# An analysis of continuous and discrete time non-linear Kalman Filtering

## Or: How I Learned To Stop Worrying And Love Statistics

Adam Attarian [1]

[1] Department of Mathematics
North Carolina State University

2 December 2008

# Outline

# Problem Context

The Kalman Filter (KF) was developed in the early 60's and was first applied to trajectory and estimation problems. There were however, some immediate drawbacks to the original formulation:

- Assumes a linear state-space model.
- Assumes the underlying distribution to be Gaussian.
- Original filter was in discrete time, requiring an a-priori discretization of the state space model.

As it happens, many (interesting) problems are:

- Nonlinear.
- Nongaussian.
- Continuous time.

Where to go from here?

# Problem Context

The Kalman Filter (KF) was developed in the early 60's and was first applied to trajectory and estimation problems. There were however, some immediate drawbacks to the original formulation:

- Assumes a linear state-space model.
- Assumes the underlying distribution to be Gaussian.
- Original filter was in discrete time, requiring an a-priori discretization of the state space model.

As it happens, many (interesting) problems are:

- Nonlinear.
- Nongaussian.
- Continuous time.

Where to go from here?

# Problem Context

Suboptimal filters were developed to handle these situations. These filters employ

- Linearizations of the state model (Extended KF)
- Approximations of the underlying distribution to a Gaussian pdf (Unscented KF)
- Monte Carlo sampling techniques (Ensemble KF, Particle Filtering)

Questions to consider:

- Which filter to use when?
- Discretize the model first and use a discrete filter?
- (or) Use a continuous time filter and numerically integrate?

# Outline



NC STATE UNIVERSITY

# Discrete Filtering

Assume that initially, we have a continuous dynamic system given by

$$\dot{x} = f(t, x; q) + w(t) \tag{1}$$
$$y = h(t, x) + v(t). \tag{2}$$

We then discretize the system into a discrete-in-time model given by

$$x_{k+1} = f(t_k, x_k; q) + w_k, \quad w \sim \mathcal{N}(0, V) \tag{3}$$
$$y_k = h(t_k, x_k) + v_k, \quad v \sim \mathcal{N}(0, R) \tag{4}$$

where $x_k \in \mathbb{R}^n, y_k \in \mathbb{R}^m, q \in \mathbb{R}^p$ and $w_k, v_k$ are additive white gaussian noise (AWGN) processes. Let $f, h$ be nonlinear for generality.

# Discrete Filtering

To propagate the state from $t_k$ to $t_{k+1}$ we're taking a single step, at a fixed step size. We can use

- Explicit Euler
- Fixed step size Runge-Kutta
- (insert favorite solver here)

$\Rightarrow$ This choice matters. I'm using a fourth order RK integrator with fixed step size.

We wish to estimate the $x_k$ given the noisy observations $y_k$. As $f, h$ are nonlinear, some approximation needs to be made. Option one: approximate $f, h$ with a linearization. This leads to...

# Extended Kalman Filtering

The Extended Kalman Filter (EKF) linearizes the state dynamics around the current estimate. The EKF equations are given by

Prediction Steps:

$$\hat{x}_k^- = \text{rk4}(f, h, \hat{x}_{k-1}) \tag{5}$$

$$P_k^- = \nabla f(\hat{x}_k^-) P_{k-1} \nabla f^T(\hat{x}_k^-) + V \tag{6}$$

Correction Steps:

$$K_k = P_k \nabla h^T(\hat{x}_k^-) \left[ \nabla h(\hat{x}_k^-) P_k \nabla h^T(\hat{x}_k^-) + R \right]^{-1} \tag{7}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - h_k(\hat{x}_k^-)) \tag{8}$$

$$P_k = \left[ I - K_k \nabla h(\hat{x}_k^-) \right] P_k^- \tag{9}$$

# Extended Kalman Filtering

Where $P$ is the covariance of the state estimate, $\hat{x}$. $\hat{x}_k^-$ and $P_k^-$ refer to the a-priori estimates of $x$ and $P$, respectively conditioned on all prior measurements. Remarks:

- In the EKF, the state distribution is approximated by a gaussian random variable (GRV), which is then propagated through the linearization.
- This can introduce large errors in the true posterior mean and covariance of the transformed GRV.
- The more nonlinear the system is, the worse the approximation and filter performance.
- Very sensitive to initial conditions of the filter.

$\Rightarrow$ Can we do better?

# Unscented Kalman Filtering

- The Unscented Kalman Filter (UKF) is built around the idea that it is easier to approximate the underlying distribution than it is to approximate the state dynamics.
- Uses a deterministic sampling approach to approximate the distribution.
- The state distribution is approximated by a GRV, but is represented by a set of sigma points, completely capturing the **true** mean and covariance of the state distribution.
- When propagated through the nonlinear system, the posterior mean and covariance are captured to third order.
- Computational cost is equal to the EKF.

# UKF Equations

$2n + 1$ ($n$ is the state dimension) sigma vectors are generated according to

$$\mathcal{X}_0 = \bar{x} \tag{10}$$

$$\mathcal{X}_i = \bar{x} + \left(\sqrt{(n+\lambda)P_x}\right)_i, \quad i = 1, \ldots, n \tag{11}$$

$$\mathcal{X}_i = \bar{x} - \left(\sqrt{(n+\lambda)P_x}\right)_i, \quad i = n+1, \ldots, 2n \tag{12}$$

where $\mathcal{X}_i$ denotes the $i$-th column of the matrix $\mathcal{X}$.

$\Rightarrow$ These points are where the distribution of $\hat{x}$ are sampled. In practice, Cholesky factors are used as the matrix square root.

# UKF Equations

Each sample point has an associated weight, weighting the mean estimation and the covariance estimation differently. $W^m \in \mathbb{R}^{n \times 2n+1}$ are the weights for the mean, $W^c \in \mathbb{R}^{n \times 2n+1}$ for the covariance estimate.

$$W_0^m = \lambda(n+\lambda)^{-1} \tag{13}$$

$$W_0^c = \lambda(n+\lambda)^{-1} + (1-\alpha^2+\beta) \tag{14}$$

$$W_i^m = W_i^c = \big(2(n+\lambda)\big)^{-1}, \quad i = 1 \ldots 2n \tag{15}$$

where $\lambda, \alpha, \beta$ are all tuning parameters.

# UKF Equations

Prediction

$$\mathcal{X} = \text{sigmapoints}(x_k, P_k) \tag{16}$$

$$\hat{x}_k^- = \sum_{i=0}^{2n} W_i^m \mathcal{X}_i \tag{17}$$

$$\hat{P}_k^- = V + \sum_{i=0}^{2n} W_i^c \left[\mathcal{X}_i - \hat{x}_k^-\right]\left[\mathcal{X}_i - \hat{x}_k^-\right]^T \tag{18}$$

$$\mathcal{X}_k^- = \text{sigmapoints}(x_k^-, P_k^-) \tag{19}$$

$$\mathcal{Y}_k = h(\mathcal{X}_k^-) \tag{20}$$

$$\hat{y}_k = \sum_{i=0}^{2n} W_i^m \mathcal{Y}_i \tag{21}$$

where $\mathcal{X}_i$ and $\mathcal{Y}_i$ denote the $i$-th column.

# UKF Equations

And the update/correction equations . . .

$$P_{\bar{y}_k\bar{y}_k} = R + \sum_{i=0}^{2n} W_i^c \left[\mathcal{Y}_i - \hat{y}_k\right]\left[\mathcal{Y}_i - \hat{y}_k\right]^T \tag{22}$$

$$P_{\bar{x}_k\bar{y}_k} = \sum_{i=0}^{2n} W_i^c \left[\mathcal{X}_{i,k}^- - \hat{y}_k\right]\left[\mathcal{Y}_i - \hat{y}_k\right]^T \tag{23}$$

$$K = P_{\bar{x}_k\bar{y}_k} P_{\bar{y}_k\bar{y}_k}^{-1} \tag{24}$$

$$\hat{x}_k = \hat{x}_k^- + K(z_k - \hat{y}_k) \tag{25}$$

$$P_k = P_k^- - K P_{\bar{y}_k\bar{y}_k} K^T. \tag{26}$$

Note: Matrix versions of these equations are derived and used in practice.

NC STATE UNIVERSITY

# Unscented Kalman Filtering

- The UKF is a recursive implementation of the Unscented Transform (UT), which computes the statistics of a random variable that undergoes a nonlinear transformation.
- Works well on nonlinear problems.
- Similar to particle filters, only with a deterministic sampling method.
- Further numerically robust versions available in the *Square Root Filter*.

# Discrete Filtering Considerations

- If data are sparse, the step size taken can be large, affecting the integration accuracy.
- Dynamics that affect accuracy may be missed by a single step.
- In fixed step size integrators, there is no automatic error control.
- Discretization of the model inherently changes the model to something new.
- Discrete filters are more sensitive to amount and quality of data.

$\Rightarrow$ Solution: Derive the filter equations for a continuous model and do not discretize the dynamics.

NC STATE UNIVERSITY

# Outline

# Continuous Kalman Filtering

The continuous Kalman Filter is known as the Kalman-Bucy Filter.

- Continuous filters do not require an a-priori discretization of the state space dynamics.
- The state space model is augmented with a matrix Riccati equation describing the propagation of the covariance matrix. The augmented system constitutes a system of stochastic differential equations (SDEs).
- Multistep, adaptive mesh integrators can be used for state and covariance prediction, increasing accuracy and increasing information content.
- Maintain the assumption that the observations are discrete in time.

# Extended Kalman-Bucy Filter

The Extended Kalman-Bucy Filter (EKBF) employs an augmented state space,

$$\dot{\hat{x}}(t) = f(\hat{x}(t), t) \tag{27}$$

$$\dot{P}(t) = P(t)\nabla f(\hat{x})^T + \nabla f(\hat{x})P(t) + V. \tag{28}$$

These equations are integrated from $t_k$ to $t_{k+1}$. The correction equations remain the same,

$$K_k = P^-(t_k)\nabla h(\hat{x})^T \left[\nabla h(\hat{x})P^-(t_k)\nabla h(\hat{x})^T + R\right]^{-1} \tag{29}$$

$$P_k = \left[I - K_k \nabla h(\hat{x})\right] P^-(t_k) \tag{30}$$

$$\hat{x}_k = \hat{x}_k^- + K_k\left[z_k - \nabla h(\hat{x})\hat{x}_k^-\right]. \tag{31}$$

# Extended Kalman-Bucy Filter

- The EKBF performs better than the EKF when fewer observations are available, either longitudinally or from issues arising from state observability.
- Tuning the integration tolerances will affect the tracking ability of the filter.

If the problem is too nonlinear, the EKBF will still fail. This motivates the Unscented Kalman-Bucy Filter (UKBF).

# Unscented Kalman-Bucy Filter

The UKBF is a natural extension of the UKF in continuous time. The sigma points become a function of time, and are given as

$$\mathcal{X}(t)_0 = \bar{x}(t) \tag{32}$$

$$\mathcal{X}(t)_i = \bar{x}(t) + \left(\sqrt{(n+\lambda)P(t)_x}\right)_i, \quad i = 1, \ldots, n \tag{33}$$

$$\mathcal{X}(t)_i = \bar{x}(t) - \left(\sqrt{(n+\lambda)P(t)_x}\right)_i, \quad i = n+1, \ldots, 2n \tag{34}$$

# Unscented Kalman-Bucy Filter

The augmented state space model is given by

$$\dot{\hat{x}}(t) = f\big(\mathcal{X}(t), t\big) W^m \tag{35}$$

$$\dot{P}(t) = \mathcal{X}(t) W f^T\big(\mathcal{X}(t), t\big) + f\big(\mathcal{X}(t), t\big) W \mathcal{X}^T(t) + V, \tag{36}$$

where $\mathcal{X}(t)$ is implicitly a function of $\hat{x}(t)$ and $P(t)$ and the matrix $W$ is given by

$$W = \big(I - [W_0^m \cdots W_{2n}^m]\big) \cdot \text{diag}\big(W_c^0 \cdots W_c^{2n}\big) \cdot \big(I - [W_0^m \cdots W_{2n}^m]\big)^T. \tag{37}$$

The correction equations remain the same (omitted for brevity). The state space is integrated from $t_k$ to $t_{k+1}$.

NC STATE UNIVERSITY

# Remarks

If we assume both filters use the same initial conditions and covariance matrices, then we assert:

- For sparse data sets, the continuous filters will outperform the discrete filters under the same filtering conditions.
- For more nonlinear systems, the UK(B)F will outperform the EK(B)F – this is well known.
- The UKF will track the unobserved states better than the EKF.

# Outline

# Lorenz Simulation

The first test model is the Lorenz Attractor. This is a chaotic dynamical system, and can be highly nonlinear. The state dynamics are given by

$$\dot{x}_1 = 10(x_2 - x_3) \tag{38}$$

$$\dot{x}_2 = x_1(28 - x_3) - x_2 \tag{39}$$

$$\dot{x}_3 = x_1 x_2 - \frac{8}{3} x_3 \tag{40}$$

When the solutions are plotted in phase space, you get the familiar "butterfly".

For the following comparisons, the same conditions and covariance matrices were used for both filters.

# Lorenz Dynamics

These are the states that we will attempt to estimate.

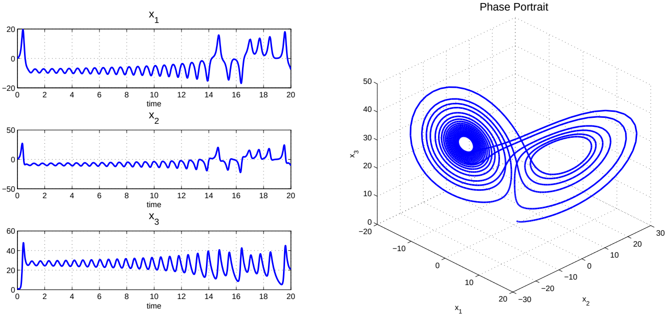


Figure: Time Dynamics of the Lorenz Attractor.

Our data set will be $x_1 +$ AWGN.

# EKF vs EKBF

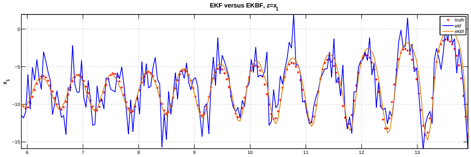


Figure: Estimation of the observed state, $x_1$ (detail), EKF vs EKBF.

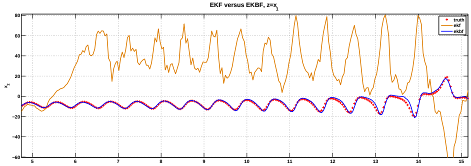


Figure: Estimation of an unobserved state, $x_2$ (detail), EKF vs EKBF.

# EKF vs UKF

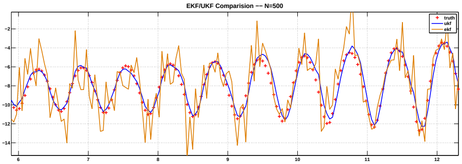


Figure: Estimation of the observed state, $x_1$ (detail), EKF vs UKF.

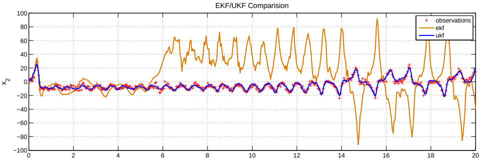


Figure: Estimation of an unobserved state, $x_2$, EKF vs UKF.

# UKF vs UKBF

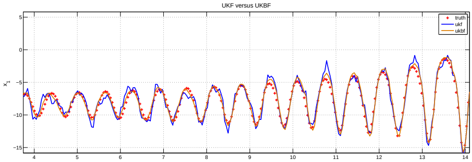


Figure: Estimation of the observed state, $x_1$ (detail), UKF vs UKBF.

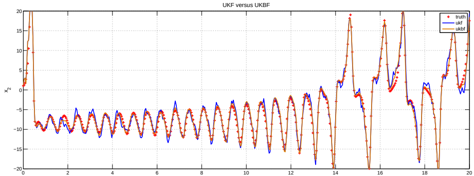


Figure: Estimation of an unobserved state, $x_2$, UKF vs UKBF.

# Indications

- The Lorenz model demonstrates the UK(B)F's superior ability in handling the nonlinearities.
- For cases where one attempts to estimate a state that is not observed, the EKF is the worst performer.
- The continuous filters tracked better in all cases.
- Improved EKF performance may result from filter tuning.

# HIV Dynamics

An acute HIV infection with no treatment can be modeled as

$$\begin{aligned} \dot{T} &= \lambda - dT - kVT \\ \dot{T}^* &= kTV - \delta T^* \\ \dot{V} &= N\delta T^* - cV, \end{aligned} \tag{41}$$

where $T^*$ is infected T-cells, $V$ is free viron particles, $\lambda$ is the recruitment of uninfected T-cells, $d$ is the per capita death rate of uninfected cells, $k$ is the infection rate, $\delta$ is the death rate of uninfected cells, $N$ is the number of new HIV virons and $c$ is the clearance rate.

Collected data could be a combination of viral load ($V$) and healthy T-cell count ($T$).

# EKF/EKBF

The frequency of data collection significantly affects filter performance. For these simulations, we assume weekly blood draws over 100 days. At this collection interval, the EKF diverges, however the EKBF tracks well.

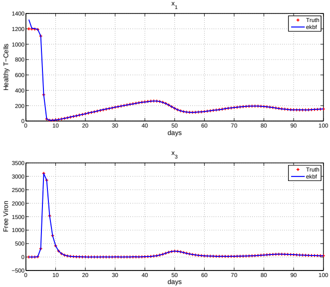


Figure: Estimation of observed states at weekly collection intervals, EKBF.

# UKF/UKBF

The UKF also fails for this collection interval. The UKBF however tracks well.

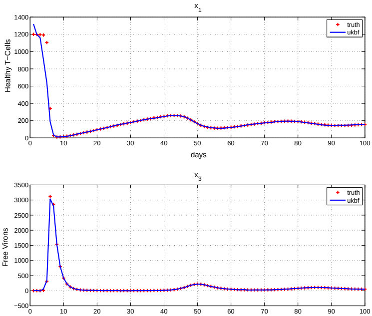


Figure: Estimation of observed states at weekly collection intervals, UKF.

# Indications

- The HIV model demonstrates the superior ability of the continuous filters to perform state estimation when the data are sparsely collected.
- The relatively large step size taken in the discrete models introduces too much error; the continuous time integrators are more robust.
- The EKF and UKF perform comparably on the HIV model, as the model quickly reaches the chronic regime where linearizations are sufficient to model the dynamics.

# Outline



NC STATE UNIVERSITY

# Remarks

- Applying a discrete filter to a discretized system isn't always the best approach.
- The Kalman-Bucy filters can perform substantially better than their discrete counterparts on some models, and are more robust with respect to the data collected.
- The UKF nearly always matches or out performs the EKF.
- The UKBF is a good starting point for many filtering applications.