WEEK 7. 상태공간모형

7-1. 상태공간모형의 표현

- 관측되는 변수와 관측되지 않는 변수를 구분함

- 사건에 따라 변하는 관측되지 않는 변수를 상태변수라 함

- 관측방정식, 상태방정식으로 구성됨

- 회귀, ARMA를 포함하는 광범위한 모형임

모형의 표현

- (예 1) 시간 t 관측치를 Y_t 라 할 때 이는 상태변수 μ_t 와 관측오차 w_t로 부터 얻어지고 상태변수는 이전 시간의 상태와 오차 v_t가 더해져 변한다고 하자. 이 때, 상태공간모형은 다음으로 표현된다.
  - (관측방정식) $Y_t = \mu_t + w_t$
  - (상태방정식) $\mu_t = \mu_{t-1} + v_t$
  - 오차항 가정: $w_t \sim Nor(0, \sigma_w^2)$, $v_t \sim Nor(0, \sigma_v^2)$ 이며 다른 변수와 독립 → 상태변수
- (예 2) 시간 t 관측치를 Y_t 라 할 때 이는 수준 (level)변수 μ_t 와 관측오차 ε_t로 부터 얻어진다. 수준변수는 이전 시간의 수준, 추세(trend)와 오차 a_t가 더해져 변하고, 추세는 이전 추세와 오차 b_t가 더해져 변한다고 하자. 이 때, 상태공간모형은 다음으로 표현된다.
  - (관측방정식) $Y_t = \mu_t + \varepsilon_t$
  - (상태방정식) $\mu_t = \mu_{t-1} + \beta_{t-1} + a_t$ $\beta_t = \beta_{t-1} + b_t$
  - ⇒ 상태변수가 2개임. 이 경우 벡터/행렬로 모형을 표현

-> 각각의 오차항들은 정규분포를 따른다고 가정

(예7-1) 다음 모형을 일반적 형태(벡터/행렬)로 표현하라.

- (관측방정식) $Y_t = \mu_t + \varepsilon_t; \varepsilon_t \sim Nor(0, \sigma_\varepsilon^2)$
- (상태방정식) $\mu_t = \mu_{t-1} + \beta_{t-1} + a_t$ $\beta_t = \beta_{t-1} + b_t; a_t \sim Nor(0, \sigma_a^2); b_t \sim Nor(0, \sigma_b^2)$

(풀이)

- 상태방정식은 다음과 같이 표현된다.

$$\begin{pmatrix}\mu_t \\ \beta_t\end{pmatrix} = \begin{bmatrix}1 & 1 \\ 0 & 1\end{bmatrix}\begin{pmatrix}\mu_{t-1} \\ \beta_{t-1}\end{pmatrix} + \begin{pmatrix}a_t \\ b_t\end{pmatrix}$$

- 상태변수 벡터를 정의하면

$$\boldsymbol{x}_t = \begin{pmatrix}\mu_t \\ \beta_t\end{pmatrix} \Rightarrow \boldsymbol{x}_t = F\boldsymbol{x}_{t-1} + \boldsymbol{v}_t;\ Q = \begin{bmatrix}\sigma_a^2 & 0 \\ 0 & \sigma_b^2\end{bmatrix}$$

- 관측방정식

$$Y_t = (1 \quad 0)\begin{pmatrix}\mu_t \\ \beta_t\end{pmatrix} + \varepsilon_t; \mathrm{R} = \sigma_\varepsilon^2$$

- (관측방정식)
  $\boldsymbol{y}_t = G\boldsymbol{x}_t + \boldsymbol{w}_t, \boldsymbol{w}_t \sim WN(\boldsymbol{0}, R)$
- (상태방정식)
  $\boldsymbol{x}_t = F\boldsymbol{x}_{t-1} + \boldsymbol{v}_t, \boldsymbol{v}_t \sim WN(\boldsymbol{0}, Q)$

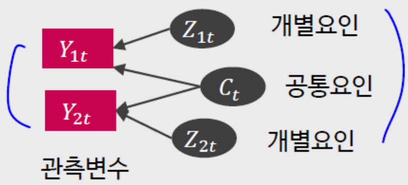
- 상태 변수가 2개이므로, 벡터 형식으로 표현해야함

- 뮤_t와 베타_t를 결합하여 x_t라는 상태변수 벡터를 정의한 것임

- x_t에 대한 x_t-1을 생각하면, 계수 F에 대해 식을 쓸 수 있음

- 관측방정식은 관측 변수가 1개이므로 y_t가 상태 변수 x_t로 표현됨

(예7-2) 다음 그림과 같은 모형을 고려하자.



상태변수 3개

$$\boldsymbol{y}_t = \begin{pmatrix} Y_{1t} \\ Y_{2t} \end{pmatrix}, \boldsymbol{x}_t = \begin{pmatrix} C_t \\ Z_{1t} \\ Z_{2t} \end{pmatrix}$$

- **관측방정식**

  $Y_{1t} = \gamma_1 C_t + Z_{1t}; Y_{2t} = \gamma_2 C_t + Z_{2t}$

- **상태방정식**
  - (공통요인) $C_t = \phi C_{t-1} + v_t, v_t \sim Nor(0,1)$
  - (개별요인1) $Z_{1t} = \alpha_1 Z_{1,t-1} + \varepsilon_{1t}, \varepsilon_{1t} \sim Nor(0, \sigma_1^2)$
  - (개별요인2) $Z_{2t} = \alpha_2 Z_{2,t-1} + \varepsilon_{2t}, \varepsilon_{2t} \sim Nor(0, \sigma_2^2)$

(관측방정식)

$$\boldsymbol{y}_t = G\boldsymbol{x}_t, \ G = \begin{bmatrix} \gamma_1 & 1 & 0 \\ \gamma_2 & 0 & 1 \end{bmatrix}$$

(상태방정식)

$$\boldsymbol{x}_t = F\boldsymbol{x}_{t-1} + \boldsymbol{v}_t, \boldsymbol{v}_t \sim WN(\boldsymbol{0}, Q)$$

$$F = \begin{bmatrix} \phi & 0 & 0 \\ 0 & \alpha_1 & 0 \\ 0 & 0 & \alpha_2 \end{bmatrix} \ Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sigma_1^2 & 0 \\ 0 & 0 & \sigma_2^2 \end{bmatrix}$$

- 관측 변수가 2개인 경우임

- 각각의 관측 변수에 개별적으로 영향을 미치는 요인이 하나씩 있고 공통요인 C_t가 존재함

- 상태변수는 총 3개임

7-2. 칼만 필터

- 최적선형예측식: 상태공간모형에서는 미래의 값을 예측하기 위해 우선적으로 상태 변수 x_t를 예측해야 함

- 새로운 관측치가 발생할 때, 상태변수의 예측식을 예측식을 이전 예측식으로부터 갱신하는 것을 칼만 필터라고 함

- 최적선형예측식은 베이지안 기법에 의거하여, 사전확률분포로부터 사후확률분포의 기댓값과 분산을 구하는 것임

- 목적: 확률변수 L를 예측하는데 관측치 Y의 선형식 사용
- 예측식 형태: $E[L|Y] = \alpha + \beta Y$
  - α, β: 추정필요 계수
- 계수추정을 위해 다음 제곱합 기대치 최소화
  - $Q = E[(L - \alpha - \beta Y)^2]$
  - Q를 풀면 다음과 같다
  - $Q = Var[L] + E^2[L] + \beta^2\{Var[Y] + E^2[Y]\} - 2\beta\{Cov[L,Y] + E[L]E[Y]\} + \alpha^2 - 2\alpha E[L] + 2\alpha\beta E[Y]$
  - $\frac{\partial Q}{\partial \alpha} = 0; \frac{\partial Q}{\partial \beta} = 0$ 풀면
  - $\alpha^* = E[L] - \beta^* E[Y]$
  - $\beta^* = \frac{Cov[L,Y]}{Var[Y]}$
- 최적 선형예측식
  - $E^*[L|Y] = \alpha^* + \beta^* Y = E[L] + \frac{Cov[L,Y]}{Var[Y]}(Y - E[Y])$
  - $Q^* = Var[L] - \frac{Cov^2[L,Y]}{Var[Y]}$

최적선형예측식과 칼만필터

- 최적선형예측식
  - $E[L|Y] = \alpha^* + \beta^* Y = E[L] + \frac{Cov[L,Y]}{Var[Y]}(Y - E[Y])$
- 시계열 $Y_t, Y_{t-1}, Y_{t-2}, \dots$ 이 있는경우
  - $E[L|Y] = E[L|Y_t, Y_{t-1}, \dots]$: 가장 최근 관측치 Y_t를 포함한 조건부 기대치
  - $E[L] = E[L|Y_{t-1}, \dots]$: 최근 관측치 Y_t가 없고 과거치만의 조건부 기대치

베이지안 관점에서
  - $E[L|Y_{t-1}, \dots]$: 사전확률분포의 기대치
  - $E[L|Y_t, Y_{t-1}, \dots]$: 사후확률분포의 기대치
- 아래 최적선형예측식을 칼만필터라 함
  - $E[L|Y_t, Y_{t-1}, \dots] = E[L|Y_{t-1}, \dots] + \frac{Cov[L, Y_t|Y_{t-1}, \dots]}{Var[Y_t|Y_{t-1}, \dots]}\underbrace{(Y_t - E[Y_t|Y_{t-1}, \dots])}_{\text{예측오차}}$

- 주어진 시계열에 대해서, L(상태 변수)을 예측해야 함

- 가장 최근 데이터 Y_t로부터 t 시점이 포함된 경우와 그렇지 않은 경우로 나누어 L의 조건부 기댓값을 구함

- 베이지안 관점에서 사전/사후확률분포를 정의함

- 사전확률분포에 예측오차의 일정 부분을 더하여 사후확률분포의 기댓값을 구함

(예7-3) 다음 모형에서 상태변수들의 칼만 필터식을 유도하라.

- (관측방정식) $Y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim Nor(0, \sigma_\varepsilon^2)$
- (상태방정식) $\mu_t = \mu_{t-1} + \beta_{t-1} + a_t, \quad a_t \sim Nor(0, \sigma_a^2)$

  $\beta_t = \beta_{t-1} + b_t, \quad b_t \sim Nor(0, \sigma_b^2)$

(풀이)

- 기호정의
- $l_t = E[\mu_t|Y_t, \dots], m_t = E[\beta_t|Y_t, \dots]$
- $p_t = Var[\mu_t|Y_t, \dots], q_t = Var[\beta_t|Y_t, \dots], r_t = Cov[\mu_t, \beta_t|Y_t, \dots]$
- 상태변수 μ_t의 최적선형예측식 (칼만필터)

$$l_t = E[\mu_t|Y_t, \dots] = E[\mu_t|Y_{t-1}, \dots] + \frac{Cov[\mu_t, Y_t|Y_{t-1}, \dots]}{Var[Y_t|Y_{t-1}, \dots]}(Y_t - E[Y_t|Y_{t-1}, \dots])$$

$$E[\mu_t|Y_{t-1}, \dots] = E[\mu_{t-1} + \beta_{t-1} + a_t|Y_{t-1}, \dots] = l_{t-1} + m_{t-1}$$

$$l_t = l_{t-1} + m_{t-1} + K_t^\mu(Y_t - l_{t-1} - m_{t-1}), K_t^\mu = \frac{p_{t-1} + q_{t-1} + 2r_{t-1} + \sigma_a^2}{p_{t-1} + q_{t-1} + 2r_{t-1} + \sigma_a^2 + \sigma_\varepsilon^2}$$

- 상태 변수 2개에 대한 조건부 기대값을 구하는 것이 목표임

- 뮤_t, 베타_t 각각에 대해 칼만 필터를 적용함

(예7-3 계속)

- 상태변수 β_t의 최적선형예측식 (칼만필터)

$$m_t = E[\beta_t|Y_t,\dots] = E[\beta_t|Y_{t-1},\dots] + \frac{Cov[\beta_t, Y_t|Y_{t-1},\dots]}{Var[Y_t|Y_{t-1},\dots\}}(Y_t - E[Y_t|Y_{t-1},\dots])$$

$$= m_{t-1} + K_t^{\beta}(Y_t - l_{t-1} - m_{t-1}),\quad K_t^{\beta} = \frac{q_{t-1}+r_{t-1}}{p_{t-1}+q_{t-1}+2r_{t-1}+\sigma_a^2+\sigma_\varepsilon^2}$$

- 분산의 갱신 공식

$$p_t = Var[\mu_t|Y_t,\dots] = Var[\mu_t|Y_{t-1},\dots] - \frac{Cov^2[\mu_t, Y_t|Y_{t-1},\dots]}{Var[Y_t|Y_{t-1},\dots\}} = K_t^{\mu}\sigma_\varepsilon^2$$

$$q_t = Var[\beta_t|Y_t,\dots] = q_{t-1} + \sigma_b^2 - K_t^{\beta}(q_{t-1} + r_{t-1})$$

$$r_t = Cov[\mu_t, \beta_t|Y_t,\dots] = K_t^{\beta}\sigma_\varepsilon^2$$

- 일반적 상태공간모형
  - (관측방정식) $\boldsymbol{y}_t = G\boldsymbol{x}_t + \boldsymbol{w}_t,\ \boldsymbol{w}_t \sim WN(\boldsymbol{0}, R)$
  - (상태방정식) $\boldsymbol{x}_t = F\boldsymbol{x}_{t-1} + \boldsymbol{v}_t,\ \boldsymbol{v}_t \sim WN(\boldsymbol{0}, Q)$

- 상태변수 기대치 및 공분산
  - 기대치 벡터: $\boldsymbol{m}_t = E[\boldsymbol{x}_t|\boldsymbol{y}_t,\dots]$
  - 공분산행렬: $P_t = Var[\boldsymbol{x}_t|\boldsymbol{y}_t,\dots]$

- 베이지안 기법
  - $\boldsymbol{x}_t$의 사전확률분포 $\boldsymbol{x}_t|\boldsymbol{y}_{t-1}$로 부터 관측치 $\boldsymbol{y}_t$ 발생시 사후확률분포 $\boldsymbol{x}_t|\boldsymbol{y}_t$ 도출
  - 칼만 필터 공식

$$\boldsymbol{m}_t = F\boldsymbol{m}_{t-1} + K_t(\boldsymbol{y}_t - GF\boldsymbol{m}_{t-1})$$

$$P_t = B_t - K_t G B_t$$

$$K_t = B_t G^T (G B_t G^T + R)^{-1},\ B_t = F P_{t-1} F^T + Q$$

7-3. 모형의 추정 및 예측

추정: MLE 방식으로 계수를 추정하고 예측함

(예 7-5) 다음의 MA(1) 모형을 상태공간모형으로 표현하고 최우추정법으로 계수 및 오차분산을 추정하라.

$$Y_t = a_t - \theta a_{t-1},\ a_t \sim Nor(0, \sigma^2)$$

(풀이)

- 상태변수 정의 $\boldsymbol{x}_t = \begin{pmatrix} a_t \\ a_{t-1} \end{pmatrix}$
- (관측방정식) $Y_t = (1 \quad -\theta)\begin{pmatrix} a_t \\ a_{t-1} \end{pmatrix} \Rightarrow G = (1 \quad -\theta),\ R = 0$
- (상태방정식) $\begin{pmatrix} a_t \\ a_{t-1} \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{pmatrix} a_{t-1} \\ a_{t-2} \end{pmatrix} + \begin{pmatrix} a_t \\ 0 \end{pmatrix} \Rightarrow F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, Q = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$
- 칼만필터식

$$m_t = E[a_t|Y_t] = \frac{\sigma^2}{\sigma^2 + \theta^2 p_{t-1}}(Y_t + \theta m_{t-1})$$

$$p_t = Var[a_t|Y_t] = \frac{\sigma^2\theta^2 p_{t-1}}{\sigma^2 + \theta^2 p_{t-1}}$$

(예 7-5 계속)

- 관측치 조건부 기대치 및 분산
  - $\mu_t = E[Y_t|Y_{t-1}] = -\theta m_{t-1}$
  - $C_t = Var[Y_t|Y_{t-1}] = \sigma^2 + \theta^2 p_{t-1}$
- 로그 우도함수

$$logL(\boldsymbol{y}_1, \dots, \boldsymbol{y}_T) = -\frac{Td}{2}log2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|C_t| - \frac{1}{2}\sum_{t=1}^{T}(\boldsymbol{y}_t - \boldsymbol{\mu}_t)^T C_t^{-1}(\boldsymbol{y}_t - \boldsymbol{\mu}_t)$$

$$= -\frac{1}{2}\left[Tlog2\pi + \sum_{t=1}^{T}\log(\sigma^2 + \theta^2 p_{t-1}) + \sum_{t=1}^{T}\frac{(Y_t + \theta m_{t-1})^2}{\sigma^2 + \theta^2 p_{t-1}}\right]$$ → 최대가 되는 σ^2, θ 찾기

- 관측 데이터

	1	2	3	4	5	6	7	8	9	10	11	12
Y_t	8	10	-9	13	-5	-15	24	6	-21	20	-7	-24

- 최우추정치
- 초기치: $m_0 = 0, p_0 = \sigma^2$ $\hat{\theta} = 0.85, \widehat{\sigma^2} = 140$

(예 7-6) 다음 모형에서 관측치를 예측하라.

- (관측방정식) $Y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim Nor(0, \sigma_\varepsilon^2)$
- (상태방정식) $\mu_t = \mu_{t-1} + \beta_{t-1} + a_t, \quad a_t \sim Nor(0, \sigma_a^2)$

  $\beta_t = \beta_{t-1} + b_t, \quad b_t \sim Nor(0, \sigma_b^2)$

추정 필요

- (풀이)

$$l_t = E[\mu_t|Y_t, \dots], m_t = E[\beta_t|Y_t, \dots]$$

$$p_t = Var[\mu_t|Y_t, \dots], q_t = Var[\beta_t|Y_t, \dots], r_t = Cov[\mu_t, \beta_t|Y_t, \dots]$$

$$f_{t,1} = E[Y_{t+1}|Y_t, \dots] = l_t + m_t$$

$$V_{t,1} = Var[Y_{t+1}|Y_t, \dots] = p_t + 2r_t + q_t + \sigma_a^2 + \sigma_\varepsilon^2$$