Target Tracking: Lecture 4
Maneuvering Target Tracking

Umut Orguner
umut@metu.edu.tr
room: EZ-12 tel: 4425

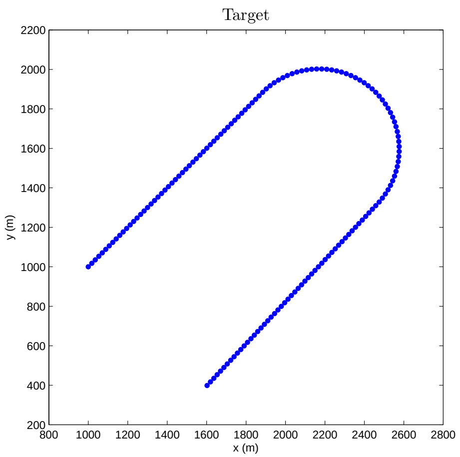
Department of Electrical & Electronics Engineering
Middle East Technical University
Ankara, Turkey

Lecture Outline

- Maneuver Detection
- Detection Based Methods
  - Adjustable level process noise
  - Variable state dimension
  - Input estimation
- Multiple Model Approaches
  - Non-switching multiple models
  - Switching multiple models
    - Generalized pseudo Bayesian (GPB) methods
    - Interacting multiple model (IMM) algorithm

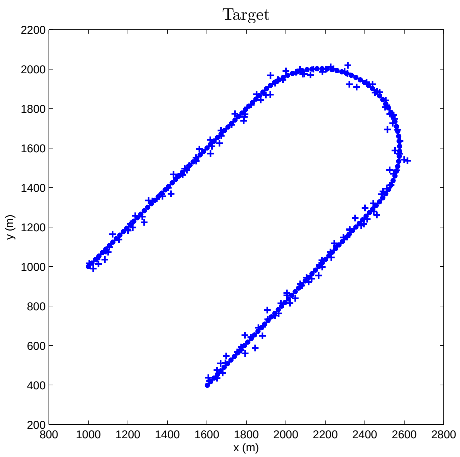
Maneuver Illustration

- A simple illustration of the maneuver problem with simplistic parameters.
- $P_D = 1.$
- $P_G = 1$
- $P_{FA} = 0$
- KF with CV model
- We try different process noise standard deviations $\sigma_a = 0.1, 1, 10 \text{m/s}^2$.

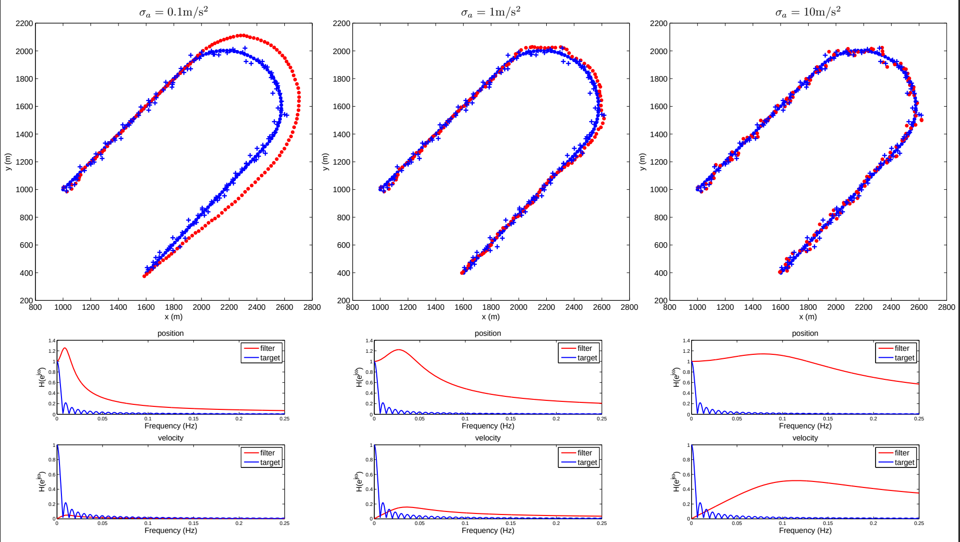


Maneuver Illustration

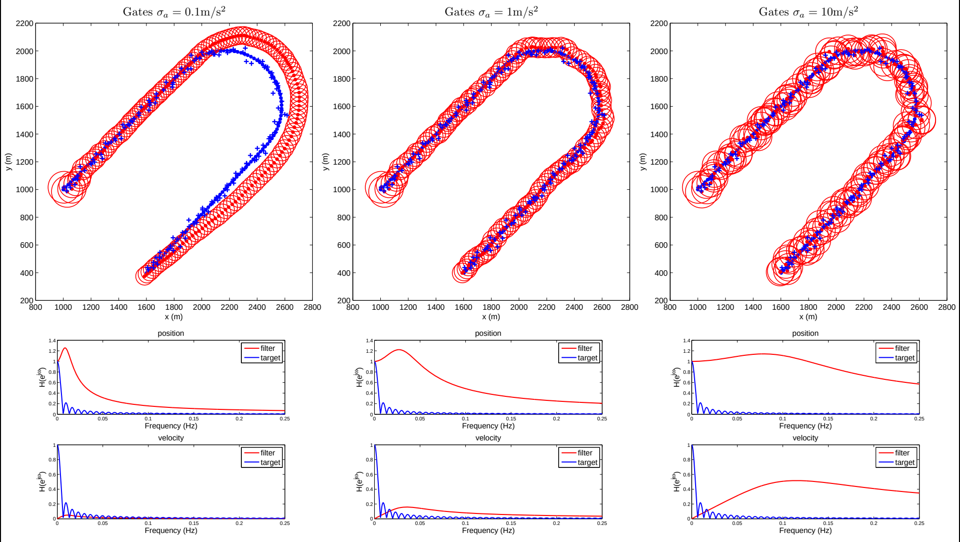
- A simple illustration of the maneuver problem with simplistic parameters.
- $P_D = 1.$
- $P_G = 1$
- $P_{FA} = 0$
- KF with CV model
- We try different process noise standard deviations $\sigma_a = 0.1, 1, 10 \text{m/s}^2$.



Maneuver Illustration



Maneuver Illustration



Maneuvers

- Maneuvers are the model mismatch problem in target tracking.
- Using a high order kinematic model that allows versatile tracking all the time is not a solution in the case where data origin uncertainty is present.
  This can instead make the gates unnecessarily large and makes the tracker sensitive to clutter.
- Hence maneuvers should be detected and compensated.
- A maneuver should be detected both when the target switches to a higher order model than we use in our KF, and when it switches to a lower order model than we use in the KF.

Maneuver Detection: Low-Pass $\rightarrow$ High-Pass

- Normalized innovation square again comes into the picture.

$$\epsilon_{\tilde{y}_k} = \tilde{y}_k^T S_{k|k-1}^{-1} \tilde{y}_k$$

- We know that $\epsilon_{\tilde{y}_k} \sim \chi^2_{n_y}$.
- This is also the gating statistics. So we should check this quantity in a window to avoid false alarms.
- Use a sliding window or a recursive forgetting.

$$\epsilon_k^s = \sum_{i=k-N+1}^{k} \epsilon_{\tilde{y}_i} \qquad \text{or} \qquad \epsilon_k^r = \alpha \epsilon_{k-1}^r + \epsilon_{\tilde{y}_k}$$

where $\alpha < 1$.

Maneuver Detection: Low-Pass → High-Pass

- Use one of the statistics

$$\epsilon_k^s = \sum_{i=k-N+1}^{k} \epsilon_{\tilde{y}_i} \qquad \text{or} \qquad \epsilon_k^r = \alpha \epsilon_{k-1}^r + \epsilon_{\tilde{y}_k}$$

- In the case of perfect model match, we have

$$\epsilon_k^s \sim \chi^2_{N n_y} \qquad \text{and} \qquad \epsilon_k^r \sim \chi^2_{\frac{1}{1-\alpha} n_y}$$

  where the second distribution is an approximation at the steady state (effective window length $\approx \frac{1}{1-\alpha}$).
- A maneuver is declared when the maneuver statistics ϵ_k exceeds a threshold $\epsilon_{\max}$.
- The threshold $\epsilon_{\max}$ is adjusted such that in the case of no maneuver

$$P(\epsilon_k \leq \epsilon_{\max}) = 1 - \underbrace{P_{FA}^{maneuver}}_{\ll 1}$$

Maneuver Detection: Low-Pass → High-Pass

- During a low-pass filter to high-pass filter transition detected from an ϵ_k that is obtained by summing $\epsilon_{\tilde{y}_i}$ over a window of length N (or effective window length $\frac{1}{1-\alpha}$), there accumulates considerable amount of error in the estimates.
- These should be compensated when such a detection happens.
- Generally last estimates in the (effective) window are recalculated.
- For this purpose, some previous history of estimates and measurements are kept in memory.

Maneuver Detection: High-Pass → Low-Pass

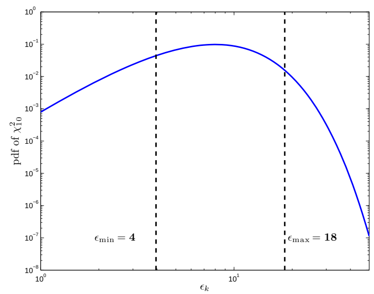
- The decision in the reverse direction (from high-pass to low-pass) can be given with the same statistics if the statistics ϵ_k gets lower than a threshold $\epsilon_{\min}$.
- The threshold $\epsilon_{\min}$ is adjusted such that in the case of correct model

$$P(\epsilon_k \leq \epsilon_{\min}) = P_{miss}^{maneuver} \ll 1$$

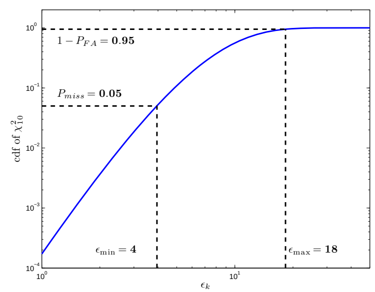
Maneuver Detection: High-Pass → Low-Pass

- The decision in the reverse direction (from high-pass to low-pass) can be given with the same statistics if the statistics ϵ_k gets lower than a threshold $\epsilon_{\min}$.
- The threshold $\epsilon_{\min}$ is adjusted such that in the case of correct model

$$P(\epsilon_k \leq \epsilon_{\min}) = P_{miss}^{maneuver} \ll 1$$



chi2pdf(.,n)



chi2cdf(.,n)

Adjustable level process noise

- Kalman filters bandwidth depends on the noise covariances.
- Process noise covariance
  - Small → Low bandwidth
  - Big → High bandwidth
- Measurement noise covariance
  - Small → High bandwidth
  - Big → Low bandwidth
- The measurement noise covariance is generally selected to represent the sensor characteristics.
- Process noise covariance determines our belief on how smooth the target trajectory is and hence can be adjusted to account for maneuvers.

Continuous Process Noise Level Adjustment

- Innovation covariance is given in the standard Kalman filter as

$$S_{k|k-1} = C\left[AP_{k-1|k-1}A^T + \quad BQB^T\right]C^T + R$$

- Add a scaling factor ρ_k to Q.

Continuous Process Noise Level Adjustment

- Innovation covariance is given in the standard Kalman filter as

$$S_{k|k-1} = C\left[AP_{k-1|k-1}A^T + \rho_k BQB^T\right]C^T + R$$

- Add a scaling factor ρ_k to Q.

Continuous Process Noise Level Adjustment

- Innovation covariance is given in the standard Kalman filter as

$$S_{k|k-1} = C\left[AP_{k-1|k-1}A^T + \rho_k BQB^T\right]C^T + R$$

- Add a scaling factor ρ_k to Q.

CALPN algorithm

- Time $k = 0$: Initialize $\rho_0 = 1$.
- Time $k > 0$:
  - If $\epsilon_k > \epsilon_{\max}$, increase ρ_k such that $\epsilon_k \leq \epsilon_{\max}$ and recalculate previous estimates in the (effective) detection window.
  - If $\epsilon_k < \epsilon_{\min}$, decrease ρ_k such that $\epsilon_k \geq \epsilon_{\min}$.
  - Otherwise, keep $\rho_k = \rho_{k-1}$.

Detection Based Methods: ALPN

Discrete Process Noise Level Adjustment

- One can also select a predetermined number of process noise matrices $\{Q_1, Q_2 \ldots, Q_{n_Q}\}$ in increasing order and design a similar procedure.

DALPN algorithm

- Time $k = 0$: Initialize $i_0 = n$ where $1 \leq n \leq n_Q$ and Q_n represents a nominal process noise covariance.
- Time $k > 0$:
  - If $\epsilon_k > \epsilon_{\max}$, change $Q_{i_{k-1}}$ to $Q_{i_k} = Q_{i_{k-1}+1}$ (if $i_{k-1} < n_Q$).
  - If $\epsilon_k < \epsilon_{\min}$, change $Q_{i_{k-1}}$ to $Q_{i_k} = Q_{i_{k-1}-1}$ (if $i_{k-1} > 1$).
  - Otherwise, keep $Q_{i_k} = Q_{i_{k-1}}$.

Detection Based Methods: VSD

Variable State Dimension method: We can use

- (nearly) constant velocity (CV) (2nd order in 1D and 4th order in 2D)
- (nearly) constant acceleration models (CA) (3rd order in 1D and 6th order in 2D)

interchangeably with suitable detection rules.

VSD Main Idea

- Use CV model.
- Keep always a buffer with the last N measurements, state estimates, predictions and their covariances.
- Check ϵ_k ((effective) window size N).
- If $\epsilon_k > \epsilon_{\mathrm{max}}$
  - Go back to time $k - N + 1$.
  - Initialize a CA model.
  - Recalculate estimates for times $k - N + 1, \ldots, k$.
  - Continue with a CA model until some other condition is satisfied.

Detection Based Methods: VSD

CA Model Initialization:

- Suppose we are tracking with a CV model which has $x_k^{\mathrm{CV}} = \begin{bmatrix} \mathrm{p}_k & \mathrm{v}_k \end{bmatrix}^T$ using the measurements $y_k = \mathrm{p}_k + v_k$.
- Suppose $\epsilon_k > \epsilon_{\mathrm{max}}$, then
  - Define $x_k^{\mathrm{CA}} = \begin{bmatrix} \mathrm{p}_k & \mathrm{v}_k & \mathrm{a}_k \end{bmatrix}^T$.
  - Go back to time $k - N + 1$ where we have the estimate

$$\hat{x}^{\mathrm{CV}}_{k-N+1|k-N+1} = \begin{bmatrix} \hat{\mathrm{p}}^{\mathrm{CV}}_{k-N+1|k-N+1} & \hat{\mathrm{v}}^{\mathrm{CV}}_{k-N+1|k-N+1} \end{bmatrix}^T$$

  - Set

$$\hat{x}^{\mathrm{CA}}_{k-N+1|k-N+1} = \begin{bmatrix} y_{k-N+1} \\ \hat{\mathrm{v}}^{\mathrm{CV}}_{k-N+1|k-N+1} + T\hat{\mathrm{a}}^{\mathrm{CA}}_{k-N+1|k-N+1} \\ \hat{\mathrm{a}}^{\mathrm{CA}}_{k-N+1|k-N+1} \end{bmatrix}$$

  where

$$\hat{\mathrm{a}}^{\mathrm{CV}}_{k-N+1|k-N+1} \triangleq \frac{2}{T^2}\left(y_{k-N+1} - \hat{y}^{\mathrm{CV}}_{k-N+1|k-N}\right)$$

  - Initialization of $P^{\mathrm{CA}}_{k-N+1|k-N+1}$ is complicated and given in [Bar-Shalom (1982)].

Detection Based Methods: VSD

VSD algorithm

- Time $k = 0$: Start using CV model.
- Time $k > 0$:
  - If currently using CV model, calculate ϵ_k at each step.
    - If $\epsilon_k > \epsilon_{\max}$
      -Go back to the beginning of the detection window (which can have length N or $\frac{1}{1-\alpha}$) depending on whether you use ϵ_k^s or ϵ_k^r respectively.
      -Initialize CA model and recalculate the estimates in the window. Continue using CA model.
  - If currently using CA model, check acceleration estimate $\hat{a}_k$ and covariance P_{a_k}.
    - If $\hat{a}_k^T P_{a_k}^{-1} \hat{a}_k < \gamma_{\min}$
      -Initialize CV model from last position and velocity estimates of CA model. Continue using CV model.

Detection Based Methods: IE

Input estimation method: Suppose we consider two different state equations, one with input the other not.

$$x_k^1 = Ax_{k-1}^1 + Gw_k \quad \text{(actual model we use currently)}$$
$$x_k^2 = Ax_{k-1}^2 + Bu_k + Gw_k \quad \text{(hypothetical model with maneuver)}$$

where u_k is a deterministic input.

- We consider only the case where $u_i = u$ for $i = k-N+1, \ldots, k$ and $u_i = \mathbf{0}$ otherwise. This period will correspond to the maneuvering period.
- Suppose we currently run a KF for the 1st model and would like to check the hypothesis that actually the second model is true, i.e., that the target is maneuvering.

Detection Based Methods: IE

KF Equations for Model 1

- Prediction Update

$$\hat{x}_{k|k-1}^1 = A\hat{x}_{k-1|k-1}^1$$
$$P_{k|k-1}^1 = AP_{k-1|k-1}^1 A^T + GQG^T$$

- Measurement Update

$$\hat{x}_{k|k}^1 = \hat{x}_{k|k-1}^1 + K_k^1 \underbrace{(y_k - \hat{y}_{k|k-1}^1)}_{\tilde{y}_k^1}$$
$$P_{k|k}^1 = P_{k|k-1}^1 - K_k^1 S_{k|k-1}^1 \left(K_k^1\right)^T$$

$$\hat{y}_{k|k-1}^1 = C\hat{x}_{k|k-1}^1$$
$$S_{k|k-1}^1 = CP_{k|k-1}^1 C^T + R$$
$$K_k^1 = P_{k|k-1}^1 C^T (S_{k|k-1}^1)^{-1}$$

KF Equations for Model 2

- Prediction Update

$$\hat{x}_{k|k-1}^2 = A\hat{x}_{k-1|k-1}^2 + Bu_k$$
$$P_{k|k-1}^2 = AP_{k-1|k-1}^2 A^T + GQG^T$$

- Measurement Update

$$\hat{x}_{k|k}^2 = \hat{x}_{k|k-1}^2 + K_k^2 \underbrace{(y_k - \hat{y}_{k|k-1}^2)}_{\tilde{y}_k^2}$$
$$P_{k|k}^2 = P_{k|k-1}^2 - K_k^2 S_{k|k-1}^2 \left(K_k^2\right)^T$$

$$\hat{y}_{k|k-1}^2 = C\hat{x}_{k|k-1}^2$$
$$S_{k|k-1}^2 = CP_{k|k-1}^2 C^T + R$$
$$K_k^2 = P_{k|k-1}^2 C^T (S_{k|k-1}^2)^{-1}$$

Detection Based Methods: IE

- Kalman filter covariances are equal for the two models.

$$P_{k|k-1}^1 = P_{k|k-1}^2 \triangleq P_{k|k-1} \qquad S_{k|k-1}^1 = S_{k|k-1}^2 \triangleq S_{k|k-1}$$
$$P_{k|k}^1 = P_{k|k}^2 \triangleq P_{k|k} \qquad K_k^1 = K_k^2 \triangleq K_k$$

- Suppose, $\hat{x}_{k-N|k-N}^1 = \hat{x}_{k-N|k-N}^2$, then

$$\hat{x}_{k-N+1|k-N}^2 = \hat{x}_{k-N+1|k-N}^1 + Bu$$
$$\hat{x}_{k-N+1|k-N+1}^2 = \hat{x}_{k-N+1|k-N+1}^1 + (I - K_{k-N+1}C)Bu$$
$$\hat{x}_{k-N+2|k-N+1}^2 = \hat{x}_{k-N+2|k-N+1}^1 + A(I - K_{k-N+1}C)Bu + Bu$$
$$\vdots$$
$$\hat{x}_{k|k-1}^2 = \hat{x}_{k|k-1}^1 + \underbrace{\left[\sum_{i=0}^{N-1}\prod_{j=0}^{i-1}\left(A(I - K_{k-N+1+j}C)\right)\right] B}_{\triangleq F_k} u$$

Detection Based Methods: IE

- The innovations $\tilde{y}_k^1$ and $\tilde{y}_k^2$ corresponding to the KFs using 1st and 2nd models with the same measurement model can be related as

$$\underbrace{\tilde{y}_k^1}_{\substack{\text{innovations we}\\\text{currently calculate}}} = CF_k u + \underbrace{\tilde{y}_k^2}_{\substack{\text{hypothetical}\\\text{innovations}}}$$

where C is the measurement matrix and F_k can be calculated using the system model and Kalman filter gains corresponding to 1st model.

- We now stack the last N innovations and F_k matrices

$$\tilde{\mathbf{y}}_\mathbf{k} = \begin{bmatrix} \tilde{y}_k^T & \tilde{y}_{k-1}^T & \cdots & \tilde{y}_{k-N+1}^T \end{bmatrix}^T$$
$$\mathbf{F}_\mathbf{k} = \begin{bmatrix} F_k^T & F_{k-1}^T & \cdots & F_{k-N+1}^T \end{bmatrix}^T$$

Detection Based Methods: IE

We get the model

$$\tilde{\mathbf{y}}_{\mathbf{k}}^{\mathbf{1}} = C\mathbf{F}_{\mathbf{k}}u + \tilde{\mathbf{y}}_{\mathbf{k}}^{\mathbf{2}}$$

which can be solved by WLS (or with Maximum Likelihood Estimation (MLE)) for u with the result

$$\hat{u}_k = \left(\mathbf{F_k}^T C^T \mathbf{S_{k|k-1}}^{-1} C\mathbf{F_k}\right)^{-1} \mathbf{F_k}^T C^T \mathbf{S_{k|k-1}}^{-1} \tilde{\mathbf{y}}_{\mathbf{k}}^{\mathbf{1}}$$
$$P_{u_k} = \left(\mathbf{F_k}^T C^T \mathbf{S_{k|k-1}}^{-1} C\mathbf{F_k}\right)^{-1}$$

where

$$\mathbf{S_{k|k-1}} \triangleq \text{blkdiag}\left(S^1_{k|k-1}, \ldots, S^1_{k-N+1|k-N}\right)$$

is the covariance of $\tilde{\mathbf{y}}_{\mathbf{k}}^{\mathbf{1}}$.

Detection Based Methods: IE

We are going to check whether $\hat{u}$ is statistically significant.

Input Estimation Method

- Make estimation with the first model i.e., calculate $\hat{x}^1_{k|k}$, $P^1_{k|k}$.
- With the arrival of each measurement y_k, calculate $\hat{u}_k$ and P_{u_k} using the input estimation procedure.
    - If $\hat{u}_k^T P_{u_k}^{-1} \hat{u}_k > \gamma_{\max}$,
        - Declare a maneuver and compensate the estimation errors by updating the predicted quantities as

$$\hat{x}^{1+}_{k|k-1} = \hat{x}^1_{k|k-1} + F_k \hat{u}_k$$
$$\hat{P}^{1+}_{k|k-1} = P^1_{k|k-1} + F_k P_{u_k} F_k^T$$

        - Calculate $\hat{x}^1_{k|k}$, $P^1_{k|k}$ from updated quantities $\hat{x}^{1+}_{k|k-1}$, $P^{1+}_{k|k-1}$

$\gamma_{\max}$ can be calculated from the statistics of $\chi^2_{n_u}$.

Detection Based Methods

- ALPN uses only the covariance of the white process noise to compensate maneuvers.
    - This might not be sufficient when the maneuvers are long and persistent.
- VSD seems limited to the Constant Acceleration model but the opposite is claimed in [Bar-Shalom, Li, Kirubarajan (2001)].
    - How to initialize the maneuvering model in the case of e.g. a Coordinated Turn Model is not known.
- IE assumes a constant acceleration profile during both detection and compensation procedure.
    - This can make it unable to compensate the maneuvers well if the accelerations change fast.
- A case study in [Bar-Shalom, Li, Kirubarajan (2001)] shows that

$$\text{MSE}_{\text{VSD}} \lessapprox \text{MSE}_{\text{ALPN}} < \text{MSE}_{\text{IE}}$$

Multiple Model Approaches

- Detection based methods are in general too slow to compensate the maneuvers.
- Target motions can generally be classified into a number of predefined number of modes e.g.
    - Constant velocity
    - Coordinated turn (circular motion with constant speed and angular rate)
    - Constant acceleration
- Using maneuver detection is a type of making a hard decision between these models i.e., serial use of models (use one model first then switch to another one etc.).
- The soft version uses all the models at the same time (parallel use of models) and combines their results to the extent that they suit to the measurements collected probabilistically.

Multiple Model Approaches: JMLS

Jump Markov linear systems (JMLS): give a useful framework for using multiple models

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

- x_k is the state that we would like estimate from y_k. This state is called as **base state**.
- $r_k \in \{1, 2, \ldots, N_r\}$ represents model number and is called as **mode (or modal) state**. Note that r_k is also unknown and must be estimated from measurements y_k.
- $A(\cdot)$, $B(\cdot)$, $C(\cdot)$ and $D(\cdot)$ are mode dependent parameters.

Multiple Model Approaches: JMLS

Multiple model approaches can be classified into two broad categories as

- Non-switching models
- Switching models

Non-Switching case:

- The underlying model r_k is unknown but fixed for all times, i.e., $r_k = r$, $k = 1, 2, \ldots$.
- This type of approaches is useful in *system identification* with finite number of model alternatives but not very suitable for TT.

Switching case:

- The underlying model r_k can jump between different values in $\{1, 2, \ldots, N_r\}$
- The time behavior of r_k is generally modeled as first order homogeneous Markov chain with a fixed transition probability matrix.

Multiple Model Approaches: Optimal Solution

Suppose we started estimation at time 0 and now we are at time k.

- There are a total of N_r^k different model histories $r_{1:k}$ that might have occurred in this period. We show these by $\{r_{1:k}^i\}_{i=1}^{N_r^k}$.
- When a specific model history $r_{1:k}^i$ is given we can calculate the estimated density of the state x_k as

$$p(x_k|y_{1:k}, r_{1:k}^i) = \mathcal{N}(x_k; \hat{x}_{k|k}^i, \Sigma_{k|k}^i)$$

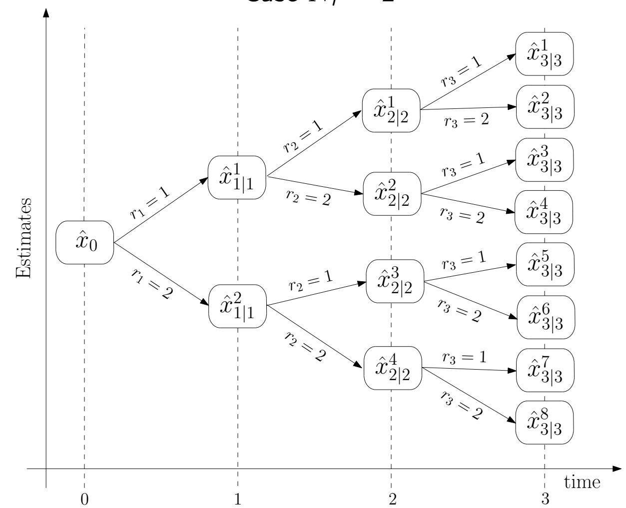
which is given by a KF that is matched to the model history.
- The overall MMSE estimate $\hat{x}_{k|k}$ is then given as

$$\hat{x}_{k|k} = \sum_{i=1}^{N_r^k} \mu_k^i \hat{x}_{k|k}^i$$

where $\mu_k^i \triangleq P(r_{1:k}^i|y_{1:k})$.

Multiple Model Approaches: Optimal Solution

Case $N_r = 2$



Multiple Model Approaches: Optimal Solution

- Storage and computation requirements of the optimal filter increase exponentially.
- The posterior density of the state at time k is given as

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

- The number of components in the Gaussian mixture should be decreased.
- Some approaches use **pruning** (discarding low probability terms).
- We here will consider the most popular approach **merging**.

Multiple Model Approaches: Mixture Reduction

The Gaussian mixture given by

$$p(x_k) = \sum_{i=1}^{N} \pi_i \mathcal{N}(x_k; \hat{x}_k^i, \Sigma_k^i)$$

can be approximated as

$$p(x_k) \approx \mathcal{N}(x_k; \hat{x}_k, \Sigma_k)$$

where

$$\hat{x}_k \triangleq \sum_{i=1}^{N} \pi_i \hat{x}_k^i \qquad \Sigma_k \triangleq \sum_{i=1}^{N} \pi_i \left[\Sigma_k^i + (\hat{x}_k^i - \hat{x}_k)(\hat{x}_k^i - \hat{x}_k)^T\right]$$

This is a moment matching approximation and called as **merging**. The second term in the covariance approximation (brackets) is called as the **spread of the means**.

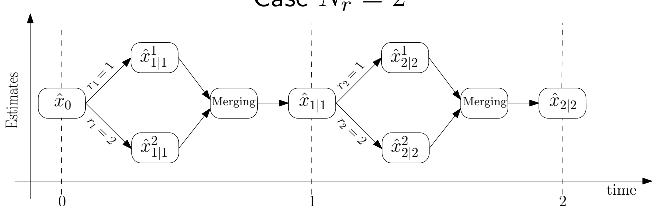
Multiple Model Approaches: GPB1

Generalized pseudo Bayesian algorithms (GPB)

GPB1 Approximation: $p(x_k|y_{1:k}) \approx \mathcal{N}(x_k; \hat{x}_{k|k}, \Sigma_{k|k})$

- Storage: 1 mean and covariance
- Computation: N_r Kalman filters
- Merge with probabilities $\mu_k^i \triangleq P(r_k = i|y_{1:k})$.

Case $N_r = 2$



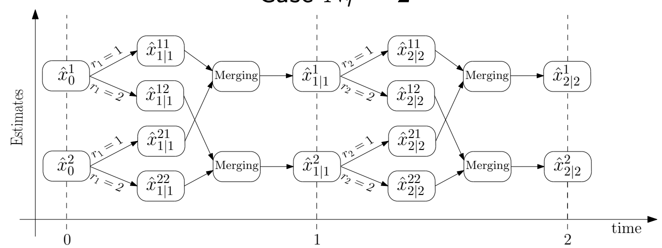
Multiple Model Approaches: GPB2

Generalized pseudo Bayesian algorithms (GPB)

GPB2 Approximation: $p(x_k|y_{1:k}) \approx \sum_{i=1}^{N_r} \mu_k^i \underbrace{\mathcal{N}(x_k; \hat{x}_{k|k}^i, \Sigma_{k|k}^i)}_{=p(x_k|y_{0:k}, r_k=i)}$

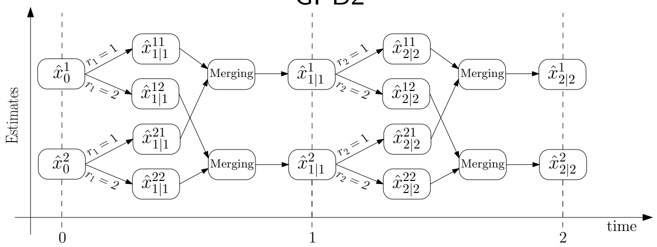
- Storage: N_r means and covariances
- Computation: N_r^2 Kalman filters

Case $N_r = 2$



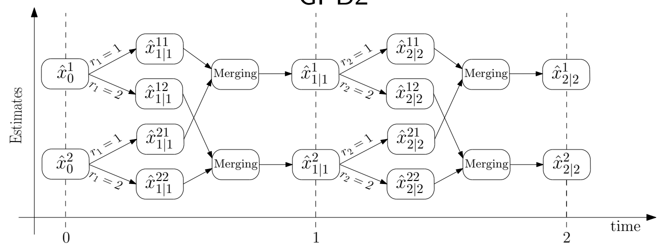
Multiple Model Approaches: GPB2 vs. IMM

GPB2

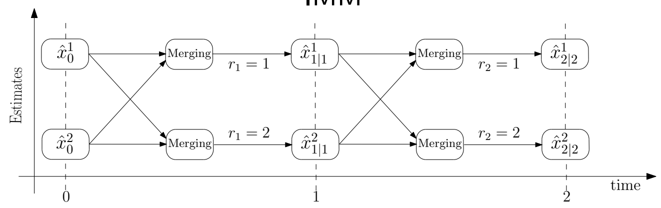


Multiple Model Approaches: GPB2 vs. IMM

GPB2



IMM



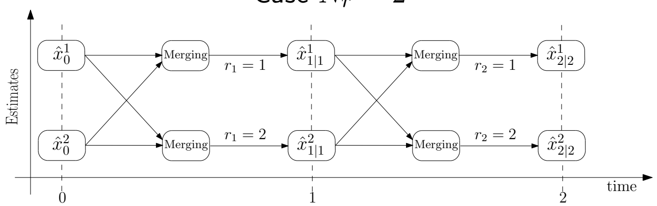
Multiple Model Approaches: IMM

Interacting Multiple Models

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

- Same approximation as GPB2
- Storage: N_r means and covariances
- Computation: N_r Kalman filters

Case $N_r = 2$



Multiple Model Approaches: IMM

One Step of IMM Algorithm

-Suppose we have the previous summary statistics $\{x_{k-1|k-1}^j, \Sigma_{k-1|k-1}^j, \mu_{k-1}^j\}_{j=1}^{N_r}$.

-We would like to obtain the new sufficient statistics $\{x_{k|k}^j, \Sigma_{k|k}^j, \mu_k^j\}_{j=1}^{N_r}$.

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

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

  - 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

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

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

Multiple Model Approaches: IMM

- **Mode Matched Prediction Update:** For $i = 1, \ldots, N_r$, calculate $\hat{x}^i_{k|k-1}$ and $\Sigma^i_{k|k-1}$ from $\hat{x}^{0i}_{k-1|k-1}$ and $\Sigma^{0i}_{k-1|k-1}$ as

$$\hat{x}^i_{k|k-1} = A(i)\hat{x}^{0i}_{k-1|k-1},$$
$$\Sigma^i_{k|k-1} = A(i)\Sigma^{0i}_{k-1|k-1}A^T(i) + B(i)QB^T(i).$$

- **Mode Matched Measurement Update:** For $i = 1, \ldots, N_r$,
  - Calculate $\hat{x}^i_{k|k}$ and $\Sigma^i_{k|k}$ from $\hat{x}^i_{k|k-1}$ and $\Sigma^i_{k|k-1}$ as

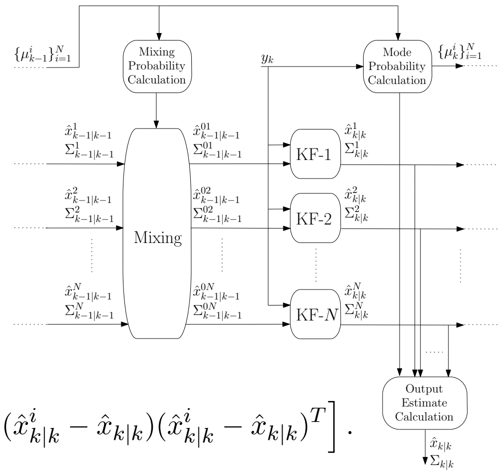
$$\hat{x}^i_{k|k} = \hat{x}^i_{k|k-1} + K^i_k(y_k - \hat{y}^i_{k|k-1}), \quad \hat{y}^i_{k|k-1} = C(i)\hat{x}^i_{k|k-1},$$
$$\Sigma^i_{k|k} = \Sigma^i_{k|k-1} - K^i_k S^i_k K^{iT}_k, \quad S^i_{k|k-1} = C(i)\Sigma^i_{k|k-1}C^T(i) + D(i)RD^T(i),$$
$$K^i_k = \Sigma^i_{k|k-1}C^T(i)(S^i_{k|k-1})^{-1}.$$

  - Calculate the updated mode probability μ^i_k as

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

Multiple Model Approaches: IMM

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



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

-Output estimate is only calculated for output purposes and it is not in the main IMM recursion.

Multiple Model Approaches: IMM

Gating and Data Association with IMM

- At each step, one can just calculate the following overall predicted measurement $\hat{y}_{k|k-1}$ and innovation covariance $S_{k|k-1}$

$$\hat{y}_{k|k-1} = \sum_{i=1}^{N_r}\mu^i_{k|k-1}\hat{y}^i_{k|k-1} \qquad \mu^i_{k|k-1} \triangleq \sum_{j=1}^{N_r}\pi_{ji}\mu^j_{k-1}$$
$$S_{k|k-1} = \sum_{i=1}^{N_r}\mu^i_{k|k-1}\left[S^i_{k|k-1} + (\hat{y}^i_{k|k-1} - \hat{y}_{k|k-1})(\cdot)^T\right]$$

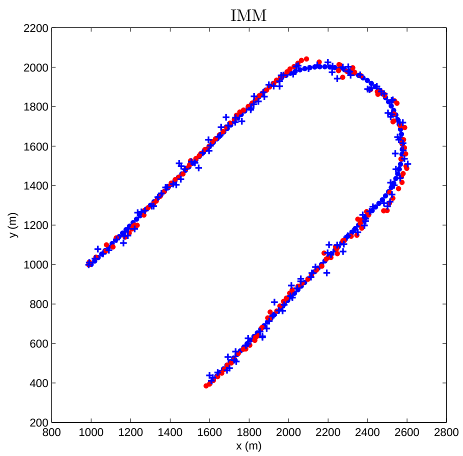
We can do the gating and data association with these quantities.

- An alternative is to do individual gating for each model and then to take the union of the gated measurements from all models. In this case, the overall likelihood for association is formed from individual likelihoods as

$$p(y_k|y_{1:k-1}) = \sum_{i=1}^{N_r}\mu^i_{k|k-1}\underbrace{p(y_k|y_{1:k-1}, r_k = i)}_{\substack{\text{individual likelihood from } i\text{th KF} \\ = \mathcal{N}(y_k; \hat{y}^i_{k|k-1}, S^i_{k|k-1})}}$$

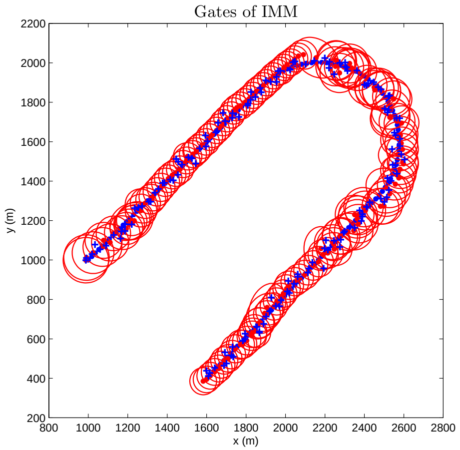
IMM Illustration

- The same illustration with now IMM filter.
- $P_D = 1.$
- $P_G = 1$
- $P_{FA} = 0$
- IMM uses two CV models same except for
  - $\sigma^1_a = 0.1\text{m/s}^2.$
  - $\sigma^2_a = 10\text{m/s}^2.$



IMM Illustration

- The same illustration with now IMM filter.
- $P_D = 1$.
- $P_G = 1$
- $P_{FA} = 0$
- IMM uses two CV models same except for
  - $\sigma_a^1 = 0.1\text{m/s}^2$.
  - $\sigma_a^2 = 10\text{m/s}^2$.



References

- Y. Bar-Shalom and K. Birmiwal, "Variable Dimension Filter For Maneuvering Target Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol.18, no.5, pp. 621–629, Sep. 1982.
- Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: Wiley, 2001. (Chapter 11 is about Maneuvering Target Tracking)
- X. Rong Li and V.P. Jilkov , "Survey of maneuvering target tracking. Part I. Dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol.39, no.4, pp. 1333–1364, Oct. 2003.
- X. Rong Li and V.P. Jilkov , "Survey of maneuvering target tracking. Part V. Multiple-model methods," *IEEE Transactions on Aerospace and Electronic Systems*, vol.41, no.4, pp. 1255–1321, Oct. 2005.