Exercise sheet 5

Lars Kühmichel
Nicolas Wolf

Task 2. M-Step in a linear Gaussian state space model

Consider a linear Gaussian state space model,

$$z_t = Az_{t-1} + Cs_t + \epsilon, \qquad \epsilon \sim N(0,\Sigma)$$
$$x_t = Bz_t + \eta, \qquad \eta \sim N(0,\Gamma)$$

In the lecture we derived the M-step to determine the transition matrix A and last week you derived the M-step for the covariance matrix Σ. Now, derive the M-step for the latent control variable C, the factor loading matrix B and for the observation noise covariance matrix Γ by maximizing the expected log-likelihood $\mathbb{E}[\log p(X,Z)]$, with respect to C, B, and Γ where $X = \{x_t \mid t \in 1 \dots T\}$ and $Z = \{z_t \mid t \in 1 \dots T\}$ are the sets of all latent states and observations from time 1 to T.
(Hints: First spell out the ELBO for this model and then identify the parts in the ELBO that explicitly depends on the parameters and ignore the rest. Keep in mind that you want all maximization to depend only on $\mathbb{E}[z_t]$, $\mathbb{E}[z_t z_t^T]$, and $\mathbb{E}[z_t z_{t-1}^T]$. The matrix cookbook by Petersen and Pedersen (2012)
https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
provides a helpful summary for matrix algebra).

2

$$\mathbb{E}[\log p(x,z|\Theta)] = \mathbb{E}\Big[\log\Big[p(z_1)p(x_1|z_1)\prod_{t=2}^{T}p(z_t|z_{t-1})p(x_t|z_t)\Big]\Big]$$
$$= \mathbb{E}_q[\log p(z_1)] + \mathbb{E}_q\Big[\sum_{t=2}^{T}\log p(z_t|z_{t-1})\Big] + \mathbb{E}_q\Big[\sum_{t=1}^{T}\log p(x_t|z_t)\Big]$$
$$= \mathbb{E}_q\Big[-\frac{M}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(z_1-\mu_0)^T\Sigma^{-1}(z_1-\mu_0)\Big]$$
$$+ \mathbb{E}\Big[\sum_{t=2}^{T}\Big(-\frac{M}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2}(z_t - Az_{t-1} - Cs_t)^T\Sigma^{-1}(z_t - Az_{t-1} - Cs_t)\Big)\Big]$$
$$+ \mathbb{E}\Big[-\frac{MT}{2}\log(2\pi) - \frac{T}{2}\log|\Gamma| - \frac{1}{2}\sum_{t=1}^{T}(x_t - Bz_t)^T\Gamma^{-1}(x_t - Bz_t)\Big]$$

a) C:

$$\rightsquigarrow \mathbb{E}[\log p(x,z)] = -\frac{1}{2}\mathbb{E}\Big[\sum_{t=2}^{T}(z_t - Az_{t-1} - Cs_t)^T\Sigma^{-1}(z_t - Az_{t-1} - Cs_t)\Big] + \text{const}$$
$$= -\frac{1}{2}\mathbb{E}\Big[\sum_{t=2}^{T}\Big(-z_t^T\Sigma^{-1}Cs_t + z_{t-1}^TA^T\Sigma^{-1}Cs_t + s_t^TC^T\Sigma^{-1}Cs_t - s_t^TC^T\Sigma^{-1}z_t + s_t^TC^T\Sigma^{-1}Az_{t-1}\Big)\Big] + \text{const}$$

$x^TAy = tr(Ayx^T)$

$$= -\frac{1}{2}\Big[\sum_{t=2}^{T}\Big(-\mathbb{E}[z_t^T]\Sigma^{-1}Cs_t + \mathbb{E}[z_{t-1}^T]A^T\Sigma^{-1}Cs_t + s_t^TC^T\Sigma^{-1}Cs_t - s_t^TC^T\Sigma^{-1}\mathbb{E}[z_t] + s_t^TC^T\Sigma^{-1}A\,\mathbb{E}[z_{t-1}]\Big)\Big] + \text{const}$$

- $\partial_X(a^TXb) \overset{(70)}{=} ab^T$
- $\partial_X(a^TX^Tb) \overset{(71)}{=} ba^T$
- $\partial_X(b^TX^TDXc) \overset{(82)}{=} D^TXbc^T + DXcb^T$

$\Rightarrow$
- $\partial_C(\mathbb{E}[z_t^T]\Sigma^{-1}Cs_t) = (\mathbb{E}[z_t^T]\Sigma^{-1})^T\cdot s_t^T = \Sigma^{-T}\mathbb{E}[z_t]s_t^T = \Sigma^{-1}\mathbb{E}[z_t]s_t^T$
- $\partial_C(\mathbb{E}[z_{t-1}^T]A^T\Sigma^{-1}Cs_t) = (\mathbb{E}[z_{t-1}^T]A^T\Sigma^{-1})^T\cdot s_t^T = \Sigma^{-1}A\,\mathbb{E}[z_{t-1}]\cdot s_t^T$
- $\partial_C(s_t^TC^T\Sigma^{-1}Cs_t) = \Sigma^{-1}C\,s_ts_t^T + \Sigma^{-1}C\,s_ts_t^T = 2\Sigma^{-1}Cs_ts_t^T$
- $\partial_C(s_t^TC^T\Sigma^{-1}\mathbb{E}[z_t]) = \Sigma^{-1}\mathbb{E}[z_t]s_t^T$
- $\partial_C(s_t^TC^T\Sigma^{-1}A\,\mathbb{E}[z_{t-1}]) = \Sigma^{-1}A\,\mathbb{E}[z_{t-1}]\cdot s_t^T$

$$\Rightarrow 0 \overset{!}{=} \partial_C\,\mathbb{E}[\log p(x,z|\Theta)]$$
$$= -\frac{1}{2}\sum_{t=2}^{T}\Big[-\Sigma^{-1}\mathbb{E}[z_t]s_t^T + \Sigma^{-1}A\,\mathbb{E}[z_{t-1}]s_t^T + 2\Sigma^{-1}Cs_ts_t^T - \Sigma^{-1}\mathbb{E}[z_t]s_t^T + \Sigma^{-1}A\,\mathbb{E}[z_{t-1}]s_t^T\Big]$$

$$\overset{\Sigma\cdot 2}{\Rightarrow} 0 = \sum_{t=2}^{T} -\mathbb{E}[z_t]s_t^T + A\,\mathbb{E}[z_{t-1}]s_t^T + Cs_ts_t^T$$
$$\Rightarrow C\sum_{t=2}^{T}s_ts_t^T = \sum_{t=2}^{T}(\mathbb{E}[z_t] - A\,\mathbb{E}[z_{t-1}])s_t^T$$
$$\Rightarrow C = \Big(\sum_{t=2}^{T}(\mathbb{E}[z_t] - A\,\mathbb{E}[z_{t-1}])s_t^T\Big)\cdot\Big(\sum_{t=2}^{T}s_ts_t^T\Big)^{-1}$$

b) $\mathcal{B}$:

$$\rightsquigarrow \mathbb{E}[\log p(x,z|\Theta)] = -\frac{1}{2}\mathbb{E}\Big[\sum_{t=2}^{T}(x_t-\mathcal{B}z_t)^T\Gamma^{-1}(x_t-\mathcal{B}z_t)\Big] + \text{const}$$

$$= -\frac{1}{2}\mathbb{E}\Big[\sum_{t=2}^{T} -x_t^T\Gamma^{-1}\mathcal{B}z_t - z_t^T\mathcal{B}^T\Gamma^{-1}x_t + z_t^T\mathcal{B}^T\Gamma^{-1}\mathcal{B}z_t\Big] + \text{const}$$

$$= -\frac{1}{2}\sum_{t=2}^{T}\Big[-x_t^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t] - \mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1}x_t + tr\big(\mathcal{B}^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]\big)\Big] + \text{const}$$

- $\partial_X(a^TXb) \overset{(70)}{=} ab^T$ • $\partial_X(a^TX^Tb) \overset{(71)}{=} ba^T$
- $\partial_X tr(X^TBXC) \overset{(119)}{=} BXC + B^TXC^T$ $\quad \Gamma^T = \Gamma$

$\Rightarrow$
- $\partial_\mathcal{B}(x_t^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t]) = (x_t^T\Gamma^{-1})^T\,\mathbb{E}[z_t^T] = \Gamma^{-1}x_t\,\mathbb{E}[z_t^T]$
- $\partial_\mathcal{B}(\mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1}x_t) = \Gamma^{-1}x_t\,\mathbb{E}[z_t^T]$
- $\partial_\mathcal{B}\big(tr(\mathcal{B}^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T])\big) = \Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T] + \underbrace{\Gamma^{-T}}_{=\Gamma^{-1}}\mathcal{B}\,\underbrace{\mathbb{E}[z_tz_t^T]^T}_{=\mathbb{E}[z_tz_t^T]} = 2\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]$

$$\Rightarrow 0 \overset{!}{=} \partial_\mathcal{B}\,\mathbb{E}[\log p(x,z|\Theta)]$$

$$= -\frac{1}{2}\sum_{t=2}^{T}\Big[-\Gamma^{-1}x_t\,\mathbb{E}[z_t^T] - \Gamma^{-1}x_t\,\mathbb{E}[z_t^T] + 2\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]\Big]$$

$\Gamma\cdot$:

$$\Rightarrow \mathcal{B}\sum_{t=2}^{T}\mathbb{E}[z_tz_t^T] = \sum_{t=2}^{T}x_t\,\mathbb{E}[z_t^T]$$

$$\Rightarrow \mathcal{B} = \Big(\sum_{t=2}^{T}x_t\,\mathbb{E}[z_t^T]\Big)\cdot\Big(\sum_{t=2}^{T}\mathbb{E}[z_tz_t^T]\Big)^{-1}$$

c) Γ:

$$\mathbb{E}[\log p(x,z|\Theta)] = -\frac{1}{2}\mathbb{E}\Big[T\log|\Gamma| + \sum_{t=2}^{T}(x_t-\mathcal{B}z_t)^T\Gamma^{-1}(x_t-\mathcal{B}z_t)\Big]$$

$$= -\frac{T}{2}\log|\Gamma| - \frac{1}{2}\mathbb{E}\Big[\sum_{t=2}^{T}x_t^T\Gamma^{-1}x_t - x_t^T\Gamma^{-1}\mathcal{B}z_t - z_t^T\mathcal{B}^T\Gamma^{-1}x_t + z_t^T\mathcal{B}^T\Gamma^{-1}\mathcal{B}z_t\Big]$$

$$= -\frac{T}{2}\log|\Gamma| - \frac{1}{2}\sum_{t=2}^{T}\Big[x_t^T\Gamma^{-1}x_t - x_t^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t] - \mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1}x_t + tr\big(\mathcal{B}^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]\big)\Big]$$

- $\partial_X\log|X| \overset{(57)}{=} X^{-T}$ • $\partial_X(a^TX^{-1}b) \overset{(61)}{=} -X^{-T}ab^TX^{-T}$
- $\partial_X\big(tr(AX^{-1}B)\big) \overset{(63)}{=} -(X^{-1}BAX^{-1})^T = -X^{-T}A^TB^TX^{-T}$

$\Rightarrow$
- $\partial_\Gamma\log|\Gamma| = \Gamma^{-T} = \Gamma^{-1}$
- $\partial_\Gamma(x_t^T\Gamma^{-1}x_t) = -\Gamma^{-T}x_tx_t^T\Gamma^{-T} = -\Gamma^{-1}x_tx_t^T\Gamma^{-1}$
- $\partial_\Gamma(x_t^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t]) = -\Gamma^{-T}x_t(\mathcal{B}\,\mathbb{E}[z_t])^T\Gamma^{-T} = -\Gamma^{-1}x_t\,\mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1}$
- $\partial_\Gamma(\mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1}x_t) = -\Gamma^{-T}(\mathbb{E}[z_t^T]\mathcal{B}^T)^Tx_t^T\Gamma^{-T} = -\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t]x_t^T\Gamma^{-1}$
- $\partial_\Gamma\big(tr(\mathcal{B}^T\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T])\big) = -\Gamma^{-T}\mathcal{B}(\mathcal{B}\,\mathbb{E}[z_tz_t^T])^T\Gamma^{-T} = -\Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]\mathcal{B}^T\Gamma^{-1}$

$$\Rightarrow 0 \overset{!}{=} \partial_\Gamma\,\mathbb{E}[\log p(x,z|\Theta)]$$

$$= -\frac{T}{2}\Gamma^{-1} - \frac{1}{2}\sum_{t=2}^{T}\Big[-\Gamma^{-1}x_tx_t^T\Gamma^{-1} + \Gamma^{-1}x_t\,\mathbb{E}[z_t^T]\mathcal{B}^T\Gamma^{-1} + \Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_t]x_t^T\Gamma^{-1} - \Gamma^{-1}\mathcal{B}\,\mathbb{E}[z_tz_t^T]\mathcal{B}^T\Gamma^{-1}\Big]$$

$-2\Gamma\cdot, \cdot\Gamma$

$$\Rightarrow 0 = T\Gamma - \sum_{t=2}^{T}\Big[x_tx_t^T - x_t\,\mathbb{E}[z_t^T]\mathcal{B}^T - \mathcal{B}\,\mathbb{E}[z_t]x_t^T + \mathcal{B}\,\mathbb{E}[z_tz_t^T]\mathcal{B}^T\Big]$$

$$\Rightarrow \Gamma = \frac{1}{T}\sum_{t=2}^{T}\Big[x_tx_t^T - \big(x_t\,\mathbb{E}[z_t^T]\mathcal{B}^T + (x_t\,\mathbb{E}[z_t^T]\mathcal{B}^T)^T\big) + \mathcal{B}\,\mathbb{E}[z_tz_t^T]\mathcal{B}^T\Big]$$