MULTIPLE MODELS AND ADAPTIVE ESTIMATION

The multiple model (MM) algorithms; Hybrid systems — the system behaves according to one of a finite number of models, it is in one of several modes (operating regimes), both discrete (structure/parameters) and continuous uncertainties

1. The static MM algorithm —for fixed (nonswitching) models

2. The optimal dynamic MM algorithm — for switching models —Markov chain, two suboptimal approaches: Generalized pseudo-Bayesian (GPB) ; Interacting multiple model (IMM)

Adaptive estimation algorithms — in many practical situations the "parameters of the problem" are partially unknown and possibly time-varying. The state estimation techniques that can "adapt" themselves to certain types of uncertainties

11/1/2022

THE MULTIPLE MODEL APPROACH

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

A **Bayesian framework** : Starting with prior probabilities of each

model being correct (i.e., the system is in a particular mode), the corresponding posterior probabilities are calculated.

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 (timeinvariant 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

Notation

The model is one of r **possible models** (the system is in one of r **modes**).

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

The prior probability that M_j is correct, the system is in **mode j**

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

the correct model is among the assumed r possible models.
It will be assumed that all models are **linear-Gaussian**.

Subsequently, the dynamic situation of switching models or mode jumping is considered: the system undergoes transitions from one mode to another; the **dynamic MM estimator.**

The Static Multiple Model Estimator

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

$$\mu_j(k) \triangleq P\{M_j|Z^k\} = P\{M_j|z(k), Z^{k-1}\} = \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{p[z(k)|Z^{k-1}]}$$

$$= \frac{p[z(k)|Z^{k-1}, M_j]P\{M_j|Z^{k-1}\}}{\sum_{i=1}^{r} p[z(k)|Z^{k-1}, M_i]P\{M_i|Z^{k-1}\}}$$

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

the **likelihood function of mode j at time k**, which, under the linear-Gaussian assumptions

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

where v_j and S_j are the innovation and its covariance from the **mode-matched filter** corresponding to **mode j**

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 based on its likelihood function relative to the other filters' likelihood functions.

In a nonlinear situation the filters are EKF instead of KF

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.

The **static multiple model estimator** for r = 2 fixed models, a bank of filters
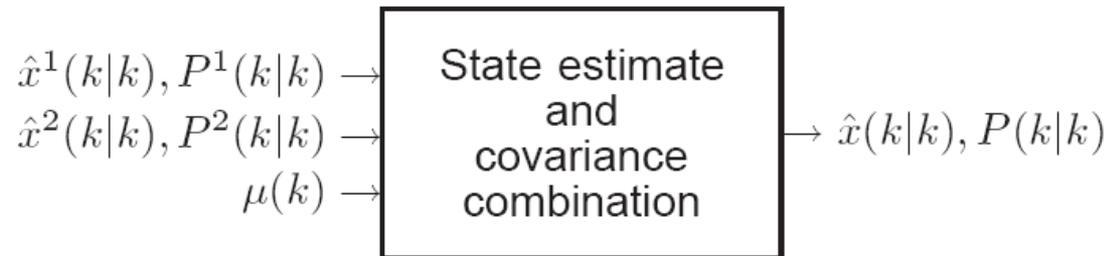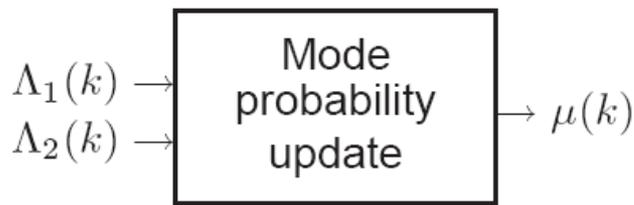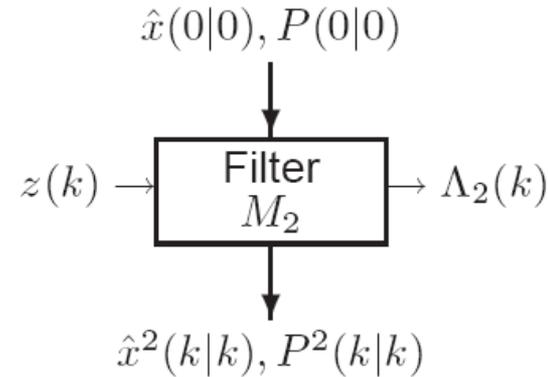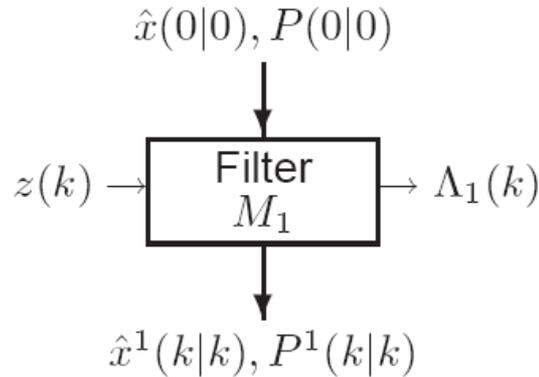
Under the above assumptions the pdf of the state of the system is a **Gaussian mixture** with r terms

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

The **combination of the mode-conditioned estimates**

$$\hat{x}(k|k) = \sum_{j=1}^{r} \mu_j(k)\hat{x}^j(k|k)$$

the **covariance of the combined estimate** is

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

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 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.

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.

A shortcoming of the static MM estimator when used (wrongly) 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

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

$$x(k) = F[M(k)]x(k-1) + v[k-1, M(k)]$$

$$z(k) = H[M(k)]x(k) + w[k, M(k)]$$

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.

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

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

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

$$F[M_j] = F_j$$

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

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

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

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

$$1 \leq i_{\kappa,l} \leq r \qquad \kappa = 1, \ldots, k$$

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)

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**

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

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 is a generalized version of a **hidden
Markov model**.

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

$$M_j(k) \triangleq \{M(k) = M_j\}$$

The **conditional probability of the lth history** will be evaluated next

$$\mu^{k,l} \triangleq P\{M^{k,l}|Z^k\}$$

The lth sequence of models through time k

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

where sequence s through k−1 is its parent sequence and M$_j$ is its last element, in view of the **Markov property**

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

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, as a **Gaussian mixture with an exponentially increasing number of terms**

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

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

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

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

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

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

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

Impossible to implement in Practice, approximations are needed

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.

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, indicated as assumed to be known, are actually estimator design parameters to be selected in the design process of the algorithm.

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.

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

Thus at time k−1 there is a single lumped estimate and $\hat{x}(k-1|k-1)$ 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,

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

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

After the update, the estimates are combined with the weightings $\mu_j(k)$, 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.** 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

$$(N_e; N_f) = (1; r)$$

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 GPB1 MM estimator for r = 2 switching models (one cycle).
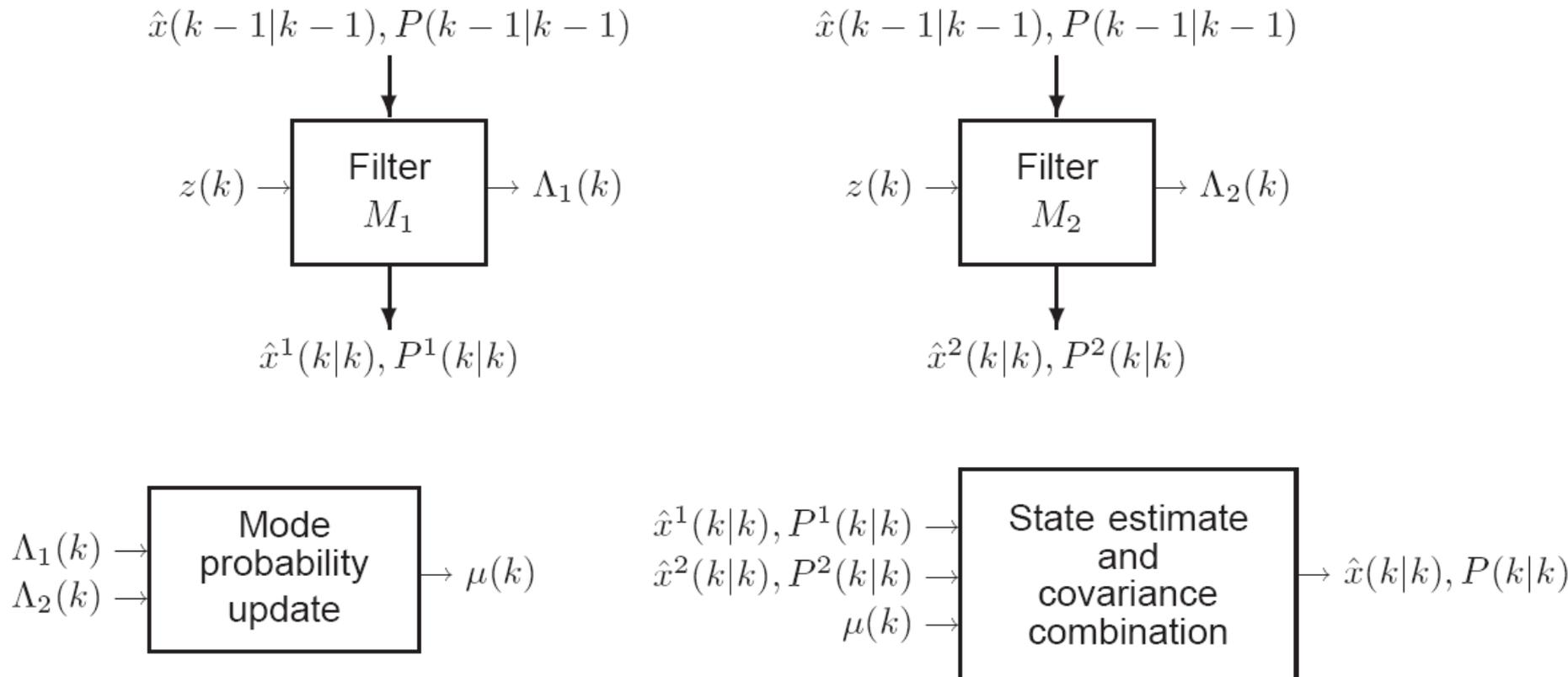
The GPB1 MM Algorithm

1. **Mode-matched filtering** (j = 1, . . . , 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

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

corresponding to these r filters are evaluated as Λ_j

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

2. **Mode probability update** (j = 1, . . . , r).

$$\mu_j(k) \triangleq P\{M_j(k)|Z^k\}$$

$$= P\{M_j(k)|z(k), Z^{k-1}\}$$

$$= \frac{1}{c} p[z(k)|M_j(k), Z^{k-1}]P\{M_j(k)|Z^{k-1}\}$$

$$= \frac{1}{c} \Lambda_j(k) \sum_{i=1}^{r} P\{M_j(k)|M_i(k-1), Z^{k-1}\}$$

$$\cdot P\{M_i(k-1)|Z^{k-1}\}$$

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

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

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

3. **State estimate and covariance combination**.

$$\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)$$

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

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.

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

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

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

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

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 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.

The GPB2 MM estimator for r = 2 models (one cycle).

$\hat{x}^1(k-1|k-1)$ $\quad$ $\hat{x}^1(k-1|k-1)$ $\quad$ $\hat{x}^2(k-1|k-1)$ $\quad$ $\hat{x}^2(k-1|k-1)$
$P^1(k-1|k-1)$ $\quad$ $P^1(k-1|k-1)$ $\quad$ $P^2(k-1|k-1)$ $\quad$ $P^2(k-1|k-1)$

$z(k) \rightarrow$ | Filter M_1 | $\qquad$ $z(k) \rightarrow$ | Filter M_2 | $\qquad$ $z(k) \rightarrow$ | Filter M_1 | $\qquad$ $z(k) \rightarrow$ | Filter M_2 |

$\hat{x}^{11}(k|k)$ $\qquad\qquad$ $\hat{x}^{12}(k|k)$ $\qquad\qquad$ $\hat{x}^{21}(k|k)$ $\qquad\qquad$ $\hat{x}^{22}(k|k)$
$P^{11}(k|k)$ $\qquad\qquad$ $P^{12}(k|k)$ $\qquad\qquad$ $P^{21}(k|k)$ $\qquad\qquad$ $P^{22}(k|k)$
$\Lambda_{11}(k)$ $\qquad\qquad$ $\Lambda_{12}(k)$ $\qquad\qquad$ $\Lambda_{21}(k)$ $\qquad\qquad$ $\Lambda_{22}(k)$

$\hat{x}^{11}(k|k), P^{11}(k|k)$ $\qquad\qquad\qquad$ $\hat{x}^{12}(k|k), P^{12}(k|k)$
$\hat{x}^{21}(k|k), P^{21}(k|k)$ $\qquad\qquad\qquad$ $\hat{x}^{22}(k|k), P^{22}(k|k)$

$\mu(k-1|k) \rightarrow$ | Merging | $\qquad\qquad$ $\mu(k-1|k) \rightarrow$ | Merging |

$\hat{x}^1(k|k), P^1(k|k)$ $\qquad\qquad\qquad$ $\hat{x}^2(k|k), P^2(k|k)$

$\Lambda_{11}(k) \rightarrow$ | Mode probability update and merging probability calculation | $\rightarrow \mu(k)$
$\Lambda_{12}(k) \rightarrow$ | | $\rightarrow \mu(k-1|k)$
$\Lambda_{21}(k) \rightarrow$ | |
$\Lambda_{22}(k) \rightarrow$ | |

$\hat{x}^1(k|k), P^1(k|k) \rightarrow$ | State estimate and covariance combination | $\rightarrow \hat{x}(k|k)$
$\hat{x}^2(k|k), P^2(k|k) \rightarrow$ | | $\rightarrow P(k|k)$
$\mu(k) \rightarrow$ | |

The structure of the GPB2 algorithm is

$$(N_e; N_f) = (r; r^2)$$

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 GPB2 MM Algorithm

1. **Mode-matched filtering** (j = 1, . . . , 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

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

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

2. **Calculation of the merging probabilities** (i, j = 1, . . . , 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,

$$\mu_{i|j}(k-1|k) \triangleq P\{M_i(k-1)|M_j(k), Z^k\}$$

$$= P\{M_i(k-1)|z(k), M_j(k), Z^{k-1}\}$$

$$= \frac{1}{c_j}P[z(k), M_j(k)|M_i(k-1), Z^{k-1}]P\{M_i(k-1)|Z^{k-1}\}$$

$$= \frac{1}{c_j}p[z(k)|M_j(k), M_i(k-1), Z^{k-1}]$$

$$\cdot P\{M_j(k)|M_i(k-1), Z^{k-1}\}P\{M_i(k-1)|Z^{k-1}\}$$

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

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

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

3. *Merging* ($j = 1, \ldots, r$).

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

$$P^j(k|k) = \sum_{i=1}^r \mu_{i|j}(k-1|k)\{P^{ij}(k|k)$$
$$+[\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)][\hat{x}^{ij}(k|k) - \hat{x}^j(k|k)]'\}$$

4. **Mode probability updating** ($j = 1, \ldots, r$).

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

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

5. *State estimate and covariance combination*

The latest state estimate and covariance *for output only* are

$$\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)$$

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

The Interacting Multiple Model Estimator (IMM)

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

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

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

The model-conditioned posterior pdf of the state, given above, reflects one cycle of the state estimation filter matched to model $Mj(k)$ starting with the prior, which is the last term above. The total probability theorem is now applied to the prior

$$p[x(k)|M_j(k), Z^{k-1}] = \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), Z^{k-1}]$$

$$\cdot P\{M_i(k-1)|M_j(k), Z^{k-1}\}$$

$$\approx \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), \{\hat{x}^l(k-1|k-1), P^l(k-1|k-1)\}_{l=1}^{r}]$$

$$\cdot \mu_{i|j}(k-1|k-1)$$

$$= \sum_{i=1}^{r} p[x(k)|M_j(k), M_i(k-1), \hat{x}^i(k-1|k-1), P^i(k-1|k-1)]$$

$$\cdot \mu_{i|j}(k-1|k-1)$$

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 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 :

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

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. More specifically, the r hypotheses, instead of "fanning out" into r2 hypotheses (as in the GPB2), are "mixed" into a new set of r hypotheses. This is the key feature that yields r hypotheses with r filters, rather than with r2 filters as in the GPB2 algorithm.

The structure of the IMM algorithm is

$$(N_e; N_f) = (r; r)$$

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 IMM estimator (one cycle).

$\hat{x}^1(k-1|k-1), P^1(k-1|k-1)$ $\hat{x}^2(k-1|k-1), P^2(k-1|k-1)$

Interaction/mixing $\leftarrow \mu(k-1|k-1)$

$\hat{x}^{01}(k-1|k-1), P^{01}(k-1|k-1)$ $\hat{x}^{02}(k-1|k-1), P^{02}(k-1|k-1)$

$z(k) \rightarrow$ Filter M_1 $\rightarrow \Lambda_1(k)$ $z(k) \rightarrow$ Filter M_2 $\rightarrow \Lambda_2(k)$

$\hat{x}^1(k|k), P^1(k|k)$ $\hat{x}^2(k|k), P^2(k|k)$

$\Lambda_1(k) \rightarrow$ Mode probability update and mixing probability calculation $\rightarrow \mu(k|k)$ $\rightarrow \mu(k)$

$\Lambda_2(k) \rightarrow$

$\hat{x}^1(k|k), P^1(k|k) \rightarrow$ State estimate and covariance combination $\rightarrow \hat{x}(k|k)$
$\hat{x}^2(k|k), P^2(k|k) \rightarrow$ $\rightarrow P(k|k)$
$\mu(k) \rightarrow$
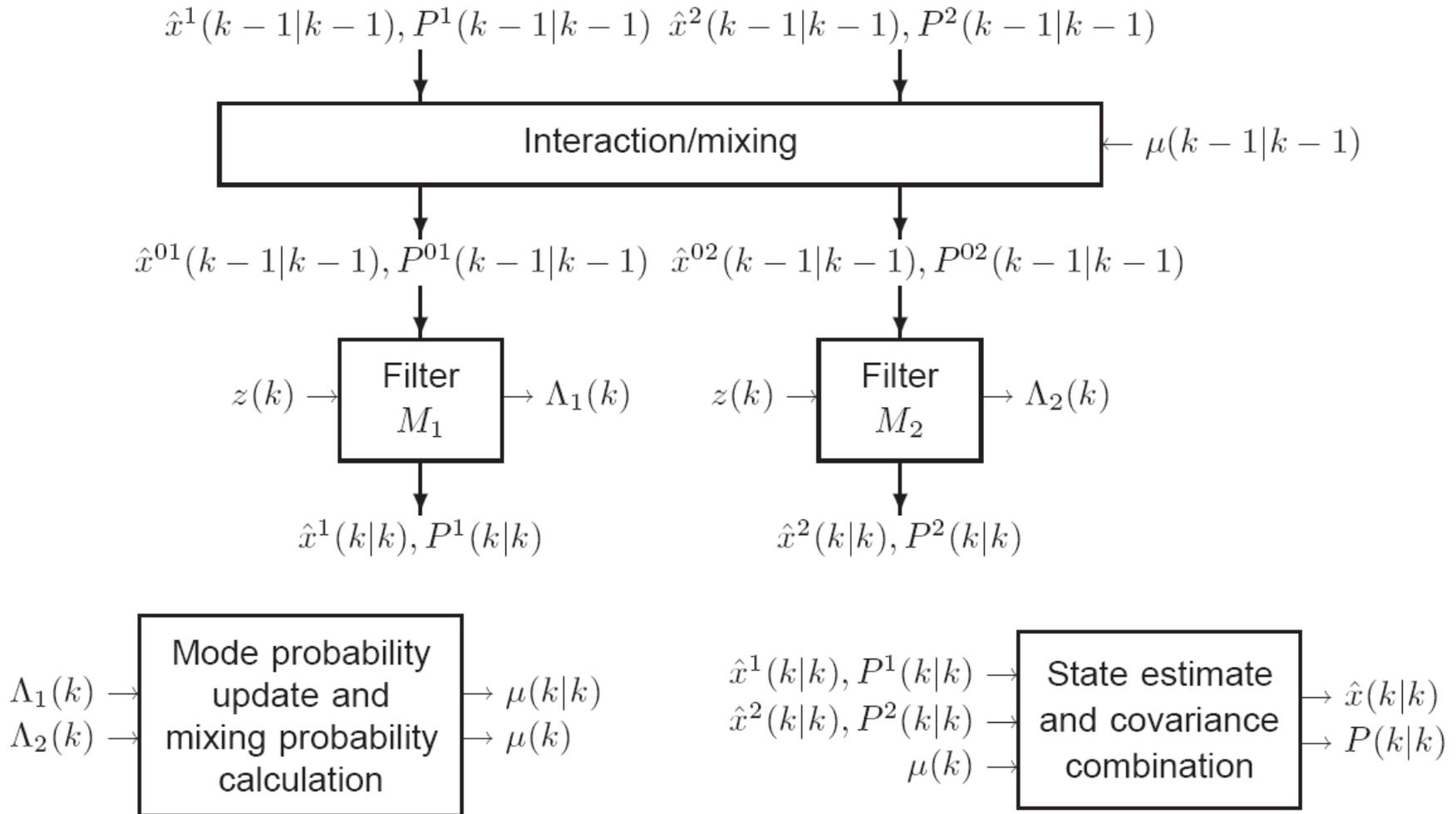
The IMM Algorithm

1. **Calculation of the mixing probabilities** (i, j = 1, . . . , 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

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

The above are the mixing probabilities

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

the normalizing constants are

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

the mixing at the beginning of the cycle, rather than the standard merging at the end of the cycle.

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

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

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

3. Mode-matched filtering (j = 1, . . . , r). The estimate and covariance are used as input to the filter matched to *Mj(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

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

are computed using the mixed initial condition and the associated covariance as

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

$$j = 1, \ldots, r$$

$$\Lambda_j(k) = \mathcal{N}[z(k); \hat{z}^j[k|k-1; \hat{x}^{0j}(k-1|k-1)], S^j[k; P^{0j}(k-1|k-1)]]$$

$$j = 1, \ldots, r$$

4. **Mode probability update** (j = 1, . . . , r).

$$\mu_j(k) \triangleq P\{M_j(k)|Z^k\}$$

$$= \frac{1}{c} p[z(k)|M_j(k), Z^{k-1}] P\{M_j(k)|Z^{k-1}\}$$

$$= \frac{1}{c} \Lambda_j(k) \sum_{i=1}^{r} P\{M_j(k)|M_i(k-1), Z^{k-1}\} P\{M_i(k-1)|Z^{k-1}\}$$

$$= \frac{1}{c} \Lambda_j(k) \sum_{i=1}^{r} p_{ij}\mu_i(k-1) \qquad j = 1, \ldots, r$$

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

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

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

$$\hat{x}(k|k) = \sum_{j=1}^{r} \hat{x}^j(k|k)\mu_j(k)$$

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

This combination is only for output purposes

One possible generalization of the IMM estimator is the "second-order IMM" with an extra period depth. It has been reported that this algorithm is identical to the GPB2

The Multiple Model Approach — Summary

The multiple model or hybrid system approach assumes the system to be in one of a finite number of modes.

Each model is characterized by its parameters — the models can differ in the level of the process noise (its variance), a deterministic input, and/or any other parameter (different dimension state vectors are also possible).

For the **fixed model case** the estimation algorithm consists of the following:

- For each model a filter "matched" to its parameters is yielding model conditioned estimates and covariances.

- A mode probability calculator — a Bayesian model comparator — updates the probability of each mode using the likelihood function (innovations) of each filter the prior probability of each model

- An Estimate combiner computes the overall estimate and the associated covariance as the weighted sum of the model-conditioned estimates and the corresponding covariance — via the (Gaussian) mixture equations.

The Multiple Model Approach — Summary continues

For systems that undergo changes in their mode during their operation —

mode jumping (model switching) — one can obtain the optimal multiple model estimator which, however, consists of an **exponentially increasing number of filters**. This is because the optimal approach requires conditioning on each mode history, and their number is increasing exponentially.

Thus, suboptimal algorithms are necessary.

The first-order **generalized pseudo-Bayesian (GPB1)** MM approach computes the state estimate accounting for each possible current model.

The **second-order generalized pseudo-Bayesian (GPB2)** MM approach computes the state estimate accounting for

- Each possible current model
- Each possible model at the previous time

The **interacting multiple model (IMM)** approach computes the state estimate that accounts for each possible current model using a suitable mixing of the previous model-conditioned estimates depending on the current model.

USE OF EKF FOR SIMULTANEOUS STATE AND PARAMETER ESTIMATION

Augmentation of the State.

Denoting the unknown parameters as a vector θ, the augmented state will be the stacked vector consisting of the base state x and θ

$$y(k) \triangleq \begin{bmatrix} x(k) \\ \theta \end{bmatrix}$$

The linear dynamic equation of x

$$x(k+1) = F(\theta)x(k) + G(\theta)u(k) + v(k)$$

and the "dynamic equation" of the parameter vector

$$\theta(k+1) = \theta(k)$$, better is ! $$\theta(k+1) = \theta(k) + v_\theta(k)$$

can be rewritten as a nonlinear dynamic equation for the augmented state

$$y(k+1) = f[y(k), u(k)] + v(k)$$

Any nonzero variance of the process noise for the parameter will prevent the filter-calculated variances of the parameter estimates from converging to zero. Furthermore, this also gives the filter the ability to estimate slowly varying parameters

The **choice of the variance of the artificial process noise for the parameters — the tuning of the filter** — can be done as follows:

1. Choose the standard deviation of the process noise as a few percent of the (estimated/guessed) value of the parameter.

2. Simulate the system and the estimator with random initial estimates (for the base state as well as the parameters) and monitor the normalized estimation errors.

3. Adjust the noise variances until, for the problem of interest, the filter is consistent — it yields estimation errors commensurate with the calculated augmented state covariance matrix.

An Example of Use of the EKF for Parameter estimation

Consider the scalar system, that is, its base state x is a scalar, given by

$$x(k+1) = a(k)x(k) + b(k)u(k) + v_1(k)$$

where $v_1(k)$ is the base state process noise and the two unknown parameters

are $a(k)$ and $b(k)$, possibly time-varying.

The observations are

$$z(k) = x(k) + w(k)$$

the augmented state is

$$y(k) \triangleq \begin{bmatrix} y_1(k) \\ y_2(k) \\ y_3(k) \end{bmatrix} \triangleq \begin{bmatrix} x(k) \\ a(k) \\ b(k) \end{bmatrix}$$

With this the nonlinear dynamic equation

$$y_1(k+1) = f^1[y(k), u(k)] + v_1(k) \triangleq y_1(k)y_2(k) + y_3(k)u(k) + v_1(k)$$

$$y_i(k+1) = f^i[y(k), u(k)] + v_i(k) \triangleq y_i(k) + v_i(k) \qquad i = 2, 3$$

The augmented state equation is then

$$y(k+1) = f[y(k), u(k)] + v(k)$$

with the augmented process noise

$$v(k) \triangleq \begin{bmatrix} v_1(k) \\ v_2(k) \\ v_3(k) \end{bmatrix}$$

assumed zero mean and with covariance

$$Q = \text{diag}(q_1, q_2, q_3)$$

The **second-order EKF** will use the following augmented state prediction equations.

$$\hat{y}_1(k+1|k) = \hat{y}_2(k|k)\hat{y}_1(k|k) + \hat{y}_3(k|k)u(k) + P_{21}(k|k)$$

since

$$E[a(k)x(k)|Z^k] = E[y_2(k)y_1(k)|Z^k] = \hat{y}_2(k|k)\hat{y}_1(k|k) + \text{cov}[y_2(k), y_1(k)|Z^k]$$

The predicted values of the remaining two components of the augmented state, which are the system's unknown parameters

$$\hat{y}_i(k+1|k) = \hat{y}_i(k|k) \qquad i = 2,3$$

evaluation of the Jacobian of the vector f and the Hessians of its components

$$F(k) = \begin{bmatrix} \hat{y}_2(k|k) & \hat{y}_1(k|k) & u(k) \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$f_{yy}^1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad f_{yy}^2 = 0 \qquad f_{yy}^3 = 0$$

for the augmented state

$$\hat{y}(k+1|k) = f[\hat{y}(k|k), u(k)] + \frac{1}{2}\sum_{i=1}^{n_x} e_i \text{tr}[f_{yy}^i P(k|k)]$$

The prediction covariance of the base state

$$P_{11}(k+1|k) = \hat{y}_2(k|k)^2 P_{11}(k|k) + 2\hat{y}_2(k|k)\hat{y}_1(k|k)P_{21}(k|k)$$

$$+2\hat{y}_2(k|k)u(k)P_{13}(k|k) + \hat{y}_1(k|k)^2 P_{22}(k|k) + 2\hat{y}_1(k|k)u(k)P_{23}(k|k)$$

$$+u(k)^2 P_{33}(k|k) + P_{21}(k|k)^2 + P_{22}(k|k)P_{11}(k|k) + q_1$$

EKF for Parameter Estimation — Summary

The EKF can be used to estimate simultaneously

• the base state and

• the unknown parameters

of a system.

This is accomplished by stacking them into an augmented state and carrying out the series expansions of the EKF for this augmented state.

Since the EKF is a suboptimal technique, significant care has to be exercised to avoid filter inconsistencies. The filter has to be tuned with artificial process noise so that its estimation errors are commensurate with the calculated variances.