Adaptive Kalman Filters - group 4

Cameron Penne, Giacomo Melloni, Bjørnar Kaarevik, Aria Alinejad, Andreas Sitorus, Dimitris Voskakis, Aafan Ahmad Toor

Outline

- Kalman filter theory
  - Errors
  - Optimization
- Adaptive Kalman methods
  - Maybeck’s adaptive estimation approach
  - Forgetfulness
- Examples
- Advantages and disadvantages
- When to use

Kalman Filters

- System Model:

$$\mathbf{X}(k) = \mathbf{F}(k)\, \mathbf{X}(k\text{-}1) + w(k)$$

  X: state variables

  F: state transition model

  w: noise term

- Measurement Process:

$$\mathbf{z}(k) = \mathbf{H}^T(k)\, \mathbf{X}(k) + v(k)$$

  z: observation or response

  H: observation model

  v: noise term

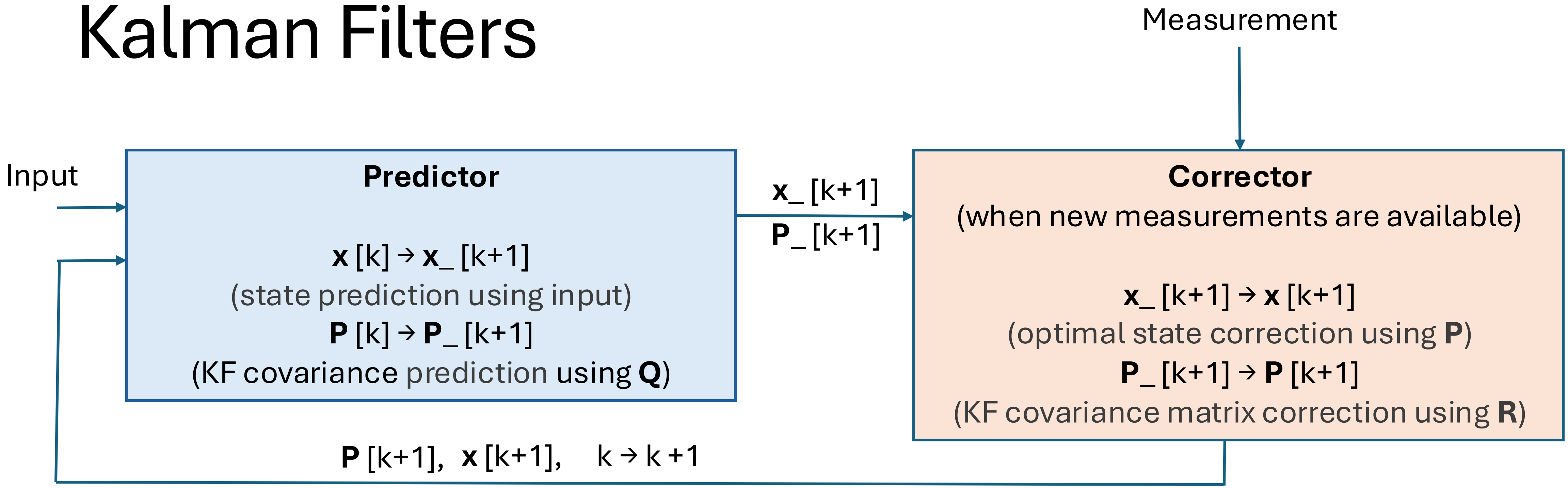
Kalman Filters

$$\mathbf{X}(k) = \mathbf{F}(k)\, \mathbf{X}(k\text{-}1) + w(k)$$

$$\mathbf{z}(k) = \mathbf{H}^{T}(k)\, \mathbf{X}(k) + v(k)$$

- Noise terms w(k) and v(k
  - Assumed to be zero mean white noise
- **Q**(k): covariance of noise from the system model
- **R**(k): covariance of noise from the measurement process

Kalman Filters



- Regular Kalman Filter **assumes stationary processes**

  Q: Process covariance　　　　**R**: Measurement covariance

- What if the system degrades, or the environment changes?

  Q(t)　　　　**R**(t)

Kalman Filter Errors

- Types of errors:
  - There are time dependent states but time invariance is assumed
  - Incorrect values for the observation model **H**(k)
  - Incorrect number of values for the observation model **H**(k)

- Want to obtain an optimal filter by reducing the errors
  - Measure optimality with an innovation sequence v(k):
    - $v(k) = \mathbf{z}(k) - \mathbf{H}^{T}(k)\ \mathbf{X}(k \mid k-1)$
  - From the measurement process

Kalman Filter Optimization

- Kalman filters can be optimized by modifying:
  - State transition model $\mathbf{F}(k)$
  - Observation model $\mathbf{H}(k)$
- Multiple methods exist for the optimization problem
- Precise method depends on knowledge of the noise covariance terms $\mathbf{Q}(k)$ and $\mathbf{R}(k)$

There are many methods for adapting Q and R

Multiple Model Adaptive Estimation	• Use different models for different circumstances and switch between them!
Maximum Likelihood Estimation	• Set up a likelihood function $L(Q, R)$ for the matrices and optimize!
Covariance matching	• Eliminate the state and *use the measurements* to estimate either Q or R!
Correlation methods	• Use the autocorrelation of *the innovation* which should be Gaussian!

Adaptive Kalman Filters – Problem definition

- Kalman filters use **two main model errors**:
  - The system model errors
  - The measurement model errors
- Errors in system model occurs if the system model assumes things stay the same, but in actuality not.
- Errors in measurement model happens when it does not fully capture the reality of the system.
- To compensate for these errors, adaptive methods use **innovation sequence.**

AKF with Maybeck's adaptive estimation approach

Key Variables and Initial Conditions

- $\hat{x}(0)$: Initial estimate of the state vector, set to the expected initial state $E[x(0)]$
- $\boldsymbol{P}(0)$: Initial error covariance matrix, representing the uncertainty of the initial state estimate
- $\widehat{\boldsymbol{Q}}(0)$ and $\widehat{\boldsymbol{R}}(0)$: Initial estimates for the noises covariances.

AKF beigns with these initial conditions and updates them at each time step k

Initialize with

$$\hat{\mathbf{x}}(0) = \mathbf{E}[\mathbf{x}(0)],$$

$$\mathbf{P}(0) = \mathbf{E}\left[(\mathbf{x}(0) - \hat{\mathbf{x}}(0))(\mathbf{x}(0) - \hat{\mathbf{x}}(0))^{\mathrm{T}}\right],$$

$$\widehat{\mathbf{Q}}(0) = \mathbf{Q}\_0,$$

$$\widehat{\mathbf{R}}(0) = \mathbf{R}\_0.$$

AKF with Maybeck's adaptive estimation approach

UPDATING Q

- $\hat{Q}(k)$ is updated using the previous estimate $\hat{Q}(k-1)$, the secondary state error $\tilde{x}(k)\tilde{x}(k)^T$, and the current error covariance $\hat{P}(k)$.
- $\hat{Q}$ **estimate 2nd term** represents a model correction term.
- The term $\frac{1}{L_Q}$ acts as a smoothing actor or window size for averaging.

For k = 1,2, …∞, Maybeck's estimator is

$$\tilde{\mathbf{x}}(k) = \hat{\mathbf{x}}(k) - \hat{\mathbf{x}}^{-}(k),$$

$\hat{Q}$ update

$$\hat{\mathbf{Q}}(k) = \hat{\mathbf{Q}}(k-1) + \frac{1}{L_{\mathbf{Q}}}\Big(\tilde{\mathbf{x}}(k)\tilde{\mathbf{x}}(k)^{\mathbf{T}} + \mathbf{P}(k)$$

$\hat{Q}$ estimate 2nd term

$$-\boldsymbol{\Phi}_{k,k-1}\mathbf{P}(k-1)\boldsymbol{\Phi}^{\mathbf{T}}_{k,k-1} - \hat{\mathbf{Q}}(k-1)\Big),$$

$$\tilde{\mathbf{z}}(k) = \mathbf{z}(k) - \mathbf{H}(k)\hat{\mathbf{x}}(k),$$

$$\hat{\mathbf{R}}(k) = \hat{\mathbf{R}}(k-1) + \frac{1}{L_{\mathbf{R}}}\Big(\tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}(k)^{\mathbf{T}}$$

$$-\mathbf{H}(k)\mathbf{P}(k)\mathbf{H}(k)^{\mathbf{T}} - \hat{\mathbf{R}}(k-1)\Big),$$

AKF with Maybeck's adaptive estimation approach

UPDATING R

- $\tilde{z}(k) = z(k) - H(k)\hat{x}(k)$ is the innovation sequence
- $\hat{R}(k)$ is updated using the previous estimate $\hat{R}(k-1)$, the mesurement residual $\tilde{z}(k)\tilde{z}(k)^T$, and the adjustment term involving $P(k)$ and $H(k)$.
- The term $\frac{1}{L_R}$ acts as a smoothing actor or window size for averaging.

For k = 1,2, …∞, Maybeck's estimator is

$$\bar{\mathbf{x}}(k) = \hat{\mathbf{x}}(k) - \bar{\mathbf{x}}^{-}(k),$$

$$\widehat{\mathbf{Q}}(k) = \widehat{\mathbf{Q}}(k-1) + \frac{1}{L_{\mathbf{Q}}}\Big(\bar{\mathbf{x}}(k)\,\bar{\mathbf{x}}(k)^{\mathbf{T}} + \mathbf{P}(k)$$

$$-\boldsymbol{\Phi}_{k,k-1}\mathbf{P}(k-1)\,\boldsymbol{\Phi}_{k,k-1}^{\mathbf{T}} - \widehat{\mathbf{Q}}(k-1)\Big),$$

$$\tilde{\mathbf{z}}(k) = \mathbf{z}(k) - \mathbf{H}(k)\,\hat{\mathbf{x}}(k),$$

$\hat{R}$ update

$$\widehat{\mathbf{R}}(k) = \widehat{\mathbf{R}}(k-1) + \frac{1}{L_{\mathbf{R}}}\Big(\tilde{\mathbf{z}}(k)\,\tilde{\mathbf{z}}(k)^{\mathbf{T}}$$

$\hat{R}$ estimate 2nd term

$$-\mathbf{H}(k)\,\mathbf{P}(k)\,\mathbf{H}(k)^{\mathbf{T}} - \widehat{\mathbf{R}}(k-1)\Big),$$

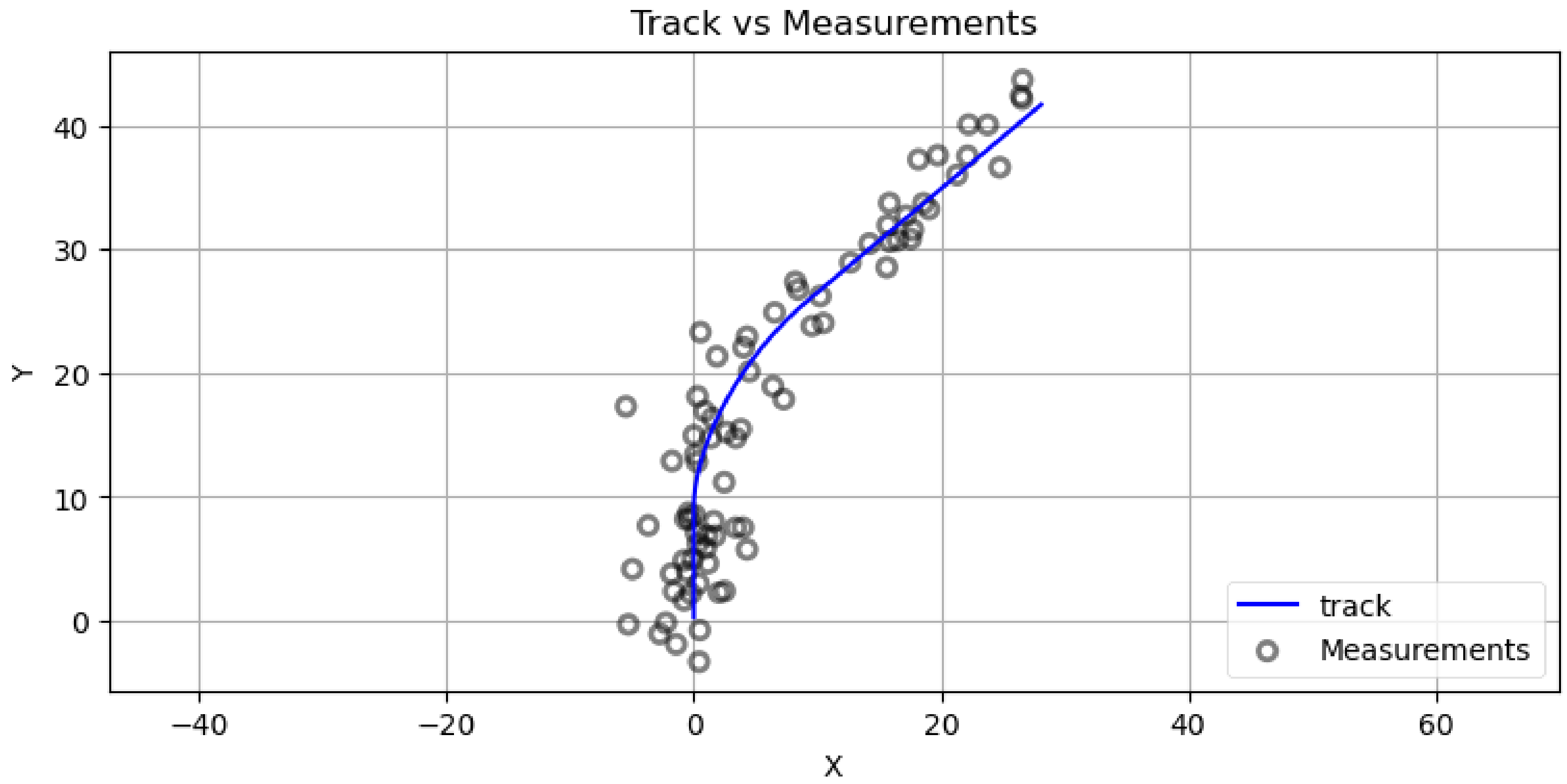
In the example, we will use «forgetfulness»

- Q and R with the innovation and prediction update
  - Innvoation: $v_k = z_k - H_k^T \hat{x}_{k|k-1}$
  - Update: $K_k v_k = \hat{x}_{k|k} - \hat{x}_{k|k-1}$
- Forgetting factor λ which is «close to 1»
  - We use $\lambda = 0.95$

- $Q_k = \lambda Q_{k-1} + (1-\lambda)(K_k v_k)(K_k v_k)^T$
- $R_k = \lambda R_{k-1} + (1-\lambda) v_k v_k^T$

Example - Model manouvering Target

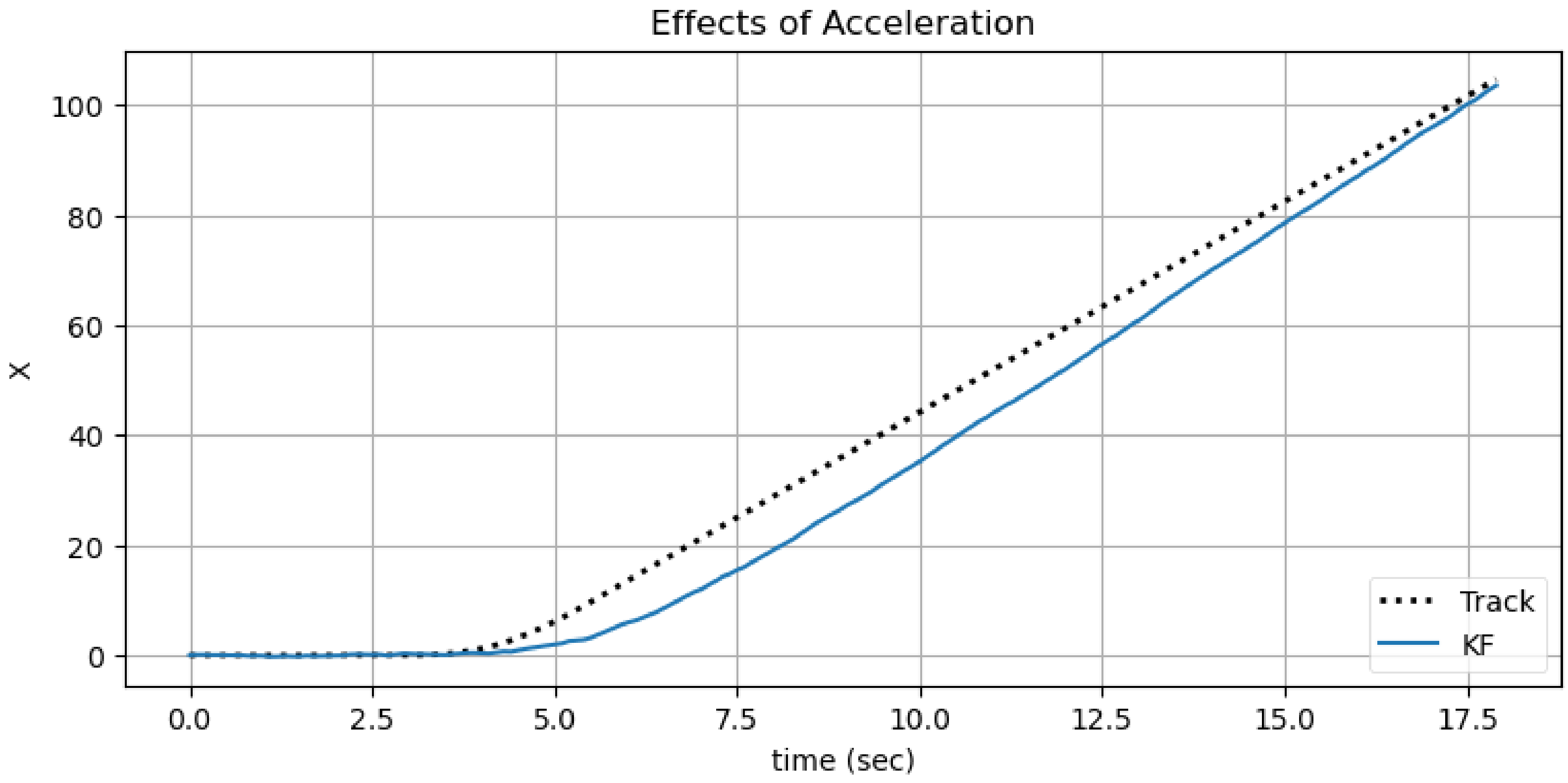
- We only look at x-position for simplicity



Source: [Kalman-and-Bayesian-Filters-in-Python/14-Adaptive-Filtering.ipynb at master · rlabbe/Kalman-and-Bayesian-Filters-in-Python]()

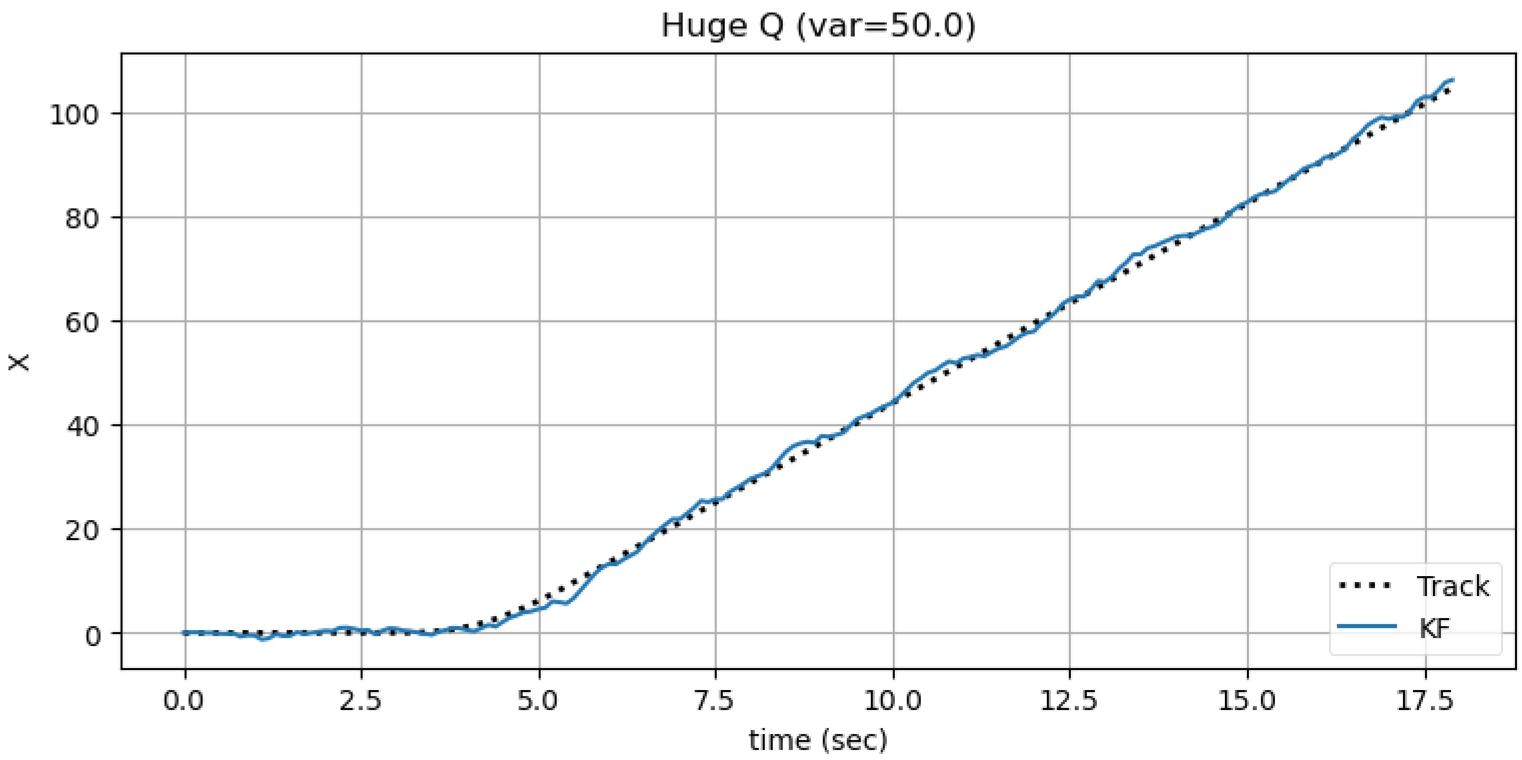
Model with constant velocity

- Kalman predictions with this assumtions are off,
  - But converge back as the velocity goes back to constant
- How to fix?

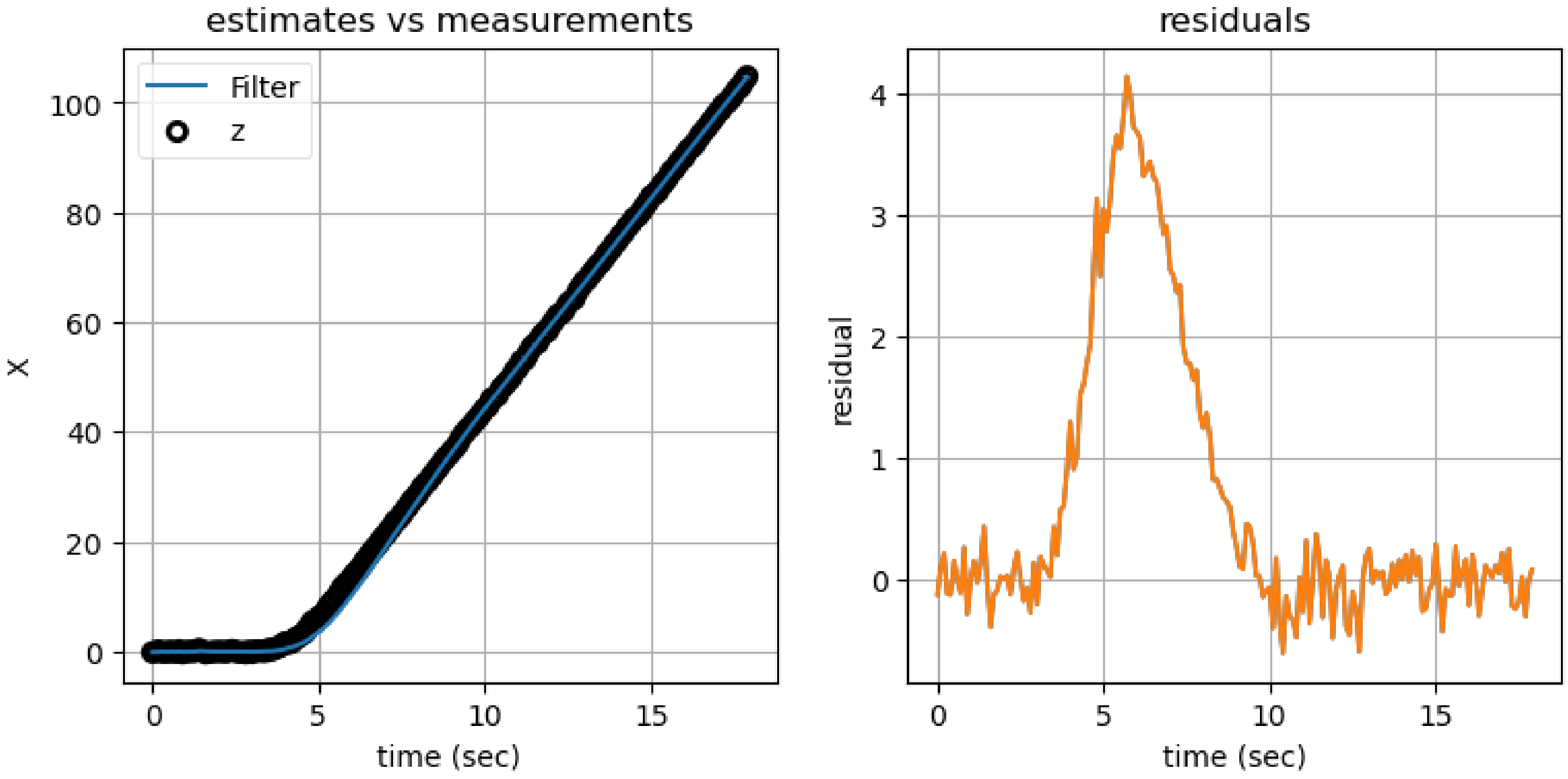


Increase Q

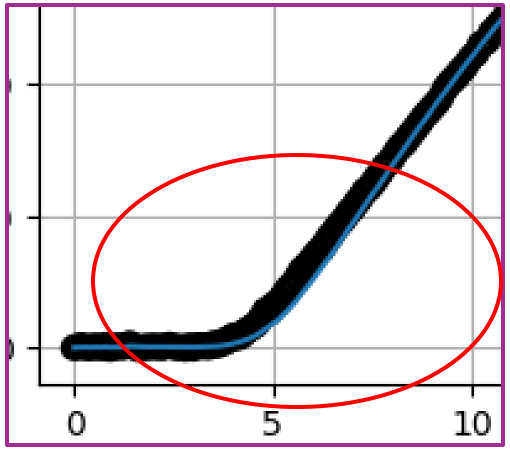
- Handles change better
- More noisy

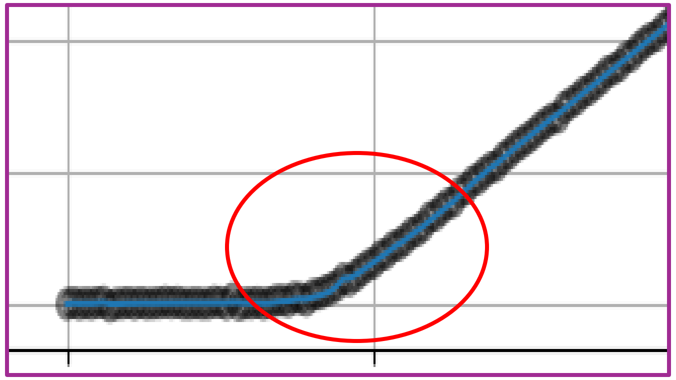


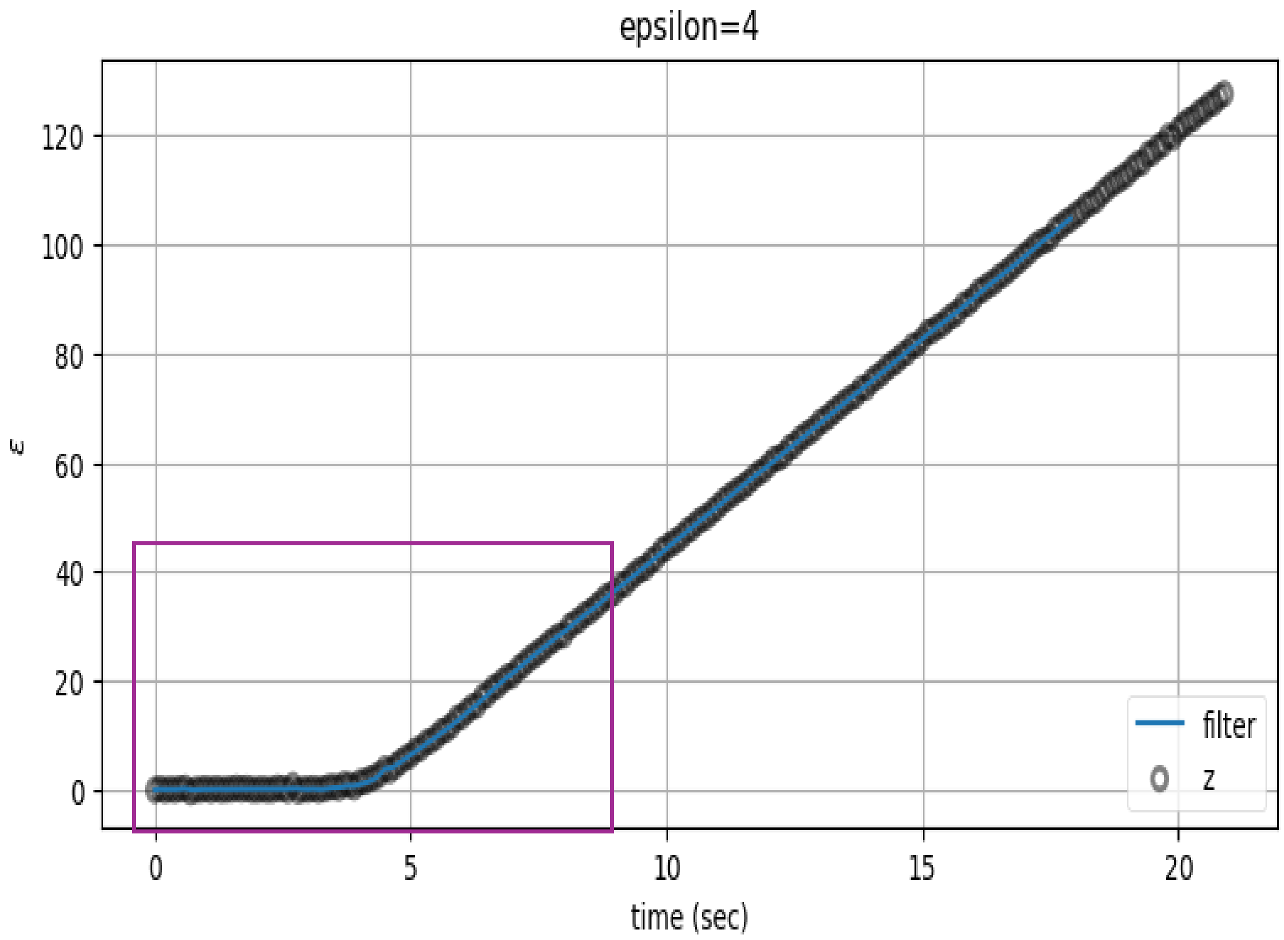
Detect change in model



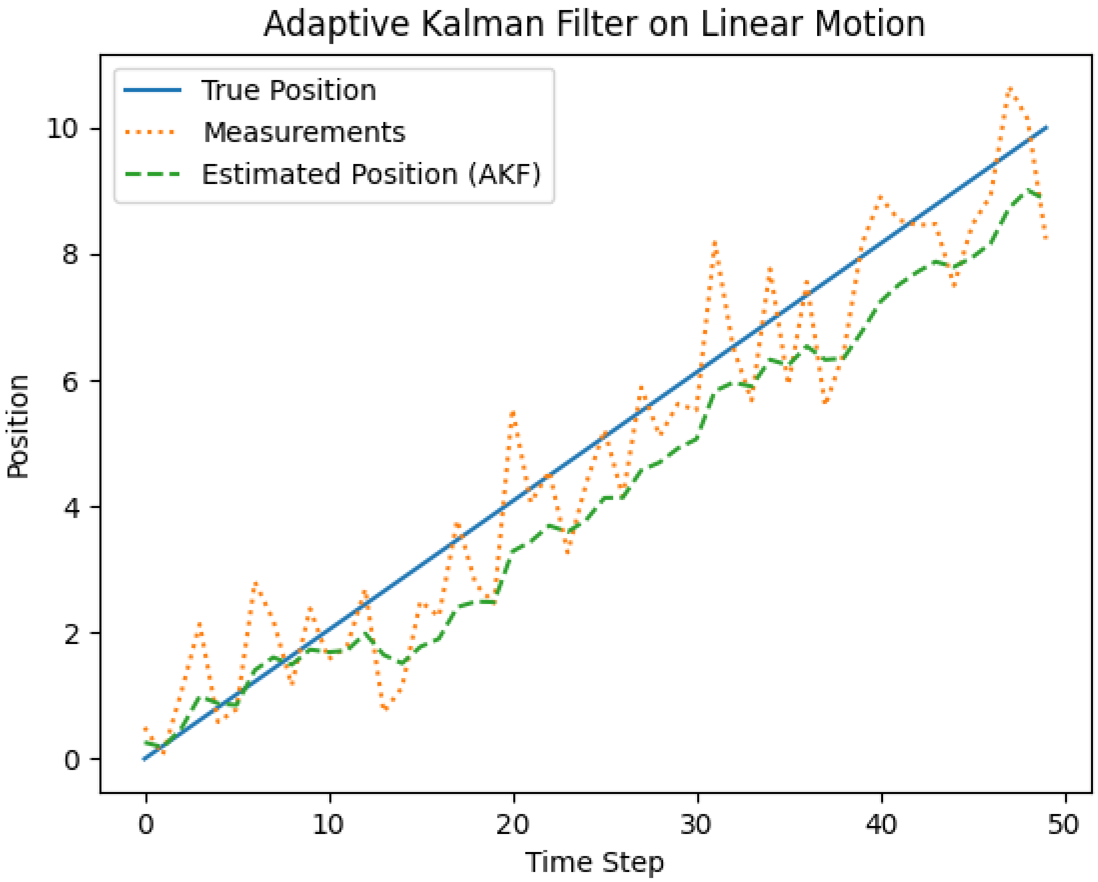
Adjust Q based on residual

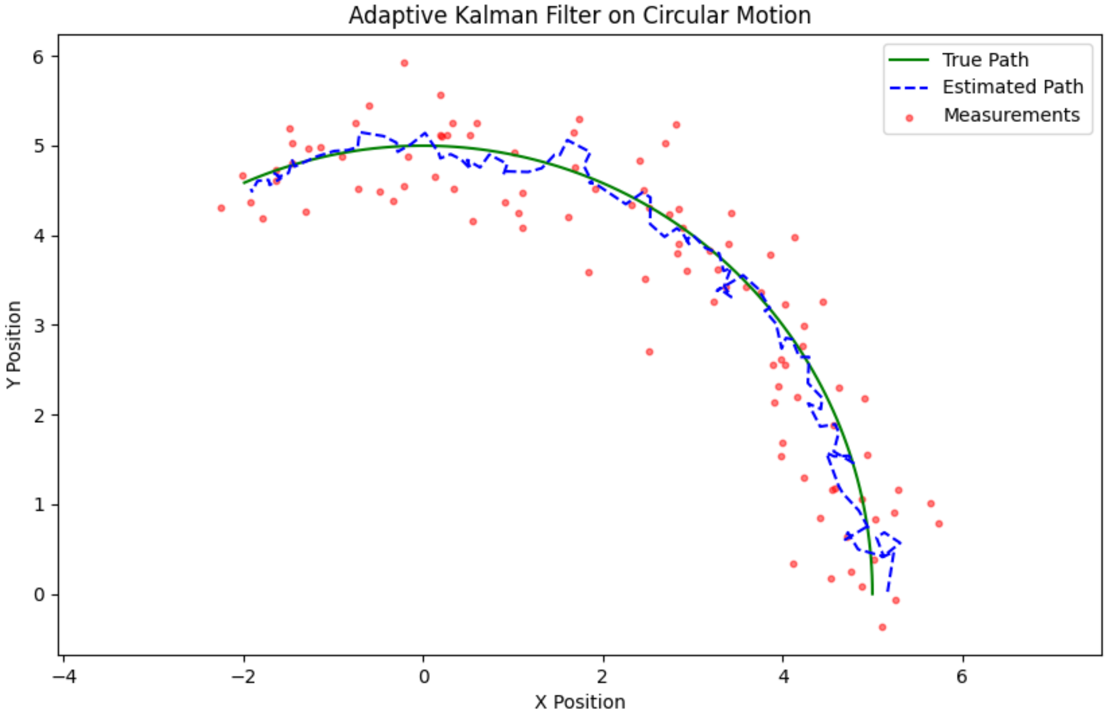






AKF on linear and circular motion





Advantages & Disadvantages

Advantages

- Useful for systems where models are poorly know
- Can handle systems with changing dynamics
- Good results from empirical tests

Disadvantages

- Needs initial guess
  - Should be run with different guesses
- Bias variance tradeoff
- Stability proofs are hard to come by

When to use

- Changing model dynamics
- Poorly known model of system

Sources

- adaptive kalman paper.pdf
- root-1.pdf
- Kalman-and-Bayesian-Filters-in-Python/14-Adaptive-Filtering.ipynb at master · rlabbe/Kalman-and-Bayesian-Filters-in-Python
- https://onlinelibrary.wiley.com/doi/10.1155/2015/381478#bib-0024