Distributed Kalman Filtering under Model Uncertainty

Mattia Zorzi

Proposition 4.5: Assume that for some t the distribution of $x_{k,t}$ given Y_{t-1} at node k is fixed and it is the same for RKF diff, KF diff, that is $f_{k,t}(x_t|Y_{t-1})$ and $\tilde{f}_{k,t}(x_t|Y_{t-1})$ coincide. Then, for c sufficiently large we have that

$$\mathbb{D}_{KL}(\tilde{p}_t, \tilde{p}_t^{loc}) \ll \mathbb{D}_{KL}(\tilde{p}_t, p_t^{loc}). \tag{1}$$

Proof: Let $\tilde{f}_{k,t} \sim \mathcal{N}(\hat{x}_{k,t}, V_{k,t})$ with $V_{k,t} > 0$ which is fixed and thus it does not depend on c. First, notice that $p_t(z_t|Y_{t-1}) = \bar{p}_t^{loc}(z_t|Y_{t-1})$ because the distribution of $x_{k,t}$ given Y_{t-1} is the same for RKF diff and KF diff. Accordingly,

$$\begin{aligned} p_t^{loc}(z_t|Y_{t-1}) &\sim \mathcal{N}(\mu_t^{loc}, K_t^{loc}) \\ \tilde{p}_t^{loc}(z_t|Y_{t-1}) &\sim \mathcal{N}(\mu_t^{loc}, \tilde{K}_t^{loc}) \\ \tilde{p}_t(z_t|Y_{t-1}) &\sim \mathcal{N}(\mu_t, \tilde{K}_t) \end{aligned} \tag{2}$$

where

$$\mu_t^{loc} = \begin{bmatrix} A \\ C_k^{loc} \\ 0 \end{bmatrix} \hat{x}_{k,t}, \quad \mu_t = \begin{bmatrix} A \\ C_k^{loc} \\ \breve{C}_k^{loc} \end{bmatrix} \hat{x}_{k,t}, \tag{3}$$

$$\begin{aligned} K_t^{loc} &= \begin{bmatrix} A \\ C_k^{loc} \\ 0 \end{bmatrix} V_{k,t} \begin{bmatrix} A^T & (C_k^{loc})^T & 0 \end{bmatrix} + \begin{bmatrix} BB^T & 0 & 0 \\ 0 & R_k^{loc} & 0 \\ 0 & 0 & Q_{k,t}^{loc} \end{bmatrix}, \\ \tilde{K}_t^{loc} &= K_t^{loc} + \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} (V_{k,t+1} - P_{k,t+1}) \begin{bmatrix} I & 0 & 0 \end{bmatrix}, \\ K_t &= \begin{bmatrix} A \\ C_k^{loc} \\ \breve{C}_k^{loc} \end{bmatrix} V_{k,t} \begin{bmatrix} A^T & (C_k^{loc})^T & (\breve{C}_k^{loc})^T \end{bmatrix} + \begin{bmatrix} BB^T & 0 & 0 \\ 0 & R_k^{loc} & 0 \\ 0 & 0 & \breve{R}_k^{loc} \end{bmatrix}, \\ \tilde{K}_t &= K_t + \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} (V_{t+1} - P_{t+1}) \begin{bmatrix} I & 0 & 0 \end{bmatrix}; \end{aligned}$$

M. Zorzi is with the Department of Information Engineering, University of Padova, Padova, Italy; e-mail: `zorzimat@dei.unipd.it`
.

$\breve{C}_k^{loc}$ and $\breve{R}_k^{loc}$ are the matrices obtained by using C_l and R_l, respectively, with $l \notin \mathcal{N}_k$. It is worth noting that the relation between $\tilde{K}_t^{loc}$ and K_t^{loc} is given by [2, Theorem 1]. The same observation holds between $\tilde{K}_t$ and K_t where the latter represents the covariance matrix of z_t given Y_{t-1} in the nominal model. Moreover,

$$P_{k,t+1} = AV_{k,t}A^T - AV_{k,t}(C_k^{loc})^T \left(C_k^{loc}V_{k,t}(C_k^{loc})^T + R_k^{loc}\right)^{-1} C_k^{loc}V_{k,t}A^T + BB^T$$

$$V_{k,t+1} = (P_{k,t+1}^{-1} - \theta_{k,t}I)^{-1}$$

$$P_{t+1} = AV_{k,t}A^T - AV_{k,t}\begin{bmatrix}(C_k^{loc})^T & (\breve{C}_k^{loc})^T\end{bmatrix} \left(\begin{bmatrix}C_k^{loc}\\ \breve{C}_k^{loc}\end{bmatrix} V_{k,t}\begin{bmatrix}(C_k^{loc})^T & (\breve{C}_k^{loc})^T\end{bmatrix} + \begin{bmatrix}R_k^{loc} & 0\\ 0 & \breve{R}_k^{loc}\end{bmatrix}\right)^{-1} \begin{bmatrix}C_k^{loc}\\ \breve{C}_k^{loc}\end{bmatrix} V_{k,t}A^T + BB^T$$

$$V_{t+1} = (P_{t+1}^{-1} - \theta_t I)^{-1}$$

and $\theta_{k,t}$, θ_t are the solution to $\gamma(P_{k,t+1}, \theta_{k,t}) = c$, $\gamma(P_{t+1}, \theta_t) = c$, respectively. Recall that

$$\gamma(P, \theta) := \log\det(I - \theta P) + \operatorname{tr}((I - \theta P)^{-1} - I). \tag{4}$$

In view of (2), it is not difficult to see that

$$\mathbb{D}_{KL}(\tilde{p}_t, \tilde{p}_t^{loc}) = \mathbb{D}_{KL}(\tilde{p}_t, p_t^{loc}) + \frac{1}{2}d_\Delta \tag{5}$$

where

$$\begin{aligned} d_\Delta &= \delta^T((\tilde{K}_t^{loc})^{-1} - (K_t^{loc})^{-1})\delta + \log\det(\tilde{K}_t^{loc}) \\ &\quad - \operatorname{tr}(\tilde{K}_t(K_t^{loc})^{-1}) + \operatorname{tr}(\tilde{K}_t(\tilde{K}_t^{loc})^{-1}) - \log\det(K_t^{loc}) \\ &\leq \log\det(\tilde{K}_t^{loc}) + \operatorname{tr}\left[\tilde{K}_t\left((\tilde{K}_t^{loc})^{-1} - (K_t^{loc})^{-1}\right)\right] - \log\det(K_t^{loc}) \end{aligned} \tag{6}$$

where $\delta = \mu_t - \mu_t^{loc}$ and we have exploited the fact that $(\tilde{K}_t^{loc})^{-1} - (K_t^{loc})^{-1} \leq 0$ because $P_{k,t+1} < V_{k,t+1}$ and thus $\tilde{K}_t^{loc} \geq K_t^{loc}$. Moreover, after some algebraic manipulations we obtain

$$d_\Delta \leq n\log\|V_{k,t+1}\| - \beta_{k,t}\|V_{t+1}\| + \nu_{k,t} \tag{7}$$

where

$$\beta_{k,t} = \lambda_{min}(P_{k,t+1}^{-1}[P_{k,t+1}^{-1} + (V_{k,t+1} - P_{k,t+1})^{-1}]^{-1}P_{k,t+1}^{-1})^{-1}\operatorname{tr}(\bar{V}_{t+1} - \|V_{t+1}\|^{-1}P_{t+1})$$

$$\nu_{k,t} = -\log\det K_t^{loc} + (Np+n)\log\lambda_{max}(K_t^{loc}) + \log\det(\|V_{k,t+1}\|^{-1}I_n + \lambda_{max}(K_t^{loc})^{-1}\bar{V}_{k,t+1})$$

$\lambda_{max}(K_t^{loc})$ denotes the maximum eigenvalue of K_t^{loc}, $\bar{V}_{k,t+1} := \|V_{k,t+1}\|^{-1}V_{k,t+1}$ and $\bar{V}_{t+1} := \|V_{t+1}\|^{-1}V_{t+1}$.

It [1] it has been shown that the mapping $c \mapsto \|V_{k+1,t}\|$ has singular value which is positive. Accordingly, if we take a sequence $c^{(m)}$, $m \in \mathbb{N}$, such that $c^{(m)} > 0$ and $c^{(m)} \to \infty$ as $m \to \infty$, then $\|V_{k,t+1}^{(m)}\| \to \infty$. The same reasoning holds for the mapping $c \mapsto \|V_{t+1}\|$ and thus $\|V_{t+1}^{(m)}\| \to \infty$. Consider the sequences $\bar{V}_{k,t+1}^{(m)} := \|V_{k,t+1}^{(m)}\|^{-1}V_{k,t+1}^{(m)}$ and $\bar{V}_{t+1}^{(m)} := \|V_{t+1}^{(m)}\|^{-1}V_{t+1}^{(m)}$ which belong to the compact set $\mathcal{U} := \{\, V \text{ s.t. } \|V\| = 1\,\}$. Therefore, there exist the subsequences $\bar{V}_{k,t+1}^{(m_l)}$, $l \in \mathbb{N}$ and $\bar{V}_{t+1}^{(m_l)}$, $l \in \mathbb{N}$, converging to $\bar{V}_{k,t+1}^{(\infty)}$ and $\bar{V}_{t+1}^{(\infty)}$, respectively. It is worth noting that $\bar{V}_{k,t+1}^{(\infty)}, \bar{V}_{t+1}^{(\infty)} \geq 0$ and different from the null matrix because $\bar{V}_{k,t+1}^{(\infty)}, \bar{V}_{t+1}^{(\infty)} \in \mathcal{U}$. Accordingly, if we consider the corresponding subsequences for $\beta_{k,t}$ and $\nu_{k,t}$, we have: $\beta_{k,t}^{(m_l)} \to \lambda_{min}(P_{k,t+1}^{-1})^{-1}\operatorname{tr}(\bar{V}_{t+1}) > 0$ and $\nu_{k,t}^{(m_l)}$ is bounded above.

Next we show that $\|V_{t+1}^{(m_l)}\|/\|V_{k,t+1}^{(m_l)}\| \to \zeta > 0$. First, we recall that $V_{k,t+1}^{(m_l)}$ and $V_{t+1}^{(m_l)}$ are given by $\theta_{k,t}^{(m_l)}$ and $\theta_t^{(m_l)}$, respectively. In particular, we have $\gamma(P_{t+1}^{(m_l)}, \theta_t^{(m_l)}) = c^{(m_l)}$. Notice that we can rewrite the latter as

$$\sum_{i=1}^{n} \log(1 - d_i\theta_t^{(m_l)}) + (1 - \theta d_i^{(m_l)})^{-1} - 1 = c^{(m_l)} \tag{8}$$

where $d_i \geq d_{i+1}$ denotes the eigenvalues of P_{t+1} and $0 < \theta_t^{(m_l)} < d_1^{-1}$. In what follows we assume that the eigenvalue d_1 has multiplicity equal to one, and thus $d_1 > d_i$ with $i \geq 2$. This assumption is not restrictive, indeed it generically holds. Then we can rewrite (8) as

$$f(d_1\theta_t^{(m_l)}) + \breve{c}^{(m_l)} = c^{(m_l)}$$

where

$$\begin{aligned} f(x) &= \log(1 - x) + (1 - x)^{-1} - 1 \\ \breve{c}^{(m_l)} &= \sum_{i=2}^{n} \log(1 - d_i\theta_t^{(m_l)}) + (1 - \theta d_i^{(m_l)})^{-1} - 1, \end{aligned}$$

$\breve{c}^{(m_l)} \to \breve{c}$ and $\breve{c}$ is a bounded value. Therefore

$$f(d_1\theta_t^{(m_l)}) = c^{(m_l)} - \breve{c}^{(m_l)}.$$

Since $c^{(m_l)} \to \infty$, we have $\breve{c}^{(m_l)} = o(c^{(m_l)})$, i.e. $\breve{c}^{(m_l)}/c^{(m_l)} \to 0$ as l tends to infinity. Accordingly,

$$f(d_1\theta_t^{(m_l)}) = c^{(m_l)} - o(c^{(m_l)}). \tag{9}$$

The same reasoning applies for $\theta_{k,t}^{(m_l)}$:

$$f(d_{k,1}\theta_{k,t}^{(m_l)}) = c^{(m_l)} - o(c^{(m_l)}) \tag{10}$$

where $d_{k,i} \geq d_{k,i+1}$ are the eigenvalues of $P_{k,t+1}$ and $d_{k,1}$ has multiplicity equal to one. Notice that $d_1\theta_t^{(m_l)}$ and $d_{k,1}\theta_{k,t}^{(m_l)}$ belong to the interval $[0, 1)$. It is not difficult to see that $f : [0, 1) \to [0, \infty)$ is monotone increasing in the interval $[0, 1)$. Accordingly, it admits the continuous inverse function $g : [0, \infty) \to [0, 1)$ and

$$\begin{aligned} \theta_t^{(m_l)} &= d_1^{-1} g\left(c^{(m_l)} - o(c^{(m_l)})\right) \\ \theta_{k,t}^{(m_l)} &= d_{k,1}^{-1} g\left(c^{(m_l)} - o(c^{(m_l)})\right). \end{aligned}$$

Notice that

$$\begin{aligned} \lim_{l\to\infty} g\left(c^{(m_l)} - o(c^{(m_l)})\right) &= g\left(\lim_{l\to\infty} c^{(m_l)} - o(c^{(m_l)})\right) \\ &= g\left(\lim_{l\to\infty} c^{(m_l)} \lim_{l\to\infty}\left(1 - \frac{o(c^{(m_l)})}{c^{(m_l)}}\right)\right) = g\left(\lim_{l\to\infty} c^{(m_l)}\right) = \lim_{l\to\infty} g\left(c^{(m_l)}\right) \end{aligned} \tag{11}$$

Finally, we have

$$
\begin{aligned}
\lim_{l\to\infty}\frac{\|V_{t+1}^{(m_l)}\|}{\|V_{k,t+1}^{(m_l)}\|} &= \lim_{l\to\infty}\sqrt{\frac{\sum_{i=1}^{n}\frac{1}{d_i^{-1}-\theta_t^{(m_l)}}}{\sum_{i=1}^{n}\frac{1}{d_{k,i}^{-1}-\theta_{k,t}^{(m_l)}}}} = \lim_{l\to\infty}\sqrt{\frac{\frac{1}{d_1^{-1}-\theta_t^{(m_l)}}+\sum_{i=2}^{n}\frac{1}{d_i^{-1}-\theta_t^{(m_l)}}}{\frac{1}{d_{k,1}^{-1}-\theta_{k,t}^{(m_l)}}+\sum_{i=2}^{n}\frac{1}{d_{k,i}^{-1}-\theta_{k,t}^{(m_l)}}}}\\
&= \lim_{l\to\infty}\sqrt{\frac{\frac{d_1}{1-g\left(c^{(m_l)}-o(c^{(m_l)})\right)}+\sum_{i=2}^{n}\frac{1}{d_i^{-1}-\theta_t^{(m_l)}}}{\frac{d_{k,1}}{1-g\left(c^{(m_l)}-o(c^{(m_l)})\right)}+\sum_{i=2}^{n}\frac{1}{d_{k,i}^{-1}-\theta_{k,t}^{(m_l)}}}} = \lim_{l\to\infty}\sqrt{\frac{\frac{d_1}{1-g\left(c^{(m_l)}\right)}+\sum_{i=2}^{n}\frac{1}{d_i^{-1}-\theta_t^{(m_l)}}}{\frac{d_{k,1}}{1-g\left(c^{(m_l)}\right)}+\sum_{i=2}^{n}\frac{1}{d_{k,i}^{-1}-\theta_{k,t}^{(m_l)}}}}\\
&= \lim_{l\to\infty}\sqrt{\frac{\frac{d_1}{1-g\left(c^{(m_l)}\right)}}{\frac{d_{k,1}}{1-g\left(c^{(m_l)}\right)}}}
\end{aligned}
\tag{12}
$$

where we exploited the fact that $\lim_{x\to\infty} g(x)=1$ in the last equality. Then, we have

$$
\lim_{l\to\infty}\frac{\|V_{t+1}^{(m_l)}\|}{\|V_{k,t+1}^{(m_l)}\|} = \lim_{l\to\infty}\sqrt{\frac{\frac{d_1}{1-g\left(c^{(m_l)}\right)}}{\frac{d_{k,1}}{1-g\left(c^{(m_l)}\right)}}} = \sqrt{\frac{d_1}{d_{k,1}}} > 0. \tag{13}
$$

Accordingly the corresponding subsequence $d_\Delta^{(m_l)}$ approaches $-\infty$ because the term $-\beta_{k,t}^{(m_l)}\|V_{t+1}^{(m_l)}\|$ dominates the logarithmic term $n\log\|V_{k,t+1}^{(m_l)}\|$. We conclude that for c sufficiently large (1) holds. □

References

[1] B. Levy and R. Nikoukhah, "Robust least-squares estimation with a relative entropy constraint," *Information Theory, IEEE Transactions on*, vol. 50, no. 1, pp. 89–104, Jan. 2004.

[2] ——, "Robust state-space filtering under incremental model perturbations subject to a relative entropy tolerance," *IEEE Trans. Automat. Control*, vol. 58, pp. 682–695, Mar. 2013.