TTK4250

Week 5

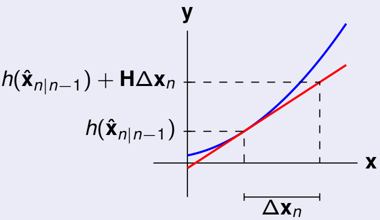
Multiple models, Gaussian mixtures and single-target tracking

Edmund Førland Brekke

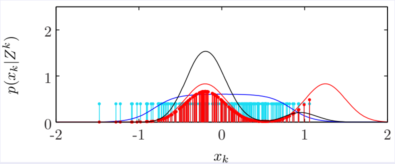
18. September 2023

Recap from last week

The EKF



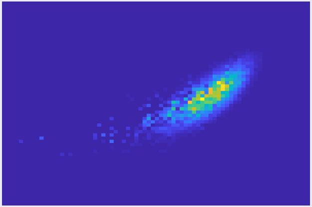
Particle filters



CV model and CT model

- CV model: A target that is expected to continue with its current velocity.
- CT model: A target whose velocity is expected to change according to a coordinated turn.

Bearing-only tracking



Selection of the proposal density

To do better than the SIR filter we can use information from the latest measurement.

The optimal importance density

For particle i at time step k, the best we can do is to sample from the posterior density of $\mathbf{x}_k$ conditional on $\mathbf{x}_{k-1}^i$ and $\mathbf{z}_k$:

$$q(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i, \mathbf{z}_k)_{\text{opt}} = p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i, \mathbf{z}_k) = \frac{p(\mathbf{z}_k \mid \mathbf{x}_{k-1}^i, \mathbf{x}_k)\, p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i)}{p(\mathbf{z}_k \mid \mathbf{x}_{k-1}^i)}$$

- The weight update is now given by

  $$w_k^i \propto w_{k-1}^i p(\mathbf{z}_k \mid \mathbf{x}_{k-1}^i) = w_{k-1}^i \int p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i)\, \mathrm{d}\mathbf{x}_k$$

  Evaluation of this integral may be very difficult.
- It may also be very difficult to sample from $p(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i, \mathbf{z}_k)$.

Selection of the proposal density

Optimal density for nonlinear dynamics and Gaussian noise

Let the state space model be $\mathbf{x}_k = \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1}$, $\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k$, and let $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ be mutually independent zero-mean white Gaussian sequences with covariances $\mathbf{Q}_k$ and $\mathbf{R}_k$.

In this case

$$\begin{aligned} q(\mathbf{x}_k \mid \mathbf{x}_{k-1}^i, \mathbf{z}_k)_{\text{opt}} &= \mathcal{N}(\mathbf{x}_k \,;\, \mathbf{a}_k, \Sigma_k) \\ p(\mathbf{z}_k \mid \mathbf{x}_{k-1}) &= \mathcal{N}(\mathbf{z}_k \,;\, \mathbf{b}_k, \mathbf{S}_k) \end{aligned}$$

where

$$\begin{aligned} \mathbf{a}_k &= \mathbf{f}_{k-1}(\mathbf{x}_{k-1}) + \Sigma_k \mathbf{H}_k^{\mathsf{T}} \mathbf{R}_k^{-1} (\mathbf{z}_k - \mathbf{b}_k) \\ \Sigma_k &= \mathbf{Q}_k - \mathbf{Q}_k \mathbf{H}_k^{\mathsf{T}} \mathbf{S}_k^{-1} \mathbf{H}_k \mathbf{Q}_k \\ \mathbf{S}_k &= \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^{\mathsf{T}} + \mathbf{R}_k \\ \mathbf{b}_k &= \mathbf{H}_k \mathbf{f}(\mathbf{x}_{k-1}). \end{aligned}$$

Sampling from a multivariate Gaussian is easy:

```
xCloudNew = a(z,xCloud) + chol(Sigma)'*randn(n,m)
```

Selection of the proposal density

Local linearization

- Approximate $q(\mathbf{x}_k \mid \mathbf{x}^i_{k-1}, \mathbf{z}_k)_{\text{opt}}$ by means of the output from an EKF/UKF:

$$q(\mathbf{x}^i_k \mid \mathbf{x}^i_{k-1}, \mathbf{z}_k) = \mathcal{N}(\mathbf{x}^i_k \,;\, \hat{\mathbf{x}}^i_k, \hat{\mathbf{P}}^i_k)$$

- For each particle i we run an EKF/UKF: $(\mathbf{x}^i_{k-1}, \hat{\mathbf{P}}^i_{k-1}) \longrightarrow (\bar{\mathbf{x}}^i_k, \bar{\mathbf{P}}^i_k) \longrightarrow (\hat{\mathbf{x}}^i_k, \hat{\mathbf{P}}^i_k)$.
- Then we sample the new particles,
- and calculate their weights according to

$$\tilde{w}^i_k = w^i_{k-1} \frac{p(\mathbf{z}_k \mid \mathbf{x}^i_k) p(\mathbf{x}^i_k \mid \mathbf{x}^i_{k-1})}{\mathcal{N}(\mathbf{x}^i_k \,;\, \hat{\mathbf{x}}^i_k, \hat{\mathbf{P}}^i_k)}$$

- and finally we normalize and resample. The resampled particles inherit the covariances of their parents.

Improving the sample diversity

Resampling can cause particle depletion: Only replicates of a few good particles take over the entire particle cloud. Subsequent steps of importance sampling may not contain sufficient noise to counteract the collapse.

Solution: The regularized particle filter (RPF).

- Before resampling, the RPF calculates the empirical covariance $\mathbf{S}_k$ of the particle cloud, and its Cholesky factorization $\mathbf{D}_k$.
- After resampling, the RPF regularizes the particle cloud according to

$$\mathbf{x}_k^{i*} = \mathbf{x}_k^i + h_{\text{opt}} \mathbf{D}_k \epsilon^i$$

  where
  - ϵ^i is drawn from a scaled kernel density $K_h(\mathbf{x}) = \frac{1}{h^{n_x}} K\left(\frac{\mathbf{x}}{h}\right)$ with bandwidth h.
  - h_{opt} is the optimal bandwidth value.

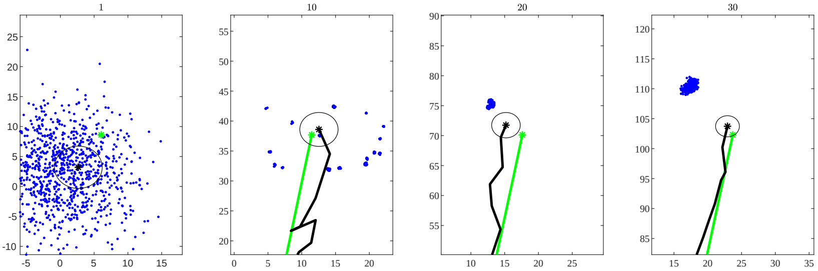
See the textbook Section 5.2.6 for further details.

To ensure convergence towards true posterior as $N \to \infty$ the regularization step $\mathbf{x}_k^{i*} = \mathbf{x}_k^i + h_{\text{opt}} \mathbf{D}_k \epsilon^i$ can be accepted or rejected according to the Metropolis-Hastings algorithm.[1]

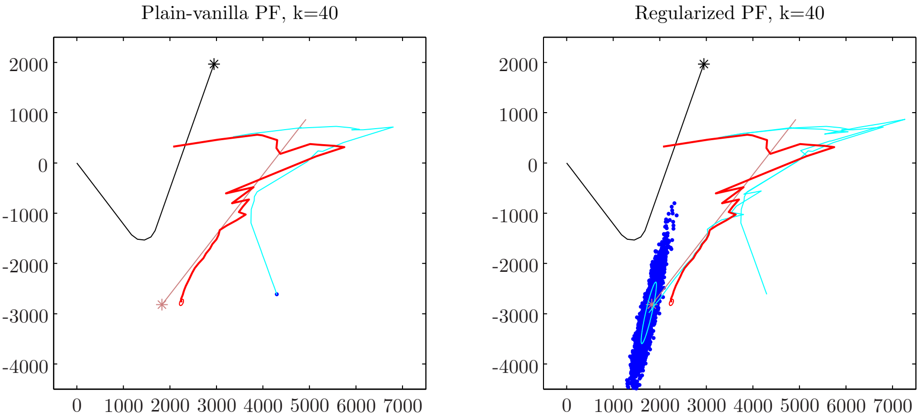
[1] Ristic et al. (2004) "Beyond the Kalman Filter: Particle Filters for Tracking Applications", Artech House.

Example of particle depletion

- We simulate a simple scenario with the CV model and position measurements.
- We use a fairly low process noise $\sigma_a = 0.05$ and typical GPS-resolution measurement noise $\sigma_z = 5$.
- Particle depletion causes the particles to form compact clusters, and this leads to divergence.



Plain-vanilla vs regularized PF for bearing-only



- Expectation and covariance of EKF (in modified coordinates) displayed in red.
- Particles, expectation and covariance of particle filter displayed in blue colors.

Representing the output from a particle filter

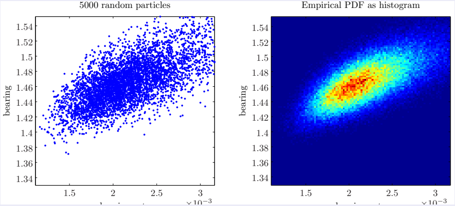
Empirical moments.

$$E[\mathbf{x}_k|\mathbf{z}_{1:k}] \approx \sum_{i=1}^{N} w_k^i \mathbf{x}_k^i \qquad \mathrm{Cov}[\mathbf{x}_k|\mathbf{z}_{1:k}] \approx \sum_{i=1}^{N} w_k^i (\mathbf{x}_k^i - \boldsymbol{\mu}_k)(\mathbf{x}_k^i - \boldsymbol{\mu}_k)^\mathsf{T}$$

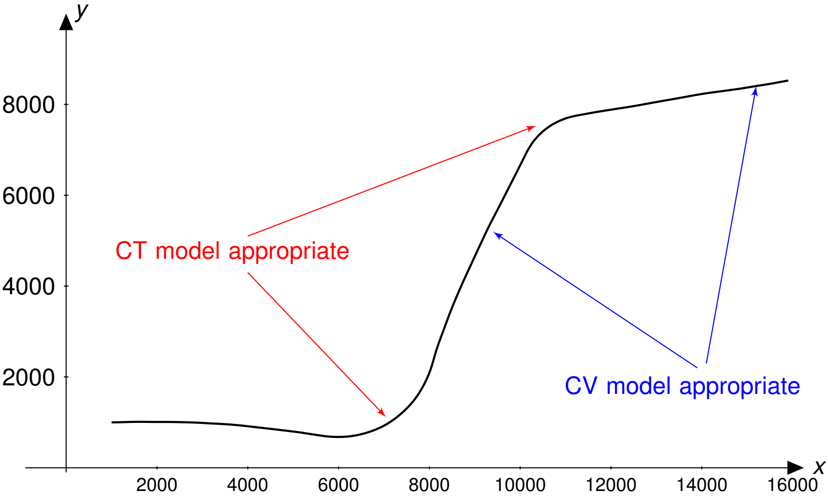
MAP estimate.

- A dubious approach is to pick the particle with the greatest weight.
- A more refined approximation of the MAP-estimator was proposed in Driessen & Boers (2008): "MAP Estimation in Particle Filter Tracking".

Empirical PDF on grid.



Multiple models: The air traffic control (ATC) scenario



On/off-phenomena such as maneuvers cannot be well modeled by white noise alone.

Why study multiple model filtering in TTK4250?

- To become familiar the kind of notation and manipulations that are common in scientific sensor fusion literature.
- As an introduction to core concepts in target tracking (continuous-discrete uncertainty, Gaussian mixtures, etc).
- Because a core challenge in autonomous vehicle systems is to estimate the course and speed of other vehicles with sufficient accuracy.

Goal of this lecture.

We will derive the Interacting Multiple Model (IMM) algorithm used in maneuvering target tracking, look at how to implement it and investigate its performance.

Multiple models framework

- We now have M possible instances of the process and measurement models, indexed by $s = 1, \dots, M$.

$$\begin{aligned} \mathbf{x}_k =& \mathbf{f}^{(s)}(\mathbf{x}_{k-1}) + \mathbf{v}_k, \;\; \mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}^{(s)}) \\ \mathbf{z}_k =& \mathbf{h}^{(s)}(\mathbf{x}_k) + \mathbf{w}_k, \;\; \mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}^{(s)}). \end{aligned}$$

- At a given time, the system is in one of these M **modes**. The system changes mode according to a Markov process with transition probabilities

$$\begin{aligned} \pi^{ij} &= \Pr\{\text{The target goes from model } j \text{ at time } k-1 \text{ to model } i \text{ at time } k\} \\ &= \Pr\{s_k = j \mid s_{k-1} = i\}. \end{aligned}$$

  We say that the modes constitute a **Hidden Markov Model (HMM)**.

To evaluate the posterior of $\mathbf{x}_k$ we both have to find the posterior of $\mathbf{x}_k$ given a particular mode[2] or mode-sequence[3], and we have to calculate probabilities of the different modes or mode-sequences.

[2] In a practical (suboptimal) method such as the IMM method.

[3] In theoretically optimal method.

Generic procedure to solving Bayesian filtering problems

The models involved

We must specify the transition model (Markov model, process model) and the measurement model (likelihood).

The previous posterior

The generic form of the posterior density that we expect to have at the beginning of the cycle.

The prediction

The density that we get after combining the prediction and the transition model in Chapman-Kolmogorov.

The posterior (measurement update)

The density that we get after combining the prediction and the likelihood in Bayes.

We aim to have the previous posterior and the new posterior of the same generic form (“conjugate prior”). Often, approximations are needed to achieve this.

The hybrid system: Models

The transition model

We assume that the transition model for the hybrid state can be written

$$p(\mathbf{y}_k \,|\, \mathbf{y}_{k-1}) = p(\mathbf{x}_k \,|\, \mathbf{x}_{k-1}, s_k)\Pr\{s_k \,|\, s_{k-1}\}$$

To make notation simpler, we also denote the transition probabilities by

$$\Pr\{s_k \,|\, s_{k-1}\} = \pi^{s_{k-1}s_k}$$

Prediction of the mode probabilities can be performed by means of matrix multiplication where the π^{ij}'s are organized in a matrix and the mode probabilities before and after prediction are organized in two vectors. See (6.5)-(6.6) in the textbook for details.

The measurement model

We assume that the measurement model for the hybrid state can be written

$$p(\mathbf{z}_k \,|\, \mathbf{y}_k) = p(\mathbf{z}_k \,|\, \mathbf{x}_k, s_k)$$

The previous posterior

We assume that our knowledge about the hybrid state vector $\mathbf{y}_{k-1}$ at the end of estimation cycle $k-1$ can be expressed as

$$p(\mathbf{y}_{k-1} \,|\, \mathbf{z}_{1:k-1}) = \Pr\{s_{k-1} \,|\, \mathbf{z}_{1:k-1}\} p(\mathbf{x}_{k-1} \,|\, s_{k-1}, \mathbf{z}_{1:k-1})$$

We can express the mode probabilities with the simpler notation

$$\Pr\{s_{k-1} \,|\, \mathbf{z}_{1:k-1}\} = \mu_{k-1}^{s_{k-1}}.$$

What is $p(\mathbf{x}_{k-1} \,|\, s_{k-1}, \mathbf{z}_{1:k-1})$, really?

- It could be a Gaussian, **and that would be really nice.**
- If the problems is nonlinear/non-Gaussian, it is probably not a Gaussian.
- Even if $p(\mathbf{x}_n \,|\, \mathbf{x}_{n-1}, s_n)$ and $p(\mathbf{z}_n \,|\, \mathbf{x}_n, s_n)$ are Gaussian-linear for all $n < k-1$, it is clear that $p(\mathbf{x}_{k-1} \,|\, s_{k-1}, \mathbf{z}_{1:k-1})$ must involve marginalization over s_{k-2}, s_{k-3}, etc. These marginalizations will inevitably destroy Gaussianity.

Nevertheless, let us not worry in advance, and for now we just assume that we have some exact or approximative expression for $p(\mathbf{x}_{k-1} \,|\, s_{k-1}, \mathbf{z}_{1:k-1})$.

Prediction in a hybrid system

The hybrid Chapman-Kolmogorov integral

$$p(\mathbf{y}_k \mid \mathbf{z}_{1:k-1}) = \sum_{s_{k-1}} \int p(\mathbf{x}_k, s_k \mid \mathbf{x}_{k-1}, s_{k-1}) p(\mathbf{x}_{k-1}, s_{k-1} \mid \mathbf{z}_{1:k-1}) \,\mathrm{d}\mathbf{x}_{k-1}.$$

can be expressed as

$$\underbrace{\left(\sum_{s_{k-1}} \pi^{s_{k-1} s_k} \mu_{k-1}^{s_{k-1}} \right)}_{\mu_{k|k-1}^{s_k}} \underbrace{\int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, s_k) \overbrace{\frac{\sum_{s_{k-1}} \pi^{s_{k-1} s_k} \mu_{k-1}^{s_{k-1}} p(\mathbf{x}_{k-1} \mid s_{k-1}, \mathbf{z}_{1:k-1})}{\sum_{s_{k-1}} \pi^{s_{k-1} s_k} \mu_{k-1}^{s_{k-1}}}}^{p(\mathbf{x}_{k-1} \mid s_k, \mathbf{z}_{1:k-1})} \mathrm{d}\mathbf{x}_{k-1}}_{p(\mathbf{x}_k \mid s_k, \mathbf{z}_{1:k-1})} .$$

- $p(\mathbf{x}_{k-1} \mid s_k, \mathbf{z}_{1:k-1})$ is the pdf of the **previous state** conditional on **current mode**.
- $\mu_{k|k-1}^{s_k}$ is the predicted mode probabilities.
- $p(\mathbf{x}_k \mid s_k, \mathbf{z}_{1:k-1})$ is the predicted pdf of the state conditional on the current mode.

Prelude to mixtures

- Assume that $p(\mathbf{x}_{k-1} \mid s_{k-1}, \mathbf{z}_{1:k-1})$ is a Gaussian.
- Then the pdf

$$p(\mathbf{x}_{k-1} \mid s_k,\, \mathbf{z}_{1:k-1}) = \frac{\sum_{s_{k-1}} \pi^{s_{k-1} s_k} \mu_{k-1}^{s_{k-1}} p(\mathbf{x}_{k-1} \mid s_{k-1}, \mathbf{z}_{1:k-1})}{\sum_{s_{k-1}} \pi^{s_{k-1} s_k} \mu_{k-1}^{s_{k-1}}}$$

  is a weighted sum of Gaussians.
- If we could approximate it by a Gaussian then we would get rid of s_{k-1} and both the prediction and the posterior would also be Gaussian[4] conditional on s_k.

  That is, we could also treat $p(\mathbf{x}_k \mid s_k, \mathbf{z}_{1:k})$ as Gaussian, and so forth.

[4] Subject to standard Gaussian-linear model assumptions, of course.

Mixtures and Gaussian mixtures

- A mixture is a pdf that is a sum of pdfs $p^i(\mathbf{x})$, i.e.,

$$f(\mathbf{x}) = \sum_{i=1}^{M} w^i p^i(\mathbf{x}) \quad \text{where} \quad \sum_{i=1}^{M} w^i = 1 \quad \text{and} \quad w^i \geq 0 \quad \text{for all} \quad i.$$

- A Gaussian mixture is a pdf of the form

$$f(\mathbf{x}) = \sum_{i=1}^{M} w^i \mathcal{N}(\mathbf{x}\,;\, \boldsymbol{\mu}^i, \mathbf{P}^i) \quad \text{where} \quad \sum_{i=1}^{M} w^i = 1 \quad \text{and} \quad w^i \geq 0 \quad \text{for all} \quad i.$$

Why are mixtures important?

- Mixtures arise whenever we have to consider multiple hypotheses.
  - Does the airplane follow coordinated turn or does it follow straight-line motion?
  - Which one of m measurements come from a landmark?
- Mixtures allow us to calculate posterior pdf's of the state vector conditional on such hypotheses, and we can **hedge** on the different hypotheses with appropriate probabilities.

Notice that we can also have a mixture where each $p^i(\mathbf{x})$ is a mixture itself.

Mixture reduction

Ever (exponentially) increasing numbers of mixture components are not computationally tractable, and an important task in sensor fusion is to reduce the number of mixture components without losing essential information.

We can approximate the Gaussian mixture $\sum_{i=1}^{M} w^i \mathcal{N}(\mathbf{x}\,;\,\boldsymbol{\mu}^i, \mathbf{P}^i)$ with a Gaussian with the same expectation and covariance.

Moment-based mixture reduction formulas

The expectation of the mixture is

$$\bar{\boldsymbol{\mu}} = \sum_{i=1}^{M} w^i \boldsymbol{\mu}^i$$

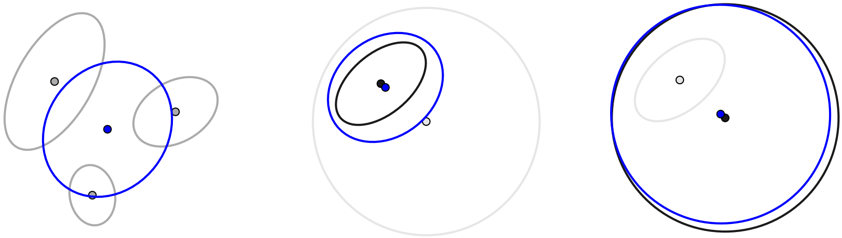
and its covariance is

$$\bar{\mathbf{P}} = \sum_{i=1}^{M} w^i \mathbf{P}^i + \tilde{\mathbf{P}}$$

where the so-called spread-of-the-innovovations term is given by

$$\tilde{\mathbf{P}} = \sum_{i=1}^{M} w^i (\boldsymbol{\mu}^i - \bar{\boldsymbol{\mu}})(\boldsymbol{\mu}^i - \bar{\boldsymbol{\mu}})^\mathsf{T} = \sum_{i=1}^{M} w^i \boldsymbol{\mu}^i (\boldsymbol{\mu}^i)^\mathsf{T} - \bar{\boldsymbol{\mu}}\bar{\boldsymbol{\mu}}^\mathsf{T}.$$

Examples of Gaussian mixtures



The reduction to a single Gaussian may or may not lead to elimination of essential information.

Prediction in a hybrid system continued

Assume that $p(\mathbf{x}_{k-1}|s_{k-1}, \mathbf{z}_{1_k-1}) = \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}^{s_{k-1}}, \mathbf{P}_{k-1}^{s_{k-1}})$.

By means of mixture reduction we can approximate

$$p(\mathbf{x}_{k-1} \mid s_k, \mathbf{z}_{1:k-1}) \approx \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{x}_{k-1}^{0,s_k}, \mathbf{P}_{k-1}^{0,s_k})$$

where

$$\begin{aligned}
\hat{\mathbf{x}}_{k-1}^{0,s_k} &= \sum_{s_{k-1}} \mu_{s_{k-1}|s_k} \hat{\mathbf{x}}_{k-1}^{s_{k-1}} \\
\mathbf{P}_{k-1}^{0,s_k} &= \sum_{s_{k-1}} \mu_{s_{k-1}|s_k} \mathbf{P}_{k-1}^{s_{k-1}} \\
\mu_{s_{k-1}|s_k} &\propto \frac{\pi^{s_k s_{k-1}} \mu_{k-1}^{s_{k-1}}}{\sum_{s_{k-1}} \pi^{s_k s_{k-1}} \mu_{k-1}^{s_{k-1}}}.
\end{aligned}$$

For each current mode s_k we have

$$p(\mathbf{x}_k \mid s_k, \mathbf{z}_{1:k-1}) \approx \int p(\mathbf{x}_k|s_k, \mathbf{x}_{k-1}) \mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{x}_{k-1}^{0,s_k}, \mathbf{P}_{k-1}^{0,s_k})\, \mathrm{d}\mathbf{x}_{k-1}$$

which is solved by standard KF/EKF prediction if the motion model $p(\mathbf{x}_k|s_k, \mathbf{x}_{k-1})$ is sufficiently Gaussian-linear.

The posterior in a hybrid system

The hybrid measurement update is

$$p(\mathbf{y}_k \mid \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k \mid \mathbf{x}_k, s_k)\mu^{s_k}_{k|k-1}\, p(\mathbf{x}_k \mid s_k,\, \mathbf{z}_{1:k-1})}{\sum\limits_{s_k} \int p(\mathbf{z}_k \mid \mathbf{x}_k, s_k)\mu^{s_k}_{k|k-1}\, p(\mathbf{x}_k \mid s_k,\, \mathbf{z}_{1:k-1})\mathrm{d}\mathbf{x}_k}$$

If we introduce the notation

$$l^{s_k}_k = \int p(\mathbf{z}_k \mid \mathbf{x}_k, s_k)p(\mathbf{x}_k \mid s_k,\, \mathbf{z}_{1:k-1})\mathrm{d}\mathbf{x}_k$$

we can express this as

$$p(\mathbf{y}_k \mid \mathbf{z}_{1:k}) = \underbrace{\frac{\mu^{s_k}_{k|k-1}\, l^{s_k}_k}{\sum\limits_{s_k} \mu^{s_k}_{k|k-1}\, l^{s_k}_k}}_{\mu^{s_k}_k} \underbrace{\frac{p(\mathbf{z}_k \mid \mathbf{x}_k, s_k)p(\mathbf{x}_k \mid s_k,\, \mathbf{z}_{1:k-1})}{l^{s_k}_k}}_{p^s_k(\mathbf{x})}$$

which is of the same form as the previous posterior.

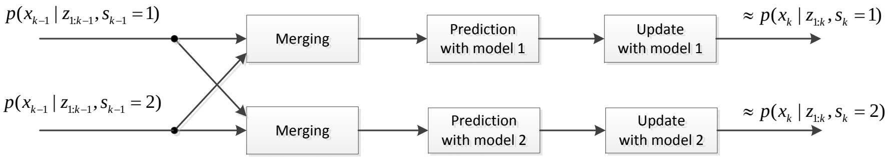
Beware

Recall that $p(\mathbf{x}_{k-1} \mid s_k,\, \mathbf{z}_{1:k-1})$ contained a sum over s_{k-1} which is inherited by $p(\mathbf{x}_k \mid s_k,\, \mathbf{z}_{1:k-1})$. We used mixture reduction to get rid of the sum, but this is an approximation.

The IMM strategy for multiple model filtering

Assume the previous posterior to be a Gaussian mixture with M components

$$p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) = \sum_{s_{k-1}} \Pr\{s_{k-1} \mid \mathbf{z}_{1:k-1}\} \mathcal{N}\left(\mathbf{x}_{k-1}\,;\, \hat{\mathbf{x}}_{k-1}^{s_{k-1}}, \mathbf{P}_{k-1}^{s_{k-1}}\right).$$



- In this case we do the mixture reduction **before** the filtering cycle.
- Thus, the filtering is done for a Gaussian mixture with M components.

The IMM method directly follows from the hybrid formulation if we approximate $p(\mathbf{x}_{k-1}|s_k, \mathbf{z}_{1:k-1})$ by a single Gaussian.

Workflow of the IMM method

The entire estimation cycle of the IMM method consists of the following 4 steps

1. Calculation of mixing probabilities.
2. Mixing.
3. Mode matched filtering.
4. Update of mode probabilities.

Step 1: Calculation of mixing probabilities

- The mixing probabilities decide how much the previous Gaussian for model s_{k-1} should influence the filtering input for model s_k.

$\Rightarrow$ They are the weights in a Gaussian mixture ranging over s_{k-1} for each s_k.

$$\begin{aligned}\mu_{s_{k-1}|s_k} &= \Pr\{s_{k-1} \mid s_k, \mathbf{z}_{1:k-1}\} \\ &= \frac{\Pr\{s_k \mid s_{k-1}, \mathbf{z}_{1:k-1}\}\Pr\{s_{k-1} \mid \mathbf{z}_{1:k-1}\}}{\Pr\{s_k \mid \mathbf{z}_{1:k-1}\}} \\ &\propto \pi^{s_{k-1}s_k}\mu_{k-1}^{s_{k-1}}.\end{aligned}$$

Workflow of the IMM method

Step 2: Mixing

For each current model s_k we want to determine a “prior” state estimate $\hat{\mathbf{x}}_{k-1}^{0,s_k}$ and a corresponding covariance matrix $\mathbf{P}_{k-1}^{0,s_k}$ so that

$$\mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}^{0,s_k},\mathbf{P}_{k-1}^{0,s_k})$$

is a good approximation of

$$p(\mathbf{x}_{k-1}\,|s_k,\mathbf{z}_{1:k-1}) = \sum_{s_{k-1}} \mu_{s_{k-1}|s_k}\,\mathcal{N}(\mathbf{x}_{k-1}\,;\,\hat{\mathbf{x}}_{k-1}^{s_{k-1}},\mathbf{P}_{k-1}^{s_{k-1}}).$$

The standard mixture reduction formulas yield

$$\begin{aligned}
\hat{\mathbf{x}}_{k-1}^{0,s_k} &= \sum_{s_{k-1}} \mu_{s_{k-1}|s_k}\,\hat{\mathbf{x}}_{k-1}^{s_{k-1}} \\
\mathbf{P}_{k-1}^{0,s_k} &= \sum_{s_{k-1}} \mu_{s_{k-1}|s_k} \left\{ \mathbf{P}_{k-1}^{s_{k-1}} + (\hat{\mathbf{x}}_{k-1}^{0,s_k} - \hat{\mathbf{x}}_{k-1}^{s_{k-1}})(\hat{\mathbf{x}}_{k-1}^{0,s_k} - \hat{\mathbf{x}}_{k-1}^{s_{k-1}})^{\mathsf{T}} \right\}.
\end{aligned}$$

Workflow of the IMM method

Step 3: Mode-matched filtering

This is just an EKF step for each model s_k:

$$[\hat{\mathbf{x}}_k^{s_k}, \mathbf{P}_k^{s_k}, \hat{\mathbf{x}}_{k|k-1}^{s_k}, \mathbf{P}_{k|k-1}^{s_k}, \mathbf{S}_k^{s_k}] = \mathrm{EKF}(\hat{\mathbf{x}}_{k-1}^{0,s_k}, \mathbf{P}_{k-1}^{0,s_k}, \mathbf{z}_k).$$

You could use a KF, UKF or any other Gaussian-based technique if appropriate.

Step 4: Update of mode probabilities

The posterior probability of the mode s_k is according to Bayes' rule given by

$$\mu_k^{s_k} = \Pr\{s_k \mid \mathbf{z}_{1:k}\} = \frac{l_k^{s_k} \mu_{k|k-1}^{s_k}}{\sum\limits_{s_k} l_k^{s_k} \mu_{k|k-1}^{s_k}}$$

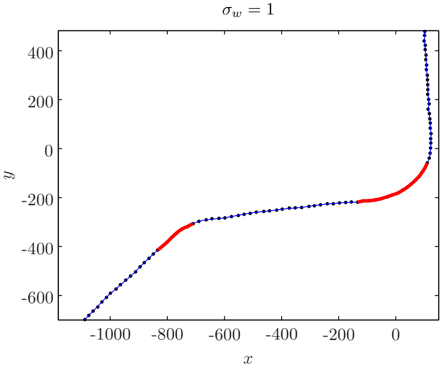
where

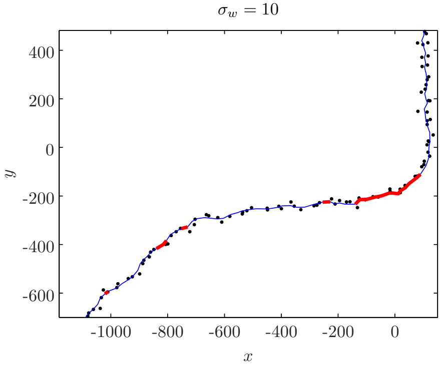
$$l_k^{s_k} = \int p(\mathbf{z}_k \mid \mathbf{x}_k, s_k) p(\mathbf{x}_k \mid s_k, \mathbf{z}_{1:k-1}) \, \mathrm{d}\mathbf{x}_k = \mathcal{N}\left(\mathbf{z}_k \,; \mathbf{h}^{(s_k)}\left(\hat{\mathbf{x}}_{k|k-1}^{s_k}\right), \mathbf{S}_k^{s_k}\right).$$

When is IMM useful?

The maneuvering index $\lambda = \dfrac{\sigma_v T^2}{\sigma_w}$ should be larger than 0.5.

- σ_v is the standard deviation of the process noise needed to explain the maneuver.
- σ_w is the standard deviation of the measurement noise.





Implementation issues

Saturation

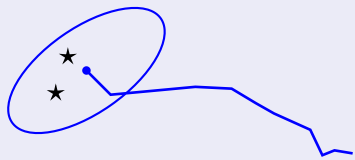
- I recommend including some max/min-rules to prevent mode probabilities from reaching zero or one.
- If they reach zero or one, the Bayes update will struggle bringing them back up when the model regime changes.

Models of different dimensions

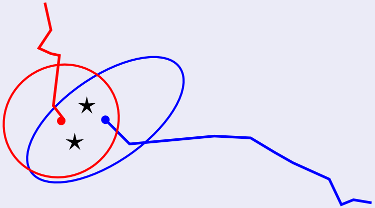
- In, e.g., a CV-CT scenario the state vectors of the two models have different dimension.
- Conceptually this is fine because only the measurement likelihood $p(\mathbf{z}_k \mid s_k, \mathbf{z}_{1:k-1})$ is used to calculate the mode probabilities.
- In practice, the easiest solution is to extend all the state vectors to the largest dimension. E.g., the CV mode gets a 5-dimensional state vector where the last state always is zero.
- Some care must then be exercised to ensure that the covariance matrices always are positive definite.

The road ahead

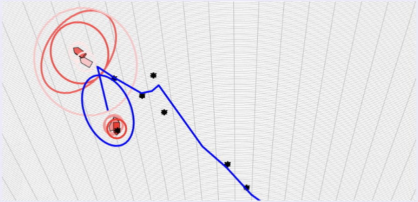
Single target tracking



Multiple target tracking



Maneuvering target tracking



Track initialization