, где $\bar{\Lambda}_t = \prod_{k=0}^{t} \bar{\lambda}_k$.
$\phi(x)$ - функции плотности вероятности для x, и $\psi(y)$ - функция плотности вероятности для y.

$$\bar{\lambda}_k = \frac{|Q_k|^{-1/2}\phi(Q_k^{-1/2}(X_k - A_{k-1}X_{k-1})}{\phi(X_k)} \cdot \frac{|R_k|^{-1/2}\psi(R_k^{-1/2}(Y_k - C_k X_k))}{\psi(Y_k)}$$

ЕМ-алгоритм.

Необходимые понятия.

$$Q(\theta, \theta') = E_{\theta'}[log \frac{dP_\theta}{dP_{\theta'}} | \mathcal{Y}_N],$$

θ соответвуется вероятность P_θ.

θ' соответсвует вероятность P'_θ.

$$E_{\theta'}[\frac{dP_\theta}{dP_{\theta'}} | \mathcal{Y}_N] = \frac{L(\theta)}{L(\theta')}$$

Тогда исходя из неравенства Йенсена:

$$LogL(\theta) - LogL(\theta') \geq Q(\theta, \theta')$$

$$\theta_{j+1} = argmax_{(\theta)} \; Q(\theta, \theta')$$

ЕМ-алгоритм.

Е-шаг.

$$Q(\theta,\theta') = E_{\theta'}[log\frac{dP_\theta}{dP_{\theta'}}|\mathcal{Y}_N]$$

$$log\frac{dP_\theta}{dP_{\theta'}} = \sum_{l=0}^{N} log\bar{\lambda}_l + constant$$

Найдем условное математическое ожидание:

$$E_{\theta'}[log\frac{dP_\theta}{dP_{\theta'}}|\mathcal{Y}_N] = -\frac{1}{2}log|\Sigma| - \frac{N+1}{2}log|R| - \frac{N}{2}log|Q|$$

$$-\frac{1}{2}trace[Q^{-1}E[\sum_{k=1}^{N}(X_k - AX_{k-1})(X_k - AX_{k-1})^*|\mathcal{Y}_N]]$$

$$-\frac{1}{2}trace[R^{-1}E[\sum_{k=1}^{N}(Y_k - CX_k)(Y_k - CX_k)^*|\mathcal{Y}_N]]$$

$$-\frac{1}{2}trace[\Sigma^{-1}E[(X_0-\mu)(X_0-\mu)^*|\mathcal{Y}_N] + \text{constant}$$

EM-алгоритм.

M-шаг.

$$\mu = \hat{X}_{0|N}$$

$$C = \sum_{k=0}^{N} Y_k \hat{X}^*_{k|N} (\sum_{k=0}^{N} [\Sigma_{k|N} + \hat{X}_{k|N} \hat{X}^*_{k|N}])^{-1}$$

$$A = (\sum_{k=1}^{N} [\Sigma_{k,k-1|N} + \hat{X}_{k|N} \hat{X}^*_{k-1|N}])(\sum_{k=1}^{N} [\Sigma_{k-1|N} + \hat{X}_{k-1|N} \hat{X}^*_{k-1|N}])^{-1}$$

$$\Sigma = \Sigma_{0|N}$$

$$R = \frac{1}{N+1} [\sum_{k=0}^{N} Y_k Y_k^* - \sum_{k=0}^{N} Y_k \hat{X}^*_{k|N} (\sum_{k=0}^{N} [\Sigma_{k|N} + \hat{X}_{k|N} X^*_{k|N}])^{-1} \hat{X}_{k|N} Y_k^*]$$

$$Q = \frac{1}{N} [\sum_{k=0}^{N} [\Sigma_{k|N} + \hat{X}_{k|N} \hat{X}^*_{k|N} - \sum_{k=1}^{N} [\Sigma_{k-1,k|N} + \hat{X}_{k-1|N} \hat{X}^*_{k|N}] (\sum_{k=0}^{N} [\Sigma_{k-1|N} + \hat{X}_{k-1|N} \hat{X}^*_{k-1|N})^{-1} \sum_{k=1}^{N} [\Sigma_{k-1,k|N} + \hat{X}_{k|N} \hat{X}^*_{k-1|N}]]$$

Сравнение методов оценивания

Верно ли оцениваются параметры модели?

Верно ли оцениваются ненаблюдаемые переменные?

Устойчив ли результат к начальным значениям?

Симуляции
Модель 1

$$x_{k+1} = x_k + \varepsilon_{k+1} \ ; \quad \varepsilon_k \sim \mathcal{N}(0, 20)$$

$$\begin{bmatrix} y_{1k} \\ y_{2k} \end{bmatrix} = \begin{bmatrix} 0.65 \\ 1.2 \end{bmatrix} x_k + \begin{bmatrix} \epsilon_k \\ \eta_k \end{bmatrix};$$

$$\begin{pmatrix} \epsilon_k \\ \eta_k \end{pmatrix} \sim \mathcal{N}(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 80 & -20 \\ -20 & 100 \end{pmatrix})$$

$$x_0 \sim \mathcal{N}(100, 30)$$

Модель 1

Начальные значения.

$$\mu = 0;\ \Sigma = 1;\ A = 0.1;\ C = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix};\ Q = 1;\ R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

Модель 1

Результаты оценивания. C - неизвестна.

Параметр	Истинное	EM	MLE
μ	100	23.59	-2.86
Σ	30	0.02	50.98
A	1	0.997	0.997
C	$\begin{pmatrix} 0.65 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 2.89 \\ 5.37 \end{pmatrix}$	$\begin{pmatrix} 248.65 \\ 462.24 \end{pmatrix}$
Q	20	0.99	0.000135
R	$\begin{pmatrix} 80 & -20 \\ -20 & 100 \end{pmatrix}$	$\begin{pmatrix} 73.14 & -20.64 \\ -20.64 & 109.78 \end{pmatrix}$	$\begin{pmatrix} 73.08 & -20.68 \\ -20.68 & 109.8 \end{pmatrix}$

Таблица : Модель 1. Результаты оценивания EM и MLE. C- неизвестна.

Модель 1

Оценка состояний. С-неизвестна.

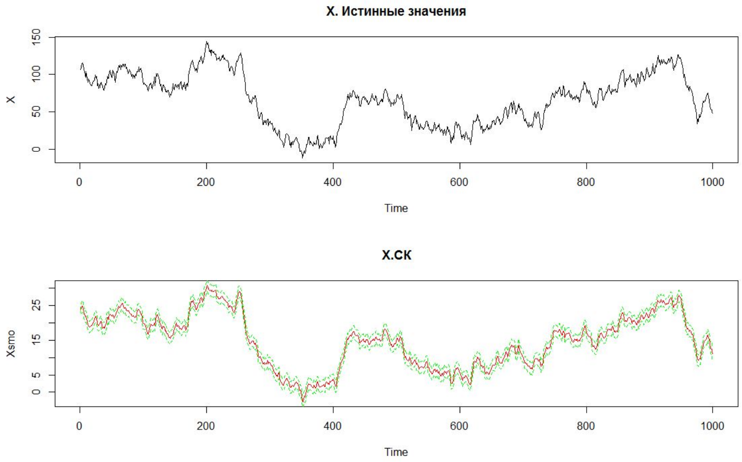


Рис. : Модель 1. Истинные и сглаженные значения при неизвестных матрицах.

Модель 1

Оценка состояний. С-неизвестна

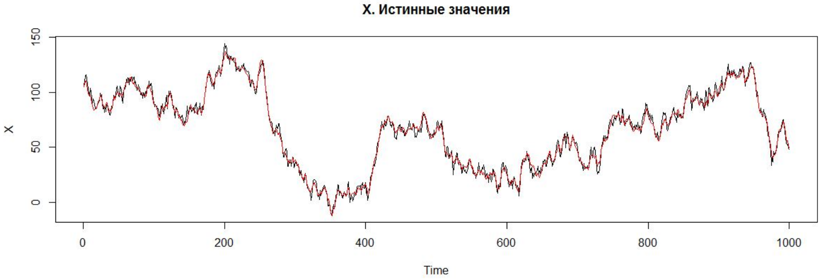


Рис. : Модель 1. Истинные и сглаженные значения при неизвестных матрицах.

Модель 1

Оценка параметров. C-известна.

Параметр	Истинное	EM	MLE
μ	100	99.48	1.74
Σ	30	0.0153	803.0674
A	1	0.997	0.997
C	$\begin{pmatrix} 0.65 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 0.65 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 0.65 \\ 1.2 \end{pmatrix}$
Q	20	19.989	19.64533
R	$\begin{pmatrix} 80 & -20 \\ -20 & 100 \end{pmatrix}$	$\begin{pmatrix} 73.06 & -20.64 \\ -20.64 & 107.06 \end{pmatrix}$	$\begin{pmatrix} 73.16 & -20.59 \\ -20.59 & 110.29 \end{pmatrix}$

Таблица : Модель 1. Результаты оценивания EM и MLE. C- известна.

Модель 1

Оценка состояний. С-известна.

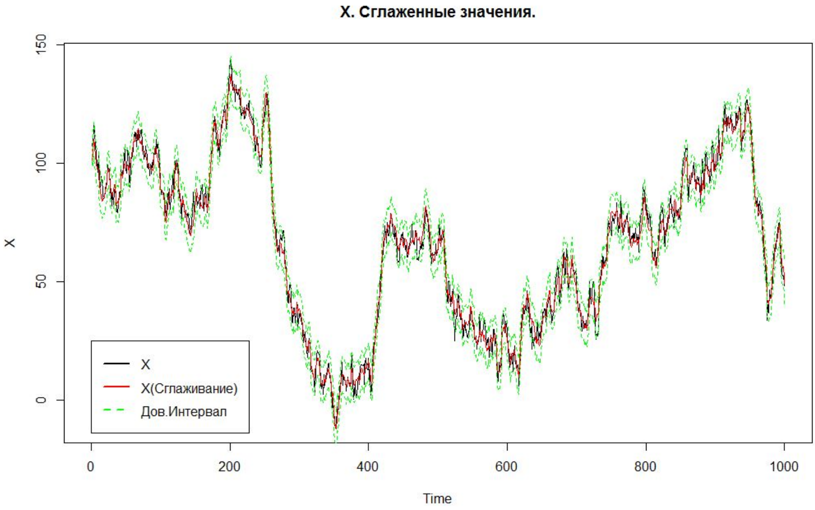
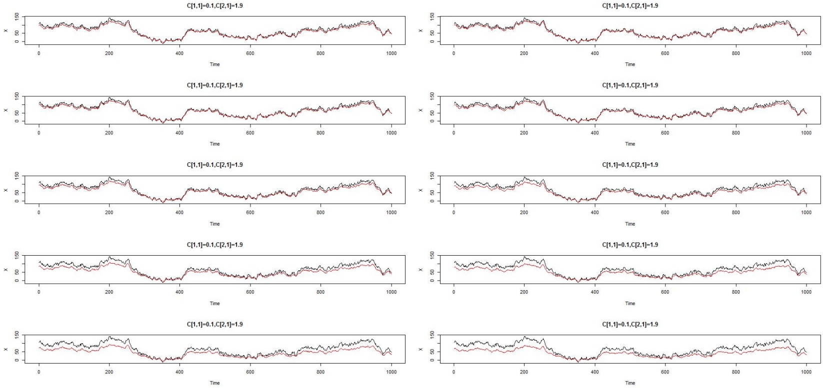


Рис. : Модель 1. Истинные и сглаженные значения при неизвестных

Модель 1.EM

Изменение начальных значений для C,при начальном $A = 1$.

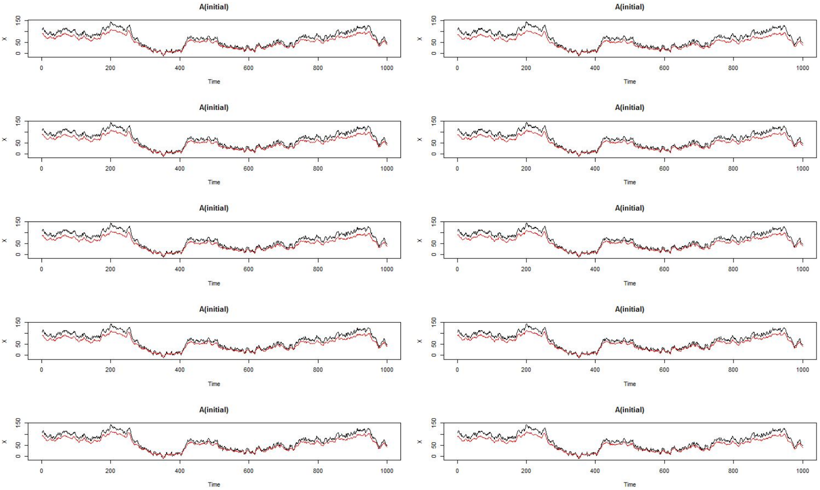
Результаты оценивания X.



Модель 1.MLE

Изменение начальных значений для C,при начальном $A = 1$.

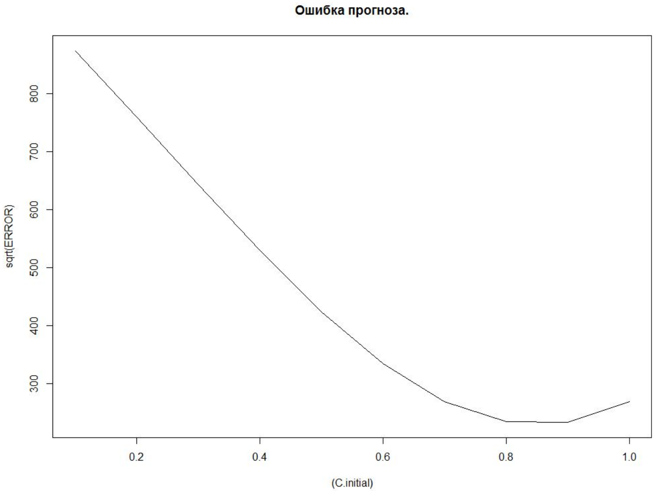
Результаты оценивания X.



Модель 1.MLE

Изменение начальных значений для C,при начальном $A = 1$.

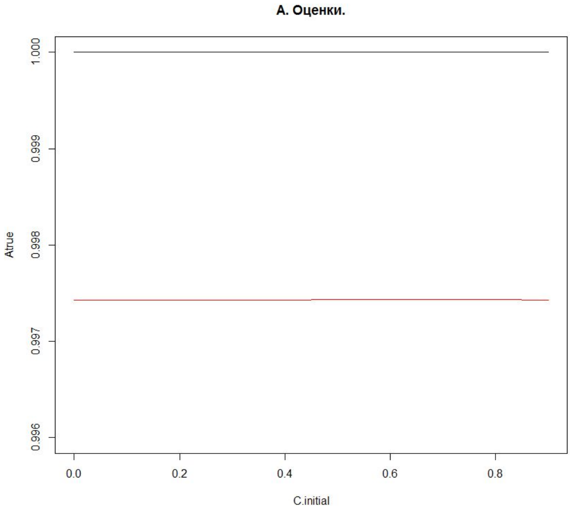
Результаты оценивания X.



Модель 1.ЕМ

Изменение начальных значений для C,при начальном $A = 1$.

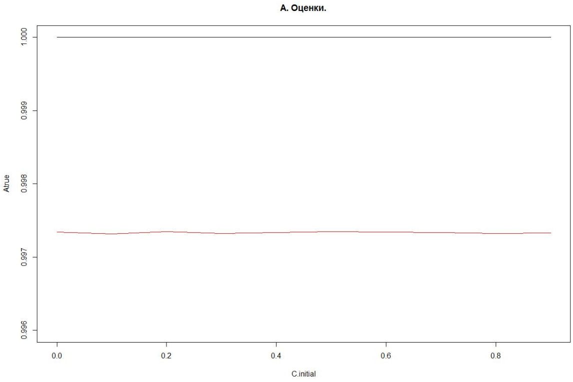
Результаты оценивания А.



Модель 1.MLE

Изменение начальных значений для C,при начальном $A = 1$.

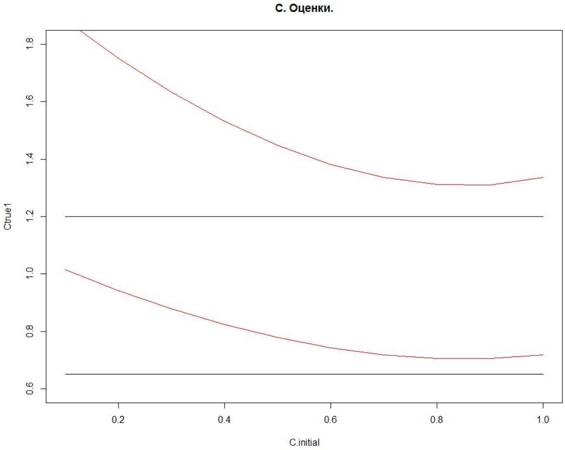
Результаты оценивания A.



Модель 1.EM

Изменение начальных значений для C,при начальном $A = 1$.

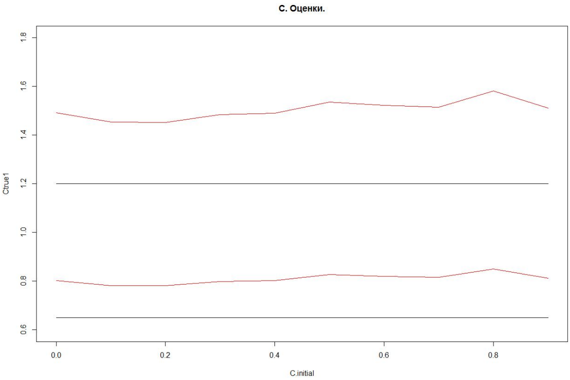
Результаты оценивания C.



Модель 1.MLE

Изменение начальных значений для C,при начальном $A = 1$.

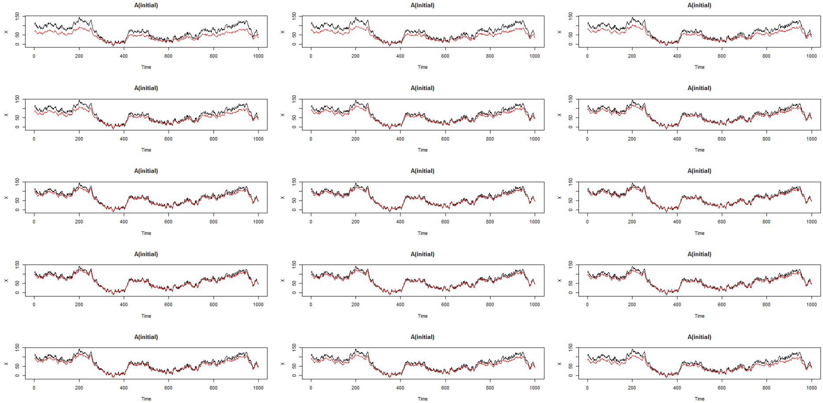
Результаты оценивания C.



Модель 1.EM

Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.

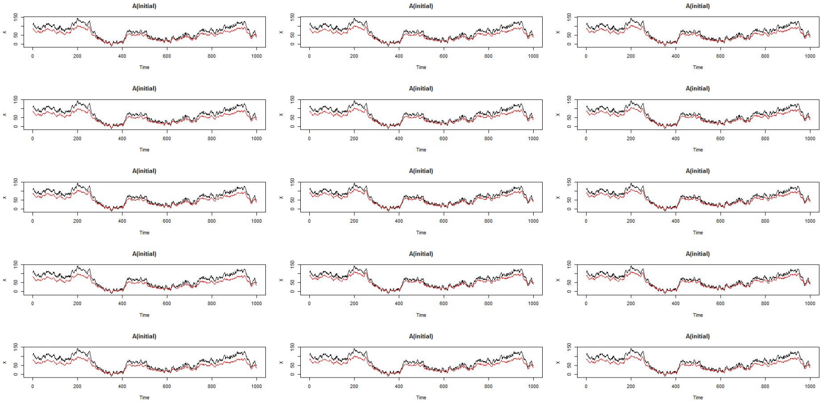
Результаты оценивания X.



Модель 1.MLE

Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.

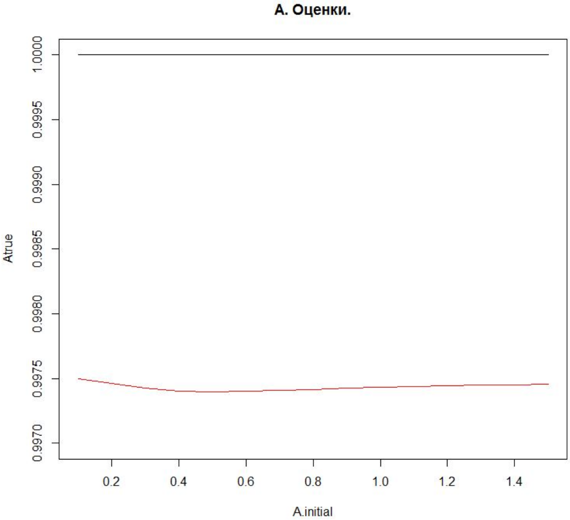
Результаты оценивания X.



Модель 1. ЕМ

Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.

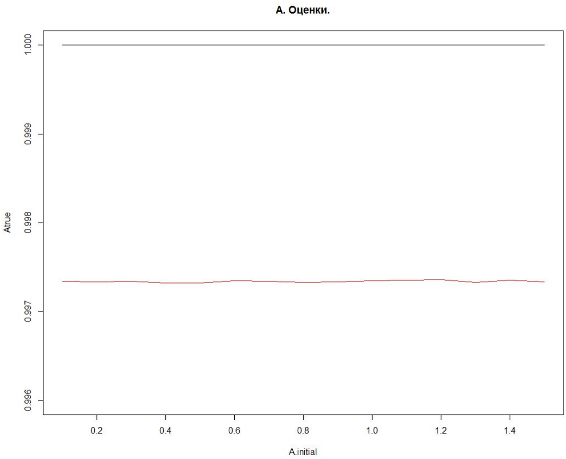
Результаты оценивания А.



Модель 1. MLE

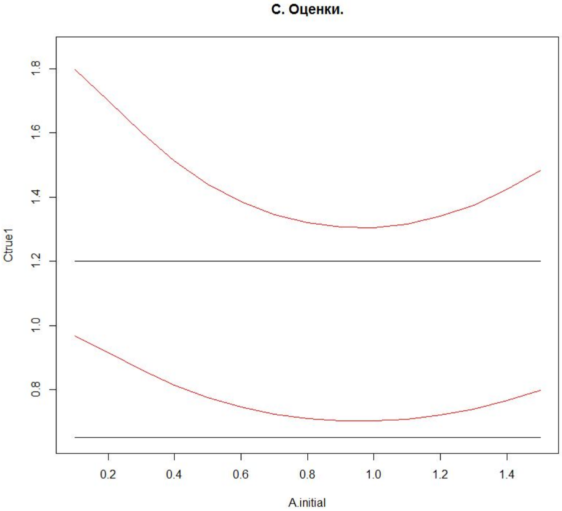
Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.

Результаты оценивания A.



Модель 1. EM

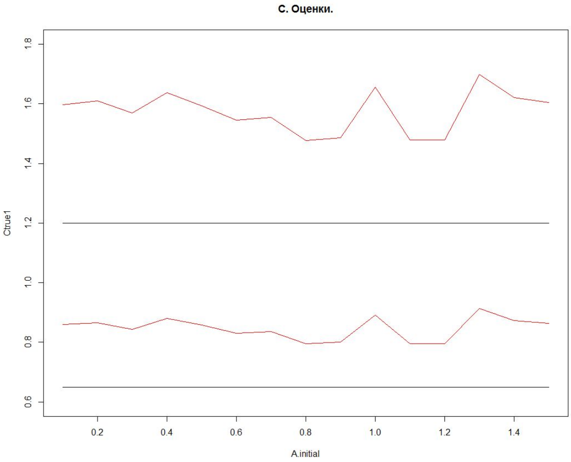
Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.
Результаты оценивания С.



Модель 1.MLE

Изменение начальных значений для A,при начальном $C = [0.9; 1.1]$.

Результаты оценивания C.



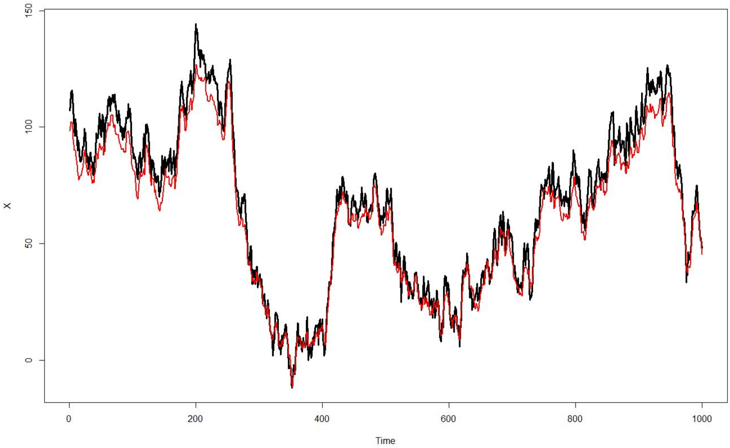
Модель 1.EM

Результаты.

Параметр	Истинное	EM
μ	100	98.31
Σ	30	0.06
A	1	0.997
C	$\begin{pmatrix} 0.65 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 0.70 \\ 1.30 \end{pmatrix}$
Q	20	16.86
R	$\begin{pmatrix} 80 & -20 \\ -20 & 100 \end{pmatrix}$	$\begin{pmatrix} 73.15 & -20.61 \\ -20.61 & 109.84 \end{pmatrix}$

Модель 1.EM

Результаты.



Модель 2

$$\begin{bmatrix} x_{1,k+1} \\ x_{2,k+1} \end{bmatrix} = \begin{bmatrix} 1.0 & 0.0 \\ 0.1 & 0.9 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + \begin{bmatrix} \epsilon_k \\ \eta_k \end{bmatrix};$$

$$\begin{pmatrix} \epsilon_k \\ \eta_k \end{pmatrix} \sim \mathcal{N}(\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} 40 & -10 \\ -10 & 60 \end{pmatrix})$$

$$\begin{pmatrix} x_{1,0} \\ x_{2,0} \end{pmatrix} \sim \mathcal{N}(\begin{pmatrix} 500 \\ 100 \end{pmatrix}; \begin{pmatrix} 30 & -5 \\ -5 & 50 \end{pmatrix})$$

$$y_k = x_{1,k} + x_{2,k} + \varepsilon_{k+1} \; ; \quad \varepsilon_k \sim \mathcal{N}(0, 100)$$

Спасибо за внимание!