Cubature Kalman Filtering: Theory & Applications

I. (Haran) Arasaratnam

Advisor: Professor Simon Haykin

Cognitive Systems Laboratory
McMaster University
April 6, 2009

Organization

- **Introduction** to Bayesian Filtering

- **Existing Algorithms**

- **Proposed Solutions**:
  - Cubature Kalman Filtering
  - Extension to Square-Root Cubature Kalman Filtering

- **Example Application**: Model-Based Signal Processing

Bayesian Filtering: Introduction

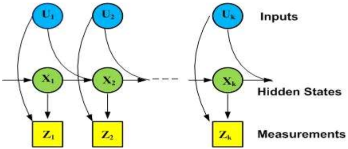


Figure 1:

- State-space model in discrete time:
  - Process equation:

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1} \ \& \ \mathbf{u}_{k-1}) + (\text{Pro. noise})_{k-1} \quad (1)$$

  - Measurement equation:

$$\mathbf{z}_k = \mathbf{h}(\mathbf{x}_k \ \& \ \mathbf{u}_k) + (\text{Meas. noise})_k \quad (2)$$

- Key Question: How do we recursively compute the posterior density of the state $\mathbf{x}_k$, given the noisy measurements up to time k, $\mathbf{z}_{1:k} = \{\mathbf{z}_1, \ \mathbf{z}_2, \ldots \mathbf{z}_k\}$?

Conceptual Recursive Solution

- Time-update step using the C-K equation:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k-1}) = \int_{\mathbb{R}^{n_x}} p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})p(\mathbf{x}_k|\mathbf{x}_{k-1})d\mathbf{x}_{k-1} \quad (3)$$

- Measurement-update step using Bayes' rule:

$$p(\mathbf{x}_k|\mathbf{z}_{1:k}) = \frac{1}{c_k}p(\mathbf{x}_k|\mathbf{z}_{1:k-1})p(\mathbf{z}_k|\mathbf{x}_k), \quad (4)$$

where the normalizing constant

$$\begin{aligned} c_k &= p(\mathbf{z}_k|\mathbf{z}_{1:k-1}) \\ &= \int_{\mathbb{R}^{n_x}} p(\mathbf{x}_k|\mathbf{z}_{1:k-1})p(\mathbf{z}_k|\mathbf{x}_k)d\mathbf{x}_k \end{aligned} \quad (5)$$

and n_x is the state-vector dimension.

Approximate Solutions: Two Approaches

- Moment-closing algorithms:
  - Kushner's nonlinear Bayesian filter (IEEE Trans. AC, 2000)
  - Grid filters
  - Particle filters
    (Gordon, Salmond & Smith, 1993)
- Innovations-based algorithms:
  - Extended Kalman filter
    (Schmidt, 1961)
  - Unscented Kalman filter
    (Julier, Ulhmann & Durrant-Whyte, 2000)
  - Central Difference Kalman filter
    (Norgaard, Poulson & Ravn, 2000)
  - Gauss-Hermite quadrature filter
    (Ito & Xiong, 2000)
- Problem statement: Develop an approximate BF that is theoretically motivated, reasonably accurate and easily extendable at a minimal cost.

Cubature Kalman Filtering

- Tradeoff global optimality for computational tractability and robustness.

- Key assumption: Represent the joint state-innovations density given the past measurement history as Gaussian.

- New problem: Compute integrals whose integrands are of the form:

$$\text{Nonlinear function} \times \text{Gaussian}$$

- The monomial-based cubature rule is chosen as a 'good' candidate for numerical computations!

Transformation to Spherical-Radial Integration

- Integral of interest:

$$I(\mathbf{f}) = \int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x})\exp(-\mathbf{x}^T\mathbf{x})d\mathbf{x} \tag{6}$$

- Key step: Transform $I(.)$ in the Cartesian coordinate into the spherical-radial coordinate.
- We may thus write the radial integral

$$I_r = \int_0^\infty S(r)r^{n-1}\exp(-r^2)dr \tag{7}$$

where $S(r)$ is defined by the spherical integral

$$S(r) = \int_{U_n} \mathbf{f}(r\mathbf{y})d\sigma(\mathbf{y}) \tag{8}$$

with $\sigma(.)$ is the spherical surface measure on the region

$$U_n = \{\mathbf{y} \in \mathbb{R}^n |\ \mathbf{y}^T\mathbf{y} = 1\}$$

Monomial-based Cubature Rules

- Fix the degree of the target cubature rule to be three.
- The 3rd degree monomial-based cubature rule for spherical integration:

$$\int_{U_n} \mathbf{f}(r\mathbf{s})d\sigma(\mathbf{s}) \approx \sum_{i=1}^{2n} \omega_s \mathbf{f}[u]_i. \tag{9}$$

- The 1st degree quadrature rule for radial integration:

$$\int_0^{\infty} f(r) r^{n-1} \exp(-r^2) dr \approx \omega_r f(x_r). \tag{10}$$

- The resulting 3rd degree spherical-radial cubature rule is written as

$$\int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x}) \exp(-\mathbf{x}^T\mathbf{x}) d\mathbf{x} \approx \sum_{i=1}^{2n} \underbrace{\omega_s \omega_r}_{\omega} \mathbf{f}(\underbrace{[x_r u]_i}_{\xi_i}). \tag{11}$$

Square-Root Filtering for Reliability

- Key idea: Reformulate the CKF so as to propagate the square-roots of covariances
- Why?
  - Preserves symmetry and positive (semi)definiteness
  - Improves numerical accuracy due to $\kappa_{\sqrt{P}} = \sqrt{\kappa_P}$
  - Doubles the order of precision
  - Makes square-roots available
- How do we do this?
  - Use triangular factorization (e.g., QR decomposition) for covariance updates
  - Replace matrix inversion with forward (backward) substitution
- Cost: 60% more computations!

Hallmark Properties of the (Square-root) CKF

- Property 1: Derivative-free
- Property 2: The number of function evaluations increases linearly with n_x
- Property 3: Computational cost grows cubically w.r.t. n_x
- Property 4: Extraction of second-order information of the hidden state embedded in the measurements at best
- Property 5: Approximation to the Bayesian filter that closely inherits the properties of the linear Kalman filter including square-root filtering for improved reliability in limited word-length systems

Model-based Signal Processing

- Goal: Given a set of noisy observations, build an empirical model for the following purposes:
  - To denoise a test signal– signal enhancement
  - To statistically decide whether the denoised test signal belongs to the empirical model– signal detection

- Experimental Setup:
  - Use the chaotic Mackey-Glass system to generate training and tests data
  - Perturb with noise such that the SNR was set to be $3dB$
  - Model the system using the 7-5R-1 recurrent neural network

- Methodology:
  - Enhance the signal using cooperative cubature filtering
  - Detect the signal based on the innovations statistic

CKF-based Cooperative Filtering

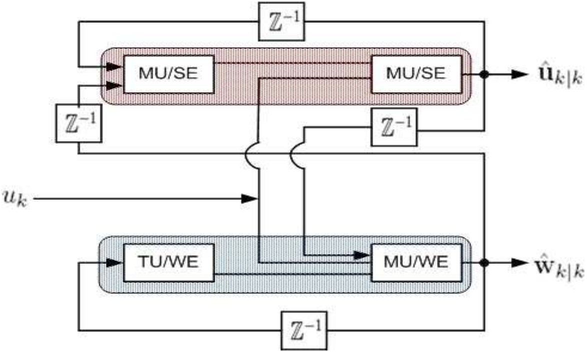


Figure 2: TU- Time Update, MU- Measurement Update, SE- Signal Estimator, WE-Weight Estimator

Representative Denoised Signals

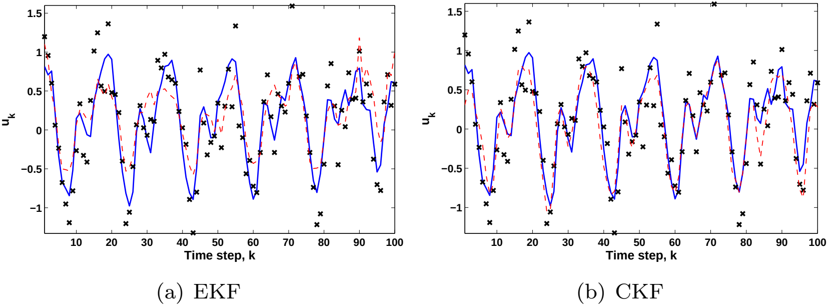
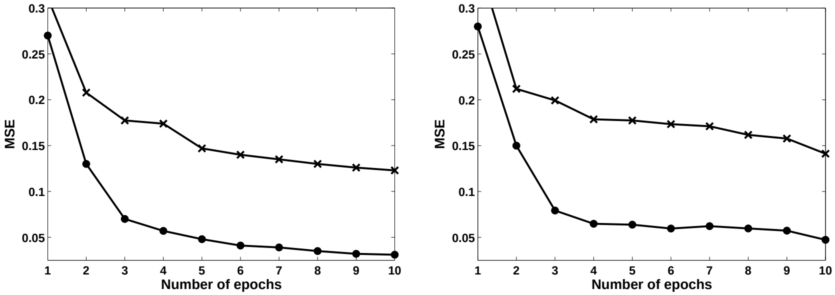


Figure 3: Representative test signal Vs. time step (thick- clean, dotted thin- Filter estimate, x- noisy measurements).

Cooperative Filtering Results



(a) Training Phase (b) Test Phase

Figure 4: Ensemble-averaged (over 50 runs) Mean Squared Error (MSE) Vs. number of epochs (x- EKF, filled circle- CKF).

Signal Detection Using the NIS Statistic

- Consider the detection index to be the normalized innovations squared (NIS), which is defined at time k as:

$$\boldsymbol{\epsilon}_k = [\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}]^T P_{zz,k|k-1}^{-1} [\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}]$$

- Under the Gaussian assumption, $\boldsymbol{\epsilon}_k$ is χ^2-distributed:

$$\boldsymbol{\epsilon}_k \sim \boldsymbol{\chi}^2_{n_\mathbf{z}}$$

- Compute 95% confidence interval to accept/reject the detection hypothesis

Signal Detection Results

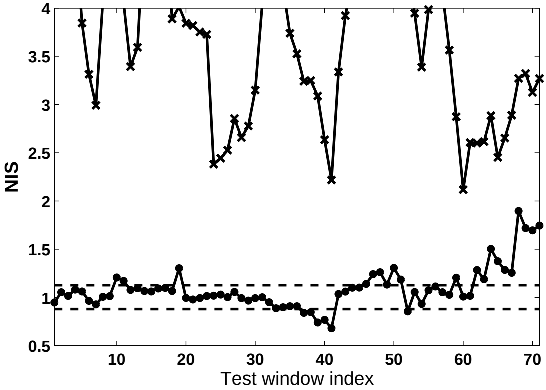


Figure 5: x- EKF, filled circle- CKF, dotted thick- 95% confidence intervals

Related Publications

- I. Arasaratnam and S. Haykin, **Cubature Kalman Filters**, forthcoming *IEEE Trans. Automatic Control*, vol. 54, June 2009.
- I. Arasaratnam and S. Haykin, **Cubature Kalman Filtering: A Powerful Tool for Aerospace Applications**, under review, *Int'l Radar Conf. 2009*, Bordeaux, France, Oct. 2009.
- I. Arasaratnam and S. Haykin, **Square-Root Quadrature Kalman Filtering**, *IEEE Trans. Signal Processing*, vol. 56, no. 6, June 2008.
- I. Arasaratnam, S. Haykin and R. Elliott, **Discrete-Time Nonlinear Filtering Algorithms Using Gauss-Hermite Quadrature**, *Proc. IEEE*, vol. 95, no. 5, pp. 953-977, May 2007.
- I. Arasaratnam and S. Haykin, **Nonlinear Bayesian Filters for Training Recurrent Neural Networks**, Book Ch., *Advances in Artificial Intelligence*, Springer, A. Gelbukh *et al.*, Eds., 2008.
- S. Haykin & I. Arasaratnam, **Nonlinear Sequential State Estimation for Solving Pattern Classification Problems**, *Adaptive Signal Processing: Next Generation Solutions*, T. Adali & S. Haykin, Eds., Book Ch. 6, Wiley, forthcoming 2009.

Thank you!