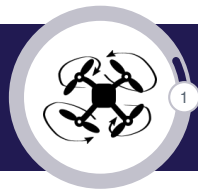
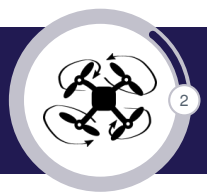


# Kalman Filter

## Nonlinear Systems



- ▶ Extended Kalman Filter (EKF)
- ▶ Unscented Kalman Filter (UKF)
- ▶ UKF with state constraints



# Kalman Filter

## Nonlinear Systems

We will consider the following generic non-linear, discrete-time dynamical system with additive noise,

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{G}\mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{v}_k\end{aligned}$$

and initial state  $\mathbf{x}_0$ . We keep the noise description the same as in the linear context,  $E[\mathbf{w}_k] = \mathbf{0}_n$ ,  $E[\mathbf{v}_k] = \mathbf{0}_m$ ,  $E[\mathbf{w}_k \mathbf{w}_k^T] = \mathbf{Q} \in \mathbb{R}^{n \times n}$ , and  $E[\mathbf{v}_k \mathbf{v}_k^T] = \mathbf{R} \in \mathbb{R}^{m \times m}$ .

# Kalman Filter

## Nonlinear Systems, Observability



- ▶ Generally input dependent (system can be observable under a set of inputs and not observable under other inputs)
- ▶ A much less straightforward concept to test for non-linear systems, but the question remains very valid. The estimation problem might be ill-posed for particular combinations of dynamics, and measurements.



# Kalman Filter

## Extended Kalman Filter

### ► Extended Kalman Filter Predict Step:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{f}(\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1})$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}_k \mathbf{P}_{k-1} \mathbf{A}_k^T + \mathbf{G} \mathbf{Q} \mathbf{G}^T$$

Where

$$\mathbf{A}_k = \left. \frac{\partial \mathbf{f}(\mathbf{x}, \mathbf{u})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k-1}, \mathbf{u}_{k-1}}$$

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\hat{\mathbf{x}}_{k|k-1}}$$

### ► Kalman Filter Update Step:

$$\hat{\mathbf{y}}_k = \mathbf{h}(\hat{\mathbf{x}}_{k|k-1}); \mathbf{P}_y = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T; \mathbf{P}_{xy} = \mathbf{P}_{k|k-1} \mathbf{H}_k^T$$

$$\mathbf{K}_k = \mathbf{P}_{xy} (\mathbf{P}_y + \mathbf{R})^{-1}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \hat{\mathbf{y}}_k)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1}$$

# Kalman Filter

## EKF vs UKF



- ▶ For the EKF, at each time step, the Jacobian of the nonlinear function is evaluated with current predicted states. These matrices can be then be used in the “standard” Kalman filter equations. The process essentially linearizes the non-linear function around the current estimate.
- ▶ Notice also, that the nonlinear functions  $\mathbf{f}$  and  $\mathbf{h}$  could be used in their original (non-linearized) form in the dynamics prediction step and for the measurement prediction. We only needed a linearized approximation for covariance estimation/calculation.
- ▶ The UKF comes with another solution to the problem of calculating these covariance terms in the non-linear context.



# Kalman Filter

## UKF, The Unscented Transform

- ▶ Consider a random variable  $\mathbf{a} \in \mathbb{R}^n$ , with mean  $E[\mathbf{a}] = \bar{\mathbf{a}}$  and (auto-)covariance  $E[(\mathbf{a} - \bar{\mathbf{a}})(\mathbf{a} - \bar{\mathbf{a}})^T] = \mathbf{C}_a$ .
- ▶ Next consider a non-linear function  $\mathbf{g} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , and consider the characteristics of the random variable  $\mathbf{b} = \mathbf{g}(\mathbf{a})$ .
- ▶ The UT is a method for approximating the mean and auto-covariance of the random variable  $\mathbf{b}$ , when we know the function  $\mathbf{g}$ , but only the mean and auto-covariance of  $\mathbf{a}$  (we would need to know the entire pdf of  $\mathbf{a}$  to attempt an exact computation).



# Kalman Filter

UKF, The Unscented Transform

- ▶ It is a type of heuristic
- ▶ The main idea is to draw a finite set of **well-selected and deterministically chosen** samples points from the random variable  $\mathbf{a}$ , apply the non-linear function  $\mathbf{g}$  to these samples, and use a type of sample mean and sample covariance as approximates for true statistics of the random variable  $\mathbf{b} = \mathbf{g}(\mathbf{a})$ .
- ▶ A main idea with the UT, is to carefully select the sample points and keep their number reduced (not doing Monte Carlo here).
- ▶ The sample points are called **sigma-points**. The name **Sigma-Points Kalman Filter** is also used.



# Kalman Filter

## UKF, The Scaled Unscented Transform

- ▶ Scalar parameters  $\alpha, \beta, \kappa$ , e.g.  $\alpha = 10^{-3}, \kappa = 0, \beta = 2$ . Derived scalar quantity  $\lambda = \alpha^2(n + \kappa) - n$ ;
- ▶ Let  $(n + \lambda)\mathbf{C}_a = \mathbf{L}\mathbf{L}^T$ , the Cholesky decomposition of the scaled covariance matrix  $\mathbf{C}_a$  where  $\mathbf{L}$  is lower triangular;
- ▶ The sigma points are defined as:  $\mathbf{s}_0 = \bar{\mathbf{a}}, \mathbf{s}_{i+1} = \bar{\mathbf{a}} + \mathbf{L}[:, i], \mathbf{s}_{i+1+n} = \bar{\mathbf{a}} - \mathbf{L}[:, i], i = 0, \dots, n$ ;
- ▶ Two sets ( set  $m$  and set  $c$ ) of weights are defined as well  $W_0^m = \lambda/(n + \lambda), W_0^c = \lambda/(n + \lambda) + (1 - \alpha^2 + \beta)$ , and  $W_j^m = W_j^c = 1/(2(n + \lambda)), j = 1 \dots 2n$ ;
- ▶ Then  $\bar{\mathbf{b}} = E[\mathbf{g}(\mathbf{a})] \approx \tilde{\mathbf{b}} = \sum_{j=0}^{2n} W_j^m \mathbf{g}(\mathbf{s}_j)$ , and  $\mathbf{C}_b = E[(\mathbf{g}(\mathbf{a}) - \bar{\mathbf{b}})(\mathbf{g}(\mathbf{a}) - \bar{\mathbf{b}})^T] \approx \tilde{\mathbf{C}}_b = \sum_{j=0}^{2n} W_j^c (\mathbf{g}(\mathbf{s}_j) - \tilde{\mathbf{b}})(\mathbf{g}(\mathbf{s}_j) - \tilde{\mathbf{b}})^T$ .



# Kalman Filter

## UKF, Practice



- Generate the sigma-points according to the unscaled unscented transform for the random variable  $\mathbf{a} \in \mathbb{R}^4$  with the following statistics:

$$\bar{\mathbf{a}} = [0 \quad 1 \quad -2 \quad 2]^T$$
$$\mathbf{C}_a = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

# Kalman Filter

## UKF, The Cubature Transform



►  $\alpha = 1$ ,  $\beta = 0$ , and  $\kappa = 0$ .



# Kalman Filter

## Unscented Kalman Filter

- Unscented Kalman Filter Predict Step:

$$\mathbf{x}_j = \mathbf{f}(\mathbf{s}_j(\hat{\mathbf{x}}_{k-1}, \mathbf{P}_{k-1})); j = 0, \dots, 2n$$

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{j=0}^{2n} W_j^m \mathbf{x}_j; \mathbf{P}_{k|k-1} = \sum_{j=0}^{2n} W_j^c (\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$$

- Kalman Filter Update Step:

$$\mathbf{x}_j = \mathbf{s}_j(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}), \mathbf{y}_j = \mathbf{h}(\mathbf{x}_j), j = 0, \dots, 2n$$

$$\hat{\mathbf{y}}_k = \sum_{j=0}^{2n} W_j^m \mathbf{y}_j, \mathbf{C}_y = \sum_{j=0}^{2n} W_j^c (\mathbf{y}_j - \hat{\mathbf{y}}_k)(\mathbf{y}_j - \hat{\mathbf{y}}_k)^T$$



# Kalman Filter

## Unscented Kalman Filter

► Kalman Filter Update Step (cont):

$$\mathbf{C}_{xy} = \sum_{j=0}^{2n} W_j^c (\mathbf{x}_j - \hat{\mathbf{x}}_{k|k-1}) (\mathbf{y}_j - \hat{\mathbf{y}}_k)^T$$

$$\mathbf{K}_k = \mathbf{C}_{xy} (\mathbf{C}_y + \mathbf{R})^{-1}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{y}_k - \hat{\mathbf{y}}_k)$$

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} - \mathbf{K}_k (\mathbf{C}_y + \mathbf{R}) \mathbf{K}_k^T$$

► with some given initial conditions  $\hat{\mathbf{x}}_0$ ,  $\mathbf{P}_0$ .

There is an alternative form where the sigma-points are extracted from an augmented state including the noise.

# Kalman Filter

UKF, Joseph Form



- ▶ Joseph form implements directly for EKF
- ▶ Joseph form for UKF makes use of the following approximation

$$\mathbf{H} = \mathbf{C}_{xy}^T \mathbf{P}_{k|k-1}^{-1};$$



# EKF and UKF

## Lecture 17

Occasional problems with divergence. Some potential fixes:

- ▶ [Perea et al. 2007] “Artificial” increases in covariance, e.g. B-EKF1:  
 $\mathbf{C}_y = \alpha \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}$ ,  $\alpha \in \mathbb{R}^+$ , e.g.  $\alpha = 2$ ; see also B-EKF2, B-EKF3, B-EKF4 in this paper;
- ▶ UKFz [Perea et al. 2007]: Replacing  $\hat{\mathbf{y}}_k$  from  $\hat{\mathbf{y}}_k = \sum_{j=0}^{2n} W_j^m \mathbf{y}_j$  to  $\mathbf{h}(\hat{\mathbf{x}}^{k|k-1})$ . Affects the calculation of  $\mathbf{C}_y$  and  $\mathbf{C}_{sy}$  and  $\hat{\mathbf{x}}_k$ ;
- ▶ IUKF [Perea et al. 2008]: Use  $\hat{\mathbf{y}}_k$  from  $\hat{\mathbf{y}}_k = \sum_{j=0}^{2n} W_j^m \mathbf{y}_j$  to  $\mathbf{h}(\hat{\mathbf{x}}^{k|k-1})$  but only in  $\hat{\mathbf{x}}_k$ ;



# Interval-Constrained UKF

## Lecture 17

- ▶ [Teixeira et. al 2010] UKF with state interval constraints (a very useful case). See especially the combination where ICUT is used in the predict step and at the end of the update step a PDF truncation procedure is performed (the TIUKF in the paper);
- ▶ ICUT is an interval constraint sigma-point generation procedure;
- ▶ The PDF truncation procedure modifies both the  $\hat{\mathbf{x}}_k$  and  $\mathbf{P}_k$  using an iterative procedure to enforce each of the interval constraints one by one.