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 \( h(\cdot) \)
\( d = h(m) + \text{noise} \)

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

Bayes’ Formula

\[
p(m, d) = p(m \mid d)p(d) = p(d \mid m)p(m)
\]

\[
p(m \mid d) = \frac{p(d \mid m)p(m)}{p(d)} = \frac{\int p(d \mid m)p(m) \, dm}{\int \int p(d \mid m)p(m) \, dm}
\]

POSTERIOR

LIKELIHOOD

PRIOR

EVIDENCE
Inversion vs. Tracking

**Inversion**

\[ d^{obs} = h(m) + e \]

Forward model

**Tracking**

Parameter evolution model

\[ x_k = f(x_{k-1}, v_k) \]  
State equation

\[ y_k = h_k(x_k, w_k) \]  
Measurement equation

\[ X_{k-1} = x_{k-1}, \ldots, x_0 \]

PPD: \[ p(m \mid d) \]

\[ Y_k = y_k, \ldots, y_0 \]

\[ p(x_k \mid X_{k-1}, Y_k) \]

\[ m \text{ : state vector} \]

\[ d^{obs} \text{ : measurement vector} \]

\[ e \text{ : measurement noise vector} \]

\[ x_k \text{ : state vector} \]

\[ y_k \text{ : measurement vector} \]

\[ v_k \text{ : process/state noise vector} \]

\[ w_k \text{ : measurement noise vector} \]
Dynamic, non-stationary system

What is a dynamic, non-stationary system?

\[ x_t = f_t(x_{t-1}, v_t) \]

state equation

\[ y_t = h_t(x_t, w_t) \]

measurement equation

Why do we care?

- **x**: position and speed of a missile,
- **x**: changing ocean properties,
- **x**: financial indicators of stock exchange,
- **x**: atmospheric refractivity profile,
- **x**: number of whales in the region,
- **y**: sensor measurement
- **y**: acoustic measurement
- **y**: stock prices
- **y**: radar clutter measurement
- **y**: visual and acoustic measurements

**vector of parameters**

**parameter evolution model**

**forward model**

- **x**: state vector
- **y**: measurement vector
- **v**: process noise vector
- **w**: measurement noise vector

Bayesian framework

\( x_t, y_t, v_t, w_t \): random variables
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.

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

\[
x_k = f_{k-1}(x_{k-1}, v_k) \quad \text{state equation}
\]

\[
y_k = h_k(x_k, w_k) \quad \text{measurement equation}
\]

\[
x_k = F_k x_{k-1} + v_k \quad \text{state equation}
\]

\[
y_k = H_k x_k + w_k \quad \text{measurement equation}
\]

\[x_k, y_k, v_k, w_k: \text{Gaussian}\]
\[F_k, H_k: \text{Linear}\]

Optimal Filter = Kalman Filter \hspace{1cm} 1963
A Single Kalman Iteration

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

\[
p(x_k | x_{k-1})
\]

\[
\hat{x}_{k|k-1} = E\{x_k | x_{k-1}\} = \int x_k p(x_k | x_{k-1}) dx_k
\]

2. Predict the covariance \( P_{k|k-1} \) using previous history.

3. Correct/update the mean using new data \( y_k \)

\[
p(x_k | Y_k)
\]

\[
\hat{x}_{k|k} = E\{x_k | Y_k\} = \int x_k p(x_k | Y_k) dx_k
\]

4. Correct/update the covariance \( P_{k|k} \) using \( y_k \)

\[\vdots \Rightarrow p(x_{k-1} | Y_{k-1}) \Rightarrow p(x_k | Y_{k-1}) \Rightarrow p(x_k | Y_k) \Rightarrow \vdots\]
The Model

Consider the discrete, linear system,

\[ x_{k+1} = M_k x_k + w_k, \quad k = 0, 1, 2, \ldots, \]  

where

- \( x_k \in \mathbb{R}^n \) is the state vector at time \( t_k \)
- \( M_k \in \mathbb{R}^{n \times n} \) is the state transition matrix (mapping from time \( t_k \) to \( t_{k+1} \)) or model
- \( \{w_k \in \mathbb{R}^n; k = 0, 1, 2, \ldots\} \) is a white, Gaussian sequence, with \( w_k \sim \mathcal{N}(0, Q_k) \), often referred to as model error
- \( Q_k \in \mathbb{R}^{n \times n} \) is a symmetric positive definite covariance matrix (known as the model error covariance matrix).
The Observations

We also have discrete, linear observations that satisfy

$$y_k = H_k x_k + v_k, \quad k = 1, 2, 3, \ldots,$$

where

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

We assume that the initial state, $x_0$ and the noise vectors at each step, $\{w_k\}, \{v_k\}$, are assumed mutually independent.
The Prediction and Filtering Problems

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

\[ x_0 \sim N(0, P_0) \]  

(3)

with \( P_0 \in \mathbb{R}^{n \times n} \) a symmetric positive definite covariance matrix.

The problem is now to compute an improved estimate of the stochastic variable \( x_k \), provided \( y_1, \ldots, y_j \) have been measured:

\[ \hat{x}_{k|j} = \hat{x}_{k|y_1,\ldots,y_j}. \]  

(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 justification and other approaches to deriving the filter are discussed by Jazwinski (1970), Chapter 7.
Product of Gaussians=$\text{Gaussian}$:

For the general linear inverse problem we would have

Prior: 
\[ p(m) \propto \exp \left\{ -\frac{1}{2}(m - m_o)^T C_m^{-1} (m - m_o) \right\} \]

Likelihood: 
\[ p(d|m) \propto \exp \left\{ -\frac{1}{2}(d - Gm)^T C_d^{-1} (d - Gm) \right\} \]

Posterior PDF
\[ \propto \exp \left\{ -\frac{1}{2}[(d - Gm)^T C_d^{-1} (d - Gm) + (m - m_o)^T C_m^{-1} (m - m_o)] \right\} \]

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

\[ S^{-1} = G^T C_d^{-1} G + C_m^{-1} \]

\[ \hat{m} = \left( G^T C_d^{-1} G + C_m^{-1} \right)^{-1} \left( G^T C_d^{-1} d + C_m^{-1} m_0 \right) \]

\[ = m_0 + \left( G^T C_d^{-1} G + C_m^{-1} \right)^{-1} G^T C_d^{-1} (d - Gm_0) \]
Prediction step

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

\[
\hat{x}_{k+1|k} = E [x_{k+1} | y_1, \ldots, y_k], \\
= E [M_k x_k + w_k], \\
= M_k \hat{x}_{k|k}
\]  

(5)
The one step prediction of the covariance is defined by,

\[ P_{k+1|k} = \mathbb{E} \left[ (x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T \mid y_1, \ldots, y_k \right]. \quad (6) \]

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

\[ P_{k+1|k} = M_k P_{k|k} M_k^T + Q_k. \quad (7) \]
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:

\[
\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - H_k \hat{x}_{k|k-1}) ,
\]

where \( K_k \in \mathbb{R}^{n \times p} \) is known as the Kalman gain and will be derived shortly.
But first we consider the covariance associated with this estimate:

\[ P_{k|k} = \mathbb{E} \left[ (x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T | y_1, \ldots, y_k \right]. \]  

(9)

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

\[
\begin{align*}
 x_k - \hat{x}_{k|k} &= x_k - \hat{x}_{k|k-1} - K_k(y_k - H_k \hat{x}_{k|k-1}) \\
 &= x_k - \hat{x}_{k|k-1} - K_k(H_k x_k + v_k - H_k \hat{x}_{k|k-1}), \\
 &\quad \text{replacing the observations with their model equivalent,} \\
 &= (I - K_k H_k)(x_k - \hat{x}_{k|k-1}) - K_k v_k. \\
\end{align*}
\]

(10)

Thus, since the error in the prior estimate, \( x_k - \hat{x}_{k|k-1} \) is uncorrelated with the measurement noise we find

\[
\begin{align*}
 P_{k|k} &= (I - K_k H_k) \mathbb{E} \left[ (x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T \right] (I - K_k H_k)^T \\
 &\quad + K_k \mathbb{E} \left[ v_k v_k^T \right] K_k^T. \\
\end{align*}
\]

(11)
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

\[ P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k R_k K_k^T = (I - K_k H_k) P_{k|k-1}. \]  

(15)

Exercise: Show this.

Note that the covariance update equation is independent of the actual measurements: so \( P_{k|k} \) could be computed in advance.
Summary of the Kalman filter

Prediction step
Mean update:
\[ \hat{x}_{k+1|k} = \mathbf{M}_k \hat{x}_{k|k} \]
Covariance update:
\[ P_{k+1|k} = \mathbf{M}_k P_{k|k} \mathbf{M}_k^T + Q_k. \]

Observation update step
Mean update:
\[ \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (y_k - H_k \hat{x}_{k|k-1}) \]
Kalman gain:
\[ K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \]
Covariance update:
\[ P_{k|k} = (I - K_k H_k) P_{k|k-1}. \]
Nonlinear Non-Gaussian

\[
x_k = F_{k-1} x_{k-1} + v_k
\]
\[
y_k = H_k x_k + w_k
\]

KF

\[
x_k = f_{k-1}(x_{k-1}, v_k)
\]
\[
y_k = h_k(x_k, w_k)
\]

Now What?

- \(h(.)\) linear
  - Gaussian \(x\) \(\Rightarrow\) Gaussian \(y\)
- \(h(.)\) non-linear
  - Gaussian \(x\) \(\Rightarrow\) non-Gaussian \(y\)
Extended Kalman Filter (EKF)

If Nonlinear $\Rightarrow$ LINEARIZE!

$x_k = f_{k-1}(x_{k-1}, v_k)$
$y_k = h_k(x_k, w_k)$

$x_k = f_{k-1}(x_{k-1}) + v_k$
$y_k = h_k(x_k) + w_k$

$x_k = f_{k-1}(x_{k-1}) + v_k \approx F_{k-1}x_{k-1} + v_k$
$y_k = h_k(x_k) + w_k \approx H_kx_k + w_k$

where

$H_k = \frac{\partial h_k}{\partial x} \bigg|_{x_k}$
$F_{k-1} = \frac{\partial f_{k-1}}{\partial x} \bigg|_{x_{k-1}}$

$x_k = F_{k-1}x_{k-1} + v_k$
$y_k = H_kx_k + w_k$

Gaussian and Linear again! $\Rightarrow$ now use KF