ASEN 5044, Fall 2018

Statistical Estimation for Dynamical Systems

Lecture 32:
Linearized KF and Extended KF:
Loose Ends + Miscellaneous Initialization/Tuning Tips

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Thurs 12/6/2018

University of Colorado Boulder

Announcements

- Midterm 2 grading will be back soon (sorry for delay)

- **Final project assignment posted**
- **Due Wed 12/19 @ 5 pm**: non-linear filtering and analysis
  - Group submission
  - Posted numbers for ground truth DT process noise covariances
  - Get started ASAP!
  - posted sanity checks for linearization steps (HW 8 problem 2)

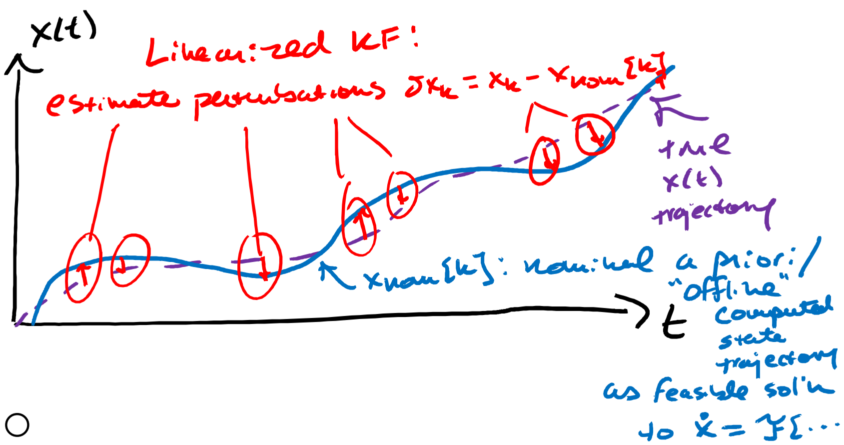
- FCQs online now – please fill out ASAP!

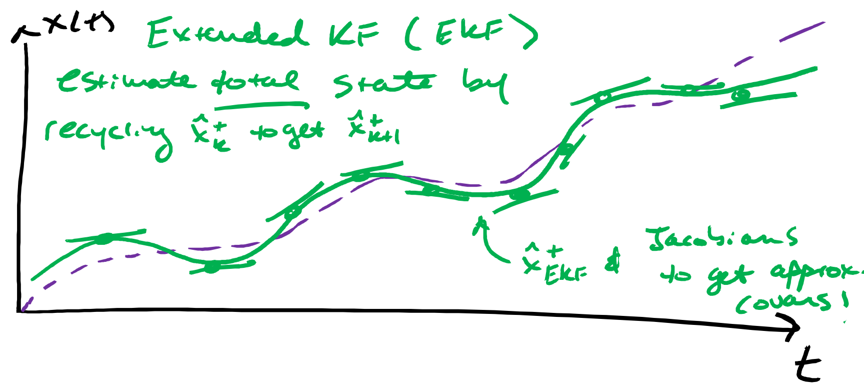
Quick Overview of Final Project Assignment

- Final assignment to be posted to Canvas
- Some data posted for each system as of today
- True nonlinear DT process noise Q and measurement noise R covariance matrices → to be used for truth model simulations
    - Observation data logs to validate linearized KF + EKF
- Note: NEES/NIS tests needed for both linearized KF and EKF
- Posted sanity checks for linearization steps (HW 8 problem 2) – make sure your Jacobians and integrators are set up correctly!

Last Time…

- Recap & Wrap-up DT Jacobian approximations for $\tilde{F}_k, \tilde{G}_k, \tilde{\Omega}_k$
- Approximately optimal DT state estimators based on linearization
  - **Linearized KF:** estimate *perturbations* around a priori nominal state trajectory:
    - Uses linearization about nominal trajectory for both mean and covariance updates





  - **Extended KF (EKF):** estimate *total state* around online estimated trajectory:
    - Uses full nonlinear model for predicted mean and predicted sensor measurements
    - Uses linearization only to approximate matrix quantities (Kalman gain and covariances)

Today…

- Quick overview of Final Project assignment

- Tie up some loose ends with linearized KF and EKF
  - Recap linearization assumptions
  - Derive covariance approx. for EKF
  - Initialization, other tips/caveats

What do we mean by “nom[k]” for linearization?

- Remember, nom[k] means two different things depending on whether you are doing linearized KF or the EKF:

For linearized KF:

- $\tilde{F}_{k+1}\big|_{nom[k]} = \left[\frac{\partial F}{\partial x} \right|_{(x^*_k,\, u^*_k,\, \tilde{w}_k = 0)}$

  where $x^*_k = x^*(t = t_k)$ & $x^*(t) = x_{nom}(t)$
  (a priori / offline computed sol'n to $\dot{x} = \mathcal{F}[x(t), u(t), \tilde{w}(t) = 0]$)

- For EKF:

  - $\tilde{F}_{k+1}\big|_{nom[k]} = \left[\frac{\partial F}{\partial x} \right|_{(\hat{x}^+_k,\, u(t_k),\, \tilde{w}(t_k) = 0)}$

    $\rightarrow x_{nom}[k] = \hat{x}^+_k$ = current (best) total state estimate

    computed online

The "1st Order" EKF Algorithm: Important Features

Useful to remember some key ideas for the EKF:

- Finding approx. Gaussian joint pdf for state and measurements from "best available guess" of total nonlinear system behavior/state at each time k
- Only use best available estimate of state at any point in time to compute required Jacobian matrices and nonlinear function evaluations at that time

  → do not need to know nominal trajectory in advance!!! (figuring it out online)

- We only need 1st order Taylor series/linearization of dynamics and measurements to get predicted covariance P_{k+1}^{-} , updated covariance P_{k+1}^{+}, and EKF gain $\tilde{K}_{k+1}$

  → all of these matrix quantities are obtained via Jacobians

  (similar to vanilla KF, except now matrices are time-varying and depend on $\hat{x}_k^+$!)

- **DO NOT use linearization/Jacobians to get predicted state $\hat{x}_{k+1}^{-}$ or measurement $\hat{y}_{k+1}$**

  → predicted vectors come directly from integrating/evaluating nonlinear CT fxns!

Origin of the Linearized Covariance Approximations

- Recall from Lec 31 that both linearized KF and EKF use similar expressions for covariance (main difference is in how Jacobians computed for each)
- e.g. for the prediction step at time k+1,

$\left(\dot{x} = \mathcal{F}[x,u] + \Gamma \tilde{w} \right.$

$$\underline{P^-_{k+1}} \; \tilde{\tilde{\approx}} \; \tilde{F}_k P^+_k \tilde{F}^T_k + \tilde{\Omega}_k Q_k \tilde{\Omega}^T_k,$$

- But how are such covariance approximations mathematically justified?
- Consider for the EKF: idea: linearize $f(\dots)$ about current best state est. $\hat{x}^+_k$

$\rightarrow \hat{x}^-_{k+1} = E[\underline{f(x(k), u(k), w(k))} | y_{1:k}]$

$\rightarrow$ Expand via Taylor series about $x(k) = \hat{x}^+_k$ & $w(k) = 0$ & some given $u(k)$

$\rightarrow \hat{x}^-_{k+1} \approx E[\underline{\underline{f}}(\hat{x}^+(k), u(k), 0) + \frac{\partial f}{\partial x}\Big|_{(\hat{x}^+(k), u(k), 0)} \cdot (x(k) - \hat{x}^+(k)) + \frac{\partial f}{\partial w}\Big|_{nom[k]} w(k) + \frac{\partial f}{\partial u}\Big|_{nom[k]} [u(k) - u_{nom}[k]] + \underline{\underline{HOTs}} | y_{1:k}]$

$E[(\dots)|y_{1:k}]$ $\quad \tilde{F}_k|_{nom[k]}$ $\quad \tilde{\Omega}_k|_{nom[k]}$ $\quad \tilde{G}_k|_{nom[k]}$

Origin of the Linearized Covariance Approximations

- So if we neglect Higher order Terms (HoTs) & hope linearization is valid near likely values of $x(k)$ [ie w.r.t. $p(x_{k+1} \mid y_{1:k})$],

we will get that $\hat{x}^-_{k+1} = E[\ldots \mid y_{1:k}]$ works out to: (using linearity of $E[\ldots]$)

ie filter working $\to 0$ if $\hat{x}^+_k = E[x_k \mid y_{1:k}]$

$$\hat{x}^-_{k+1} \approx \underbrace{f(\hat{x}^+(k), u(k), w(k)=0)}_{\text{result from non-linear ODE integration}} + \tilde{F}_k\big|_{nom[k]} \cdot E[\cancel{(x(k) - \hat{x}^+_k)} \mid y_{1:k}]$$
$$+ \tilde{G}_k\big|_{nom[k]} \cdot E[\underbrace{\cancel{(u(k) - u_{nom}(k))}}_{=0} \mid y_{1:k}]$$
$$+ \tilde{\Omega}_k\big|_{nom[k]} \cdot E[\underbrace{\cancel{w(k)}}_{0}]$$

$$\longrightarrow \boxed{\hat{x}^-_{k+1} = f(\hat{x}^+_k, u_k, w_k = 0)}$$

EKF state pred. @ time k+1

$\longrightarrow$ Now use this to compute $P^-_{k+1} = E[(x_{k+1} - \hat{x}^-_{k+1})(\cdots)^T \mid y_{1:k}]$

Origin of the Linearized Covariance Approximations

- So therefore, $P_{k+1}^- = \underline{\underline{E}}\left[(x_{k+1} - \hat{x}_{k+1}^-)(x_{k+1} - \hat{x}_{k+1}^-)^T \,|\, y_{1:k} \right]$

But since $x_{k+1} \approx \hat{x}_{k+1}^- + \tilde{F}_k|_{nom[k]} (x_k - \hat{x}_k^+) + \tilde{G}_k|_{nom[k]} (u_k - u_{nom}[k]) + \tilde{\Omega}_k|_{nom[k]} \cdot \underline{w_k}$ (by Taylor series)

$\rightarrow$ Square up the x_{k+1} approximation after substituting into $E[(\dots)(\dots)^T | y_{1:k}]$ expression & get:

$$\boxed{P_{k+1}^- \approx \tilde{F}_k|_{nom} \cdot P_k^+ \tilde{F}_k|_{nom}^T + \tilde{\Omega}_k|_{nom[k]} \cdot Q \, \tilde{\Omega}_k|_{nom[k]}^T}$$

EKF covar prediction

(can apply follow similar logic for measurement update step in EKF, & apply same ideas to linearized KF...)

Initializing and Tuning the EKF (and Linearized KF)

- One major issues for the EKF (and Linearized KF):
  - how to pick initial guess for state estimate and error covariance?
  - how to tune DT process noise covariance parameters, Q_{EKF} (Q_{LKF})?

- Poor choices can lead to filter inconsistent behavior and/or divergence!

- For nonlinear systems, cannot guarantee convergence to steady state covariance using a linearized approximation
  - different from KF for truly LTI systems, which is more “forgiving” as long as we have right model

Initializing and Tuning the EKF (and Linearized KF)

- Initialization: no silver bullet, but some general ideas and considerations (incomplete/non-exhaustive list…):
  - "Inflated" diagonal initial covariance – tricky to use for certain types of problems (e.g. how big to set Euler angle errors? Quaternions?)
  - Batch data processing to warm start, e.g. "static" initialization with linearized least squares or non-linear least squares
  - Sometimes can use LKF to initialize a "well-known"/well-observed portion of trajectory, before switching to EKF for remainder
  - Control can help quite a lot! (e.g. Skycrane system stabilization)

Initializing and Tuning the EKF (and Linearized KF)

- Process noise tuning and filter error compensation: no silver bullet, but some general ideas and considerations (incomplete/non-exhaustive list…):
  - Tuning Q_{EKF}/Q_{LKF} not too different than from linear KF, but can be non-intuitive
    - no obvious way to generalize Van Loan's method to convert CT → DT AWGN in nonlinear case
    - can still apply chi-square NEES/NIS tests to validate filter (even more important to apply NEES/NIS tests to EKF and Linearized KF – these should pass in order to validate that linearization is acceptable!)
  - Deal with presence of <u>biases</u> -- tuning Q_{EKF}/Q_{LKF} generally will <u>not be enough</u>!
    - Significant persistent biases can show up from neglected higher order terms/dynamics (linearization errors always lead to some bias…these may or may not be acceptably small…)
    - If biases are observable, can augment state vector to include and estimate/remove online
    - Depending on nature of non-linearities: can sometimes explicitly compute and remove biases (e.g. converting polar range + bearing observations to Cartesian x-y "pseudo-measurements")
  - "Desperate last resorts"/band-aids/ad-hockery to cope with covariance approximations and other stubborn covariance issues:
    - add artificial "pseudo-noise" to Q_{EKF} / Q_{LKF} to compensate for errors and tweak filter gains
    - add "fudge" terms to selectively "jack-up" predicted state covariance: $P^{-}_{k+1} = \Upsilon P^{-}_{k+1} \Upsilon^{T}$