11.6 THE MULTIPLE MODEL APPROACH

11.6.1 Formulation of the Approach

In the ***multiple model (MM) approach*** it is assumed that the system obeys one of a finite number of models. Such systems are called ***hybrid***: They have both *continuous* (noise) uncertainties and *discrete* uncertainties — ***model*** or ***mode***, or ***operating regime*** uncertainties.

A *Bayesian framework* is used: Starting with prior probabilities of each model being correct (i.e., the system is in a particular mode), the corresponding posterior probabilities are obtained.

First the static case in which the model the system obeys is *fixed*, that is, no switching from one mode to another occurs during the estimation process (time-invariant mode) is considered. This will result in the ***static MM estimator***. While the model that is in effect stays fixed, each model has its own dynamics, so *the overall estimator is dynamic*.

The model, assumed to be in effect throughout the entire process, is one of r possible models (the system is in one of r modes)

$$M \in \{M_j\}_{j=1}^{r} \tag{11.6.1-1}$$

The prior probability that M_j is correct (the system is in mode j) is

$$P\{M_j|Z^0\} = \mu_j(0) \qquad j = 1, \ldots, r \tag{11.6.1-2}$$

where Z^0 is the prior information and

$$\sum_{j=1}^{r} \mu_j(0) = 1 \tag{11.6.1-3}$$

since the correct model is among the assumed r possible models.

It will be assumed that all models are linear-Gaussian. This approach can be used for nonlinear systems as well via linearization.

Subsequently, the dynamic situation of ***switching models*** or ***mode jumping*** is considered. In the latter case, the system undergoes transitions from one mode to another. The resulting estimator is a ***dynamic MM estimator***.

11.6.2 The Static Multiple Model Estimator

The ***static MM estimator*** — for fixed models — is obtained as follows.

Using Bayes' formula, the posterior probability of model j being correct, given the measurement data up to k, is given by the recursion

$$\begin{aligned}\mu_j(k) &\triangleq P\{M_j|Z^k\} = P\{M_j|z(k), Z^{k-1}\} = \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{p[z(k)|Z^{k-1}]} \\ &= \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{\sum_{i=1}^{r} p[z(k)|Z^{k-1}, M_i]P\{M_i|Z^{k-1}\}}\end{aligned} \tag{11.6.2-1}$$

or

$$\boxed{\mu_j(k) = \frac{p[z(k)|Z^{k-1}, M_j]\mu_j(k-1)}{\sum_{i=1}^{r} p[z(k)|Z^{k-1}, M_i]\mu_i(k-1)} \qquad j = 1, \ldots, r} \tag{11.6.2-2}$$

starting with the given prior probabilities (11.6.1-2).

The first term on the right-hand side above is the ***likelihood function of mode*** j at time k, which, under the linear-Gaussian assumptions, is given by the expression (5.2.6-7)

$$\Lambda_j(k) \triangleq p[z(k)|Z^{k-1}, M_j] = p[\nu_j(k)] = \mathcal{N}[\nu_j(k); 0, S_j(k)] \tag{11.6.2-3}$$

where ν_j and S_j are the innovation and its covariance from the ***mode-matched filter*** corresponding to mode j. In a nonlinear and/or non-Gaussian situation the same Gaussian likelihood functions are used, even though they are clearly approximations.

Thus a Kalman filter matched to each mode is set up yielding ***mode-conditioned state estimates*** and the associated ***mode-conditioned covariances***. The probability of each mode being correct — the ***mode estimates*** — is obtained according to (11.6.2-2) based on its likelihood function (11.6.2-3) relative to the other filters' likelihood functions. In a nonlinear situation the filters are EKF instead of KF.

This modular estimator, which is a ***bank of filters***, is illustrated in Fig. 11.6.2-1.

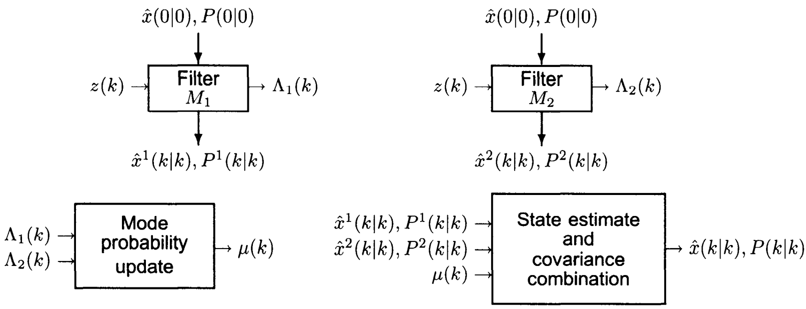

Figure 11.6.2-1: The static multiple model estimator for $r = 2$ fixed models.

The output of each mode-matched filter is the ***mode-conditioned state estimate*** $\hat{x}^j$, the associated covariance P^j and the ***mode likelihood function*** Λ_j.

After the filters are initialized, they run recursively *on their own estimates.* Their likelihood functions are used to update the mode probabilities. The latest mode probabilities are used to combine the mode-conditioned estimates and covariances.

Under the above assumptions the pdf of the state of the system is a Gaussian mixture with r terms

$$p[x(k)|Z^k] = \sum_{j=1}^{r} \mu_j(k)\mathcal{N}[x(k); \hat{x}^j(k|k), P^j(k|k)] \tag{11.6.2-4}$$

The combination of the mode-conditioned estimates is done therefore as follows

$$\boxed{\hat{x}(k|k) = \sum_{j=1}^{r} \mu_j(k)\hat{x}^j(k|k)} \tag{11.6.2-5}$$

and the covariance of the combined estimate is (see Subsection 1.4.16)

$$\boxed{P(k|k) = \sum_{j=1}^{r} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]'\}} \tag{11.6.2-6}$$

where the last term above is the ***spread of the means*** term.

The above is exact under the following assumptions:

1. The correct model is among the set of models considered,
2. The same model has been in effect from the initial time.

Assumption 1 can be considered a reasonable approximation; however, 2 is obviously not true if a maneuver has started at some time within the interval $[1, k]$, in which case a model change — ***mode jump*** — occurred.

Convergence of the Mode Estimates

If the mode set includes the correct one and no mode jump occurs, then the probability of the true mode will converge to unity, that is, this approach yields consistent estimates of the system parameters. Otherwise the probability of the model "nearest" to the correct one will converge to unity (this is discussed in detail in [Baram78]).

Ad Hoc Modifications for the Case of Switching Models

The following ad hoc modification can be made to the static MM estimator for estimating the state in the case of switching models: An artificial lower bound is imposed on the model probabilities (with a suitable renormalization of the remaining probabilities).

A shortcoming of the static MM estimator when used with switching models is that, in spite of the above ad hoc modification, the mismatched filters' errors can grow to unacceptable levels. Thus, reinitialization of the filters that are mismatched is, in general, needed. This is accomplished by using the estimate from filter corresponding to the best matched model in the other filters.

It should be pointed out that the above "fixes" are automatically (and rigorously) built into the dynamic MM estimation algorithms to be discussed next.

11.6.3 The Dynamic Multiple Model Estimator

In this case the mode the system is in can undergo switching in time. The system is modeled by the equations

$$x(k) = F[M(k)]x(k-1) + v[k-1, M(k)] \tag{11.6.3-1}$$

$$z(k) = H[M(k)]x(k) + w[k, M(k)] \tag{11.6.3-2}$$

where $M(k)$ denotes the mode or model "at time k" — in effect *during the sampling period ending at k*. Such systems are also called ***jump-linear systems***. The mode jump process is assumed ***left-continuous*** (i.e., the impact of the new model starts at t_k^+).

The mode at time k is assumed to be among the possible r modes

$$M(k) \in \{M_j\}_{j=1}^{r} \tag{11.6.3-3}$$

The continuous-valued vector $x(k)$ and the discrete variable $M(k)$ are sometimes referred to as the ***base state*** and the ***modal state***, respectively.

For example,

$$F[M_j] = F_j \tag{11.6.3-4}$$

$$v(k-1, M_j) \sim \mathcal{N}(u_j, Q_j) \tag{11.6.3-5}$$

that is, the structure of the system and/or the statistics of the noises might be different from model to model. The mean u_j of the noise can model a maneuver as a deterministic input.

The lth ***mode history*** — or ***sequence of models*** — through time k is denoted as

$$M^{k,l} = \{M_{i_{1,l}}, \ldots, M_{i_{k,l}}\} \qquad l = 1, \ldots, r^k \tag{11.6.3-6}$$

where $i_{\kappa,l}$ is the model index at time κ from history l and

$$1 \leq i_{\kappa,l} \leq r \qquad \kappa = 1, \ldots, k \tag{11.6.3-7}$$

Note that the number of histories increases *exponentially with time*.

For example, with $r = 2$ one has at time $k = 2$ the following $r^k = 4$ possible sequences (histories) as shown below:

l	$i_{1,l}$	$i_{2,l}$
1	1	1
2	1	2
3	2	1
4	2	2

It will be assumed that the ***mode (model) switching*** — that is, the ***mode jump process*** — is a Markov process (Markov chain) with known mode transition probabilities

$$\boxed{p_{ij} \triangleq P\{M(k) = M_j | M(k-1) = M_i\}} \tag{11.6.3-8}$$

These ***mode transition probabilities*** will be assumed time-invariant and independent of the base state. In other words, this is a ***homogeneous Markov chain.***

The system (11.6.3-1), (11.6.3-2), and (11.6.3-8) is a generalized version of a ***hidden Markov model.***

The event that model j is in effect at time k is denoted as

$$M_j(k) \triangleq \{M(k) = M_j\} \tag{11.6.3-9}$$

The conditional probability of the lth history

$$\mu^{k,l} \triangleq P\{M^{k,l}|Z^k\} \tag{11.6.3-10}$$

will be evaluated next.

The lth sequence of models through time k can be written as

$$M^{k,l} = \{M^{k-1,s}, M_j(k)\} \tag{11.6.3-11}$$

where sequence s through $k-1$ is its *parent sequence* and M_j is its last element.

Then, in view of the Markov property,

$$P\{M_j(k)|M^{k-1,s}\} = P\{M_j(k)|M_i(k-1)\} \triangleq p_{ij} \tag{11.6.3-12}$$

where $i = s_{k-1}$, the index of the last model in the parent sequence s through $k-1$.

The conditional pdf of the state at k is obtained using the total probability theorem with respect to the mutually exclusive and exhaustive set of events (11.6.3-6), as a ***Gaussian mixture*** with an *exponentially increasing number of terms*

$$\boxed{p[x(k)|Z^k] = \sum_{l=1}^{r^k} p[x(k)|M^{k,l}, Z^k]P\{M^{k,l}|Z^k\}} \tag{11.6.3-13}$$

Since *to each mode sequence one has to match a filter*, it can be seen that an exponentially increasing number of filters are needed to estimate the (base) state, which makes the optimal approach impractical.

The probability of a mode history is obtained using Bayes' formula as

$$\begin{aligned}
\mu^{k,l} &= P\{M^{k,l}|Z^k\} \\
&= P\{M^{k,l}|z(k), Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M^{k,l}|Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k), M^{k-1,s}|Z^{k-1}\} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M^{k-1,s}, Z^{k-1}\}\mu^{k-1,s} \\
&= \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M^{k-1,s}\}\mu^{k-1,s}
\end{aligned} \tag{11.6.3-14}$$

where c is the normalization constant.

If the current mode depends only on the previous one (i.e., it is a Markov chain), then

$$\mu^{k,l} = \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]P\{M_j(k)|M_i(k-1)\}\mu^{k-1,s} \tag{11.6.3-15}$$

or

$$\boxed{\mu^{k,l} = \frac{1}{c}\, p[z(k)|M^{k,l}, Z^{k-1}]p_{ij}\mu^{k-1,s}} \tag{11.6.3-16}$$

where $i = s_{k-1}$ is the last model of the parent sequence s.

The above equation shows that *conditioning on the entire past history* is needed even if the random parameters are Markov.

Practical Algorithms

The only way to avoid the exponentially increasing number of histories, which have to be accounted for, is by going to suboptimal techniques.

A simple-minded suboptimal technique is to keep, say, the N histories with the largest probabilities, discard the rest, and renormalize the probabilities such that they sum up to unity.

The ***generalized pseudo-Bayesian (GPB)*** approaches combine histories of models that differ in "older" models. The first-order GPB, denoted as GPB1, considers only the possible models in the last sampling period. The second-order version, GPB2, considers all the possible models in the last two sampling periods. These algorithms require r and r^2 filters to operate in parallel, respectively.

Finally, the ***interacting multiple model (IMM)*** estimation algorithm will be presented. This algorithm is conceptually similar to GPB2, but requires only r filters to operate in parallel.

The Mode Transition Probabilities

The ***mode transition probabilities*** (11.6.3-8), indicated as assumed to be known, are actually ***estimator design parameters*** to be selected in the design process of the algorithm. This will be discussed in detail in Subsections 11.6.7 and 11.7.3.

Note

The GPB1 and IMM algorithms have approximately the same computational requirements as the static (fixed model) MM algorithm, but do not require ad hoc modifications as the latter, which is actually obsolete for switching models.

11.6.4 The GPB1 Multiple Model Estimator for Switching Models

In the ***generalized pseudo-Bayesian estimator of first order (GPB1)***, at time k the state estimate is computed under each possible current model — a total of r possibilities (hypotheses) are considered. All histories that differ in "older" models are combined together.

The total probability theorem is thus used as follows:

$$\begin{aligned} p[x(k)|Z^k] &= \sum_{j=1}^{r} p[x(k)|M_j(k), Z^k] P\{M_j(k)|Z^k\} \\ &= \sum_{j=1}^{r} p[x(k)|M_j(k), z(k), Z^{k-1}]\mu_j(k) \\ &\approx \sum_{j=1}^{r} p[x(k)|M_j(k), z(k), \hat{x}(k-1|k-1), P(k-1|k-1)]\mu_j(k) \end{aligned} \tag{11.6.4-1}$$

Thus at time $k-1$ there is a single ***lumped estimate*** $\hat{x}(k-1|k-1)$ and the associated covariance that summarize (approximately) the past Z^{k-1}. From this, one carries out the prediction to time k and the update at time k under r hypotheses, namely,

$$\hat{x}^j(k|k) = \hat{x}[k|k; M_j(k), \hat{x}(k-1|k-1), P(k-1|k-1)] \qquad j = 1, \ldots, r \tag{11.6.4-2}$$

$$P^j(k|k) = P[k|k; M_j(k), P(k-1|k-1)] \qquad j = 1, \ldots, r \tag{11.6.4-3}$$

After the update, the estimates are combined with the weightings $\mu_j(k)$ (detailed later), resulting in the new combined estimate $\hat{x}(k|k)$. In other words, *the r hypotheses are merged into a single hypothesis at the end of each cycle.*

Finally, the mode (or model) probabilities are updated.

Figure 11.6.4-1 describes this estimator, which requires r filters in parallel.

The output of each model-matched filter is the ***mode-conditioned state estimate*** $\hat{x}^j$, the associated covariance P^j and the ***mode likelihood function*** Λ_j.

After the filters are initialized, they run recursively using *the previous combined estimate.* Their likelihood functions are used to update the mode probabilities. The latest mode probabilities are used to combine the model-conditional estimates and covariances.

The structure of this algorithm is

$$(N_e; N_f) = (1; r) \tag{11.6.4-4}$$

where N_e is the *number of estimates* at the start of the cycle of the algorithm and N_f is the *number of filters* in the algorithm.

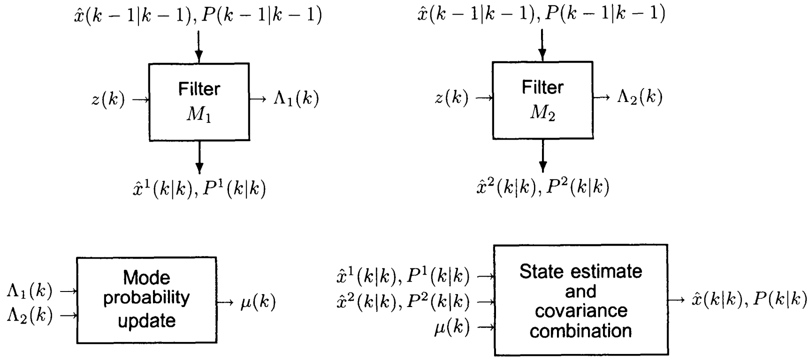

Figure 11.6.4-1: The GPB1 MM estimator for $r = 2$ switching models (one cycle).

The Algorithm

One cycle of the algorithm consists of the following:

1. ***Mode-matched filtering*** ($j = 1, \ldots, r$). Starting with $\hat{x}(k-1|k-1)$, one computes $\hat{x}^j(k|k)$ and the associated covariance $P^j(k|k)$ through a filter matched to $M_j(k)$. The likelihood functions

$$\Lambda_j(k) = p[z(k)|M_j(k), Z^{k-1}] \tag{11.6.4-5}$$

corresponding to these r filters are evaluated as

$$\boxed{\Lambda_j(k) = p[z(k)|M_j(k), \hat{x}(k-1|k-1), P(k-1|k-1)]} \tag{11.6.4-6}$$

2. ***Mode probability update*** ($j = 1, \ldots, r$). This is done as follows:

$$\begin{aligned} \mu_j(k) &\triangleq P\{M_j(k)|Z^k\} \\ &= P\{M_j(k)|z(k), Z^{k-1}\} \\ &= \frac{1}{c}\, p[z(k)|M_j(k), Z^{k-1}]P\{M_j(k)|Z^{k-1}\} \\ &= \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{r} P\{M_j(k)|M_i(k-1), Z^{k-1}\} \\ &\quad \cdot P\{M_i(k-1)|Z^{k-1}\} \end{aligned} \tag{11.6.4-7}$$

which yields with p_{ij} the known *mode transition probabilities*,

$$\boxed{\mu_j(k) = \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{r} p_{ij}\mu_i(k-1)} \tag{11.6.4-8}$$

where c is the normalization constant

$$c = \sum_{j=1}^{r} \Lambda_j(k) \sum_{i=1}^{r} p_{ij}\mu_i(k-1) \tag{11.6.4-9}$$

3. ***State estimate and covariance combination.*** The latest combined state estimate and covariance are obtained according to the summation (11.6.4-1) as

$$\boxed{\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)} \tag{11.6.4-10}$$

$$\boxed{P(k|k) = \sum_{j=1}^{r} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]'\}} \tag{11.6.4-11}$$

11.6.5 The GPB2 Multiple Model Estimator for Switching Models

In the ***generalized pseudo-Bayesian estimator of second order*** (or ***GPB2***), at time k the state estimate is computed under *each possible current and previous model* — a total of r^2 hypotheses (histories) are considered. All histories that differ only in "older" models are merged.

The total probability theorem is thus used as follows:

$$\begin{aligned} p[x(k)|Z^k] &= \sum_{j=1}^{r}\sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), Z^k]P\{M_i(k-1)|M_j(k), Z^k\} \\ &\quad \cdot P\{M_j(k)|Z^k\} \\ &= \sum_{j=1}^{r}\sum_{i=1}^{r} p[x(k)|M_j(k), z(k), M_i(k-1), Z^{k-1}]\mu_{i|j}(k-1|k)\mu_j(k) \\ &\approx \sum_{j=1}^{r}\sum_{i=1}^{r} p[x(k)|M_j(k), z(k), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)] \\ &\quad \cdot \mu_{i|j}(k-1|k)\mu_j(k) \end{aligned} \tag{11.6.5-1}$$

that is, the past $\{M_i(k-1), Z^{k-1}\}$ is approximated by the ***mode-conditioned estimate*** $\hat{x}^i(k-1|k-1)$ and associated covariance.

Thus at time $k-1$ there are r estimates and covariances, each to be predicted to time k and updated at time k under r hypotheses, namely,

$$\hat{x}^{ij}(k|k) \triangleq \hat{x}[k|k; M_j(k), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)] \qquad i,j = 1,\ldots,r \tag{11.6.5-2}$$

$$P^{ij}(k|k) \triangleq P[k|k; M_j(k), P^i(k-1|k-1)] \qquad i,j = 1,\ldots,r \tag{11.6.5-3}$$

After the update, the estimates corresponding to the same latest model hypothesis are combined with the weightings $\mu_{i|j}(k-1|k)$, detailed later, resulting

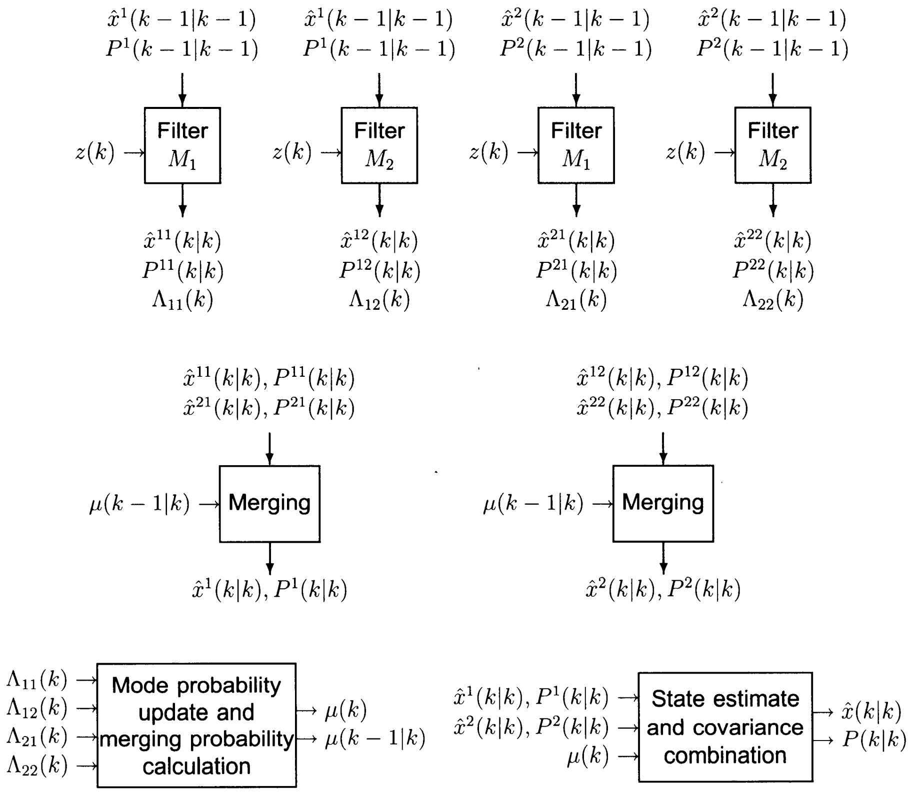

Figure 11.6.5-1: The GPB2 MM estimator for $r = 2$ models (one cycle).

in r estimates $\hat{x}^j(k|k)$. In other words, the r^2 hypotheses are *merged* into r *at the end of each estimation cycle.*

Figure 11.6.5-1 describes this algorithm, which requires r^2 parallel filters.

The structure of the GPB2 algorithm is

$$(N_e; N_f) = (r; r^2) \tag{11.6.5-4}$$

where N_e is the *number of estimates* at the start of the cycle of the algorithm and N_f is the *number of filters* in the algorithm.

The Algorithm

One cycle of the algorithm consists of the following:

1. ***Mode-matched filtering*** $(i, j = 1, \ldots, r)$. Starting with $\hat{x}^i(k-1|k-1)$, one computes $\hat{x}^{ij}(k|k)$ and the associated covariance $P^{ij}(k|k)$ through a filter matched to $M_j(k)$. The likelihood functions corresponding to these r^2 filters

$$\Lambda_{ij}(k) = p[z(k)|M_j(k), M_i(k-1), Z^{k-1}] \tag{11.6.5-5}$$

are evaluated as

$$\boxed{\Lambda_{ij}(k) = p[z(k)|M_j(k), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)] \qquad i, j = 1, \ldots, r} \tag{11.6.5-6}$$

2. ***Calculation of the merging probabilities*** $(i, j = 1, \ldots, r)$. The probability that mode i was in effect at $k-1$ if mode j is in effect at k is, conditioned on Z^k,

$$\begin{aligned}
\mu_{i|j}(k-1|k) &\triangleq P\{M_i(k-1)|M_j(k), Z^k\} \\
&= P\{M_i(k-1)|z(k), M_j(k), Z^{k-1}\} \\
&= \frac{1}{c_j} P[z(k), M_j(k)|M_i(k-1), Z^{k-1}] P\{M_i(k-1)|Z^{k-1}\} \\
&= \frac{1}{c_j} p[z(k)|M_j(k), M_i(k-1), Z^{k-1}] \\
&\quad \cdot P\{M_j(k)|M_i(k-1), Z^{k-1}\} P\{M_i(k-1)|Z^{k-1}\}
\end{aligned} \tag{11.6.5-7}$$

where $P[\cdot]$ denotes a mixed pdf-probability. Thus the ***merging probabilities*** are

$$\boxed{\mu_{i|j}(k-1|k) = \frac{1}{c_j} \Lambda_{ij}(k) p_{ij} \mu_i(k-1) \qquad i, j = 1, \ldots, r} \tag{11.6.5-8}$$

where

$$c_j = \sum_{i=1}^{r} \Lambda_{ij}(k) p_{ij} \mu_i(k-1) \tag{11.6.5-9}$$

The ***mode transition probabilities*** p_{ij} are assumed to be known — their selection is part of the algorithm design process.

3. ***Merging*** ($j = 1, \ldots, r$). The state estimate corresponding to $M_j(k)$ is obtained by combining the estimates (11.6.5-2) according to the inner summation in (11.6.5-1) as follows

$$\boxed{\hat{x}^j(k|k) = \sum_{i=1}^{r} \hat{x}^{ij}(k|k)\mu_{i|j}(k-1|k) \qquad j = 1, \ldots, r} \tag{11.6.5-10}$$

The covariance corresponding to the above is

$$\boxed{\begin{aligned} P^j(k|k) = \textstyle\sum_{i=1}^{r} \mu_{i|j}(k-1|k)\{P^{ij}(k|k) \\ +[\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)][\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)]'\} \end{aligned}} \tag{11.6.5-11}$$

4. ***Mode probability updating*** ($j = 1, \ldots, r$). This is done as follows

$$\begin{aligned} \mu_j(k) &\triangleq P\{M_j(k)|z(k), Z^{k-1}\} \\ &= \frac{1}{c} P[z(k), M_j(k)|Z^{k-1}] \\ &= \frac{1}{c} \sum_{i=1}^{r} P[z(k), M_j(k)|M_i(k-1), Z^{k-1}]P\{M_i(k-1)|Z^{k-1}\} \\ &= \frac{1}{c} \sum_{i=1}^{r} p(z(k)|M_j(k), M_i(k-1), Z^{k-1}) \\ &\quad \cdot P\{M_j(k)|M_i(k-1), Z^{k-1}\}\mu_i(k-1) \end{aligned} \tag{11.6.5-12}$$

Thus the updated ***mode probabilities*** are

$$\boxed{\mu_j(k) = \frac{1}{c} \sum_{i=1}^{r} \Lambda_{ij}(k) p_{ij} \mu_i(k-1) = \frac{c_j}{c} \qquad j = 1, \ldots, r} \tag{11.6.5-13}$$

where c_j is the expression from (11.6.5-9) and c is the normalization constant

$$c = \sum_{j=1}^{r} c_j \tag{11.6.5-14}$$

5. ***State estimate and covariance combination.*** The latest state estimate and covariance *for output only* are

$$\boxed{\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)} \tag{11.6.5-15}$$

$$\boxed{P(k|k) = \sum_{j=1}^{r} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]'\}} \tag{11.6.5-16}$$

11.6.6 The Interacting Multiple Model Estimator

In the ***interacting multiple model (IMM) estimator***, at time k the state estimate is computed under *each possible current model* using r filters, with each filter using a different combination of the previous model-conditioned estimates — *mixed initial condition.*

The total probability theorem is used as follows to yield r filters running in parallel:

$$\begin{aligned} p[x(k)|Z^k] &= \sum_{j=1}^{r} p[x(k)|M_j(k), Z^k] P\{M_j(k)|Z^k\} \\ &= \sum_{j=1}^{r} p[x(k)|M_j(k), z(k), Z^{k-1}]\mu_j(k) \end{aligned} \tag{11.6.6-1}$$

The model-conditioned posterior pdf of the state, given by

$$p[x(k)|M_j(k), z(k), Z^{k-1}] = \frac{p[z(k)|M_j(k), x(k)]}{p[z(k)|M_j(k), Z^{k-1}]} p[x(k)|M_j(k), Z^{k-1}] \tag{11.6.6-2}$$

reflects one cycle of the state estimation filter matched to model $M_j(k)$ starting with the prior, which is the last term above.

The total probability theorem is now applied to the last term above (the prior), yielding

$$\begin{aligned} p[x(k)|M_j(k), Z^{k-1}] &= \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), Z^{k-1}] \\ &\qquad \cdot P\{M_i(k-1)|M_j(k), Z^{k-1}\} \\ &\approx \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), \{\hat{x}^l(k-1|k-1), P^l(k-1|k-1)\}_{l=1}^{r}] \\ &\qquad \cdot \mu_{i|j}(k-1|k-1) \\ &= \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)] \\ &\qquad \cdot \mu_{i|j}(k-1|k-1) \end{aligned} \tag{11.6.6-3}$$

The second line above reflects the approximation that the past through $k-1$ is summarized by r model-conditioned estimates and covariances. The last line of (11.6.6-3) is a mixture with weightings, denoted as $\mu_{i|j}(k-1|k-1)$, different for each current model $M_j(k)$. This mixture is assumed to be a mixture of Gaussian pdfs (a Gaussian sum) and then approximated via moment matching by a single Gaussian (details given later):

$$\begin{aligned}
p[x(k)|M_j(k), Z^{k-1}] &= \sum_{i=1}^{r} \mathcal{N}\left[x(k); E[x(k)|M_j(k), \hat{x}^i(k-1|k-1)], \text{cov}[\cdot]\right] \\
&\quad \cdot \mu_{i|j}(k-1|k-1) \\
&\approx \mathcal{N}\left[x(k); \sum_{i=1}^{r} E\left[x(k)|M_j(k), \hat{x}^i(k-1|k-1)\right] \mu_{i|j}(k-1|k-1), \text{cov}[\cdot]\right] \\
&= \mathcal{N}\left[x(k); E[x(k)|M_j(k), \sum_{i=1}^{r} \hat{x}^i(k-1|k-1)\mu_{i|j}(k-1|k-1)], \text{cov}[\cdot]\right]
\end{aligned} \tag{11.6.6-4}$$

The last line above follows from the linearity of the Kalman filter and amounts to the following:

> The input to the filter matched to model j is obtained from an ***interaction*** of the r filters, which consists of the ***mixing*** of the estimates $\hat{x}^i(k-1|k-1)$ with the weightings (probabilities) $\mu_{i|j}(k-1|k-1)$, called the ***mixing probabilities***.

The above is equivalent to hypothesis merging *at the beginning* of each estimation cycle [Blom88]. More specifically, the r hypotheses, instead of "fanning out" into r^2 hypotheses (as in the GPB2 — see Fig. 11.6.5-1), are "mixed" into a new set of r hypotheses as shown in Fig. 11.6.6-1. This is the key feature that yields r hypotheses with r filters, rather than with r^2 filters as in the GPB2 algorithm.

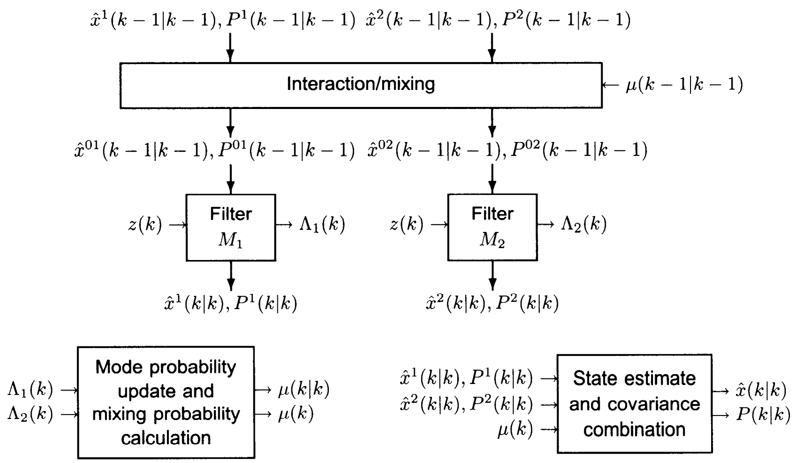

Figure 11.6.6-1: The IMM estimator (one cycle).

Figure 11.6.6-1 describes this algorithm, which consists of r interacting filters operating in parallel. The mixing is done at the input of the filters with the probabilities, detailed later in (11.6.6-6), conditioned on Z^{k-1}. In contrast to this, the GPB2 algorithm has r^2 filters and a somewhat similar mixing is done, but at their outputs, with the probabilities (11.6.5-7), conditioned on Z^k.

The structure of the IMM algorithm is

$$(N_e; N_f) = (r; r) \tag{11.6.6-5}$$

where N_e is the *number of estimates* at the start of the cycle of the algorithm and N_f is the *number of filters* in the algorithm.

The Algorithm

One cycle of the algorithm consists of the following:

1. ***Calculation of the mixing probabilities*** $(i, j = 1, \ldots, r)$. The probability that mode M_i was in effect at $k-1$ given that M_j is in effect at k conditioned on Z^{k-1} is

$$\begin{aligned} \mu_{i|j}(k-1|k-1) &\triangleq P\{M_i(k-1)|M_j(k), Z^{k-1}\} \\ &= \frac{1}{\bar{c}_j} P\{M_j(k)|M_i(k-1), Z^{k-1}\} P\{M_i(k-1)|Z^{k-1}\} \end{aligned} \tag{11.6.6-6}$$

The above are the ***mixing probabilities***, which can be written as

$$\boxed{\mu_{i|j}(k-1|k-1) = \frac{1}{\bar{c}_j}\, p_{ij}\mu_i(k-1) \qquad i, j = 1, \ldots, r} \tag{11.6.6-7}$$

where the normalizing constants are

$$\bar{c}_j = \sum_{i=1}^{r} p_{ij}\mu_i(k-1) \qquad j = 1, \ldots, r \tag{11.6.6-8}$$

Note the difference between (11.6.6-6), where the conditioning is Z^{k-1}, and (11.6.5-7), where the conditioning is Z^k. This is what makes it possible to carry out the mixing at the *beginning* of the cycle, rather than the standard merging at the *end* of the cycle.

2. ***Mixing*** $(j = 1, \ldots, r)$. Starting with $\hat{x}^i(k-1|k-1)$, one computes the mixed initial condition for the filter matched to $M_j(k)$ according to (11.6.6-4) as

$$\boxed{\hat{x}^{0j}(k-1|k-1) = \sum_{i=1}^{r} \hat{x}^i(k-1|k-1)\mu_{i|j}(k-1|k-1) \qquad j = 1, \ldots, r} \tag{11.6.6-9}$$

The covariance corresponding to the above is

$$\boxed{\begin{aligned} P^{0j}(k-1|k-1) &= \sum_{i=1}^{r} \mu_{i|j}(k-1|k-1)\Big\{P^i(k-1|k-1) \\ &\quad + [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)] \\ &\quad \cdot [\hat{x}^i(k-1|k-1) - \hat{x}^{0j}(k-1|k-1)]'\Big\} \\ &\qquad j = 1,\ldots,r \end{aligned}} \tag{11.6.6-10}$$

3. ***Mode-matched filtering*** ($j = 1,\ldots,r$). The estimate (11.6.6-9) and covariance (11.6.6-10) are used as input to the filter matched to $M_j(k)$, which uses $z(k)$ to yield $\hat{x}^j(k|k)$ and $P^j(k|k)$.

The likelihood functions corresponding to the r filters

$$\Lambda_j(k) = p[z(k)|M_j(k), Z^{k-1}] \tag{11.6.6-11}$$

are computed using the mixed initial condition (11.6.6-9) and the associated covariance (11.6.6-10) as

$$\Lambda_j(k) = p[z(k)|M_j(k), \hat{x}^{0j}(k-1|k-1), P^{0j}(k-1|k-1)] \qquad j = 1,\ldots,r \tag{11.6.6-12}$$

that is,

$$\boxed{\begin{aligned} \Lambda_j(k) = \mathcal{N}[z(k); \hat{z}^j[k|k-1; \hat{x}^{0j}(k-1|k-1)], S^j[k; P^{0j}(k-1|k-1)]] \\ j = 1,\ldots,r \end{aligned}} \tag{11.6.6-13}$$

4. ***Mode probability update*** ($j = 1,\ldots,r$). This is done as follows:

$$\begin{aligned} \mu_j(k) &\triangleq P\{M_j(k)|Z^k\} \\ &= \frac{1}{c}\, p[z(k)|M_j(k), Z^{k-1}]P\{M_j(k)|Z^{k-1}\} \\ &= \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{r} P\{M_j(k)|M_i(k-1), Z^{k-1}\}P\{M_i(k-1)|Z^{k-1}\} \\ &= \frac{1}{c}\, \Lambda_j(k) \sum_{i=1}^{r} p_{ij}\mu_i(k-1) \qquad j = 1,\ldots,r \end{aligned} \tag{11.6.6-14}$$

or

$$\boxed{\mu_j(k) = \frac{1}{c}\, \Lambda_j(k)\bar{c}_j \qquad j = 1,\ldots,r} \tag{11.6.6-15}$$

where $\bar{c}_j$ is the expression from (11.6.6-8) and

$$c = \sum_{j=1}^{r} \Lambda_j(k)\bar{c}_j \tag{11.6.6-16}$$

is the normalization constant for (11.6.6-15).

5. ***Estimate and covariance combination.*** Combination of the model-conditioned estimates and covariances is done according to the mixture equations

$$\boxed{\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)} \tag{11.6.6-17}$$

$$\boxed{P(k|k) = \sum_{j=1}^{r} \mu_j(k)\{P^j(k|k) + [\hat{x}^j(k|k) - \hat{x}(k|k)][\hat{x}^j(k|k) - \hat{x}(k|k)]'\}} \tag{11.6.6-18}$$

This combination is *only* for output purposes — it is not part of the algorithm recursions.

Note

One possible generalization of the IMM estimator is the "second-order IMM" with an extra period depth. While the derivations are rather lengthy, it has been reported that this algorithm is *identical* to the GPB2 [Barret90].

11.6.7 An Example with the IMM Estimator

The use of the IMM estimator is illustrated on the example simulated in Section 11.5 where several of the earlier techniques were compared. The results presented in the sequel, which were obtained with DynaEst$^{\text{TM}}$, deal with the turn of $90°$ over 20 sampling periods.

A two-model IMM, designated as IMM2, was first used. This algorithm consisted of

1. A constant velocity model (second-order, with no process noise); and
2. A Wiener process acceleration model (third-order model) with process noise (acceleration increment over a sampling period) $q = 10^{-3} = (0.0316\,\text{m/s}^2)^2$.

Note that the acceleration in this case ($0.075\,\text{m/s}^2$) corresponds to about $2.4\sqrt{q}$.

The Markov chain transition matrix between these models was taken as

$$[p_{ij}] = \begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix} \tag{11.6.7-1}$$

The final results were not very sensitive to these values (e.g., p_{11} can be between 0.8 and 0.98). The lower (higher) value will yield less (more) peak error during maneuver but higher (lower) RMS error during the quiescent period; that is, it has a higher (lower) bandwidth.

A three-model IMM, designated as IMM3, was also used. This consisted of the above two models plus another one:

3. A constant acceleration (third-order) model without process noise.