Newton solve for Kalman Smoother

Claire He

June 19, 2023

1 Pupil example

Observation :

$$o_t = f_v(q_{tk}) + n_t \underset{\text{linear}}{=} B_t q_t + n_t, \quad n_t \sim N(0, D_t)$$

Latent :

$$q_t = A_t q_{t-1} + e_t = q_{t-1,k} + e_t, \quad e_t \sim N(0, E_t) \text{ and } q_0 \sim N(\mu_0, \Sigma_0)$$

Maximise likelihood ie minimise - loglikelihood.

$$\begin{aligned}
\arg\max_q \log p(q) \quad &\propto \quad \arg\max_q \log p(q_0|\mu_0, \Sigma_0) + \sum_{i=1}^{T-1} \log p(q_i|q_{i-1}, A) + \sum_{i=1}^{T-1} \log p(o_t|q_t, B) \\
&\propto \quad \arg\max -\frac{1}{2}\Bigg((q_0 - \mu_0)^T \Sigma_0^{-1} (q_0 - \mu_0) + \sum_{i=1}^{T-1} (q_i - Aq_{i-1}) E^{-1} (q_i - Aq_{i-1}) \\
&\quad + \sum_{i=1}^{T-1} (o_t - Bq_t)^T D^{-1} (o_t - Bq_t) \Bigg) \\
&\propto \quad \arg\min_q \sum_{i=1}^{T-1} (q_i - Aq_{i-1}) E^{-1} (q_i - Aq_{i-1}) + \sum_{i=1}^{T-1} (o_t - Bq_t)^T D^{-1} (o_t - Bq_t) \\
&\quad + (q_0 - \mu_0)^T \Sigma_0^{-1} (q_0 - \mu_0)
\end{aligned}$$

$$\text{Gradient } \nabla = \begin{pmatrix} \frac{\partial \log p(q)}{\partial q_0} \\ \frac{\partial \log p(q)}{\partial q_1} \\ \vdots \\ \frac{\partial \log p(q)}{\partial q_{T-1}} \end{pmatrix} = \begin{pmatrix} -A^T E^{-1}(q_1 - Aq_0) + S_0^{-1}(q_0 - \mu_0) \\ \vdots \\ (A^T E^{-1} A - E^{-1}) q_t + A^T E^{-1} q_{t+1} + (E^{-1})^T A q_{t-1} \\ \vdots \\ (A^T E^{-1} A - E^{-1}) q_{T-1} + (E^{-1})^T A q_{T-2} \end{pmatrix}$$

$$\text{Then hessian is block tridiagonal } H = \begin{pmatrix} -T_0 & -R_{01}^T & \dots & 0 \\ -R_{01} & -T_1 & -R_{12}^T & \vdots \\ 0 & -R_{12} & \ddots & 0 \\ \vdots & \ddots & \ddots & -R_{T-2,T-1}^T \\ 0 & \dots & -R_{T-2,T-1} & -T_{T-1} \end{pmatrix}$$

For $i \in \{1, \dots T-2\}$,

$$\begin{aligned}
R_{i,i+1} &= \frac{\partial^2}{\partial q_i \partial q_{i+1}} \log p(q_{i+1}|q_i) = A^T E^{-1} \\
T_i &= \frac{\partial^2}{\partial q_i^2} \log p(o_i|q_i) + \frac{\partial^2}{\partial q_i^2} \log p(q_{i+1}|q_i) + \frac{\partial^2}{\partial q_i^2} \log p(q_{i-1}|q_i) = -(E^{-1} + A^T E^{-1} A + B^T D^{-1} B)
\end{aligned}$$

and $T_0 = -(\Sigma_0^{-1} + A^T E^{-1} A)$ and $R_{01} = A^T E^{-1}$, $T_{T-1} = -(E^{-1} + B^T D^{-1} B)$ at the boundary.

Solve $\arg\max_q \log p(q)$ using Newton with $\hat{Q}^{i+1} = \hat{Q}^i - H^{-1}\nabla$. With a linear map should suffice to use one pass to get exact result $\hat{Q} = Q_0 - H^{-1}\nabla$.