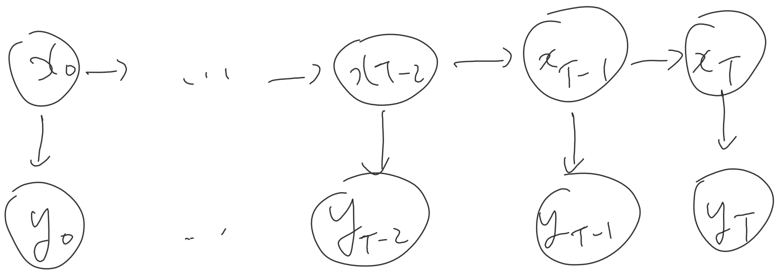
15주 Extra-3-Kalman_filter_Inference

2019년 12월 23일 월요일 오전 3:15



$$x_t = \alpha x_{t-1} + \beta + \varepsilon_x$$

$$y_t = \lambda x_t + \varepsilon_y$$

$$\begin{cases} \varepsilon_x, \varepsilon_y \sim \text{Gaussian} \\ x_0, y_0 \sim \text{Gaussian} \end{cases}$$

$\Rightarrow p(x_0, x_1, \cdots, x_T, y_0, y_1, \cdots, y_T)$: Jointly Gaussian

$$= p(x_0)\,p(y_0|x_0)\cdots p(x_t|x_{t-1})\,p(y_t|x_t)\cdots p(x_T|x_{t-1})\,p(y_T|x_T)$$

$$\boxed{\begin{array}{l}\text{Inference} \\ \quad \Rightarrow p(x_T | y_0, y_1, \cdots, y_T)\end{array}}$$

All variables are the parameters in the marginals.

undetermined parameters

determined parameters

$$\Sigma = \begin{pmatrix} \sigma_{y_0}^2 & & & & & & \\ & \sigma_{x_0}^2 & & & & & \\ & & \ddots & & & & \\ & & & \sigma_{y_t}^2 \quad \sigma_{y_t x_t} & & \sigma_{y_t x_{t+2}} & \sigma_{y_t x_T} \\ & & & \sigma_{x_t}^2 \quad \sigma_{x_t y_{t+1}} \quad \sigma_{x_t x_{t+1}} & & \sigma_{x_t x_{t+2}} \cdots & \\ & & & \sigma_{y_{t+1}}^2 & & \sigma_{y_{t+1} x_{t+2}} \cdots & \\ & & & \sigma_{x_{t+1}}^2 & & & \\ & & & \sigma_{y_{t+2}}^2 & & & \end{pmatrix}$$

$$
\begin{pmatrix}
\ddots & & & & & & \\
& x_{t+1} & & & & & \\
& & \sigma^2_{y_{t+2}} & & & & \\
& & & \sigma^2_{x_{t+2}} & & & \\
& & & & \ddots & & \\
& & & & & \sigma^2_{y_T} & \\
& & & & & & \sigma^2_{x_T}
\end{pmatrix}
$$

Note : $\mu_{a|b} = \mu_a + \Sigma_{ab}\Sigma_b^{-1}(x_b - \mu_b)$

$\Sigma_{a|b} = \Sigma_a - \Sigma_{ab}\Sigma_b^{-1}\Sigma_{ba}$

$\Bigg\{$ For the calculation of joint $p(x_0, \cdots, x_T, y_0, \cdots, y_T)$, we don't need $\sigma_{y_t x_T}$

For the calculation of conditional $p(x_T \mid y_0, \cdots, y_T)$, we need $\sigma_{y_t x_T}$.

Plan:

$$p(x_t, x_{t+1}) = \mathcal{N}\left(\begin{pmatrix} \mu_{x_t} \\ \mu_{x_{t+1}} \end{pmatrix}, \begin{pmatrix} \sigma^2_{x_t} & \sigma_{x_t x_{t+1}} \\ & \sigma^2_{x_{t+1}} \end{pmatrix} \right)$$

$$p(x_t, y_t) = \mathcal{N}\left(\begin{pmatrix} \mu_{x_t} \\ \mu_{y_t} \end{pmatrix}, \begin{pmatrix} \sigma^2_{x_t} & \sigma_{x_t y_t} \\ & \sigma^2_{y_t} \end{pmatrix} \right)$$

$\Rightarrow$ Represent everything using μ_{x_t} & $\sigma^2_{x_t}$.

$$\overset{\text{marginal}}{\mu_{x_{t+1}}} = E[\alpha x_t + \beta + \varepsilon_x]$$

$$= \alpha \mu_{x_t} + \beta$$

$\mu_{y_t} = E[\lambda x_t + \varepsilon_y] = \lambda \mu_{x_t}$

$$\cdots \mu_{x_t} + \beta$$

$$\mu_{y_t} = E[\lambda x_t + \varepsilon_y] = \lambda \mu_{x_t}$$

$$\sigma^2_{x_{t+1}} = E[x^2_{t+1}] - E[x_{t+1}]^2$$

$$= E[(\alpha x_t + \beta + \varepsilon_x)^2] - E[\alpha x_t + \beta + \varepsilon_x]^2$$

$$= \alpha^2 E[x_t^2] + \beta^2 + E[\varepsilon_x^2] + 2\alpha\beta E[x_t]$$

$$- \alpha^2 E[x_t]^2 - \beta^2 - 2\alpha\beta E[x_t]$$

$$= \alpha^2 (E[x_t^2] - E[x_t]^2) + E[\varepsilon_x^2]$$

$$= \alpha^2 \sigma^2_{x_t} + E[\varepsilon_x^2]$$

$$\sigma_{x_{t+1} x_t} = E[x_{t+1} x_t] - E[x_{t+1}] E[x_t]$$

$$= E[(\alpha x_t + \beta + \varepsilon_x) x_t] - (\alpha \mu_{x_t} + \beta) \mu_{x_t}$$

marginal means

$$= \alpha E[x_t^2] + \cancel{\beta \underbrace{E[x_t]}_{= \mu_{x_t}}} - \alpha \mu^2_{x_t} - \cancel{\beta \mu_{x_t}}$$

$$= \alpha (E[x_t^2] - \mu^2_{x_t})$$

$$= \alpha \sigma^2_{x_t}$$

$$\sigma^2_{y_t} = E[y_t^2] - E[y_t]^2$$

$$= E[(\lambda x_t + \varepsilon_y)^2] - E[\lambda x_t + \varepsilon_y]^2$$

$$= E[\lambda^2 x_t^2] + E[\varepsilon_y^2] - (\lambda \mu_{x_t})^2$$

marginal mean

$$= \lambda^2 (E[x_t^2] - \mu^2_{x_t}) + E[\varepsilon_y^2]$$

$$= \lambda^2 \sigma^2 + E[\varepsilon^2]$$

$$= \lambda^2 \sigma_{x_t}^2 + E[\varepsilon_y^2]$$

$$\sigma_{x_t y_t} = E[x_t y_t] - E[x_t]E[y_t]$$
$$= E[x_t(\lambda x_t + \varepsilon_y)] - \mu_{x_t} \cdot \lambda \mu_{x_t}$$
$$= \lambda E[x_t^2] - \lambda \mu_{x_t}^2$$
$$= \lambda \sigma_{x_t}^2$$

$$\therefore \sigma_{x_{t+1} y_{t+1}} = \lambda \sigma_{x_{t+1}}^2$$
$$= \lambda \alpha^2 \sigma_{x_t}^2 + \lambda E[\varepsilon_x^2]$$
$$\sigma_{y_{t+1}}^2 = \lambda^2 \sigma_{x_{t+1}}^2 + E[\varepsilon_y^2]$$
$$= \lambda^2 \alpha^2 \sigma_{x_t}^2 + \lambda^2 E[\varepsilon_x^2] + E[\varepsilon_y^2]$$

eee

Do I need to calculate $\sigma_{x_{t+1} y_t}$? No.

Only when we need marginals of joint $p(x_{t+1}, y_t)$.

(Due to the filtering structure.)

$$\begin{pmatrix} x_t \\ x_{t+1} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{x_t} \\ \alpha \mu_{x_t} + \beta \end{pmatrix}, \begin{pmatrix} \sigma_{x_t}^2 & \alpha \sigma_{x_t}^2 \\ \alpha \sigma_{x_t}^2 & \alpha^2 \sigma_{x_t}^2 + E[\varepsilon_{x_t}^2] \end{pmatrix} \right)$$

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_{x_t} \\ \lambda \mu_{x_t} \end{pmatrix}, \begin{pmatrix} \sigma_{x_t}^2 & \lambda \sigma_{x_t}^2 \\ \lambda \sigma_{x_t}^2 & \lambda^2 \sigma_{x_t}^2 + E[\varepsilon_y^2] \end{pmatrix} \right)$$

$\mu_{x_t}, \mu_{y_t}, \sigma_{x_t}^2, \sigma_{x_t y_t}, \sigma_{y_t}^2$ in the marginal density are defined recursively.

$$p(x_t, x_{t+1}, y_t) = p(x_t, x_{t+1})\, p(y_t | x_t)$$

Conditional

$$\mu_{x_{t+1}|y_0 \cdots y_t} = E[\alpha x_t + \beta + \varepsilon_x | y_0, \cdots, y_t]$$

$$= a\mu_{x_t|y_0 \cdots y_t} + \beta$$

Time update

$$\mu_{x_{t+1}|y_0, \cdots, y_{t+1}} = \mu_{x_{t+1}} + \Sigma_{x_{t+1}, y_0 \cdots y_{t+1}} \Sigma^{-1}_{y_0 \cdots y_{t+1}} \left(\begin{pmatrix} y_0 \\ \vdots \\ y_{t+1} \end{pmatrix} - E\left[\begin{pmatrix} y_0 \\ \vdots \\ y_{t+1} \end{pmatrix} \right] \right)$$

(방법 1) - Not good.

covariances in the marginals

Observation update

$O((t+2)^3) \leftarrow$ one update

(방법 2) - recursive

$$\mu_{x_{t+1}|y_0 \cdots y_{t+1}}$$

$$= \mu_{x_{t+1}|y_0 \cdots y_t} + \sigma_{x_{t+1} y_{t+1}|y_0 \cdots y_t} \sigma^{-2}_{y_{t+1}|y_0 \cdots y_t} \left(y_{t+1} - E[\lambda x_{t+1} + \varepsilon_y | y_0 \cdots y_t] \right)$$

$$= \mu_{x_{t+1}|y_0 \cdots y_t} + \sigma_{x_{t+1} y_{t+1}|y_0 \cdots y_t} \sigma^{-2}_{y_{t+1}|y_0 \cdots y_t} \left(y_{t+1} - \lambda \mu_{x_{t+1}|y_0 \cdots y_t} \right) \quad \cdots (1)$$

$O(T^3)$ vs. $O(T)$

Observation update

$O(1) \leftarrow$ one update

$$\sigma^2_{x_{t+1}|y_0, \cdots, y_t} = E[x_{t+1}^2 | y_0, \cdots, y_t] - E[x_{t+1} | y_0, \cdots, y_t]^2$$

$$= E\left[(\alpha x_t + \beta + \varepsilon_x)^2 \mid y_0, \cdots, y_t\right]$$
$$- E\left[\alpha x_t + \beta + \varepsilon_x \mid y_0 \cdots, y_t\right]^2$$
$$= \alpha^2 E[x_t^2 \mid y_0, \cdots, y_t] + \beta^2 + \underbrace{E[\varepsilon_x^2 \mid y_0, \cdots, y_t]}_{= E[\varepsilon_x^2]}$$
$$+ 2\alpha\beta E[x_t \mid y_0, \cdots, y_t] + 2\beta \underbrace{E[\varepsilon_x \mid y_0 \cdots, y_t]}_{=0}$$
$$+ 2\alpha E[x_t \mid y_0, \cdots, y_t] \underbrace{E[\varepsilon_x \mid y_0, \cdots, y_t]}_{=0}$$
$$- \left(\alpha E[x_t \mid y_0, \cdots, y_t] + \beta\right)^2$$
$$= \alpha^2 \left(E[x_t^2 \mid y_0, \cdots, y_t] - E[x_t \mid y_0, \cdots, y_t]^2\right)$$
$$+ E[\varepsilon_x^2]$$
$$= \alpha^2 \sigma^2_{x_t \mid y_0, \cdots, y_t} + E[\varepsilon_x^2]$$

Time update

$$\sigma^2_{y_{t+1} \mid y_0 \cdots y_t} = E\left[(\lambda x_{t+1} + \varepsilon_y)^2 \mid y_0, \cdots, y_t\right]$$
$$- E\left[\lambda x_{t+1} + \varepsilon_y \mid y_0, \cdots, y_t\right]^2$$
$$= \lambda^2 E[x_{t+1}^2 \mid y_0, \cdots, y_t] + E[\varepsilon_y^2 \mid y_0, \cdots, y_t]$$
$$- \lambda^2 E[x_{t+1} \mid y_0, \cdots, y_t]^2$$
$$= \lambda^2 \sigma^2_{x_{t+1} \mid y_0 \cdots y_t} \quad E[\varepsilon_y^2]$$

$$\sigma_{x_{t+1} y_{t+1} \mid y_0, \cdots, y_t} = E\left[x_{t+1}(\lambda x_{t+1} + \varepsilon_y) \mid y_0, \cdots, y_t\right]$$
$$- E\left[x_{t+1} \mid y_0, \cdots, y_t\right] E\left[\lambda x_{t+1} + \varepsilon_y \mid y_0 \cdots y_t\right]$$
$$= \lambda E\left[x_{t+1}^2 \mid y_0, \cdots, y_t\right] - \lambda E\left[x_{t+1} \mid y_0, \cdots, y_t\right]^2$$

$,2$

$$= \lambda E[x_{t+1}^2 | y_0, \cdots, y_t] - \lambda E[x_{t+1} | y_0, \cdots, y_t]^2$$

$$= \lambda \sigma^2_{x_{t+1}|y_0, \cdots, y_t}$$

$$\sigma^2_{x_{t+1}|y_0, \cdots, y_{t+1}} = \sigma^2_{x_{t+1}|y_0, \cdots, y_t}$$

$$- \sigma_{x_{t+1} y_{t+1} | y_0 \cdots y_t} \sigma^{-2}_{y_{t+1}|y_0 \cdots y_t} \sigma_{y_{t+1} x_{t+1} | y_0 \cdots y_t}$$

$$= \sigma^2_{x_{t+1}|y_0, \cdots, y_t} - \frac{\lambda^2 \sigma^4_{x_{t+1}|y_0 \cdots y_t}}{\lambda^2 \sigma^2_{x_{t+1}|y_0, \cdots, y_t} + E[\varepsilon_y^2]}$$

$$= \sigma^2_{x_{t+1}|y_0, \cdots, y_t} \left\{ 1 - \frac{1}{1 + \frac{1}{\lambda^2} \cdot \frac{E[\varepsilon_y^2]}{\sigma^2_{x_{t+1}|y_0 \cdots y_t}}} \right\}$$

$\Rightarrow$ Variance Reduction after observation

Observation update

From (1),

$$\mu_{x_{t+1}|y_0 \cdots y_{t+1}} = \mu_{x_{t+1}|y_0 \cdots y_t}$$

$$+ \sigma_{x_{t+1} y_{t+1} | y_0 \cdots y_t} \sigma^{-2}_{y_{t+1}|y_0 \cdots y_t} \left(y_{t+1} - \lambda \mu_{x_{t+1}|y_0 \cdots y_t} \right)$$

$$= \mu_{x_{t+1}|y_0 \cdots y_t}$$

$$+ \frac{\lambda \sigma^2_{x_{t+1}|y_0, \cdots, y_t}}{\lambda^2 \sigma^2_{x_{t+1}|y_0, \cdots, y_t} + E[\varepsilon_y^2]} \left(y_{t+1} - \lambda \mu_{x_{t+1}|y_0, \cdots, y_t} \right)$$

$$= \mu_{x_{t+1}|y_0, \cdots, y_t} + \frac{1}{\cdots 1 \; E[\varepsilon \; 2]} \left(\frac{y_{t+1}}{} - \mu_{x_{t+1}|y_0, \cdots, y_t} \right)$$

$$= \mu_{x_{t+1}|y_0,\cdots,y_t} + \frac{1}{1+\frac{1}{\lambda^2}\frac{E[\varepsilon_y^2]}{\sigma^2_{x_{t+1}|y_0,\cdots,y_t}}}\left(\frac{y_{t+1}}{\lambda} - \mu_{x_{t+1}|y_0,\cdots,y_t}\right)$$

$\Rightarrow$ Mean Update after observation.

Observation update

If $\frac{E[\varepsilon_y^2]}{\lambda^2} \uparrow$, the observation update is small.

Summary:

Time update

$$\mu_{x_{t+1}|y_0,\cdots,y_t} = \alpha\,\mu_{x_t|y_0,\cdots,y_t} + \beta$$

$$\sigma^2_{x_{t+1}|y_0,\cdots,y_t} = \alpha^2\sigma^2_{x_t|y_0,\cdots,y_t} + E[\varepsilon_x^2]$$

Observation update

$$\mu_{x_{t+1}|y_0,\cdots,y_{t+1}} = \mu_{x_{t+1}|y_0,\cdots,y_t} + \frac{1}{1+\frac{1}{\lambda^2}\frac{E[\varepsilon_y^2]}{\sigma^2_{x_{t+1}|y_0,\cdots,y_t}}}\left(\frac{y_{t+1}}{\lambda} - \mu_{x_{t+1}|y_0\cdots y_t}\right)$$

$$\sigma^2_{x_{t+1}|y_0,\cdots,y_{t+1}}$$

$$= \sigma^2_{x_{t+1}|y_0,\cdots,y_t} \left\{ 1 - \frac{1}{1 + \frac{1}{\lambda^2} \frac{E[\varepsilon_y^2]}{\sigma^2_{x_{t+1}|y_0\cdots y_t}}} \right\}$$