Convergence of extended Kalman Filter to Locate a Moving target in Wild Life Telemetry

Samanmalee Sugathadasa, samanmal@math.ttu.edu
Clyde Martin [1], martin@math.ttu.edu
W. P. Dayawansa [2], daya@math.ttu.edu
Department of Mathematics
Texas Tech University
Lubbock, Tx 79409

Abstract

In wildlife telemetry and several other applications, some or all of the measurements consist of angles made with very poor accuracy. In these cases classical extended Kalman filtering, where the angles are treated as linear variables, tend to introduce errors in estimation on a sporadic basis due to incorrect handling of the angles. Here we propose a correction term in Kalman filtering equation which will ensure that the angular terms are treated in an invariant fashion.

1 Introduction

Let us consider a randomly moving target being tracked by a set of fixed receivers, each of which (nominally) measures the angular position of the target. Each of the angle measurements is corrupted by additive noise. The problem considered here is, how to estimate the location of the moving target with increasing accuracy using newly available measurements. We assume that measurements are taken at discrete time steps, $T = 0, 1, 2, \cdots$. In realistic situations occurring in wild life telemetry (and many other applications) the angular measurements are extremely inaccurate (with standard deviations up to 20 degrees), and after a few measurements, angles cannot be treated as variables that lie strictly between two set limits. This can be easily understood by recalling that in Gaussian random walk standard deviation increases as the square root of the number of time steps. This has an interesting implication in Kalman filtering. Recall that both in standard Kalman filtering and all its variants the state updates are done using a linear correction term involving the innovations in measurements. In our case, the measurement is an angle, and the concept of linear functions of an angle is by itself faulty since angles are defined modulo 2π only. This wouldn't be a problem if the angular measurements are to remain strictly within some open interval. However, this isn't the case here as we have already described, and care must be taken to make sure that Kalman filter updates are done using formulae which are well defined on the circle. We treat this problem here and describe how to construct a sensible extended Kalman filter.

A representative problem in wild life telemetry is described below. Let us agree upon some notation first. A superscript n above a vector or a scalar indicates association with time step n. Let,

x^n	$=$	True position vector of the target at T = n.
θ_i^n	$=$	Angular location of the target as indicated by the i^{th} receiver at T = n.
α_i^n	$=$	True angular position of the target from the i^{th} receiver at T = n.

Let us assume that the target is going through a random walk. Thus, dynamics of x_n can be expressed as,

$$x^{n+1} \ = \ x^n \ + \ \sigma^n, \ \ n = 0, 1, 2, \cdots, \tag{1}$$

where, $\{\sigma^n\}_{n=0,1,2,\dots}$ are random vectors representing the random movement of the target.

Let us express measurements as,

$$\theta_i^n \ = \ \alpha_i^n \ + \ \eta_i^n, \ \ i = 1, 2, \cdots, K, \ n = 0, 1, 2, \cdots, \tag{2}$$

where, $\{\eta_i^n\}$ are random variables representing measurement inaccuracies.

For the sake of simplicity let us assume that $\{\sigma^n\}_{n=0,1,2,\dots}$ have zero means and the covariances and,

$$Cov(\sigma^n, \sigma^m) \ = \ \delta_{n,m} S, \tag{3}$$

where, S is a positive definite matrix, and $\delta_{n,n}$ is the Kronecker delta function. Similarly, let us assume that η_i^n have zero means and,

$$Cov(\eta_i^n, \eta_j^m) \ \ = \ \ R\delta_{i,j}\delta_{n,m}, \tag{4}$$

[1] Research was funded through NSF Grants 9705312 & 9720357

[2] Research was funded through NSF Grants ECS 9707927 & 9720357 and through a State of Texas ARP Grant 003644-0402-1999.

where, R is a positive constant. The objective is to estimate the state x^n from available measurements $\theta_j,\ j = 0, \cdots, n-1$.

Figure 1 illustrate the difficulty with classical extended Kalman filtering in this case. Here, the standard deviation of the target movement, i.e. S is diag[0.1, 0.1], and the standard deviation of each of the angular measurement errors is 20^0. Three symmetrically placed devices measure the angular locations. In wildlife telemetry it is typical to use least square estimates at *each instant in time*, and for comparison purposes these estimates are included as well. It is clear from the graph that, even though extended Kalman filtering does a better job than instantaneous estimates, its predictions are not very good either since they seem to deviate sporadically from reasonable estimates. We believe that the reason for this is the faulty use of angular measurements.

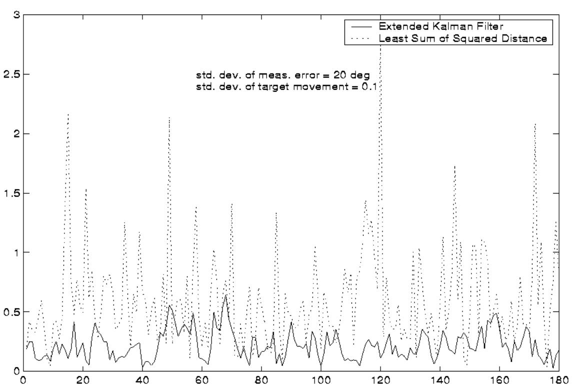

Figure 1: Estimation Error in Kalman Filtering

2 Gaussian Random Variables on the Circle and the Conditional Expectation

As mentioned in the introduction, extended Kalman filtering and its various generalizations fail to be directly applicable in our case since taking linear functions of angles don't make any sense here due to large measurement noise. Here we propose to treat a class of random variables on the circle as representative of measurements and use them in computing conditional expectations and their updates in state estimation.

Let us first consider a random variable Θ which takes values on the circle S^1. By definition Θ is a Borel measurable function from some probability space into S^1. We wish Θ to be representative of a typical angular measurement corrupted by noise. Treating noise as an additive Gaussian random variable, we will construct Θ as the random variable induced by a Gaussian distributed random variable Z upon projection of $\mathbb{R}$ on S^1 via $x \to x \bmod 2\pi$. Treatment of such random variables is described in detail in [1]. For the sake of completeness we will describe derivations very briefly here. Let Z has mean $\bar{z}$ and variance σ^2. Then, the density of Θ is given by,

$$f_\Theta(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{n \to -\infty}^{\infty} e^{-\frac{(\theta - \bar{z} + 2n\pi)^2}{2\sigma^2}}. \qquad (5)$$

This can be expressed using the Riemann theta function (see e.g. [4],[1]). Recall that the Riemann theta function ϑ is defined by,

$$\vartheta(v, \tau) = \sum_{-\infty}^{\infty} exp\{2i\pi(\frac{1}{2}n^2\tau + nv)\}. \qquad (6)$$

Here, τ is treated as a parameter, and v is treated as a complex variable. It is known that this series defines an analytic function of v whenever the imaginary part of τ is positive (see e.g. [4]) .

Expression in (5) can be rearranged and simplified as,

$$f_\Theta(\theta) \;=\; \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\theta-\bar{\theta})^2}{2\sigma^2}} \vartheta(i\frac{(\theta-\bar{\theta})}{\sigma^2}, i\frac{2\pi}{\sigma^2}), \qquad (7)$$

where, $\bar{\theta} \;=\; \bar{z} \bmod 2\pi$.

Notice that the parameter τ is equal to $i\frac{2\pi}{\sigma^2}$ in our case, which obviously has positive imaginary part, hence the formula (7) defines an analytic function of θ.

We will treat the density function of Θ given in (5) as the analog of the Gaussian density on the circle S^1, and simply say that Θ is Gaussian distributed on S^1. Its mean is equal to $\bar{\theta} \bmod 2\pi$, and variance is equal to σ^2.

Now let us turn our attention to defining a jointly distributed Gaussian random variable on $\mathbb{R}^n \times S^1$. Again, we will start with a jointy Gaussian random vector (X, Z) defined on $\mathbb{R}^n \times \mathbb{R}$, and project Z down to S^1 via the projection modulo 2π. Let us denote the mean and the covariance of (X, Z) by $(\bar{x}, \bar{z})$, and the covariance matrix by $\Sigma \;=\; \begin{bmatrix} \Sigma_{XX} & \Sigma_{XZ} \\ \Sigma_{ZX} & \Sigma_{ZZ} \end{bmatrix}$. Thus, the joint density function of (X, Z) is,

$$f_{X,Z}(x, z) = \frac{1}{(2\pi)^{\frac{n+1}{2}}\sqrt{det(\Sigma)}} exp(-\frac{1}{2} \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix}' \Sigma^{-1} \begin{bmatrix} x - \bar{x} \\ z - \bar{z} \end{bmatrix}).$$

Now, since $\Theta \;=\; Z \bmod 2\pi$, we get

$$f_{X,\Theta}(x, \theta) \;=\; \frac{1}{(2\pi)^{\frac{n+1}{2}}\sqrt{det(\Sigma)}} \sum_{n=-\infty}^{\infty} exp(-\frac{1}{2} \begin{bmatrix} x - \bar{x} \\ \theta - \bar{z} + 2n\pi \end{bmatrix}' \Sigma^{-1} \begin{bmatrix} x - \bar{x} \\ \theta - \bar{z} + 2n\pi \end{bmatrix}).$$

We will refer to (X, Θ) as jointy Gaussian distributed on $\mathbb{R}^n \times S^1$.

Let us now compute the conditional expectation $E(X|\Theta)$ and the covariance. We emphasize that the answer for $E(X|\Theta)$ should be a periodic function in Θ of period 2π.

By definition,

$$E(X|\Theta) \quad = \quad \int_{\mathbb{R}^n} \frac{x f_{(X,\Theta)}(x,\theta)}{f_\Theta(\theta)} dx, \tag{8}$$

where,

$$f_\Theta(\theta) \quad = \quad \int_{\mathbb{R}^n} f_{X,\Theta}(x,\theta) dx. \tag{9}$$

From standard computation involving Gaussian distributed random variables (see e.g. [3]) it follows that,

$$f_\Theta(\theta) \;=\; \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\frac{(\theta - \bar{z} - 2n\pi)^2}{2\sigma^2}},$$

where, $\sigma = \sqrt{\Sigma_{ZZ}}$.

Again, using the definition of the theta function ϑ we express this as,

$$f_\Theta(\theta) \;=\; \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\theta-\bar{\theta})^2}{2\sigma^2}} \vartheta(i\frac{(\theta-\bar{\theta})}{\sigma^2}, i\frac{2\pi}{\sigma^2}), \tag{10}$$

where, $\bar{\theta} = \bar{z} \bmod 2\pi$.

Also,

$$\int_{\mathbb{R}^n} x f_{(X,\Theta)}(x,\theta) dx \;=\; \frac{1}{\sqrt{2\pi}\sigma} \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\frac{(\theta-\bar{z}+2n\pi)^2}{\sigma^2}}$$
$$(\bar{x} + \Sigma_{XZ}\Sigma_{ZZ}^{-1}(\theta - \bar{z} + 2n\pi))$$
$$= f_\Theta \theta \bar{x} - [\frac{d}{d\theta} f_\Theta \theta]\Sigma_{XZ}.$$

Therefore,

$$E(X|\Theta) \;=$$
$$\bar{x} + \Sigma_{X\Theta}(\frac{\theta - \bar{\theta}}{\sigma^2} - \frac{\partial}{\partial\theta} \log \vartheta(i\frac{(\theta - \bar{\theta}}{\sigma^2}, i\frac{2\pi}{\sigma^2}). \tag{11}$$

The conditional variance is given by $\Sigma_{XX} - \Sigma_{X\Theta}\Sigma_{\Theta\Theta}^{-1}\Sigma_{\Theta X}$, as in the case of jointly distributed Gaussian random variables.

Notice the additive term $\frac{\partial}{\partial\theta} \ln \vartheta(i\frac{(\theta-\bar{\theta})}{\sigma^2}, i\frac{2\pi}{\sigma^2})$. It ensures that $E(X|\Theta)$ is defined Θ modulo 2π. To verify this all we need to do is to use the quasi periodicity property of the Riemann theta function (see e.g. [4]),

$$\vartheta(v + \tau, \tau) = e^{-i\tau\pi} e^{-i2\pi v} \vartheta(v, \tau). \tag{12}$$

In our case (5)) reads,

$$\vartheta(i\frac{(\theta - \bar{\theta} + 2\pi)}{\sigma^2}, i\frac{2\pi}{\sigma^2}) = e^{\frac{2\pi^2}{\sigma^2}} e^{\frac{2\pi(\theta-\bar{\theta})}{\sigma^2}} \vartheta(i\frac{(\theta - \bar{\theta})}{\sigma^2}, i\frac{2\pi}{\sigma^2}).$$

Therefore,

$$\ln \vartheta(i\frac{(\theta - \bar{\theta} + 2\pi)}{\sigma^2}, i\frac{2\pi}{\sigma^2}) =$$
$$\frac{2\pi^2}{\sigma^2} + \frac{2\pi(\theta - \bar{\theta})}{\sigma^2} + \ln \vartheta(i\frac{(\theta - \bar{\theta})}{\sigma^2}, i\frac{2\pi}{\sigma^2}).$$

It now follows at once that the right hand side of (11) takes the same value if we add an integral multiple of 2π to θ or $\bar{\theta}$ or both.

2.1 Case of several angles

Let us briefly consider the case of several angular variables, i.e. θ takes values in T^m, the m-dimensional torus, considered here as an m-fold cartesean product of S^1. Let X be a random vector on $\mathbb{R}^n$ and Θ be a random variable on T^m. As before, we will treat Θ as the projection modulo the lattice $2\pi\mathbb{Z}^m$ on $\mathbb{R}^m$. (Details can be found in [1]). Let Σ denote the covariance matrix of (X, Θ) and denote its inverse by $S = \begin{bmatrix} S_{XX} & S_{X\Theta} \\ S_{\Theta X} & S_{\Theta\Theta} \end{bmatrix}$. Analogous to the scalar case we may compute the joint density as,

$$f_{X\Theta}(x,\theta) =$$
$$exp\{-\frac{1}{2}\begin{bmatrix} x - \bar{x} \\ \theta - \bar{\theta} \end{bmatrix}' S \begin{bmatrix} x - \bar{x} \\ \theta - \bar{\theta} \end{bmatrix}\}\Theta(i(S_{X\Theta}(x - \bar{x}) +$$
$$S_{\Theta\Theta}(\theta - \bar{\theta})), i2\pi S),$$

where, ϑ denote the Riemann theta function on $\mathbb{C}^m$ (see e.g. [4] for details).

We may also compute,

$$f_\Theta(\theta) =$$
$$exp\{-\frac{1}{2}(\theta - \bar{\theta})'\Sigma_\Theta^{-1}(\theta - \bar{\theta})\}\vartheta(i\Sigma_\Theta^{-1}(\theta - \bar{\theta}), i2\pi\Sigma_\Theta^{-1}),$$
$$E(X|\Theta = \theta) =$$
$$\bar{x} + \Sigma_{X\Theta}\Sigma_{Th}^{-1}(\theta - \bar{\theta}) - \Sigma_{X\Theta}\frac{\partial}{\partial\theta}$$
$$\ln \Theta(i\Sigma_\Theta^{-1}(\theta - \bar{\theta}), i2\pi\Sigma_\Theta^{-1}). \tag{13}$$

Once again it is easily seen using properties of the Riemann theta function that $E(X|\Theta = \theta) = E(X|\Theta = \theta + 2\pi N)$ for arbitrary $N \in \mathbb{Z}^m$.

3 Kalman Filtering with Angular Measurements

Let us now consider a control system,

$$X_{k+1} \;=\; A_k X_k + B_k U_k + D_k \xi_k$$

$$Y_k = \rho(C_k X_k + \omega_k), \tag{14}$$

where, X denote the state ($X_k \in \mathbb{R}^n$), Y denote the output ($Y_k \in T^m$), U denote the input ($U_k \in \mathbb{R}^p$), and ξ_k and ω_k are Gaussian noise processes. Here $\rho : \mathbb{R}^m \to T^m$ denote the projection modulo the lattice $2\pi\mathbb{Z}^m$.

We make standard simplifying assumptions on noise processes, e.g. $E(\xi_k) = 0$, $E(\omega_k) = 0$, $E(\xi_k, \xi_j) = \delta_{j,k}\Xi_k$ and $E(\omega_k, \omega_j) = \delta_{j,k}\Omega_k$. We also assume that the initial state X_0 is uncorrelated with ξ and ω processes.

For the sake of convenience let us define $\Theta_k = (Y_0, \cdots, Y_k) \in T^{(k+1)m}$. Ideally, we would like to compute $E(X_{k+1}|\Theta_k)$, $k = 0, 1, \cdots$. In principle this is provided by (13). However, for practical considerations it is much more convenient to have a recursive formula to compute $E(X_{k+1}|\Theta_k)$ which updates $E(X_{k+1}|\Theta_{k-1})$ using new information available in Θ_k. Unfortunately no such formula exists here. Below we propose an approximation.

Recall the recursion in the classical Kalman filter (see e.g.[3, 2]),

$$\begin{aligned} \hat{X}_{k+1|k} &= A_k\hat{X}_{k|k-1} + B_k U_k + L_k(Z_k - C_k\hat{X}_{k|k-1}), \\ \Sigma_{k+1} &= A_k(\Sigma_k - \Sigma_k C_k^{'}(C_k\Sigma_k C_k^{'} + \Omega_k)^{-1} C_k \Sigma_k) A_k^{'} \\ &\quad + D_k \Xi_k D_k^{'}, \end{aligned} \tag{15}$$

with initial conditions, $\hat{X}_{0|-1} = E(X_0)$ and $\Sigma_0 = cov(X_0, X_0)$, and where $L_k = A_k\Sigma_k C_k^{'}[C_k\Sigma_k C_k^{'} + \Omega_k]^{-1}$.

The crucial point for our purposes is that $L_k(Z_k - C_k\hat{X}_{k|k-1})$ represents update using new information available in the k^{th} measurement. In view of (13) we propose to replace this by,

$$\begin{aligned} &L_k(\Theta_k - \rho(C_k\hat{X}_{k|k-1})) - A_k\Sigma_k C_k^{'}\frac{\partial}{\partial\theta_k}\ln \\ &\vartheta\{i(C_k\Sigma_k C_k^{'} + \Omega_k]^{-1}(\Theta_k - \rho(C_k\hat{X}_{k|k-1})), \\ &i2\pi[C_k\Sigma_k C_k^{'} + \Omega_k]^{-1}\}. \end{aligned}$$

This correction ensures that the update formula does not depend upon the choices of angles modulo 2π, hence allow us to define Kalman filter equations in such a way that it is well defined on T^m.

This approximation turns out to be reasonable if the update at *each step* is relatively small in comparison to 2π. This is a reasonable assumption in cases such as ours when the relative change in the angular measurement is of the order of 20^0 or so. A detailed analysis will be reported shortly elsewhere.

We reiterate that the noise of the magnitude discussed here quickly adds up causing serious problems in the classical Kalman filter, and we have successfully dealt with this problem here.

Thus, the proposed approximate Kalman filter for (14) is,

$$\begin{aligned} \hat{X}_{k+1|k} &= A_k\hat{X}_{k|k-1} + B_k U_k + \\ &\quad Q_k R_k^{-1}(\Theta_k - \rho(C_k\hat{X}_{k|k-1})) - \\ &\quad Q_k\frac{\partial}{\partial\theta_k} \\ &\quad \ln\vartheta\{i(R_k^{-1}(\Theta_k - \rho(C_k\hat{X}_{k|k-1})), i2\pi R_k^{-1}\}. \\ \Sigma_{k+1} &= A_k(\Sigma_k - \Sigma_k C_k^{'}(C_k\Sigma_k C_k^{'} + \Omega_k)^{-1}C_k\Sigma_k)A_k^{'} \\ &\quad + D_k\Xi_k D_k^{'}, \end{aligned} \tag{16}$$

with initial conditions, $\hat{X}_{0|-1} = E(X_0)$ and $\Sigma_0 = cov(X_0, X_0)$, where, $Q_k = A_k\Sigma_k C_k^{'}$, $R_k = (C_k\Sigma_k C_k^{'} + \Omega_k)$, and $L_k = Q_k R_k^{-1}$.

4 Extended Kalman Filter in Wildlife Telemetry

Dynamic equations that describe our problem is not of the form described in (14) since the output function cannot be represented as the projection of a linear function. In extended Kalman filtering, at each time step the system is replaced by its linear approximation at the current estimate of the state, and updated using this linear model. In our work we have implemented the extended Kalman filter using classical Kalman filter algorithm (15), and the modified algorithm (16). It is our experience that at relatively large levels of noise measurements the classical Kalman Filter tend to loose its ability to track once in a while. This problem is alleviated in the modified extended Kalman filter. Interested reader can obtain further information at the web site www.math.ttu.edu\ ~samanmal\kalman.

References

[1] K. V. Mardia, **Statistics of Directional Data," Academic Press**, 1972.

[2] I. B. Rhodes, "A Tutorial Introduction to Estimation and Filtering," *IEEE Transactions of Automatic Control,* Vol. AC-16, No. 6, December 1971, pp 688-706.

[3] B. D. O. Anderson and J. B. Moore, **Optimal Filtering**, Prentice Hall, 1979.

[4] H. M. Farkas, **Riemann Surfaces**, Springer-Verlag, 1980.

[5] S. J. Julier, J. K. Uhlman and H. F. Durrant-Whyte, "A New Approach for Filtering Nonlinear Systems," *Proc. American Control Conf.*, Seattle, Washington, 1628-1632, 1995.