Модели пространства состояний и фильтр Калмана

Оценка максимального правдоподобия

Статистическая модель с вектором параметров θ размерности k определяет совместное распределение для вектора наблюдений: $\tilde{y}_T = [y_1 \quad y_2 \quad \dots \quad y_T]'$:

$$\text{Совместная функция плотности: } p(\tilde{y}_T|\theta) \tag{1}$$

При проведении исследований у нас есть реализация вектора $\tilde{y}_T$, но мы не знаем вектор параметров θ. В этом случае совместная плотность (1) представляет собой функцию θ при заданном $\tilde{y}_T$:

$$\text{Функция правдоподобия: } L(\theta|\tilde{y}_T) \tag{2}$$

$$\hat{\theta}_{ML} = \arg\max_{\theta} \quad \ln \mathcal{L}_t(\theta|\tilde{y}_T)$$

Декомпозиция ошибки предсказания

- Для независимых наблюдений:

$$\mathcal{L}_t(\theta|\tilde{y}_T) = \prod_{t=1}^{T} p(y_t|\theta), \tag{3}$$

  где $p(y_t|\theta)$-предельная плотность (marginal density) индивидуального наблюдения

- Для зависимых наблюдений

$$\mathcal{L}_t(\theta|\tilde{y}_T) = \prod_{t=1}^{T} p(y_t|\tilde{y}_{t-1},\theta)p(y_1|\theta), \tag{4}$$

  где $\tilde{y}_t = [y_1 \quad y_2 \quad \dots \quad y_t], p(y_t|\tilde{y}_{t-1},\theta), t = 2, 3, \dots, T$-условная плотность (conditional density), $p(y_1|\theta)$ -предельная плотность y_1

Декомпозиция ошибки предсказания (2)

Если $\tilde{y}_T$ может быть представлено многомерным нормальным распределением:

$$\tilde{y}_T \sim N(\mu, \Omega), \tag{5}$$

то функция правдоподобия имеет вид:

$$\mathcal{L}_t(\theta|\tilde{y}_T) = (2\pi)^{-\frac{T}{2}}|\Omega|^{-\frac{1}{2}} \exp\{-\frac{1}{2}(\tilde{y}_T - \mu)'\Omega^{-1}(\tilde{y}_T - \mu)\} \tag{6}$$

Проблема: μ и Ω являются функциями параметров модели, при этом Ω имеет размерность $T \times T$

Декомпозиция ошибки предсказания (3)

Решение: Harvey(1980)[1]

$$\mathcal{L}_t(\theta|\tilde{y}_T) = \prod_{t=1}^{T}\left(\frac{1}{\sqrt{2\pi f_t}}\exp\{-\frac{1}{2}\frac{\nu_t^2}{f_t}\}\right), \qquad (7)$$

где $\nu_t = y_t - y_{t|t-1}$ - ошибка прогноза, f_t - дисперсия ошибки прогноза В многомерном случае (7) заменяется на:

$$\mathcal{L}_t(\theta|\tilde{y}_T) = \prod_{t=1}^{T}\left(\frac{1}{\sqrt{(2\pi)^n|f_t|}}\exp\{-\frac{1}{2}\nu_t' f_t^{-1}\nu_t\}\right), \qquad (8)$$

где n - размерность y_t

[1] Доказательство см. в Kim and Nelson (1999, p.12)

Модель пространства состояний (state-space model)- модель, в которой наблюдаемая переменная представляет собой сумму линейной функции переменной состояния и ошибки. **Переменная состояния**(state variable), в свою очередь, меняется в соответствии со стохастическим разностным уравнением, зависящим от параметров, которые в экономических приложениях обычно неизвестны. Таким образом как динамика переменной состояния, так и параметры должны быть определены на основе имеющихся данных.

Определения (2)

Основной инструмент, используемый при анализе моделей пространства состояний, - **фильтр Калмана**. Это рекурсивная процедура, позволяющая оценить ненаблюдаемую компоненту (вектор состояния) на основе информации, доступной к периоду t. Когда шоки в модели и ненаблюдаемые переменные в начальный момент времени нормально распределены, фильтр Калмана позволяет посчитать **функцию правдоподобия** (likelihood function) с помощью декомпозиции ошибки предсказания.

Уравнение перехода (transition equation)

$$\alpha_t = T\alpha_{t-1} + Ru_t \tag{9}$$

Уравнение измерения (measurement equation)

$$y_t = S\alpha_t + \epsilon_t \tag{10}$$

где

$$u_t \sim N(0, Q) \tag{11}$$

$$\epsilon_t \sim N(0, H) \tag{12}$$

$$\epsilon_t \perp u_t \tag{13}$$

$$\alpha_0 \sim N(a_0, P_0) \tag{14}$$

$$\alpha_t \mid \underbrace{\{y_1, \ldots, y_{t-1}\}}_{\tilde{y}_{t-1}} \sim N(\alpha_{t|t-1}, P_{t|t-1}) \tag{15}$$

Формулировка модели

$$\alpha_{t|t-1} = E[\alpha_t \mid \underbrace{\{y_1, \dots, y_{t-1}\}}_{\tilde{y}_{t-1}}] = \tag{16}$$

$$= E[T\alpha_{t-1} + Ru_t \mid \tilde{y}_{t-1}] = \tag{17}$$

$$= TE[\alpha_{t-1} \mid \tilde{y}_{t-1}] = \tag{18}$$

$$= T\alpha_{t-1|t-1} \tag{19}$$

$$P_{t|t-1} = Var(\alpha_t \mid \tilde{y}_{t-1}) = \tag{20}$$

$$= Var(T\alpha_{t-1} + Ru_t \mid \tilde{y}_{t-1}) = \tag{21}$$

$$= TVar(\alpha_{t-1} \mid \tilde{y}_{t-1})T' + RQR' = \tag{22}$$

$$= TP_{t-1|t-1}T' + RQR' \tag{23}$$

Свойства многомерного нормального распределения

Пусть Z_1 и Z_2 при условии I_{t-1} нормально распределены:

$$\begin{matrix} Z_1 \\ Z_2 \end{matrix} \mid I_{t-1} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$

Тогда распределение Z_1 при заданных Z_2 и I_{t-1} задается как:

$$Z_1 \mid Z_2, I_{t-1} \sim N\left(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(Z_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\right) \tag{24}$$

Ошибка прогноза

Определим **ошибку прогноза** ν_t, ортогональную $\tilde{y}_{t-1}$:

$$\nu_{t|t-1} = y_t - E[y_t \mid \tilde{y}_{t-1}] = \tag{25}$$
$$= y_t - E[S\alpha_t + \epsilon_t \mid \tilde{y}_{t-1}] = \tag{26}$$
$$= y_t - S\alpha_{t|t-1} \tag{27}$$

Определим:

$$M_{t|t-1} = Cov(\alpha_{t|t-1}, \nu_{t,t-1}) \tag{28}$$
$$F_{t|t-1} = Var(\nu_{t|t-1}) \tag{29}$$

При этом:

$$M_{t|t-1} = Cov(\alpha_t, \nu_{t|t-1}) = Cov(\alpha_t, y_t - S\alpha_{t,t-1}) =$$
$$= Cov(\alpha_t, S\alpha_t + \epsilon_t - S\alpha_{t|t-1}) = P_{t|t-1}S'$$
$$F_{t|t-1} = Var(\nu_{t|t-1}) = Var(y_t - S\alpha_{t|t-1}) =$$
$$= Var(S\alpha_t + \epsilon_t - S\alpha_{t|t-1}) = SP_{t|t-1}S' + H$$

Условное ожидание может быть записано как:

$$E[\alpha_t \mid Y_t] = E[\alpha_t \mid \tilde{y}_{t-1}, \nu_{t|t-1}] = \tag{30}$$

$$= \tag{31}$$

$$= a_{t|t-1} + M_{t|t-1} F_{t|t-1}^{-1} \nu_{t|t-1} \tag{32}$$

По аналогии с (24), где $Z_2 = \nu_t, I_{t-1} = \tilde{y}_{t-1}, \mu_2 = 0$

$$a_{t|t} = a_{t|t-1} + \underbrace{M_{t|t-1} F_{t|t-1}^{-1}}_{K_t} \nu_{t|t-1} \tag{33}$$

Рассмотрим условную дисперсию (см. уравнение (22)):

$$
\begin{aligned}
Var(\alpha_t \mid \tilde{y}_t) &= Var(\alpha_t \mid \tilde{y}_{t-1}, \nu_t | t-1) = \\
&= P_{t|t-1} - (P_{t|t-1}S')F^{-1}_{t|t-1}(SP_{t,t-1}) = \\
&= P_{t|t-1} - K_t S_t P'_{t|t-1}
\end{aligned} \tag{34}
$$

Этапы фильтра Калмана

При условии, что нам известны начальное значение переменной состояния и матрицы коэффициентов, фильтр Калмана включает в себя два этапа:

- **Предсказание:** В начале периода t мы хотим сформировать оптимальный прогноз y_t на базе всей доступной к моменту $t-1$ информации: $y_{t|t-1}$. Для этого нужно посчитать $a_{t|t-1}$.
- **Корректировка:** Когда y_t становится известно в конце периода t, можно посчитать ошибку прогноза: $\nu_{t|t-1} = y_t - y_{t|t-1}$. Эта ошибка прогноза содержит новую информацию о α_t в дополнение к той, что содержалась в $a_{t|t-1}$. Таким образом, информация о y_t позволяет сделать более точные выводы относительно α_t

- **Предсказание**

$$a_{t|t-1} = Ta_{t-1|t-1} \tag{35}$$
$$P_{t|t-1} = TP_{t-1|t-1}T' + RQR' \tag{36}$$
$$\nu_{t|t-1} = y_t - y_{t|t-1} = y_t - S\alpha_{t|t-1} \tag{37}$$
$$F_{t|t-1} = SP_{t|t-1}S' + H \tag{38}$$

- **Корректировка**

$$K_t = P_{t|t-1}S'F^{-1}_{t|t-1} \tag{39}$$
$$a_{t|t} = a_{t|t-1} + K_t\nu_{t|t-1} \tag{40}$$
$$P_{t|t} = P_{t|t-1} - K_tSP_{t|t-1} = \tag{41}$$
$$= P_{t|t-1} - P_{t,t-1}S'F'_{t|t-1}SP_{t,t-1} \tag{42}$$

Сглаживающая схема

- **Сглаживание** На этапах предсказания и корректировки используется вся прошлая информация и текущая реализация y_t, но не будущая информация о выборке, которая приходит позже момента t. В некоторых случаях имеет смысл использовать всю информацию о выборке вплоть до последнего наблюдения. Это процедура сглаживания, которая корректирует переменные состояния в направлении от конца выборки к началу, а не наоборот начиная с $T-1$ наблюдения.[2]

$$a_{t,T} = a_{t,t} + V_t(a_{t+1,T} - a_{t+1,t}) \tag{43}$$

$$P_{t|T} = P_{t|t} + V_t(P_{t+1|T} - P_{t+1|t})V_t' \tag{44}$$

$$\text{где } V_t = P_{t|t}F'P_{t+1|t}^{-1} \tag{45}$$

[2] Доказательство см. в Hamilton, 1994, p. 394 или в Цыплаков (2011, стр.23)

Сглаживающая схема (2)

$a_{t|T}$ - оценка α_t при условии всей известной выборки, а $P_{t|T}$ - ее ковариационная матрица. Так как оценки вектора переменных состояния и его ковариационной матрицы в момент $T, a_{T,T}$ и $P_{T|T}$ получены из уравнений (40) и (41), то все оценки $a_{t|T}$ и $P_{t|T}$ могут быть определены рекурсивно из уравнений (43) и (44)

Оценка параметров

До этого предполагалось, что матрицы коэффициентов известны. На практике они неизвестны и требуется получить их оценки. Предполагается, что все матрицы коэффициентов и ковариационные матрицы модели (9)-(15) могут зависеть от вектора неизвестных параметров θ, который и необходимо оценить. Для этого используется метод максимального правдоподобия.

$$f(y) = \prod_{t=1}^{T} f(y_t \mid \tilde{y}_{t-1}) \tag{46}$$

В нашем случае (гауссовская линейная модель пространства состояния) предполагается, что ошибки имеют совместное нормальное распределение, поэтому вклады в функцию правдоподобия представляют собой условные нормальные распределения:

$$y_t \mid \tilde{y}_{t-1} \sim N(y_{t|t-1}, F_t) \tag{47}$$

Отсюда по формуле плотности многомерного нормального распределения:

$$\ln f(y_t \mid \tilde{y}_{t-1}) = -\frac{k}{2}\ln(2\pi) - \frac{1}{2}\ln \mid F_t \mid - \frac{1}{2}(y_t - y_{t|t-1})' F_t^{-1}(y_t - y_{t|t-1})$$

Заметим, что $y_t - y_{t|t-1}$ - ошибки прогноза на один шаг ν_t.

$$\mathcal{L}_t(\theta) = -\frac{k}{2}\ln(2\pi) - \frac{1}{2}\ln|F_t(\theta)| - \frac{1}{2}\nu_t(\theta)' F_t^{-1}\nu_t(\theta) \qquad (48)$$

В целом логарифмическая функция правдоподобия - это сумма таких вкладов:

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} \mathcal{L}_t(\theta) \Rightarrow \qquad (49)$$

$$\hat{\theta} \in \arg\max_{\theta} \mathcal{L}_t(\theta) \qquad (50)$$

$\hat{\theta}$ - ML-оценка гауссовской линейной модели пространства состояний.

Пример использования state-space models (1)

- Регрессия с меняющимися коэффициентами

$$y_t = x_t \beta_t + \epsilon_t, \quad t = 1, 2, \ldots, T \tag{51}$$

$$\beta_t = \mu + F\beta_{t-1} + u_t \tag{52}$$

$$\epsilon_t \sim N(0, R) \tag{53}$$

$$u_t \sim N(0, Q) \tag{54}$$

Если $\mu = 0$ и F - единичная матрица, то $\beta_i \sim RW$. Если F - диагональная матрица с элементами по главной диагонали меньше единицы по абсолютной величине, то $\beta_i \sim AR(1)$