Derivation of the IMM filter

Umut Orguner

April 16, 2013

We consider the following jump Markov linear system (JMLS).

$$x_k = A(r_k)x_{k-1} + B(r_k)w_k \tag{1}$$
$$y_k = C(r_k)x_k + D(r_k)v_k \tag{2}$$

where

- $x_k \in \mathbb{R}^{n_x}$ is the base state;
- $r_k \in \{1, 2, \ldots, N_r\}$ is the mode state;
- $y_k \in \mathbb{R}^{n_y}$ is the measurement;
- $w_k \in \mathbb{R}^{n_w}$ is the process noise distributed with $w_k \sim \mathcal{N}(w_k; 0, Q)$;
- $v_k \in \mathbb{R}^{n_v}$ is the measurement noise distributed with $v_k \sim \mathcal{N}(v_k; 0, R)$;
- The properly sized matrices $A(\cdot)$, $B(\cdot)$, $C(\cdot)$ and $D(\cdot)$ are known functions of the mode state.

The time behavior of the mode state $r_k \in \{1, 2, \ldots, N_r\}$ is modeled as a homogeneous (time invariant) Markov chain with transition probability matrix (TPM) $\Pi = [\pi_{ij} \triangleq P(r_k = j|r_{k-1} = i)]$. The problem is to estimate or approximate the posterior distribution $p(x_k|y_{0:k})$ in a computationally tractable way.

We have seen in the class that the optimal $p(x_k|y_{0:k})$ is a mixture of Gaussians with an exponentially growing number of components. Hence approximations are necessary. The so-called interacting multiple model (IMM) filter [1] makes the approximation

$$p(x_k|y_{0:k}) \approx \sum_{i=1}^{N_r} \mu_k^i \mathcal{N}(x_k; \hat{x}_{k|k}^i, \Sigma_{k|k}^i) \tag{3}$$

where

$$\mu_k^i \triangleq P(r_k = i|y_{0:k}) \tag{4}$$

are the posterior mode probabilities. When such an approximation is given, one can calculate the overall posterior mean $\hat{x}_{k|k}$ and covariance $\Sigma_{k|k}$ using the standard Gaussian mixture mean and covariance formulas as

$$\hat{x}_{k|k} = \sum_{i=1}^{N_r} \mu_k^i \hat{x}_{k|k}^i \tag{5}$$
$$\Sigma_{k|k} = \sum_{i=1}^{N_r} \mu_k^i \left[\Sigma_{k|k}^i + (\hat{x}_{k|k}^i - \hat{x}_{k|k})(\hat{x}_{k|k}^i - \hat{x}_{k|k})^{\mathrm{T}}\right] \tag{6}$$

This estimate and covariance can be given to the user as the output. The mode conditional means $\{\hat{x}_{k|k}^i\}_{i=1}^{N_r}$, covariances $\{\Sigma_{k|k}^i\}_{i=1}^{N_r}$ and mode probabilities $\{\mu_k^i\}_{i=1}^{N_r}$ must be calculated recursively from their previous values $\{\hat{x}^i_{k-1|k-1}, \Sigma^i_{k-1|k-1}, \mu^i_{k-1}\}_{i=1}^{N_r}$. For deriving such a recursion, we write the posterior density as follows.

$$
\begin{aligned}
p(x_k|y_{0:k}) \triangleq& \sum_{i=1}^{N_r} p(x_k|y_{0:k}, r_k = i) \underbrace{P(r_k = i|y_{0:k})}_{\triangleq \mu^i_k} && (7)\\
=& \sum_{i=1}^{N_r} \mu^i_k p(x_k|y_{0:k}, r_k = i) && (8)\\
=& \sum_{i=1}^{N_r} \mu^i_k \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} p(x_k|y_{0:k-1}, r_k = i) && (9)\\
=& \sum_{i=1}^{N_r} \mu^i_k \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} \int p(x_k|x_{k-1}, r_k = i) p(x_{k-1}|y_{0:k-1}, r_k = i) \, \mathrm{d}x_{k-1} && (10)\\
=& \sum_{i=1}^{N_r} \mu^i_k \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} \int p(x_k|x_{k-1}, r_k = i)\\
&\times \sum_{j=1}^{N_r} p(x_{k-1}|y_{0:k-1}, r_{k-1} = j) \underbrace{P(r_{k-1} = j|r_k = i, y_{0:k-1})}_{\mu^{ji}_{k-1|k-1}} \, \mathrm{d}x_{k-1} && (11)\\
=& \sum_{i=1}^{N_r} \mu^i_k \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} \int p(x_k|x_{k-1}, r_k = i)\\
&\times \sum_{j=1}^{N_r} \mu^{ji}_{k-1|k-1} p(x_{k-1}|y_{0:k-1}, r_{k-1} = j) \, \mathrm{d}x_{k-1} && (12)
\end{aligned}
$$

We know from the previous step that

$$
p(x_{k-1}|y_{0:k-1}, r_{k-1} = j) = \mathcal{N}(x_{k-1}; \hat{x}^i_{k-1|k-1}, \Sigma^i_{k-1|k-1}) \quad (13)
$$

Substituting this into (12) would clearly yield a Gaussian mixture with N_r^2 components for $p(x_k|y_{0:k})$. Therefore we make the following approximation

$$
\begin{aligned}
p(x_{k-1}|y_{0:k-1}, r_k = i) \triangleq& \sum_{j=1}^{N_r} \mu^{ji}_{k-1|k-1} p(x_{k-1}|y_{0:k-1}, r_{k-1} = j) && (14)\\
=& \sum_{j=1}^{N_r} \mu^{ji}_{k-1|k-1} \mathcal{N}(x_{k-1}; \hat{x}^i_{k-1|k-1}, \Sigma^i_{k-1|k-1}) && (15)\\
\approx& \mathcal{N}(x_{k-1}; \hat{x}^{0i}_{k-1|k-1}, \Sigma^{0i}_{k-1|k-1}) && (16)
\end{aligned}
$$

where the merged mean $\hat{x}^{0i}_{k-1|k-1}$ and covariance $\Sigma^{0i}_{k-1|k-1}$ are obtained by moment matching as

$$
\begin{aligned}
\hat{x}^{0i}_{k-1|k-1} =& \sum_{j=1}^{N_r} \mu^{ji}_{k-1|k-1} \hat{x}^j_{k-1|k-1}, && (17)\\
\Sigma^{0i}_{k-1|k-1} =& \sum_{j=1}^{N_r} \mu^{ji}_{k-1|k-1} \left[\Sigma^j_{k-1|k-1} + (\hat{x}^j_{k-1|k-1} - \hat{x}^{0i}_{k-1|k-1})(\hat{x}^j_{k-1|k-1} - \hat{x}^{0i}_{k-1|k-1})^{\mathrm{T}}\right]. && (18)
\end{aligned}
$$

In order to separate the merging in (17) and (18) from the merging in output calculation (i.e., in (5) and (6)), the merging in (17) and (18) is called "mixing" in the literature [1]. The estimate

$\hat{x}^{0i}_{k-1|k-1}$ and covariance $\Sigma^{0i}_{k-1|k-1}$ are called the mixed estimate and covariance accordingly. Now substituting the approximation (16) into (12), we get

$$p(x_k|y_{0:k}) = \sum_{i=1}^{N_r} \mu_k^i \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} \int p(x_k|x_{k-1}, r_k = i)\mathcal{N}(x_{k-1}; \hat{x}^{0i}_{k-1|k-1}, \Sigma^{0i}_{k-1|k-1})dx_{k-1} \tag{19}$$

Noticing that

$$p(x_k|x_{k-1}, r_k = i) = \mathcal{N}(x_k; A(i)x_{k-1}, B(i)QB^{\mathrm{T}}(i)) \tag{20}$$

the integral in (19) represents a prediction update with ith model. Hence, we see that

$$p(x_k|y_{0:k}) = \sum_{i=1}^{N_r} \mu_k^i \frac{p(y_k|x_k, r_k = i)}{p(y_k|y_{0:k-1}, r_k = i)} \mathcal{N}(x_k; \hat{x}^i_{k|k-1}, \Sigma^i_{k|k-1}) \tag{21}$$

where

$$\hat{x}^i_{k|k-1} = A(i)\hat{x}^{0i}_{k-1|k-1} \tag{22}$$

$$\Sigma^i_{k|k-1} = A(i)\Sigma^{0i}_{k-1|k-1}A^{\mathrm{T}}(i) + B(i)QB^{\mathrm{T}}(i) \tag{23}$$

Again noticing that

$$p(y_k|x_k, r_k = i) = \mathcal{N}(y_k; C(i)x_k, D(i)RD^{\mathrm{T}}(i)), \tag{24}$$

the multiplication in (21) represents a measurement update with the ith model. Hence

$$p(x_k|y_{0:k}) = \sum_{i=1}^{N_r} \mu_k^i \mathcal{N}(x_k; \hat{x}^i_{k|k}, \Sigma^i_{k|k}) \tag{25}$$

where

$$\hat{x}^i_{k|k} = \hat{x}^i_{k|k-1} + K_k^i(y_k - \hat{y}^i_{k|k-1}) \tag{26}$$

$$\Sigma^i_{k|k} = \Sigma^i_{k|k-1} - K_k^i S_k^i K_k^{iT} \tag{27}$$

$$\hat{y}^i_{k|k-1} = C(i)\hat{x}^i_{k|k-1} \tag{28}$$

$$S_k^i = C(i)\Sigma^i_{k|k-1}C^{\mathrm{T}}(i) + D(i)RD^{\mathrm{T}}(i) \tag{29}$$

$$K_k^i = \Sigma^i_{k|k-1}C^{\mathrm{T}}(i)(S_k^i)^{-1} \tag{30}$$

We have obtained the recursions for the estimates and covariances. In order to complete the recursion, all we need to do is to give recursions for the probabilities $\{\mu_k^i\}_{i=1}^{N_r}$ in terms of $\{\mu_{k-1}^j\}_{j=1}^{N_r}$ and we have to find an expression for $\mu^{ji}_{k-1|k-1}$.

The following gives a recursion for $\{\mu_k^i\}_{i=1}^{N_r}$.

$$\mu_k^i \triangleq P(r_k = i|y_{0:k}) \tag{31}$$

$$\propto \underbrace{p(y_k|y_{0:k-1}, r_k = i)}_{=\mathcal{N}(y_k; \hat{y}^i_{k|k-1}, S_k^i)} P(r_k = i|y_{0:k-1}) \tag{32}$$

$$= \mathcal{N}(y_k; \hat{y}^i_{k|k-1}, S_k^i) \sum_{j=1}^{N_r} \underbrace{P(r_k = i|r_{k-1} = j)}_{\pi_{ji}} \underbrace{P(r_{k-1} = j|y_{0:k-1})}_{\mu_{k-1}^j} \tag{33}$$

$$= \mathcal{N}(y_k; \hat{y}^i_{k|k-1}, S_k^i) \sum_{j=1}^{N_r} \pi_{ji}\mu_{k-1}^j \tag{34}$$

Hence, we have

$$\mu_k^i = \frac{\mathcal{N}(y_k; \hat{y}_{k|k-1}^i, S_k^i) \sum_{j=1}^{N_r} \pi_{ji} \mu_{k-1}^j}{\sum_{\ell=1}^{N_r} \mathcal{N}(y_k; \hat{y}_{k|k-1}^\ell, S_k^\ell) \sum_{j=1}^{N_r} \pi_{j\ell} \mu_{k-1}^j}. \tag{35}$$

We can calculate the mixing probabilities $\mu_{k-1|k-1}^{ji}$ as

$$\begin{aligned}
\mu_{k-1|k-1}^{ji} &\triangleq P(r_{k-1} = j | r_k = i, y_{0:k-1}) && (36)\\
&\propto P(r_k = i | r_{k-1} = j, y_{0:k-1}) \underbrace{P(r_{k-1} = j | y_{0:k-1})}_{\mu_{k-1}^j} && (37)\\
&= \underbrace{P(r_k = i | r_{k-1} = j)}_{\pi_{ji}} \mu_{k-1}^j && (38)\\
&= \pi_{ji} \mu_{k-1}^j && (39)
\end{aligned}$$

Hence, we have

$$\mu_{k-1|k-1}^{ji} = \frac{\pi_{ji} \mu_{k-1}^j}{\sum_{\ell=1}^{N_r} \pi_{\ell i} \mu_{k-1}^\ell}. \tag{40}$$

Now we can give a verbal description of one step of the IMM filter as follows.

Algorithm 1 ***Single Step of IMM filter:*** *Suppose we have the previous sufficient statistics $\{x_{k-1|k-1}^j, \Sigma_{k-1|k-1}^j, \mu_{k-1}^j\}_{j=1}^{N_r}$. Then, a single step of the IMM algorithm to obtain current sufficient statistics $\{x_{k|k}^i, \Sigma_{k|k}^i, \mu_k^i\}_{i=1}^{N_r}$ is given as follows.*

- ***Mixing:***
  - *Calculate the mixing probabilities $\{\mu_{k-1|k-1}^{ji}\}_{i,j=1}^{N_r}$ as*

$$\mu_{k-1|k-1}^{ji} = \frac{\pi_{ji} \mu_{k-1}^j}{\sum_{\ell=1}^{N_r} \pi_{\ell i} \mu_{k-1}^\ell}. \tag{41}$$

  - *Calculate the mixed estimates $\{\hat{x}_{k-1|k-1}^{0i}\}_{i=1}^{N_r}$ and covariances $\{\Sigma_{k-1|k-1}^{0i}\}_{i=1}^{N_r}$ as*

$$\hat{x}_{k-1|k-1}^{0i} = \sum_{j=1}^{N_r} \mu_{k-1|k-1}^{ji} \hat{x}_{k-1|k-1}^j, \tag{42}$$

$$\Sigma_{k-1|k-1}^{0i} = \sum_{j=1}^{N_r} \mu_{k-1|k-1}^{ji} \left[\Sigma_{k-1|k-1}^j + (\hat{x}_{k-1|k-1}^j - \hat{x}_{k-1|k-1}^{0i})(\hat{x}_{k-1|k-1}^j - \hat{x}_{k-1|k-1}^{0i})^{\mathrm{T}}\right]. \tag{43}$$

- ***Mode Matched Prediction Update:*** *For ith model, $i = 1, \ldots, N_r$, calculate the predicted estimate $\hat{x}_{k|k-1}^i$ and covariance $\Sigma_{k|k-1}^i$ from the mixed estimate $\hat{x}_{k-1|k-1}^{0i}$ and covariance $\Sigma_{k-1|k-1}^{0i}$ as*

$$\begin{aligned}
\hat{x}_{k|k-1}^i &= A(i) \hat{x}_{k-1|k-1}^{0i}, && (44)\\
\Sigma_{k|k-1}^i &= A(i) \Sigma_{k-1|k-1}^{0i} A^{\mathrm{T}}(i) + B(i) Q B^{\mathrm{T}}(i). && (45)
\end{aligned}$$

- ***Mode Matched Measurement Update:*** *For ith model, $i = 1, \ldots, N_r$,*

– *Calculate the updated estimate* $\hat{x}^i_{k|k}$ *and covariance* $\Sigma^i_{k|k}$ *from the predicted estimate* $\hat{x}^i_{k|k-1}$ *and covariance* $\Sigma^i_{k|k-1}$ *as*

$$\hat{x}^i_{k|k} = \hat{x}^i_{k|k-1} + K^i_k(y_k - \hat{y}^i_{k|k-1}), \tag{46}$$

$$\Sigma^i_{k|k} = \Sigma^i_{k|k-1} - K^i_k S^i_k K^{iT}_k, \tag{47}$$

$$\hat{y}^i_{k|k-1} = C(i)\hat{x}^i_{k|k-1}, \tag{48}$$

$$S^i_k = C(i)\Sigma^i_{k|k-1}C^{\mathrm{T}}(i) + D(i)RD^{\mathrm{T}}(i), \tag{49}$$

$$K^i_k = \Sigma^i_{k|k-1}C^{\mathrm{T}}(i)(S^i_k)^{-1}. \tag{50}$$

– *Calculate the updated mode probability* μ^i_k *as*

$$\mu^i_k = \frac{\mathcal{N}(y_k; \hat{y}^i_{k|k-1}, S^i_k)\sum_{j=1}^{N_r}\pi_{ji}\mu^j_{k-1}}{\sum_{\ell=1}^{N_r}\mathcal{N}(y_k; \hat{y}^\ell_{k|k-1}, S^\ell_k)\sum_{j=1}^{N_r}\pi_{j\ell}\mu^j_{k-1}}. \tag{51}$$

- ***Output Estimate Calculation:*** *Calculate the overall estimate* $\hat{x}_{k|k}$ *and covariance as*

$$\hat{x}_{k|k} = \sum_{i=1}^{N_r}\mu^i_k\hat{x}^i_{k|k}, \tag{52}$$

$$\Sigma_{k|k} = \sum_{i=1}^{N_r}\mu^i_k\left[\Sigma^i_{k|k} + (\hat{x}^i_{k|k} - \hat{x}_{k|k})(\hat{x}^i_{k|k} - \hat{x}_{k|k})^{\mathrm{T}}\right]. \tag{53}$$

■

Remark 1 *The output calculation step is done only for output purposes and therefore does not affect the sufficient statistics.* ■

A block diagram of a single step of the IMM filter is given in Figure 1 below.

References

[1] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001.

[2] Y. Bar-Shalom, S. Challa, and H. A. P. Blom, "IMM estimator versus optimal estimator for hybrid systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 41, no. 3, pp. 986–991, Jul. 2005.

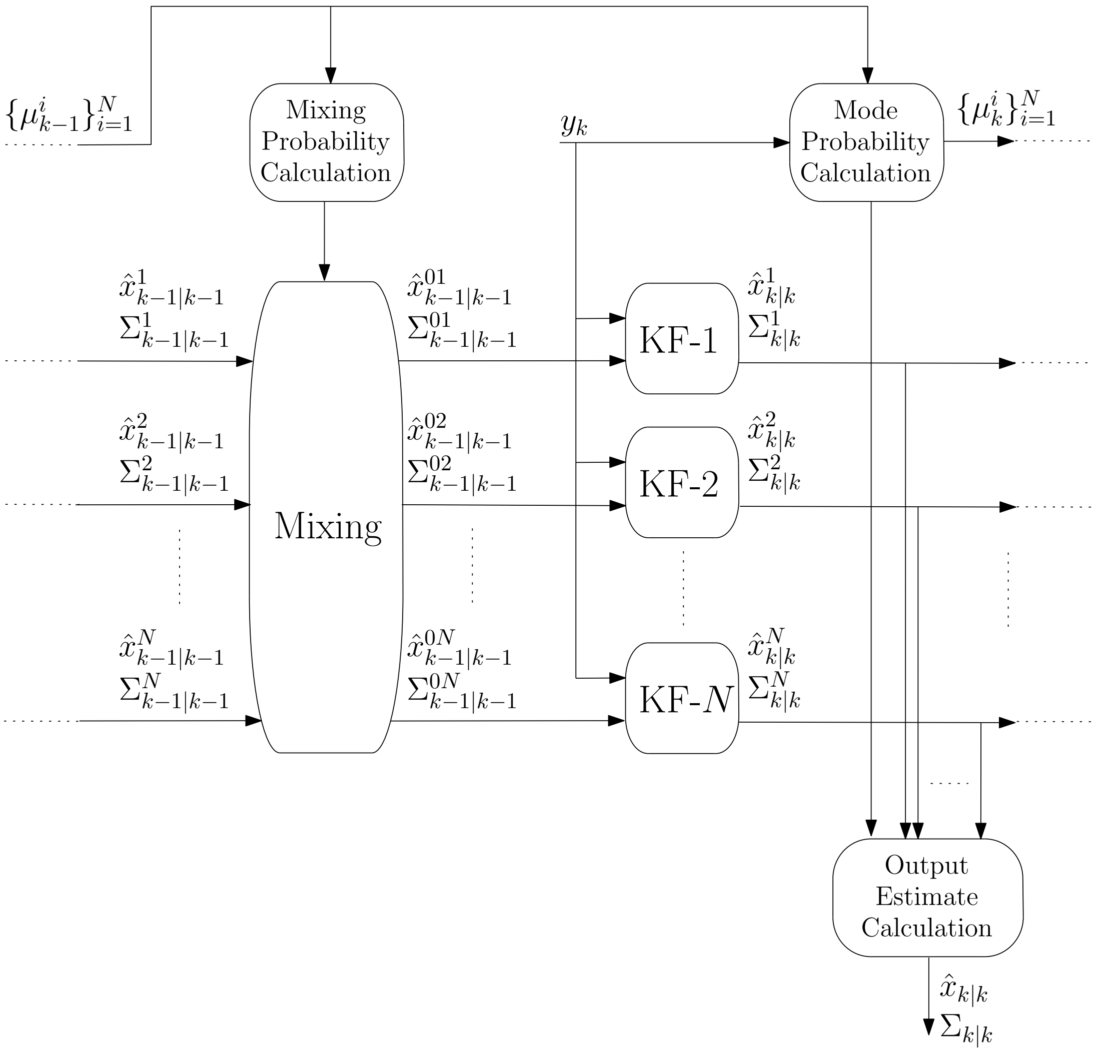

Figure 1: Block diagram of a single step of the IMM algortihm for N-models.