TSSL: Exercise session 2

Johan Alenlöv

Kalman Filter for the Linear Gaussian State Space Model

In the lectures we went over the Kalman filter for the local-level model. In the next laboration we will need to apply the Kalman filter on the linear Gaussian state-space model,

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

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

We will do this in an inductive fashion, we start by assuming that we have that

$$\alpha_t \mid y_{1:t-1} \sim \mathcal{N}[\hat{\alpha}_{t|t-1}, P_{t|t-1}].$$

We will then find the updating equations to time $t+1$ for both the filter and one step ahead predictor.

1. Find the filter distribution of α_t given $y_{1:t}$.
   (a) Let $v_t = y_t - \mathbb{E}[y_t \mid y_{1:t-1}]$. Find the joint distribution of (v_t, α_t) conditioned on the observations $y_{1:t-1}$.
   (b) Using the joint distribution above, find the conditional distribution of $\alpha_t \mid v_t, y_{1:t-1}$. Is this the same as the distribution $\alpha_t \mid y_{1:t}$?

   hint: the joint distribution (v_t, α_t) is multivariate Gaussian.
2. Find the predictive distribution of α_{t+1} given $y_{1:t}$.

EM-Algorithm

Now we focus on parameter estimation in the same model (1) as previously. In this exercise we will derive the parameter updating equations for the parameters $\theta = (Q, \sigma_\epsilon^2)$. Let $\tilde{\theta}$ be the current parameter values.

As we saw in the lecture we had that,

$$\mathcal{Q}(\theta, \tilde{\theta}) = \text{const.} - \frac{1}{2} \sum_{t=1}^{n} \big[\log |\sigma_\epsilon^2| + \log |Q| \\ + \{\hat{\varepsilon}_{t|n}^2 + \mathrm{Var}[\varepsilon_t \,|\, y_{1:n}]\}\sigma_\epsilon^{-2} + \mathrm{tr}[\{\hat{\eta}_{t|n}\hat{\eta}_{t|n}^\mathsf{T} + \mathrm{Var}[\eta_t \,|\, y_{1:n}]\}Q^{-1}]\big],$$

where the smoothing distributions are calculated using the current parameter values $\tilde{\theta}$. In the EM algorithm we are now tasked with maximizing this quantity.

3. Find the equations for θ that maximizes the intermediate quantity $\mathcal{Q}(\theta, \tilde{\theta})$.

   (a) Calculate the derivatives of the intermediate quantities,

   $$\frac{\partial}{\partial \theta} \mathcal{Q}(\theta, \tilde{\theta}).$$

   (b) Solve

   $$\frac{\partial}{\partial \theta} \mathcal{Q}(\theta, \tilde{\theta}) = 0.$$

   for θ.

Useful Results

Conditional Gaussian

Let (x, y) be jointly Gaussian with distribution

$$\begin{pmatrix} x \\ y \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{xy} & \Sigma_{yy} \end{pmatrix} \right].$$

Then the conditional distribution $x \,|\, y$ is also Gaussian with distribution

$$x \,|\, y \sim \mathcal{N}[\mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)\,,\, \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^\mathsf{T}].$$

Derivatives

We have that, if X is a square matrix then

$$\frac{\partial \log |X|}{\partial X} = (X^\mathsf{T})^{-1} = (X^{-1})^\mathsf{T}$$

$$\frac{\partial\, \mathrm{tr}[AX^{-1}B]}{\partial X} = -(X^{-1}BAX^{-1})^\mathsf{T}$$