Лекция 9.

Применения фильтра Калмана – Бьюси.

Примеры применения.

Резюмируем то, что мы назвали фильтром Калмана-Бьюси. Имеется модель, содержащая векторы X_t, называемые «векторами состояний». Предполагается, что эти векторы описывают текущее состояние рассматриваемой системы. Эти вектора состояний обычно ненаблюдаемы, другой важной составной частью системы являются наблюдаемые переменные Y_t. Первый шаг состоит в том, чтобы записать модель в виде линейной системы, состоящей из двух множеств линейных уравнений. Первое множество уравнений описывает эволюцию системы и называется «уравнением состояний» (state equation):

$$X_{t+1} = AX_t + \varepsilon_t, \qquad t = 1,2, \ldots$$

где A - матрица констант и $\varepsilon_t \sim N(0, Q)$. Второе множество уравнений описывает связь между состояниями системы и наблюдениями и называется «уравнением наблюдений» (observation equation):

$$Y_t = CX_t + \tau_t,$$

где $\tau_t \sim N(0, R)$, $X_1, \varepsilon_1, \varepsilon_2, \ldots, \tau_1, \tau_2, \ldots$ независимые гауссовские векторы.

Оказывается, что многие модели могут быть приведены к рассмотренной форме. Главные ограничения, конечно, это линейность модели и гауссовское распределение остатков. Существуют различные обобщения этой модели, в том числе на случай матриц, зависящих от времени t, и на нелинейный случай. Предполагается, что мы последовательно наблюдаем значения $Y_1, Y_2, \ldots, Y_T$. Одной из целей могла бы быть оценка неизвестных параметров A, C, Q, R по наблюдениям. При выводе алгоритма Калмана мы предполагаем, однако, что эти параметры известны точно. Вопрос о том, как эти параметры могут быть оценены по наблюдениям, мы обсудим позже. Заметим также, что ряд авторов обходит стороной вопрос о невырожденности ковариационной матрицы R в уравнении наблюдений, в частности, допускается, чтобы эта матрица была нулевой, как это будет в примерах ниже. В этом случае существование обратной матрицы $(CP_t^- C^* + R)^{-1}$, которая используется в алгоритме, требует дополнительного обоснования. Рассмотрим несколько конкретных примеров.

1) Рассмотрим модель, в которой экономика состоит из двух секторов, производящих однородный продукт, причём мы можем наблюдать только ***агрегированный*** выход продукта, подверженный случайной ошибке измерения. Предположим, что выход этих двух продуктов (X_t^1, X_t^2) может быть описан векторной авторегрессионной моделью первого порядка VAR (1). Тогда система запишется следующим образом. Уравнение состояний:
$$\begin{pmatrix} X_{t+1}^1 \\ X_{t+1}^2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1}^1 \\ \varepsilon_{t+1}^2 \end{pmatrix},$$
   и уравнение наблюдений:
$$Y_t = (1,1) \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} + \tau_t.$$

2) Рассмотрим одномерную модель MA(1): $y_t = \varepsilon_t + \theta \varepsilon_{t-1}$, где ε_t - процесс белого шума, $t = 2,3, \ldots$. Введём двумерный процесс состояний
$$X_t = \begin{pmatrix} y_t \\ \theta \varepsilon_t \end{pmatrix}.$$
   Наблюдаемой является только первая компонента процесса X_t.
   Уравнение состояний запишется в виде
$$X_{t+1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} X_t + \begin{pmatrix} 1 \\ \theta \end{pmatrix} \varepsilon_{t+1},$$
   а уравнение наблюдений в виде
$$y_t = (1,0) X_t$$

В этом примере $\tau_t \equiv 0$ (а, следовательно, и $R = 0$), матрицы A,C,Q известны с точностью до параметра θ, который мы пока тоже будем полагать известным.

Следует сказать, что представление модели в виде фильтра Калмана ***неоднозначно***. Например, для модели MA(1) можно предложить другое представление:
$$X_t = \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix}.$$

В этом случае обе компоненты вектора состояний X_t ненаблюдаемые. Уравнение состояний:

$$X_{t+1} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} X_t + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \varepsilon_{t+1}.$$

Наблюдаемая с.в. y_t. Уравнение наблюдений имеет вид

$$y_t = (1, \theta) \begin{pmatrix} \varepsilon_t \\ \varepsilon_{t-1} \end{pmatrix} = (1, \theta) X_t.$$

3) Рассмотрим модель MA(2): $y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, t = 3,4, \dots$ где ε_t – процесс белого шума. Введём трёхмерный процесс состояний

$$X_t = \begin{pmatrix} y_t \\ \theta_1 \varepsilon_t + \theta_2 \varepsilon_{t-1} \\ \theta_2 \varepsilon_t \end{pmatrix}.$$

Наблюдаемой снова является только первая компонента процесса X_t. Уравнение состояний запишется в виде

$$X_{t+1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X_t + \begin{pmatrix} 1 \\ \theta_1 \\ \theta_2 \end{pmatrix} \varepsilon_{t+1}$$

а уравнение наблюдений в виде

$$y_t = (1,0,0) X_t.$$

4) Рассмотрим одномерный AR(p) процесс:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t.$$

Уравнение состояний (это уравнение VAR (1)):

$$\begin{pmatrix} y_{t+1} \\ \vdots \\ y_{t-p+2} \end{pmatrix} = \begin{pmatrix} \phi_1 & \phi_2 & \dots & \phi_p \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Уравнение наблюдений:

$$y_t = (1,0, \dots ,0) \begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}.$$

Заметим, что в такой записи все компоненты вектора состояний наблюдаемы.

Вектор состояний здесь $X_t = \begin{pmatrix} y_t \\ \vdots \\ y_{t-p+1} \end{pmatrix}$, а вектор наблюдений $Y_t = y_t$, скаляр, первая компонента вектора X_t. Можно записать систему Калмана и иным способом. Возьмём в качестве вектора состояний вектор X_t

$$X_t = \begin{pmatrix} y_t \\ \phi_2 y_{t-1} + \dots + \phi_p y_{t-p+1} \\ \phi_3 y_{t-1} + \dots + \phi_p y_{t-p+2} \\ \vdots \\ \phi_{p-1} y_{t-1} + \phi_p y_{t-2} \\ \phi_p y_{t-1} \end{pmatrix}.$$

Тогда

$$X_{t+1} = AX_t + \begin{pmatrix} \varepsilon_{t+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

$$Y_t = y_t = (1,0,\dots,0)X_t.$$

Матрица A, как и ранее, равна

$$A = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots \; 1 & 0 \end{pmatrix}^T.$$

5) Рассмотрим одномерный ARMA (p,q) процесс

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}.$$

Положим $r = \max(p, q+1)$, $\phi_j = 0$ для $j > p$ и $\theta_j = 0$ для $j > q$. Тогда этот процесс может быть записан единообразно в виде

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_r y_{t-r} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_{r-1} \varepsilon_{t-r+1}. \tag{1}$$

Пусть X_t – ненаблюдаемый вектор состояния системы. Рассмотрим следующие уравнения состояний и наблюдений.

Уравнение состояний:

$$X_{t+1}=\begin{pmatrix}\phi_1 & \phi_2 & \dots & \phi_{r-1} & \phi_r\\ 1 & 0 & \dots & 0 & 0\\ 0 & 1 & \dots & 0 & 0\\ \vdots & \vdots & \dots & \vdots & \vdots\\ 0 & 0 & \dots & 1 & 0\end{pmatrix}X_t+\begin{pmatrix}\varepsilon_{t+1}\\ 0\\ 0\\ \vdots\\ 0\end{pmatrix}.$$

Уравнение наблюдений:

$$y_t=(1,\theta_1,\theta_2,\dots,\theta_{r-1})X_t.$$

Пусть X_{jt} – j-я компонента вектора X_t. Из второй строки уравнения состояния следует, что $X_{2.t+1}=X_{1.t}$, а из третьей строки следует, что $X_{3.t+1}=X_{2.t}=X_{1.t-1}$. Продолжая, получим, что $X_{j.t+1}=L^{j-1}X_{1.t+1}$. Таким образом, все компоненты получаются из первой компоненты применением степеней оператора лага L. Из первой строки уравнения состояния получаем

$$X_{1.t+1}=(\phi_1+\phi_2L+\phi_3L^2+\dots+\phi_rL^{r-1})X_{1.t}+\varepsilon_{t+1}.$$

или

$$(1-\phi_1L-\phi_2L^2-\dots-\phi_rL^r)X_{1.t+1}=\varepsilon_{t+1}. \qquad (2)$$

Из уравнения наблюдения следует, что

$$y_t=(1+\theta_1L+\theta_2L^2+\dots+\theta_{r-1}L^{r-1})X_{1t} \qquad (3)$$

Умножая (3) на $(1-\phi_1L-\phi_2L^2-\dots-\phi_rL^r)$, используя (2) с t вместо $t+1$ и переставляя местами полиномы лага, получим

$$(1-\phi_1L-\phi_2L^2-\dots-\phi_rL^r)y_t=$$

$$(1+\theta_1L+\theta_2L^2+\dots+\theta_{r-1}L^{r-1})\varepsilon_t.$$

Последнее уравнение совпадает с (1). То есть наблюдаемая компонента y_t является ARMA (p,q) процессом.

Замечание об обозначениях.

Разные авторы используют разные обозначения. Различия в обозначениях матриц A,C, Q, R фильтра Калмана-Бьюси несущественны, мы на этом останавливаться не будем. Более существенным является различие в обозначениях процессов, условных математических ожиданий и ковариационных матриц. Некоторые авторы иначе определяют и саму матрицу усиления Калмана. В обозначениях мы следуем, в основном, курсу Ph. Bougerol “Modèles Stochastiques et Applications à la Finance”, читаемому в магистратуре 1-го года обучения в Университете «Париж-6 ». Весьма популярной монографией по временным рядам является книга J. Hamilton “Time Series Analysis”, Princeton University, 1994. Для удобства чтения сравним обозначения этих двух источников

	матрица усиления	процесс состояний	проекция к моменту $t-1$	проекция к моменту t	Ковариационная матрица ошибки	Ковариационная матрица оценeной ошибки
Boug	K_t^B	X_t	$\hat{X}_t^-$	$\hat{X}_t$	P_t	P_t^-
Ham	$K_t^H = AK_t^B$	ξ_t	$\hat{\xi}_{t\|t-1}$	$\hat{\xi}_{t\|t}$	$P_{t\|t}$	$P_{t\|t-1}$

Соответственно меняются и основные соотношения алгоритма. Сравним их.

- *Вычисление ковариационной матрицы оценки ошибки*

Bougerol:

$$P_t^- = AP_{t-1}A^* + Q$$

Hamilton:

$$P_{t|t-1} = AP_{t|t}A^* + Q$$

- *Вычисление матрицы усиления (gain matrix) Калмана*

Bougerol:

$$K_t^B = P_t^- C^*(CP_t^- C^* + R)^{-1}$$

Hamilton:

$$K_t^H = AP_{t|t-1}C^*(CP_{t|t-1}C^* + R)^{-1}$$

- *Вычисление ковариационной матрицы ошибки*

Bougerol:

$$P_t = P_t^- - P_t^- C^*(CP_t^- C^* + R)^{-1}CP_t^-$$

Hamilton:

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}C^*(CP_{t|t-1}C^* + R)^{-1}CP_{t|t-1}$$

- *Пересчёт проекции с использованием нового наблюдения* Y_t.

Bougerol:

$$\hat{X}_{t+1}^{-} = A\hat{X}_t^{-} + AK_t^B(Y_t - C\hat{X}_t^{-})$$

Hamilton:

$$\hat{\xi}_{t+1|t} = A\hat{\xi}_{t|t-1} + K_t^H\left(Y_t - C\hat{\xi}_{t|t-1}\right).$$

Вычисление прогноза на s шагов вперёд с помощью фильтра Калмана.

Будем использовать обозначения книги Hamilton. Прогноз на s шагов вперёд, $s \geq 2$, будем обозначать $\hat{Y}_{t+s|t}$ и $\hat{X}_{t+s|t}$. Прогноз для Y_t легко считается через прогноз для X_t. В самом деле,

$$Y_{t+s} = CX_{t+s} + \tau_{t+s},$$

откуда

$$\hat{Y}_{t+s|t} := E\left(Y_{t+s}\middle|\sigma(Y_1, \dots, Y_t)\right) = C\hat{X}_{t+s|t}.$$

Итерируя уравнение состояний, получим

$$X_{t+s} = A^s X_t + A^{s-1}\varepsilon_{t+1} + A^{s-2}\varepsilon_{t+2} + \dots + A\varepsilon_{t+s-1} + \varepsilon_{t+s}.$$

Переходя к условным математическим ожиданиям, получим

$$\hat{X}_{t+s|t} = A^s\hat{X}_t.$$

Таким образом, ошибка прогноза на s шагов вперёд для процесса состояний X_t равна

$$X_{t+s} - \hat{X}_{t+s|t} =$$

$$A^s\left(X_t - \hat{X}_t\right) + A^{s-1}\varepsilon_{t+1} + A^{s-2}\varepsilon_{t+2} + \dots + A\varepsilon_{t+s-1} + \varepsilon_{t+s}.$$

Ковариационная матрица этой ошибки равна

$$P_{t+s|t} = A^s P_t (A^*)^s + A^{s-1}Q(A^*)^{s-1} + A^{s-2}Q(A^*)^{s-2} + \dots + AQA^* + Q.$$

Одно важное обобщение фильтра Калмана.

Фильтр Калмана может включать в себя экзогенные переменные, т.е. переменные, значения которых не определяются внутри экономической модели, эти переменные являются объясняющими и никогда не могут быть зависимыми переменными. Модель имеет вид

$$X_{t+1} = AX_t + \varepsilon_{t+1}$$

$$Y_t = Bz_t + CX_t + \tau_t, \qquad (4)$$

где B - $p \times k$ матрица, z_t- k-мерный вектор-столбец. Вектор z_t не содержит никакой новой информации об X_{t+s} или τ_{t+s}, $s = 0,1,2,\ldots$ кроме той, которая содержится в $Y_1, \ldots, Y_{t-1}$. Например, z_t может включать лаговые значения Y_t или переменные, которые некоррелированы с X_t и τ_t для всех t. Они могут быть и не случайными, задаваемыми извне векторами. Вывод алгоритма остаётся практически тем же самым, только теперь всюду вместо σ-алгебр $\sigma(Y_1, \ldots, Y_t)$ следует рассматривать σ- алгебры $\sigma(z_1, \ldots, z_t, Y_1, \ldots, Y_t)$. Опишем рекурсивный алгоритм в этом случае. Начальное приближение выбирается так же, как и ранее. Далее:
Переход от P_t^- к P_{t+1}^- остаётся без изменений:

$$P_{t+1}^- = AP_tA^* + Q = A(P_t^- - K_tCP_t^-)A^* + Q =$$

$$A\left(P_t^- - \underbrace{P_t^-C^*(CP_t^-C^* + R)^{-1}}_{K_t}CP_t^-\right)A^* + Q.$$

Переход от $\hat{X}_t^-$ к $\hat{X}_{t+1}^-$:

$$\hat{X}_{t+1}^- = A\hat{X}_t^- + AK_t\left(Y_t - Bz_t - C\hat{X}_t^-\right).$$

На каждом шаге проекции $\hat{X}_t$ и дисперсии ошибок P_t находятся из соотношений

$$P_t = P_t^- - K_tCP_t^-$$

и

$$\hat{X}_t = \hat{X}_t^- + K_tJ_t = \hat{X}_t^- + K_t\left(Y_t - Bz_t - C\hat{X}_t^-\right).$$

Использование фильтра Калмана для оценки функции правдоподобия.

Рассмотрим фильтр Калмана (4) с экзогенными переменными. Если X_1 и $\{\varepsilon_t, \tau_t\}|_{t=1}^{T}$ гауссовские, то условное распределение Y_t при условии $(z_t, \mathcal{Y}_{t-1})$, *где* $\mathcal{Y}_{t-1} = (z_1, \dots, z_{t-1}, Y_1, \dots, Y_{t-1})$ является гауссовским со средним

$$E(Y_t|z_t, \mathcal{Y}_{t-1}) = E(Bz_t + CX_t + \tau_t|z_t, \mathcal{Y}_{t-1}) =$$

$$Bz_t + CE(X_t|z_t, \mathcal{Y}_{t-1}) = Bz_t + C\hat{X}_t^-$$

и матрицей ковариаций

$$E\left(Y_t - Bz_t - C\hat{X}_t^-\right)\left(Y_t - Bz_t - C\hat{X}_t^-\right)^* =$$

$$E\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)\left(C\left(X_t - \hat{X}_t^-\right) + \tau_t\right)^* = CP_t^- C^* + R.$$

Гауссовская условная плотность Y_t имеет вид

$$f_{Y_t|z_t,\mathcal{Y}_{t-1}}(y_t|z_t, \mathcal{Y}_{t-1}) = (2\pi)^{-p/2}\det{}^{1/2}(CP_t^- C^* + R)$$

$$\times \exp\left\{-\frac{1}{2}\left(y_t - Bz_t - C\hat{X}_t^-\right)^*(CP_t^- C^* + R)^{-1}(y_t - Bz_t - C\hat{X}_t^-)\right\} \quad (5)$$

для $t = 1,2, \dots, T$.
Чтобы выписать функцию правдоподобия по наблюдениям $y_1, \dots, y_T$, сделаем следующее замечание. Если имеется функция правдоподобия (плотность)

$$f(y_1, \dots, y_T; \theta),$$

где θ вектор параметров, то её всегда можно факторизовать

$$f(y_1, \dots, y_T; \theta) = f(y_1, \dots, y_{T-1}; \theta) \cdot f(y_T|y_1, \dots, y_{T-1}; \theta).$$

Итерируя, получим разложение

$$f(y_1,\dots,y_T;\theta)=\prod_{t=1}^{T} f(y_t|y_1,\dots,y_{t-1};\theta).$$

(при $t=1$ полагаем $f(y_1|y_0;\theta)=f(y_1)$). Переходя к логарифмам, получим для логарифмической функции правдоподобия (log - likelihood function)

$$L(y_1,\dots,y_T;\theta)=\sum_{t=1}^{T}\log[f(y_t|y_1,\dots,y_{t-1};\theta)].$$

Возвращаясь к условной плотности (5), получим для логарифмической функции правдоподобия

$$\sum_{t=1}^{T}\log\left[f_{Y_t|z_t,\mathcal{Y}_{t-1}}(y_t|z_t,\mathcal{Y}_{t-1})\right] \qquad (6)$$

Выражение (6) можно максимизировать, применяя методы численной оптимизации. Параметры, по которым проводится максимизация – это элементы матриц A, B, C, R $и$ Q.

Рассмотрим пример. Предположим, мы хотим оценить параметры двумерной регрессионной модели вида

$$Y_{1t}=a_1^* z_t+u_{1t}$$

$$Y_{2t}=a_2^* z_t+u_{2t},$$

где z_t – $(k\times 1)$ вектор экзогенных объясняющих переменных, и a_1 $и$ a_2 $(k\times 1)$ векторы коэффициентов. Считаем, что обе регрессии имеют одинаковые объясняющие переменные (этого можно добиться, добавляя нулевые коэффициенты). Предположим, что вектор ошибок, в свою очередь, является двумерным MA(1) процессом

$$\begin{bmatrix}u_{1t}\\u_{2t}\end{bmatrix}=\begin{bmatrix}\delta_{1t}\\\delta_{2t}\end{bmatrix}+\begin{bmatrix}\theta_{11}&\theta_{12}\\\theta_{21}&\theta_{22}\end{bmatrix}\begin{bmatrix}\delta_{1.t-1}\\\delta_{2.t-1}\end{bmatrix},$$

где $(\delta_1,\delta_2)^*\sim i.i.d.N(0,\Omega)$. Эту модель можно записать в виде фильтра Калмана. Уравнение состояний имеет вид

$$X_{t+1}=AX_t+\varepsilon_t,$$

где

$$X_t = \begin{bmatrix} \delta_{1t} \\ \delta_{2t} \\ \delta_{1.t-1} \\ \delta_{2.t-1} \end{bmatrix}, A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \varepsilon_{t+1} = \begin{pmatrix} \delta_{1.t+1} \\ \delta_{2.t+1} \\ 0 \\ 0 \end{pmatrix},$$

$$Q = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

а уравнение наблюдений имеет вид

$$Y_t = Bz_t + CX_t,$$

где

$$Y_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix}, B = \begin{pmatrix} a_1^* \\ a_2^* \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & \theta_{11} & \theta_{12} \\ 0 & 1 & \theta_{21} & \theta_{22} \end{pmatrix}, R = 0,$$

$$\sigma_{ij} = E(\delta_{it}\delta_{jt}), i, j = 1,2.$$

Итерации фильтра Калмана начинаются с

$$\hat{X}_1^- = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, P_1^- = \begin{pmatrix} \sigma_{11} & \sigma_{12} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 & 0 \\ 0 & 0 & \sigma_{11} & \sigma_{12} \\ 0 & 0 & \sigma_{21} & \sigma_{22} \end{pmatrix}.$$

Чтобы начать максимизацию в (6), надо сделать начальные предположения о численных значениях неизвестных параметров. Для получения начального значения для a_1 одним из естественных путей является нахождение МНК - регрессии y_{1t} на элементы z_t, появляющиеся в первом уравнении. И аналогично для a_2. Полагая сначала $\theta_{11} = \theta_{12} = \theta_{21} = \theta_{22} = 0$, первое приближение для Ω получим, вычисляя оценку ковариационной матрицы остатков регрессионной модели. Для этих начальных значений параметров мы строим матрицы A, B, C, Q и итерируем фильтр Калмана с этими значениями матриц для $t = 1,2, \dots, T$. В результате получаем последовательности $\{\hat{X}_t^-\}_{t=1}^T$ и $\{P_t^-\}_{t=1}^T$, которые используем в (5) и (6) для максимизации логарифма функции правдоподобия.

Получаем новое приближение для элементов матриц A, B, C, Q. Затем мы итерируем фильтр Калмана с новыми значениями матриц и.т.д. Мы сейчас не обсуждаем вычислительные проблемы, связанные с этой задачей максимизации и вопросы сходимости алгоритма.
В заключение напомним алгоритм и рассмотрим пример.

Алгоритм.

Покажем как по $\hat{X}_{t|t}$ и $P_{t|t}$ вычислить $\hat{X}_{t+1|t+1}$ и $P_{t+1|t+1}$:

$$\hat{X}_{t+1|t} = A\hat{X}_{t|t}$$

$$P_{t+1|t} = AP_{t|t}A^* + Q$$

$$\nu_{t+1} = Y_{t+1} - C\hat{X}_{t+1|t}$$

$$S_{k+1|k} = CP_{t+1|t}C^* + R$$

$$G_{t+1} = P_{t+1|t}C^*\left[S_{k+1|k}\right]^{-1}$$

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

$$P_{t+1|t+1} = (I - G_{t+1}C)P_{t+1|t}$$

Пример. Оценка константы по зашумленным наблюдениям.

$$X_{k+1} = X_k, \qquad X_k \in R.$$

$$Y_k = X_k + W_k,$$

$\{W_k\} -$ i.i.d., имеющие распределение $N(0, \sigma^2)$ и не зависящие от X_0. Положим $\hat{X}_{0|0} = 0$ и $P_{0|0} = \sigma_0^2 > 0$. В нашем случае $A = 1, Q = 0, C = 1, R = \sigma^2$. Имеем

$$\hat{X}_{t+1|t} = A\hat{X}_{t|t} \Rightarrow \hat{X}_{1|0} = \hat{X}_{0|0} = 0$$

$$P_{t+1|t} = AP_{t|t}A^* + Q \Rightarrow P_{1|0} = P_{0|0} + Q = P_{0|0} = \sigma_0^2$$

$$\nu_{t+1} = Y_{t+1} - C\hat{X}_{t+1|t} \Rightarrow \nu_1 = Y_1 - \hat{X}_{1|0} = Y_1$$

$$S_{k+1|k} = CP_{t+1|t}C^* + R \Rightarrow \ S_{1|0} = P_{1|0} + R = \sigma_0^2 + \sigma^2$$

$$G_{t+1} = P_{t+1|t}C^*\big[S_{k+1|k}\big]^{-1} \Rightarrow \ G_1 = P_{1|0}[\sigma_0^2 + \sigma^2]^{-1} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}$$

$$\hat{X}_{t+1|t+1} = \hat{X}_{t+1|t} + G_{t+1}\nu_{t+1} \Rightarrow \ \hat{X}_{1|1} = 0 + \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}Y_1$$

$$P_{t+1|t+1} = (I - G_{t+1}C)P_{t+1|t} \Rightarrow$$

$$P_{1|1} = \left(1 - \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}\right)\sigma_0^2 = \frac{\sigma_0^2\sigma^2}{\sigma_0^2 + \sigma^2}.$$

Первый шаг закончен. Повторим все вычисления для $k = 2$.

$$\hat{X}_{2|1} = \hat{X}_{1|1} = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}Y_1$$

$$P_{2|1} = P_{1|1} = \frac{\sigma_0^2\sigma^2}{\sigma_0^2 + \sigma^2}$$

$$\nu_2 = Y_2 - \hat{X}_{2|1} = Y_2 - \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}Y_1$$

$$S_{2|1} = P_{2|1} + R = \frac{\sigma_0^2\sigma^2}{\sigma_0^2 + \sigma^2} + \sigma^2 = \frac{(2\sigma_0^2 + \sigma^2)\sigma^2}{\sigma_0^2 + \sigma^2}$$

$$G_2 = P_{2|1}\big[S_{2|1}\big]^{-1} = \frac{\sigma_0^2\sigma^2}{\sigma_0^2 + \sigma^2} \cdot \frac{\sigma_0^2 + \sigma^2}{(2\sigma_0^2 + \sigma^2)\sigma^2} = \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}$$

$$\hat{X}_{2|2} = \hat{X}_{2|1} + G_2\nu_2 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}Y_1 + \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}\left[Y_2 - \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2}Y_1\right] =$$

$$\frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}Y_1 + \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}Y_2$$

$$P_{2|2} = \left(1 - \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}\right) P_{2|1} = \left(1 - \frac{\sigma_0^2}{2\sigma_0^2 + \sigma^2}\right) \frac{\sigma_0^2 \sigma^2}{\sigma_0^2 + \sigma^2} =$$

$$\frac{\sigma_0^2 \sigma^2}{2\sigma_0^2 + \sigma^2}$$

и.т.д. На $k -$ ом шаге получим

$$\hat{X}_{k|k} = \frac{\sigma_0^2}{k\sigma_0^2 + \sigma^2} \sum_{i=1}^{k} Y_i$$

$$P_{k|k} = \frac{\sigma_0^2 \sigma^2}{k\sigma_0^2 + \sigma^2}$$

$$G_k = \frac{\sigma_0^2}{k\sigma_0^2 + \sigma^2}$$

$$P_{k|k-1} = \frac{\sigma_0^2 \sigma^2}{(k-1)\sigma_0^2 + \sigma^2}$$

$$\nu_k = Y_k - \hat{X}_{k|k-1}$$

$$\hat{X}_{k|k-1} = \hat{X}_{k-1|k-1}$$

Мы видим, что независимо от начального выбора σ_0^2 (и σ^2)

$$\hat{X}_{k|k} \sim \frac{1}{k} \sum_{i=1}^{k} Y_i\,, \qquad k \to \infty.$$