Bayesian Methods in Positioning Applications

Vedran Dizdarević
v.dizdarevic@TUGraz.at

Graz University of Technology, Austria

24. May 2006
Outline

- Problem Statement
- Solution Alternatives
- Introduction to Bayesian Estimation
- Bayes Rule
- Recursive Bayesian Estimator
- Particle Filtering
- MATLAB Demo
- Bayesian Estimation in Positioning Applications
Problem Statement (1/2)

- System description in two steps
  - Find system model (modeling)
  - Find the particular realization of the model (identification; parameter estimation $\hat{\theta}$)

- Time **variant** system parameters require recursive estimation
  - on-line vs. off-line data processing
Problem Statement (2/2)

- Basic notation
  - Time-discrete data record \( z_{0:k} = \{ z_0, \ldots, z_k \} \)
  - With N-dimensional observation \( z_k = \{ z^1_k, \ldots, z^N_k \} \)
  - Inference of the model parameter \( z_{0:k} \rightarrow \hat{\theta}(k; z_{0:k}) \)
  - Recursive estimation \( \hat{\theta}_k = F(\hat{\theta}_{k-1}, z_k, S_{k-1}) \)

- Initial condition(s) - selection of \( \theta_0 \)
Solution Alternatives

- Modification of off-line estimation methods
  - RLS

- Stochastic approximation
  - Gradient-based methods

- Model reference techniques and pseudolinear regression
  - Adaptive adjustment through comparison of two systems
  - Extension of linear regression (LS) methods

- Bayesian inference
  - Estimation of distribution of $\theta$
Introduction to Bayesian Estimation

- Bayesian approach
  - Probabilistic framework for recursive state estimation
  - In Bayesian estimation $\theta$ is treated as a random variable with $p(\theta)$
  - By defining $p(\theta)$ estimation uncertainty is implied
  - The formalism allows introduction of prior knowledge of $p(\theta)$
  - Relying on a Bayes rule updated $p(\theta)$ is inferred from prior (predicted) $p(\theta)$ combined with incoming observations
Bayes Rule and Contemporary Science

Published in T. Bayes *Essay Towards Solving a Problem in the Doctrine of Chances* (1764)

Steam engine of J. Watt (1769)

B. Franklin defines the concept of positive and negative polarity (1752)

L. Galvani and frogs (1770)

K.F. Gauss was born (1777)
Bayes Rule

- Assume a disjoint set of $n$ discrete events $\{E_i\}_{i=1}^n$
- Probabilities $P(E_i)_{i=1}^n$ are assigned to each event
- If events are not independent observing that event $E_k$ has occurred, probabilities of all other events $\{E_i\}_{i=1}^n$ will change
- This is described using conditional probability $P(E_i|E_k)$
- $P(E_i, E_k) = P(E_i|E_k)P(E_k) = P(E_k|E_i)P(E_i)$
- $P(E_i|E_k) = \frac{P(E_k|E_i)P(E_i)}{P(E_k)}$
Bayesian Estimator (1/4)

Two equations describe the discrete-time, time-variant, d-dimensional system

- **Motion model**
  \[ \theta_{k+1} = f(\theta_k) + w_k \]
  \[ \theta_k, w_k \in \mathbb{R}^d, f : \mathbb{R}^d \rightarrow \mathbb{R}^d \]

- **Measurement model**
  \[ z_k = h_k(\theta_k) + v_k \]
  \[ z_k, v_k \in \mathbb{R}^{m_k}, h_k : \mathbb{R}^d \rightarrow \mathbb{R}^{m_k} \]
Bayesian Estimator (2/4)

Notation

- **Conditional pdf**
  \[ p_{k|l}(\theta_k) = p(\theta_k | z_1, \ldots, z_l) \]

- **Likelihood function**
  \[ L_k(\theta_k) = p(z_k | \theta_k) = p_v(k)(z_k - h_k(\theta_k)) \]

- **Transition pdf**
  \[ \phi_k(\theta_{k+1} | \theta_k) = p_w(k)(\theta_{k+1} - f(\theta_k)) \]
Bayesian Estimator (3/4)

- **Step 0**
  Select initial distribution $p_{0|0}$ and set $k = 1$

- **Step 1**
  Compute the **predictive pdf**
  $$p_{k|k-1}(\theta_k) = \int_{\mathbb{R}^d} p(\theta_k|\theta_{k-1}) p(\theta_{k-1}|z_1, \ldots, z_{k-1}) d\theta$$

- **Step 2**
  Compute the **posterior pdf**
  $$p_{k|k}(\theta_k) = \frac{p(z_k|\theta_k)p(\theta_k|z_1, \ldots, z_{k-1})}{p(z_k|z_1, \ldots, z_{k-1})} \propto p_{k|k-1}(\theta_k)L_k(\theta)$$

- **Step 3**
  Output $\hat{\theta}_k$ and variance $V(p_{k|k}(\theta_k))$

- **Step 4**
  Increment $k$ and repeat from Step 1
Bayesian Estimator (4/4)

- Calculation of $\hat{\theta}_k$
  
  $\hat{\theta}_k = E(p_{k|k}(\theta_k))$
  
  $\hat{\theta}_k = \arg\max(p_{k|k}(\theta_k))$ (MAP)

- Problem: it is difficult to find an analytical solution to this problem!

- If the noise pdfs are independent, white and zero-mean Gaussian and the measurement equation linear in $\theta$, it can be derived that the Bayesian approach reduces to the Kalman Filter

- A general algorithm which approximates the pdfs with arbitrary accuracy is the particle filter
Particle Filter (1/4)

- Two steps require intractable integration
  - Posterior pdf computation
  - Expectation of the posterior pdf
- Numerical means for the tracking the evolution of the pdfs
- The pdf is approximated as weighted sum of samples in the space state (particles)
  \[ p(\theta) \approx \sum_{m=1}^{M} w^{(m)} \delta(x - x^{(m)}) \]
- Recursive update
  - Prediction: Calculate the predictive pdf using the motion model
  - Update: Weights update with normalization
Particle Filter (2/4)

- **Initialization**
  Select an importance function $\pi$ and create a set of particles $x^{(m)}_0 \sim \pi$

- **Measurement update**
  Weight update according to the likelihood
  $$w^{(m)}_k = w^{(m)}_{k-1} \pi(\theta|z)$$
  Normalize the weights
  $$w_k := \frac{w^{(m)}_k}{\sum_m w^{(m)}_k}$$
  Calculate the estimate
  $$\hat{\theta}_k \approx \sum_{m=1}^M w^{(m)}_k \theta^{(m)}_k$$

- **Prediction**
  $$\theta_{k+1} = h(\theta_k) + v_k$$
Particle Filter (3/4)

- Measurement update
- Resampling
- Prediction

- Weights tend to concentrate in few particles
- Draw $P$ particles from the current particle with probabilities proportional to the corresponding weight; set weights for the new particle as $w_k = \frac{1}{P}$
- Optionally resampling only if a given criterion is fulfilled

$$N_{eff} = \frac{1}{\sum_i (w_k^{(i)})^2} < N_{th}$$
Particle Filter (4/4)
MATLAB example

System equations
\[ \theta_k = 0.5\theta_{k-1} + 25\frac{\theta_{k-1}}{1+\theta_{k-1}^2} + 8\cos(1.2k) + w_k \]
\[ z_k = \frac{\theta_k^2}{20} + v_k \]

Simulation parameters
\[ \sigma_w^2 = 10 \]
\[ \sigma_v^2 = 1 \]
\[ \sigma_{\pi}^2 = 5 \]
\[ M = 500 \]
Pro and Contra

This algorithm is known under many names!

Pro:
- Non-linear state model needs no approximations
- Non-Gaussian distributions $\theta$
- Combines different observations in a single model
- Variable dimension of the observation vector no problem

Contra:
- Computationally expensive
- Selection of prior distribution
Bayesian Estimation in Positioning Applications (1/2)

- Application: Positioning, Navigation, Tracking
- Type of object: Persons, Robots, Vehicles
- Supplementary information: Road map, terrain profile

- State space:
  \[ \theta = [p, \phi, v, \dot{v}, \ldots]^T \]

- Observations:
  \[ z = [p_{ext}, \phi_{ins}, \dot{v}_{ins}, \gamma_{image}, \gamma_{radar}, \mathbf{S}_{ss}, \ldots]^T \]
Current research directions

- Motion models
  Prediction of pedestrian movements
- Complexity reduction
  Employ KF for linear states
- Parallelization
  Distributed PF in a sensor network
References