Time Series and Sequence Learning

Lecture 5c – Kalman Filter

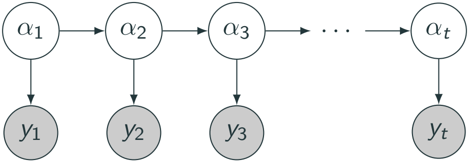
Johan Alenlöv, Linköping University

A general state space model

Def. A **Linear Gaussian State-Space (LGSS)** model is given by:

$$\alpha_t = T\alpha_{t-1} + R\eta_t, \qquad \eta_t \sim \mathcal{N}(0, Q),$$
$$y_t = Z\alpha_t + \varepsilon_t \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2),$$

and initial distribution $\alpha_1 \sim \mathcal{N}(a_1, P_1)$.



Filtering, Smoothing, and Predicting

Given a time-series $y_{1:n} = (y_1, y_2, \ldots, y_n)$ we wish to calculate the distribution of α_t conditioned on the observed time-series $y_{1:n}$.

This problem changes depending on the relationship of n and t:

$n < t$: This is known as the **forecasting** problem.

Filtering, Smoothing, and Predicting

Given a time-series $y_{1:n} = (y_1, y_2, \ldots, y_n)$ we wish to calculate the distribution of α_t conditioned on the observed time-series $y_{1:n}$.

This problem changes depending on the relationship of n and t:

$n < t$: This is known as the **forecasting** problem.

$n = t$: This is known as the **filtering** problem.

Filtering, Smoothing, and Predicting

Given a time-series $y_{1:n} = (y_1, y_2, \ldots, y_n)$ we wish to calculate the distribution of α_t conditioned on the observed time-series $y_{1:n}$.

This problem changes depending on the relationship of n and t:

$n < t$: This is known as the **forecasting** problem.

$n = t$: This is known as the **filtering** problem.

$n > t$: This is known as the **smoothing** problem.

The Kalman filter

For any s, t, denote by $\hat{\alpha}_{t|s} = \mathbb{E}[\alpha_t \mid y_{1:s}]$ and $P_{t|s} = \mathrm{Cov}(\alpha_t \mid y_{1:s})$.

The Kalman filter

For any s, t, denote by $\hat{\alpha}_{t|s} = \mathbb{E}[\alpha_t \mid y_{1:s}]$ and $P_{t|s} = \mathrm{Cov}(\alpha_t \mid y_{1:s})$.

Thm. For an LGSS model, $p(\alpha_t \mid y_{1:s}) = \mathcal{N}(\alpha_t \mid \hat{\alpha}_{t|s}, P_{t|s})$.

The Kalman filter

For any s, t, denote by $\hat{\alpha}_{t|s} = \mathbb{E}[\alpha_t \mid y_{1:s}]$ and $P_{t|s} = \mathrm{Cov}(\alpha_t \mid y_{1:s})$.

Thm. For an LGSS model, $p(\alpha_t \mid y_{1:s}) = \mathcal{N}(\alpha_t \mid \hat{\alpha}_{t|s}, P_{t|s})$.

Of particular interest are:

- Filtering distribution,

$$p(\alpha_t \mid y_{1:t}) = \mathcal{N}(\alpha_t \mid \hat{\alpha}_{t|t}, P_{t|t}).$$

- (1-step) Predictive distributions,

$$p(\alpha_t \mid y_{1:t-1}) = \mathcal{N}(\alpha_t \mid \hat{\alpha}_{t|t-1}, P_{t|t-1}),$$
$$p(y_t \mid y_{1:t-1}) = \mathcal{N}(y_t \mid \hat{y}_{t|t-1}, F_{t|t-1}).$$

Kalman filter

Kalman filter: For $t = 1, 2, \ldots$

- **Predict:**

  · Predict α_t:
  $$\begin{cases} \hat{\alpha}_{t|t-1} = T\hat{\alpha}_{t-1|t-1}, \\ P_{t|t-1} = TP_{t-1|t-1}T^{\mathsf{T}} + RQR^{\mathsf{T}} \end{cases} \qquad (\star)$$

  · Predict y_t:
  $$\begin{cases} \hat{y}_{t|t-1} = Z\hat{\alpha}_{t|t-1}, \\ F_{t|t-1} = ZP_{t|t-1}Z^{\mathsf{T}} + \sigma_\varepsilon^2 \end{cases}$$

- **Update:**

  · Kalman gain: $$K_t = P_{t|t-1}Z^{\mathsf{T}}F_{t|t-1}^{-1}$$

  · Update filter:
  $$\begin{cases} \hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + K_t(y_t - \hat{y}_{t|t-1}), \\ P_{t|t} = (I - K_tZ)P_{t|t-1} \end{cases} \qquad (\star\star)$$

($\star$) At time $t = 1$ we initialize $\hat{\alpha}_{1|0} = a_1$ and $P_{1|0} = P_1$.

($\star\star$) If y_t is missing we skip the update and set $\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1}$ and $P_{t|t} = P_{t|t-1}$.

ex) GMSL data