TRACKING ALGORITHMS:

An Introduction to Kalman & Particle Filters

- Bayesian framework
- Introduction to tracking
- Kalman Filter
- Nonlinear & Non-Gaussian Problem
- Suboptimal Kalman Filters
  - EKF
  - UKF
- Particle Filter
- Example
- Conclusions

Bayesian Framework

m : model parameter vector (unknown parameters to be estimated)
d : data vector relating to m via an equation h(.)
d = h(m) + noise

Classical parameter estimation framework: Unknown but deterministic m
Bayesian parameter estimation framework: Unknown and random variable m

Bayes' Formula

$$p(m,d) = p(m \mid d)\, p(d) = p(d \mid m)\, p(m)$$

$$\overbrace{p(\mathbf{m}|\mathbf{d})}^{\text{POSTERIOR}} = \frac{p(\mathbf{d}|\mathbf{m})\, p(\mathbf{m})}{p(\mathbf{d})} = \frac{\overbrace{p(\mathbf{d}|\mathbf{m})}^{\text{LIKELIHOOD}}\, \overbrace{p(\mathbf{m})}^{\text{PRIOR}}}{\underbrace{\int p(\mathbf{d}|\mathbf{m})\, p(\mathbf{m})\, d\mathbf{m}}_{\text{EVIDENCE}}}$$

Inversion vs. Tracking

Inversion

$$\mathbf{d}^{obs} = h(\mathbf{m}) + \mathbf{e}$$

Forward model

- $\mathbf{m}$: state vector
- $\mathbf{d}^{obs}$: measurement vector
- $\mathbf{e}$: measurement noise vector

PPD: $\mathrm{p}(\mathbf{m} \mid \mathbf{d})$

Tracking

parameter evolution model

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_k) \quad \text{state equation}$$

$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{w}_k) \quad \text{measurement equation}$$

Forward model

- $\mathbf{x}_k$: state vector
- $\mathbf{y}_k$: measurement vector
- $\mathbf{v}_k$: process/state noise vector
- $\mathbf{w}_k$: measurement noise vector

$$\mathrm{p}(\mathbf{x}_k \mid \mathbf{X}_{k-1}, \mathbf{Y}_k)$$

$$\mathbf{X}_{k-1} = \mathbf{x}_{k-1}, \cdots, \mathbf{x}_0 \qquad \mathbf{Y}_k = \mathbf{y}_k, \cdots, \mathbf{y}_0$$

Dynamic, non-stationary system

What is a dynamic, non-stationary system?

vector of parameters we want to track

parameter evolution model

$$\boxed{\begin{aligned} x_t &= f_t(x_{t-1}, v_t) \\ y_t &= h_t(x_t, w_t) \end{aligned}}$$

state equation

measurement equation

measurement

forward model

$\mathbf{x}_t$: state vector

$\mathbf{y}_t$: measurement vector

$\mathbf{v}_t$: process noise vector

$\mathbf{w}_t$: measurement noise vector

Bayesian framework

$\mathbf{x}_t$, $\mathbf{y}_t$, $\mathbf{v}_t$, $\mathbf{w}_t$: random variables

Why do we care?

x: position and speed of a missile, **y** : sensor measurement

x: changing ocean properties, **y** : acoustic measurement

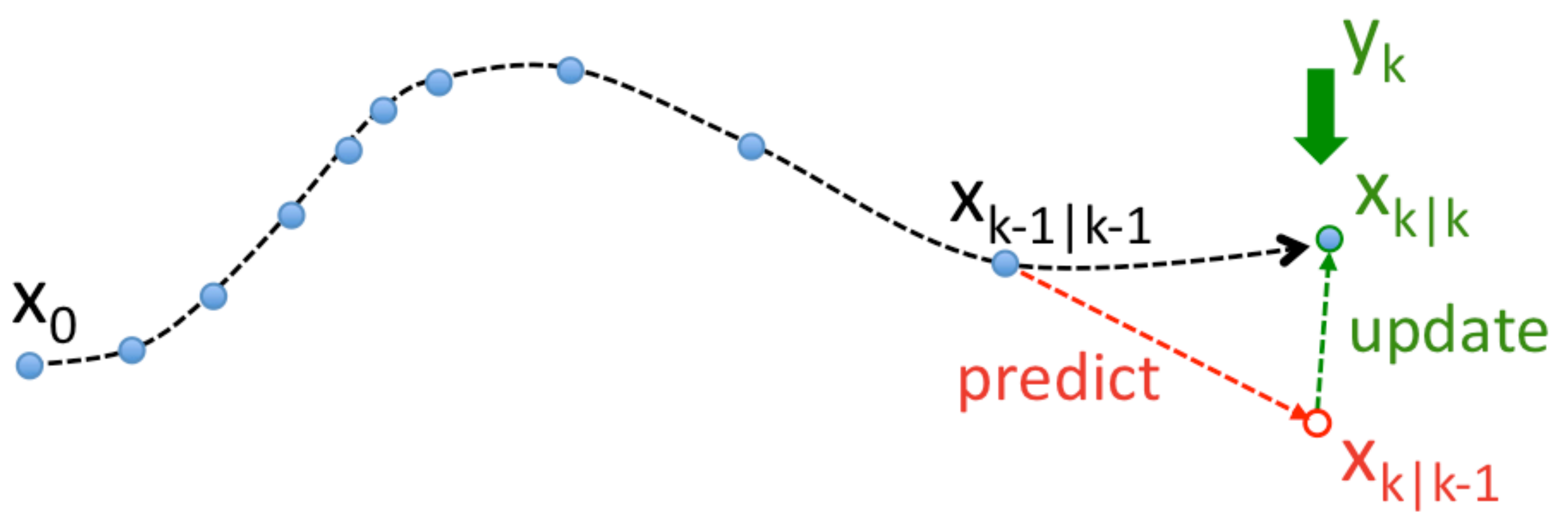
x: financial indicators of stock exchange, **y** : stock prices

x: atmospheric refractivity profile, **y** : radar clutter measurement

x: number of whales in the region, **y** : visual and acoustic measurements

How do we solve it and what does the solution look like?

KF/PFs offer solutions to dynamical systems, nonlinear in general, using prediction and update as data becomes available. Tracking in time or space offers an ideal framework for studying KF/PF.



Kalman Framework

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_k)$$ state equation

$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{w}_k)$$ measurement equation

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$ state equation

$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$ measurement equation

$\mathbf{x}_k$, $\mathbf{y}_k$, $\mathbf{v}_k$, $\mathbf{w}_k$: Gaussian

$\mathbf{F}_k$, $\mathbf{H}_k$: Linear

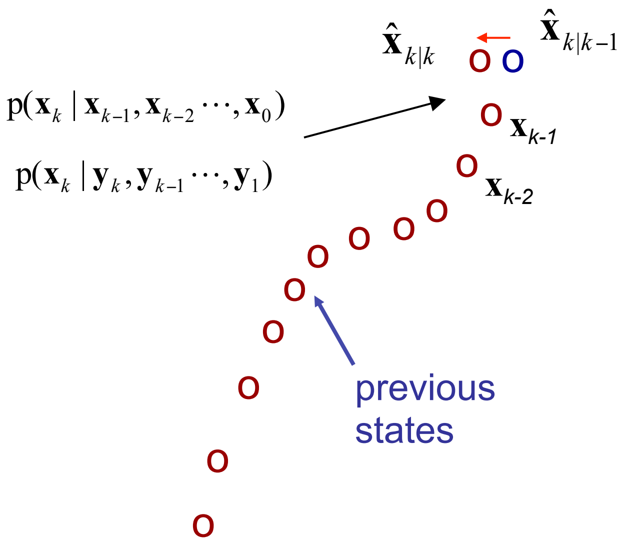
Optimal Filter = Kalman Filter

1963

A Single Kalman Iteration

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$

$$\mathbf{x}_{k|k} \sim \mathcal{N}(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k})$$



PREDICT

1. Predict the mean $\hat{\mathbf{x}}_{k|k-1}$ using previous history.

   $p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$

   $$\hat{\mathbf{x}}_{k|k-1} = \mathrm{E}\{\mathbf{x}_k \mid \mathbf{x}_{k-1}\} = \int \mathbf{x}_k \, p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) d\mathbf{x}_k$$

2. Predict the covariance $\mathbf{P}_{k|k-1}$ using previous history.

UPDATE

3. Correct/update the mean using new data $\mathbf{y}_k$

   $p(\mathbf{x}_k \mid \mathbf{Y}_k)$

   $$\hat{\mathbf{x}}_{k|k} = \mathrm{E}\{\mathbf{x}_k \mid \mathbf{Y}_k\} = \int \mathbf{x}_k \, p(\mathbf{x}_k \mid \mathbf{Y}_k) d\mathbf{x}_k$$

4. Correct/update the covariance $\mathbf{P}_{k|k}$ using $\mathbf{y}_k$

$$\cdots \Rightarrow p(\mathbf{x}_{k-1} \mid \mathbf{Y}_{k-1}) \Rightarrow p(\mathbf{x}_k \mid \mathbf{Y}_{k-1}) \Rightarrow p(\mathbf{x}_k \mid \mathbf{Y}_k) \Rightarrow \cdots$$

PREDICTOR-CORRECTOR

DENSITY PROPAGATOR

The Model

Consider the discrete, linear system,

$$\mathbf{x}_{k+1} = \mathbf{M}_k \mathbf{x}_k + \mathbf{w}_k, \quad k = 0, 1, 2, \ldots, \tag{1}$$

where

- $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector at time t_k
- $\mathbf{M}_k \in \mathbb{R}^{n \times n}$ is the state transition matrix (mapping from time t_k to t_{k+1}) or model
- $\{\mathbf{w}_k \in \mathbb{R}^n; k = 0, 1, 2, \ldots\}$ is a white, Gaussian sequence, with $\mathbf{w}_k \sim N(\mathbf{0}, \mathbf{Q}_k)$, often referred to as model error
- $\mathbf{Q}_k \in \mathbb{R}^{n \times n}$ is a symmetric positive definite covariance matrix (known as the model error covariance matrix).

Some of the following slides are from: Sarah Dance, University of Reading

The Observations

We also have discrete, linear observations that satisfy

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad k = 1, 2, 3, \ldots, \tag{2}$$

where

- $\mathbf{y}_k \in \mathbb{R}^p$ is the vector of actual measurements or observations at time t_k
- $\mathbf{H}_k \in \mathbb{R}^{n \times p}$ is the observation operator. Note that this is not in general a square matrix.
- $\{\mathbf{v}_k \in \mathbb{R}^p; k = 1, 2, \ldots\}$ is a white, Gaussian sequence, with $\mathbf{v}_k \sim N(\mathbf{0}, \mathbf{R}_k)$, often referred to as observation error.
- $\mathbf{R}_k \in \mathbb{R}^{p \times p}$ is a symmetric positive definite covariance matrix (known as the observation error covariance matrix).

We assume that the initial state, $\mathbf{x}_0$ and the noise vectors at each step, $\{\mathbf{w}_k\}$, $\{\mathbf{v}_k\}$, are assumed mutually independent.

The Prediction and Filtering Problems

We suppose that there is some uncertainty in the initial state, i.e.,

$$\mathbf{x}_0 \sim N(0, \mathbf{P}_0) \tag{3}$$

with $\mathbf{P}_0 \in \mathbb{R}^{n \times n}$ a symmetric positive definite covariance matrix.

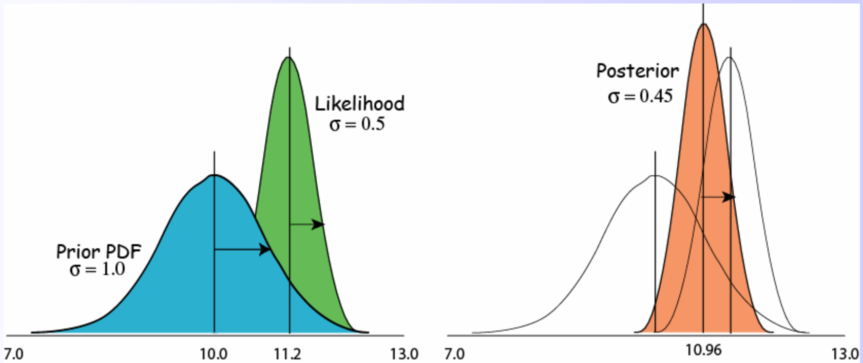
The problem is now to compute an improved estimate of the stochastic variable $\mathbf{x}_k$, provided $\mathbf{y}_1, \ldots \mathbf{y}_j$ have been measured:

$$\widehat{\mathbf{x}}_{k|j} = \widehat{\mathbf{x}}_{k|y_1,\ldots,y_j}. \tag{4}$$

- When $j = k$ this is called the filtered estimate.
- When $j = k - 1$ this is the one-step predicted, or (here) the predicted estimate.

- The Kalman filter (Kalman, 1960) provides estimates for the linear discrete prediction and filtering problem.
- We will take a minimum variance approach to deriving the filter.
- We assume that all the relevant probability densities are Gaussian so that we can simply consider the mean and covariance.
- Rigorous justifcation and other approaches to deriving the filter are discussed by Jazwinski (1970), Chapter 7.

Product of Gaussians=Gaussian:



One data point problem

For the general linear inverse problem we would have

Prior: $$p(\boldsymbol{m}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{m}-\boldsymbol{m}_o)^T C_m^{-1}(\boldsymbol{m}-\boldsymbol{m}_o)\right\}$$

Likelihood: $$p(\boldsymbol{d}|\boldsymbol{m}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{d}-G\boldsymbol{m})^T C_d^{-1}(\boldsymbol{d}-G\boldsymbol{m})\right\}$$

Posterior PDF

$$\propto \exp\left\{-\frac{1}{2}[(\boldsymbol{d}-G\boldsymbol{m})^T C_d^{-1}(\boldsymbol{d}-G\boldsymbol{m}) + (\boldsymbol{m}-\boldsymbol{m}_o)^T C_m^{-1}(\boldsymbol{m}-\boldsymbol{m}_o)]\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left[\mathbf{m}-\hat{\mathbf{m}}\right]^T \mathbf{S}^{-1}\left[\mathbf{m}-\hat{\mathbf{m}}\right]\right\}$$

$$\mathbf{S}^{-1} = \mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_m^{-1}$$

$$\hat{\mathbf{m}} = \left(\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_m^{-1}\right)^{-1}\left(\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{d} + \mathbf{C}_m^{-1}\mathbf{m}_0\right)$$

$$= \mathbf{m}_0 + \left(\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_m^{-1}\right)^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\left(\mathbf{d} - \mathbf{G}\mathbf{m}_0\right)$$

Prediction step

We first derive the equation for one-step prediction of the mean using the state propagation model (1).

$$\begin{aligned}\widehat{\mathbf{x}}_{k+1|k} &= \mathbb{E}\left[\mathbf{x}_{k+1}|\mathbf{y}_1, \ldots \mathbf{y}_k\right], \\ &= \mathbb{E}\left[\mathbf{M}_k\mathbf{x}_k + \mathbf{w}_k\right], \\ &= \mathbf{M}_k\widehat{\mathbf{x}}_{k|k} \end{aligned} \tag{5}$$

The one step prediction of the covariance is defined by,

$$\mathbf{P}_{k+1|k} = \mathbb{E}\left[(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}_{k+1|k})(\mathbf{x}_{k+1} - \widehat{\mathbf{x}}_{k+1|k})^T|\mathbf{y}_1, \ldots \mathbf{y}_k\right]. \tag{6}$$

Exercise: Using the state propagation model, (1), and one-step prediction of the mean, (5), show that

$$\mathbf{P}_{k+1|k} = \mathbf{M}_k\mathbf{P}_{k|k}\mathbf{M}_k^T + \mathbf{Q}_k. \tag{7}$$

Filtering Step

At the time of an observation, we assume that the update to the mean may be written as a linear combination of the observation and the previous estimate:

$$\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k|k-1}), \tag{8}$$

where $\mathbf{K}_k \in \mathbb{R}^{n \times p}$ is known as the Kalman gain and will be derived shortly.

But first we consider the covariance associated with this estimate:

$$\mathbf{P}_{k|k} = \mathbb{E}\left[(\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k})(\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k})^T | \mathbf{y}_1, \ldots \mathbf{y}_k\right]. \tag{9}$$

Using the observation update for the mean (8) we have,

$$\begin{aligned}
\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k} &= \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k|k-1}) \\
&= \mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k(\mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k - \mathbf{H}_k\widehat{\mathbf{x}}_{k|k-1}), \\
&\quad \text{replacing the observations with their model equivalent,} \\
&= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)(\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1}) - \mathbf{K}_k\mathbf{v}_k. \qquad (10)
\end{aligned}$$

Thus, since the error in the prior estimate, $\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1}$ is uncorrelated with the measurement noise we find

$$\begin{aligned}
\mathbf{P}_{k|k} &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbb{E}\left[(\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \widehat{\mathbf{x}}_{k|k-1})^T\right](\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T \\
&\quad + \mathbf{K}_k\mathbb{E}\left[\mathbf{v}_k\mathbf{v}_k^T\right]\mathbf{K}_k^T. \qquad (11)
\end{aligned}$$

Simplification of the a posteriori error covariance formula

Using this value of the Kalman gain we are in a position to simplify the Joseph form as

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}. \tag{15}$$

Exercise: Show this.

Note that the covariance update equation is independent of the actual measurements: so $\mathbf{P}^{k|k}$ could be computed in advance.

Summary of the Kalman filter

Prediction step

Mean update: $\widehat{\mathbf{x}}_{k+1|k} = \mathbf{M}_k \widehat{\mathbf{x}}_{k|k}$

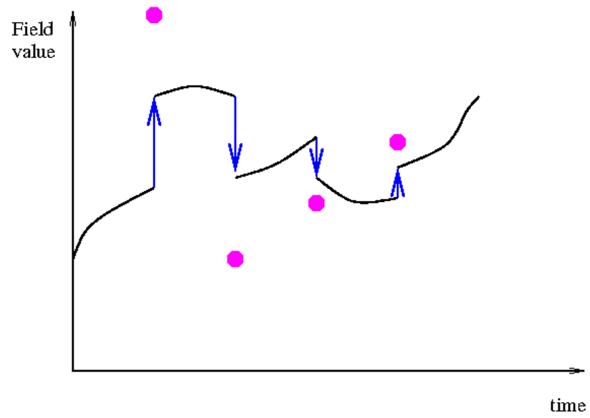
Covariance update: $\mathbf{P}_{k+1|k} = \mathbf{M}_k \mathbf{P}_{k|k} \mathbf{M}_k^T + \mathbf{Q}_k.$

Observation update step

Mean update: $\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k \widehat{\mathbf{x}}_{k|k-1})$

Kalman gain: $\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}^T + \mathbf{R}_k)^{-1}$

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



Nonlinear   Non-Gaussian

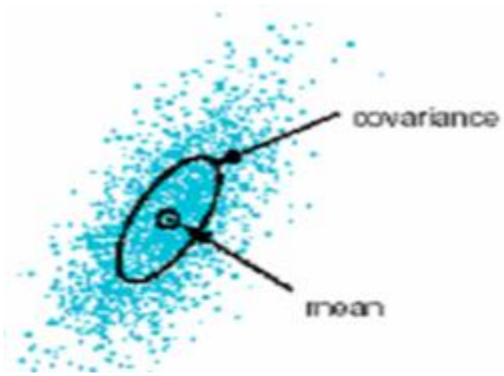
$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$

KF

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_k)$$
$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{w}_k)$$

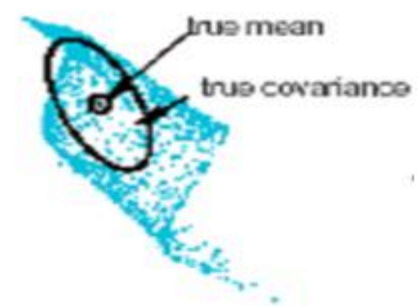
Now What?

x



$$\xrightarrow{\quad \boxed{y = h(x)} \quad}$$

y



- h(.) linear
  Gaussian x ➔ Gaussian y
- h(.) non-linear
  Gaussian x ➔ non-Gaussian y

Extended Kalman Filter (EKF)

If Nonlinear → LINEARIZE!

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}, \mathbf{v}_k)$$
$$\mathbf{y}_k = h_k(\mathbf{x}_k, \mathbf{w}_k)$$

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}) + \mathbf{v}_k$$
$$\mathbf{y}_k = h_k(\mathbf{x}_k) + \mathbf{w}_k$$

↓

$$\mathbf{x}_k = f_{k-1}(\mathbf{x}_{k-1}) + \mathbf{v}_k \cong \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$
$$\mathbf{y}_k = h_k(\mathbf{x}_k) + \mathbf{w}_k \cong \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$

where $\mathbf{H}_k = \left.\dfrac{\partial h_k}{\partial \mathbf{x}}\right|_{\mathbf{x}_k}$ & $\mathbf{F}_{k-1} = \left.\dfrac{\partial f_{k-1}}{\partial \mathbf{x}}\right|_{\mathbf{x}_{k-1}}$

↓

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_k$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{w}_k$$

Gaussian and Linear again! → now use KF