Kalman Filter

Dr Salman Aslam

Wing Commander, PAF
Associate Professor
Avionics Department
College of Aeronautical Engineering
PAF Academy Risalpur

Outline

1. Introduction
2. Details
   - Step 1. Time dynamic systems
   - Step 2. Bayes Theorem
   - Step 3. Recursion
   - Step 4. Solutions
3. Conclusion

Overview

- The Kalman filter is a solution to the recursive Bayesian equations of a linear time-dynamic system
- Ok, so working backwards, we need to know about time dynamic systems, Bayesian equations, recursion and a solution to these equations:
  1. Time dynamic systems
  2. Bayes Theorem (Prior, Likelihood, Posterior)
  3. Recursion (Predict and Update)
  4. Solutions (HMM, Kalman Filter, PDAF, JPDAF, Particle Filter)
- Let's look at these one at a time

Time dynamic systems

- Time-dynamic systems are described by two equations:

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{v}_{k-1})$$
$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{n}_k)$$

- States are denoted by $\mathbf{x}$
- Observations are denoted by $\mathbf{z}$
- Process noise is denoted by $\mathbf{v}$
- Observation noise is denoted by $\mathbf{n}$
- Time is denoted by k
- State prediction model, $f_k : R^D \times R^D \rightarrow R^D$, describes state evolution
- Observation model, $h_k : R^N \times R^N \rightarrow R^N$, relates observations to the states

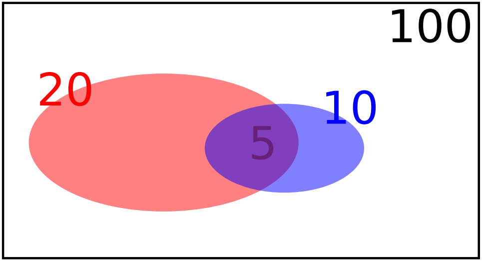
Time dynamic systems
Examples

- Robotics: motion of a robot
- Agriculture: color change of a fruit
- Human body: Secretion of insulin
- Finance: Fluctuations in inflation
- Mechanical: Spring mass system
- Radars: Motion of an aircraft

Bayes Theorem

Ok, time dynamic systems make sense but how does the Bayes theorem fit into our picture here?

Bayes Theorem
Graphical demonstration



$$p(R) = 1/5 \qquad p(B) = 1/10$$

$$p(B|R) = 1/4 \qquad p(R|B) = 1/2$$

$$p(B|R)p(R) = 1/20 \qquad p(R,B) = p(R|B)p(B) = 1/20$$

Bayes Theorem

Example from real life: Prosecutor's fallacy

- If someone is guilty, the probability of a DNA match is large, $p(D|g)$ is large
  - convict?
- Probability that a person is guilty if picked randomly from a large database is small, $p(g)$ is very small
- Probability $p(D,g) = p(D|g)p(g)$ is not so large after all
- So the prior information, $p(g)$ is important!

Bayes Theorem

Relationship with time dynamic systems

- $p(x|z) = \frac{p(z|x)p(x)}{p(z)}$

  Posterior distribution of states (our goal)
  Prior distribution of states (our belief)
  Likelihood distribution (we can compute it)
  Observation probability (we won't need it!)

- Now, what's the deal with recursion?

Recursion

- In Bayes' theorem, we talked about 3 important probability distributions (posterior, prior, likelihood)
- We want to compute the posterior distribution at every time instant
- This is recursion, incremental updates, not batch mode

Recursion

- We assume that our system and observation models are available in probabilistic form
- The probabilistic state-space formulation and the requirement for updating of information on receipt of new measurements are ideally suited for the Bayesian approach
    1. Prediction: predict states using model
    2. Update: correction applied to prediction after observation arrives

Recursion

a. Predict

x_k	state at time k
Z_{k-1}	all observations till time $k-1$

prior at time k

$$\boxed{p(x_k|Z_{k-1})} = \frac{p(x_k, Z_{k-1})}{p(Z_{k-1})}$$

nuisance variable

marginalization

$$= \frac{\int p(x_k, x_{k-1}, Z_{k-1})\,dx_{k-1}}{p(Z_{k-1})}$$

Bayes' rule

$$= \frac{\int p(x_k|x_{k-1}, Z_{k-1})\,p(x_{k-1}, Z_{k-1})\,dx_{k-1}}{p(Z_{k-1})}$$

independence,
Bayes' rule

$$= \frac{\int p(x_k|x_{k-1})\,p(x_{k-1}|Z_{k-1})\,\cancel{p(Z_{k-1})}\,dx_{k-1}}{\cancel{p(Z_{k-1})}}$$

model, posterior from time *k-1*

$$= \int \underbrace{p(x_k|x_{k-1})}_{\text{model}}\,\underbrace{p(x_{k-1}|Z_{k-1})}_{\text{posterior from time } k-1}\,dx_{k-1}$$

Recursion

b. Update

posterior

$$p(x_k|Z_k) = \frac{p(x_k, Z_k)}{p(Z_k)}$$

$p(Z_k) = p(z_k, Z_{k-1})$

$$= \frac{p(x_k, z_k, Z_{k-1})}{p(z_k, Z_{k-1})}$$

Bayes' rule

$$= \frac{p(z_k|x_k, Z_{k-1})p(x_k, Z_{k-1})}{p(z_k|Z_{k-1})p(Z_{k-1})}$$

independence,
Bayes' rule

$$= \frac{p(z_k|x_k)p(x_k|Z_{k-1})\cancel{p(Z_{k-1})}}{p(z_k|Z_{k-1})\cancel{p(Z_{k-1})}}$$

nuisance
variable

$$= \frac{p(z_k|x_k)p(x_k|Z_{k-1})}{\int p(z_k|x_k)p(x_k|Z_{k-1})dx_k}$$

likelihood prior (prediction)

$$\propto p(z_k|x_k)p(x_k|Z_{k-1})$$

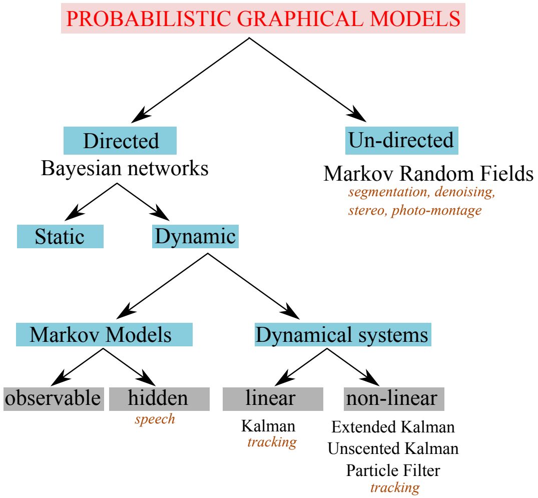
Solutions

How to implement?

- The recursive propagation of the state density is only a conceptual solution in that in general, it cannot be determined analytically.
- Solutions do exist in a restrictive set of cases, including the Kalman filter and grid-based filter
  - State space is discrete and consists of finite number of states
- In many situations of interest, the conditions of linearity and Gaussianity do not hold, and approximations are necessary.
- Three approximate nonlinear Bayesian filters are the Extended Kalman Filter (EKF), approximate grid based methods and the particle filter

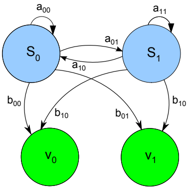
Solutions

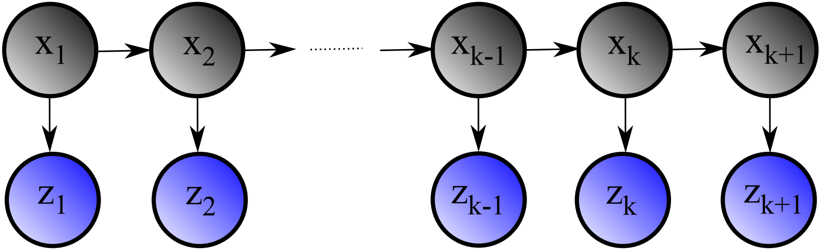
Bigger picture



Solutions

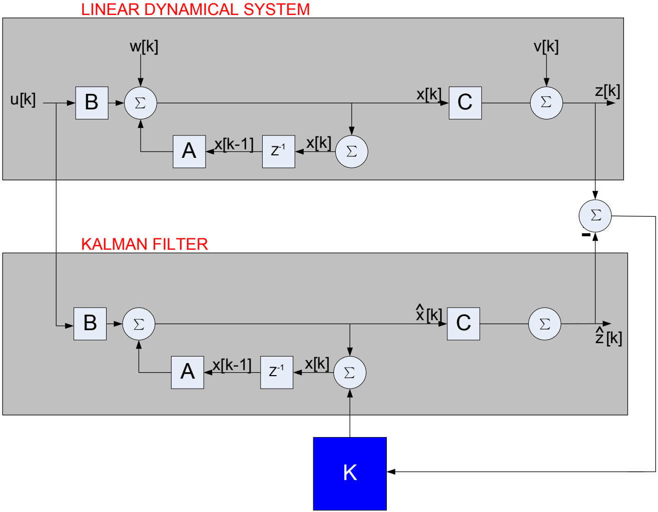
a. Hidden Markov Models





Solutions

b. Kalman Filter



Solutions

b. Kalman Filter

- Prior
  - Past information through time $k-1$
  - summarized approximately by a sufficient statistic in the form of a Gaussian posterior

$$p(x_{k-1|k-1}|Z_{k-1}) = \mathcal{N}(\hat{x}_{k-1|k-1}, P_{k-1|k-1})$$

- Prediction
  - The state prediction is distributed as,

$$p(x_{k|k-1}|Z_{k-1}) = \mathcal{N}(\hat{x}_{k|k-1}, P_{k|k-1})$$

  - The observation prediction is distributed as,

$$p(z_{k|k-1}) = \mathcal{N}(\hat{z}_{k|k-1}, S_k)$$

- Update
  - innovation
  - Kalman gain
  - aposteriori state estimate and aposteriori covariance estimate

Solutions

b. Kalman Filter

1. PREDICTION

1a.	state	$\hat{\mathbf{x}}_{k\|k-1} = \mathbf{A}_k \hat{\mathbf{x}}_{k-1} + \mathbf{B}_k \mathbf{u}_k$
1b.	state covariance	$\mathbf{M}_{k\|k-1} = \mathbf{A}_k \mathbf{M}_{k-1\|k-1} \mathbf{A}_k^T + \mathbf{Q}_k$
2a.	measurement	$\hat{\mathbf{z}}_{k\|k-1} = \mathbf{C}_k \hat{\mathbf{x}}_{k\|k-1}$
2b.	measurement covariance (called innovation covariance)	$\mathbf{S}_k = \mathbf{C}_k \mathbf{M}_{k\|k-1} \mathbf{C}_k^T + \mathbf{R}_k$
3.	gain	$\mathbf{K}_k = \mathbf{M}_{k\|k-1} \mathbf{C}_k^T \mathbf{S}_k^{-1}$

observation is received

4.	innovation	$\mathbf{y}_k = \mathbf{z}_k - \hat{\mathbf{z}}_{k\|k-1}$

2. UPDATE

5a.	state	$\hat{\mathbf{x}}_{k\|k-1} = \hat{\mathbf{x}}_{k\|k-1} + \mathbf{K}_k \mathbf{y}_k$
5b.	state covariance	$\mathbf{M}_{k\|k-1} = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{M}_{k\|k-1}$

system model covariance 'unknown'

Solutions

c. PDAF: Extension of Kalman Filter

Probabilistic Data Association Filter

- Computationally efficient
- Data association
- Single targets in clutter, high false alarm rate
- Extension of Kalman filter
  - Primary difference: computation of innovations
  - Calculates association probabilities to the target for each validated measurement

Solutions

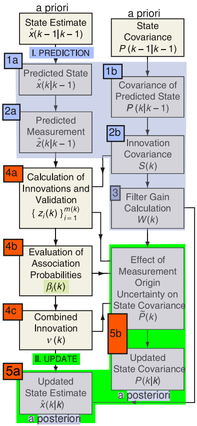
c. PDAF: Extension of Kalman Filter

Assumptions:

1. Number of targets
    - Single target tracking
2. False alarms
    - At most, one of the validated measurements can be target originated
    - Remaining measurements
        - incorrect, i.e., false alarms
        - i.i.d
        - parametric form: Poisson distribution with known spatial density λ
        - non-parametric form: uniform spatial distribution

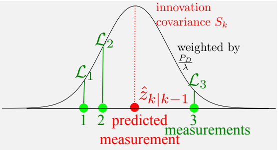
Solutions

c. PDAF: Extension of Kalman Filter



$$\beta_i(k) = p(A_i(k)|Z_k)$$

$$= \begin{cases} \dfrac{\mathcal{L}_i}{1-P_D P_G + \sum_{j=1}^{M_k} \mathcal{L}_i} & j = 1 \ldots M_k \\[2ex] \dfrac{1-P_D P_G}{1-P_D P_G + \sum_{j=1}^{M_k} \mathcal{L}_i} & j = 0 \end{cases}$$



$$\mathcal{L}_i = \frac{\mathcal{N}(z_{ik}; \hat{z}_{k|k-1}, S_k) P_D}{\lambda}$$

$$\nu_k = \sum_{i=1}^{M_k} \beta_i(k) \nu_i(k)$$

Bar-shalom, 2009

Solutions

c. PDAF: Extension of Kalman Filter

PDAF + multiple model Kalman Filters :

US Navy ROHR (Relocatable Over the Horizon Radar) for long-range surveillance for drug interdiction against aircraft and ships

Solutions

d. JPDAF: Extension of Kalman Filter

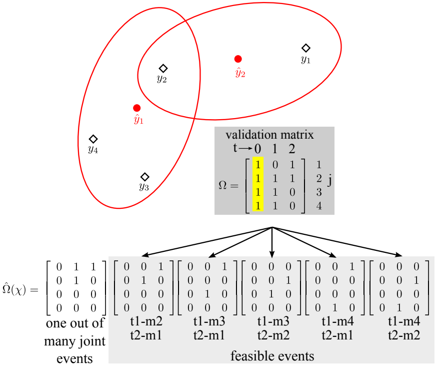
Joint Probabilistic Data Association Filter

- Extension of PDAF to multi-target tracking
  - Only difference: computation of association probabilities, β_{ik}
- Handles multiple targets by considering all measurements for all targets.
- Probability density of each candidate measurement
  - Based on all close-by targets

Solutions

d. JPDAF: Extension of Kalman Filter

Association probabilities



$$p[\chi|Z_k] = \frac{1}{c} \prod_{j:\tau_j=1} \mathcal{N}[z_j(k); \hat{z}^{t_j}(k|k-1), S^{t_j}(k)] \prod_{t:\delta_t=1} P_D^t \prod_{t:\delta_t=0} (1 - P_D^t)$$

$$\beta_j^t = \sum_{\chi} p[\chi|Z_k]\hat{\omega}_{jt}(\chi)$$

Solutions

d. JPDAF: Extension of Kalman Filter

Nearest neighbor JPDAF + EKF:

THAAD, Theater High Altitude Area Defense)

Cobra Dane, long-range surveillance against ICBMs

SBX, long-range surveillance against ICBMs

Solutions

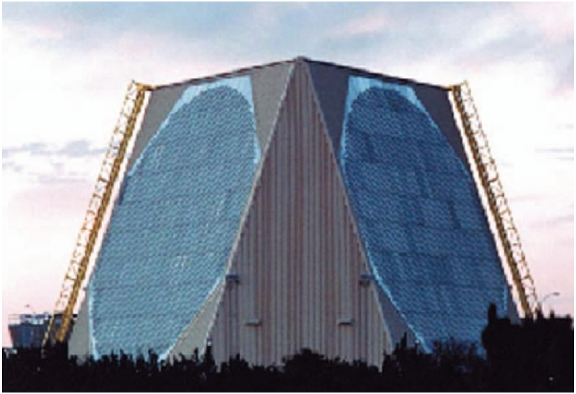
e. Multi-hypothesis tracker

1. Clustering
   - A new measurement is associated with a cluster if it falls within the validation region of any target within the cluster
2. Hypothesis Generation
   - New hypotheses are generated for the measurements associated with each cluster
   - The probability of each hypothesis is calculated
   - Target estimates are then updated
3. Reduction
   - In this step, hypotheses are combined or eliminated
4. Simplify hypothesis matrix
   - Targets that are uniquely associated are removed from the hypothesis matrix

Solutions

e. Multi-hypothesis tracker

MHT + EKF:



UEWR, Upgraded Early Warning Radar, long-range surveillance against ICBMs

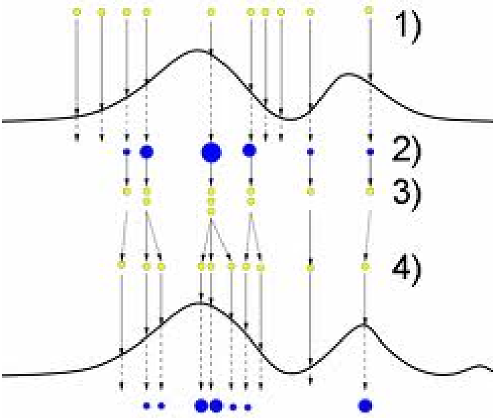
Solutions

f. Particle filter

- All densities are modeled using samples
- Two step process, predict, and then update using resampling

Solutions

f. Particle filter



http://lia.deis.unibo.it/research/SOMA/MobilityPrediction/filters.shtml

Solutions

f. Particle filter

- There is an essential duality between a sample and the density (distribution) from which it is generated
  - The density generates the sample
  - Conversely, given a sample, we can approximately recreate the density
- In terms of densities, the inference process is encapsulated in the updating of the prior density $p(x)$ to the posterior density $p(x|z)$ through the medium of the likelihood function $p(z|x)$
- In terms of samples, this corresponds to the updating of a sample from $p(x)$ to a sample from $p(x|z)$ through the likelihood function $p(z|x)$

Smith Gelfand, Bayesian Statistics Without Tears, A Sampling Resampling Perspective, 1992

Solutions

f. Particle filter

- In many application areas, elements of nonlinearity and non-Gaussianity need to be included to accurately model the underlying dynamics of a physical system
- Moreover, it is typically crucial to process data on-line as it arrives
- Particle filters are Sequential Monte Carlo (SMC) methods based on point-mass (or "particle") representations of probability densities which can be applied to any state-space model
- The particle filter generalizes the Kalman filter
- The algorithm used is Sequential Importance Sampling (SIS)

Arulampalam et al., A tutorial on particle filters for online nonlinear non-Gaussian Bayesian Tracking, 2002

Conclusion

- There are many solutions to the Bayesian recursion equations
- The Kalman Filter is optimal if these equations are linear and the models are Gaussian