ASEN 5044, Fall 2018

Statistical Estimation for Dynamical Systems

Lecture 27:
Steady state KF Behavior;
KF Consistency Evaluation

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

Thurs 11/8/2018

University of Colorado **Boulder**

Aerospace Engineering

Announcements

- **HW 7 due today**
- **Midterm 2: posted, due Nov 15 at 11 am on Canvas**
  - Will focus on HWs 5-7 (HW 7 solutions posted)
- **Final project partner sign up sheet on "Assignments" tab in Canvas**
  - Google docs sheet (**Due: Wed 11/14**) – please read + follow all instructions!!
  - Folks with a partner: enter group names
  - Folks without partner: start a new group or email each other to find a match
  - Preview system descriptions posted (will pick for HW 8 and final assignment)
  - **You must stay in your group for HW 8 (group hw) and final assignment!!! (group gets same grade on both!)**

- Final spec. topic lecture tomorrow.

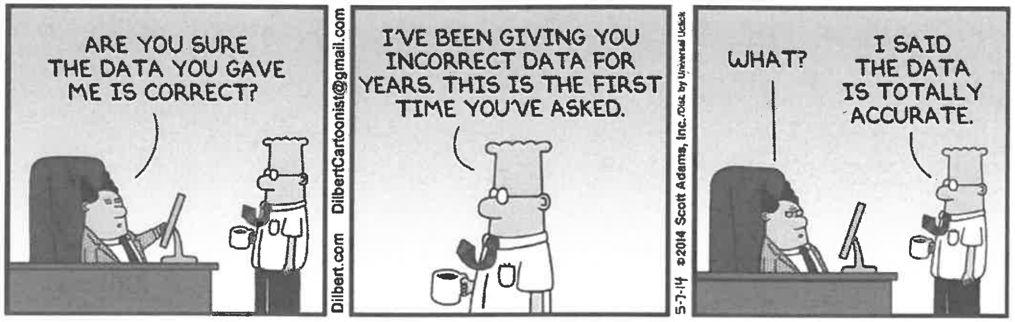
- **No Lecture on Tues 11/13, but will have class on Thurs 11/15**

Last Time…

- **The Kalman Filter (KF)**: dynamic predictor-corrector state estimator
  ("dynamic RLLS": combine state prediction with RLLS updates at each time k – i.e. handle dynamics + process noise w_k + noisy measurements y_k)
  - Algorithm
  - Example
  - Important/useful properties of the Kalman Gain

Today…

- One more generally useful property of KFs: **steady state behavior**
  - o Riccati equation, steady state error covariance, and steady state Kalman gain
- **How to tell if your (linear) KF is <u>actually</u> working correctly???**
  - o **Want to avoid GIGO systems (Garbage Input, Garbage Output)**



- **KF dynamic consistency analysis and "Truth Model Testing" (TMT)**
- **Chi-square tests (NEES/NIS)** – check if KF's state errors/measurement residuals make sense for given system + measurement + noise models
  - o **Do actual state errors/meas. residuals agree with KF's estimated error covariances?**
  - o Formal statistical tests to examine this question

Stopped here 11/6/18.

Steady-state Properties of the KF

- KF gives state estimate along with update of estimation error covariance (P_k^+)
- What is the "smallest/best possible" covariance? (i.e. what is smallest possible cost $J(k) = tr(P_k^+)$?) (How much "juice" can be squeezed from $y_{1:k}$?)

Recall: KF pred: $P_{k+1}^- = F P_k^+ F^T + Q$

KF meas. update: $P_{k+1}^+ = P_{k+1}^- - P_{k+1}^- H^T [H P_{k+1}^- H^T + R]^{-1} H P_{k+1}^-$

sub in def of KF gain K_{k+1}

→ but: @ time k: $P_k^+ = P_k^- - P_k^- H^T [H P_k^- H^T + R^{-1}] H P_k^-$

→ so plug P_k^+ expansion into P_{k+1}^- equation:

$$\boxed{P_{k+1}^- = F [P_k^- - P_k^- H^T [H P_k^- H^T + R]^{-1} H P_k^-] F^T + Q}$$

→ Non-linear DT "one-step" update to get P_{k+1}^- from P_k^- (assuming meas. update occurs @ time k)

known as → $\boxed{\text{DT Matrix Riccati Equation}}$ (DT MRE)

The Algebraic Riccati Equation (ARE)

Jacopo Riccati (1676 - 1754)

- Special case for LTI DT systems:
  - if process noise w(k) hits every state
  - AND if (F,H) is observable

then DT MRE implies convergence to a steady state a priori $P^-_\infty > 0 \text{ (posdef)}$

AND obtain the **DT Algebraic Riccati Equation (ARE)**, which is easier to solve (though still non-linear in P^-_∞):

$$\boxed{P^-_\infty = F(P^-_\infty - P^-_\infty H^T [H P^-_\infty H^T + R]^{-1} H P^-_\infty) F^T + Q}$$

Algebraic Riccati Equation (ARE)

Steady-State KF Gain

- Suppose the conditions for the ARE hold, so that

$$P_\infty^- = F(P_\infty^- - P_\infty^- H^T[HP_\infty^- H^T + R]^{-1} HP_\infty^-)F^T + Q$$

- Since the Kalman gain is

$$K_{k+1} = P_{k+1}^- H^T[HP_{k+1}^- H^T + R]^{-1}$$

→ it follows that there must also exist a steady state Kalman gain K_∞

$$K_\infty = P_\infty^- H^T[HP_\infty^- H^T + R]^{-1}$$

→ K_∞ often used in practice to save computation at each time step (since $K_{k+1} \rightarrow K_\infty$ quickly anyway, there is generally little performance loss)

*In Matlab: can use the "`dlqe.m`" (discrete linear quadratic estimator) command to find steady state KF gain, along with steady state a priori and a posteriori covariances

$$[K_\infty, P_\infty^-, P_\infty^+] = \text{dlqe}[F, \text{eye}(n), H, Q, R]$$

$\text{tr}(P_\infty^+)$ = smallest possible cost $J(k)$

Example: 1D Robot: Part Trois: Régime Permanent

- Same DT model as before:

$$x(k+1) = Fx(k) + Gu(k) + w(k)$$

$x(k) = [\xi(k), \dot{\xi}(k)]^T$

$u(k) = 2\cos(0.75 t_k)$ (ZOH)

$w(k) \sim \mathcal{N}(0, Q)$ (AWGN)

$x(0) \sim \mathcal{N}(\mu_0, P_0)$, where $\mu_0 = [0,0]^T$, $P_0 = I_{2\times 2}$

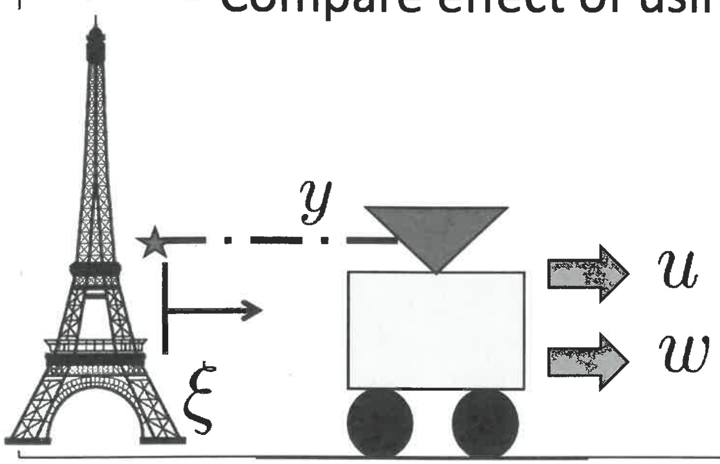
$$F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 0.5\Delta t^2 \\ \Delta t \end{bmatrix} \quad \Delta t = 0.1 \text{ sec}$$

$$Q = 1 \ (\text{m/s})^2, \ R = 0.5 \ \text{m}^2$$

- Compare effect of using dynamic KF gain to steady-state KF gain

→ i.e. usual KF updates

computed w/ dlqe



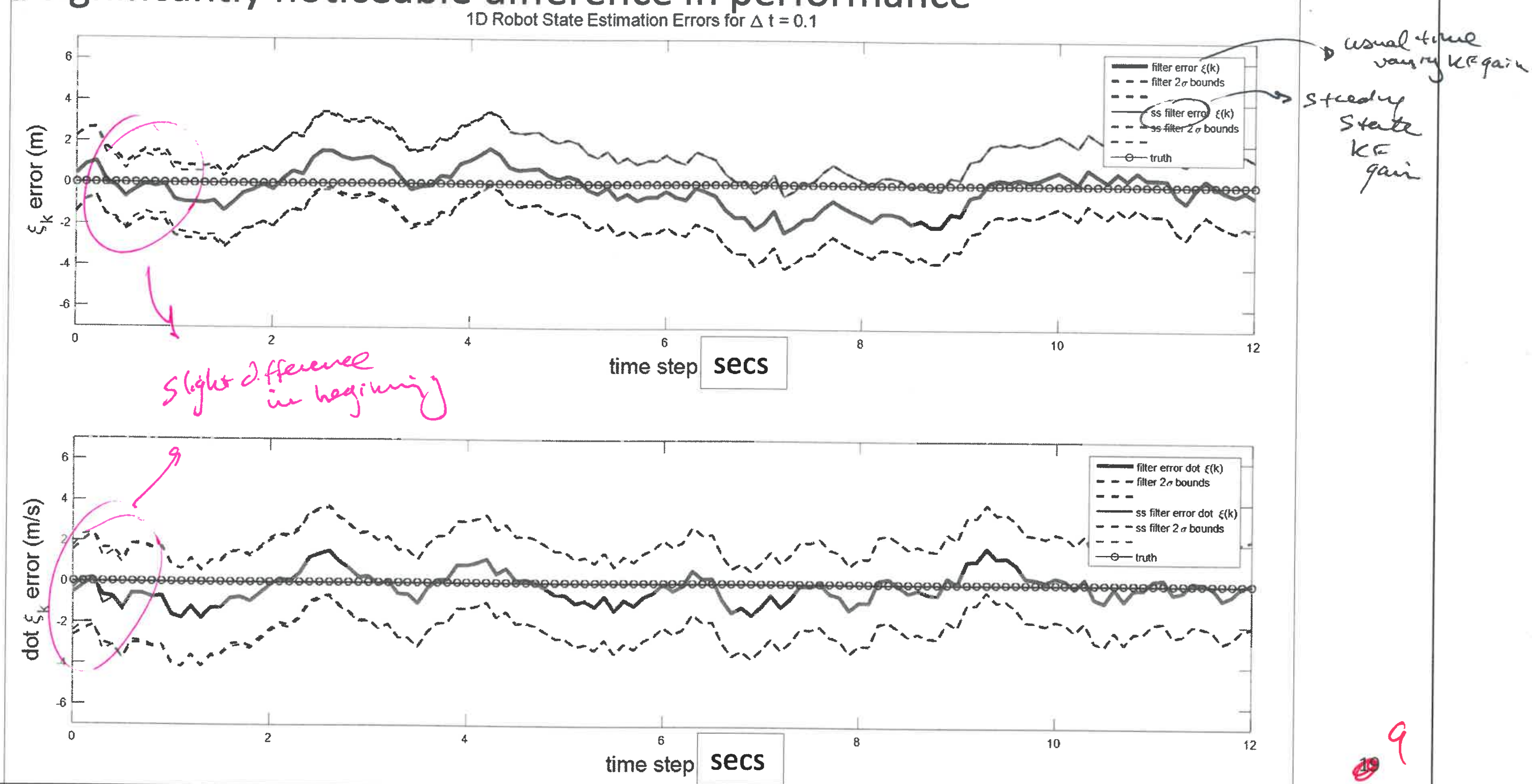
$$\rightarrow K_\infty = \begin{bmatrix} 0.1548 \\ 0.1300 \end{bmatrix}$$

$$P_\infty^- = \begin{bmatrix} 0.9157 & 0.7691 \\ & 1.2406 \end{bmatrix}$$

$$P_\infty^+ = \begin{bmatrix} 0.7740 & 0.6501 \\ & 1.1406 \end{bmatrix}$$

Results: Dynamic KF Gain vs. Steady-state KF Gain

- Not a significantly noticeable difference in performance



9

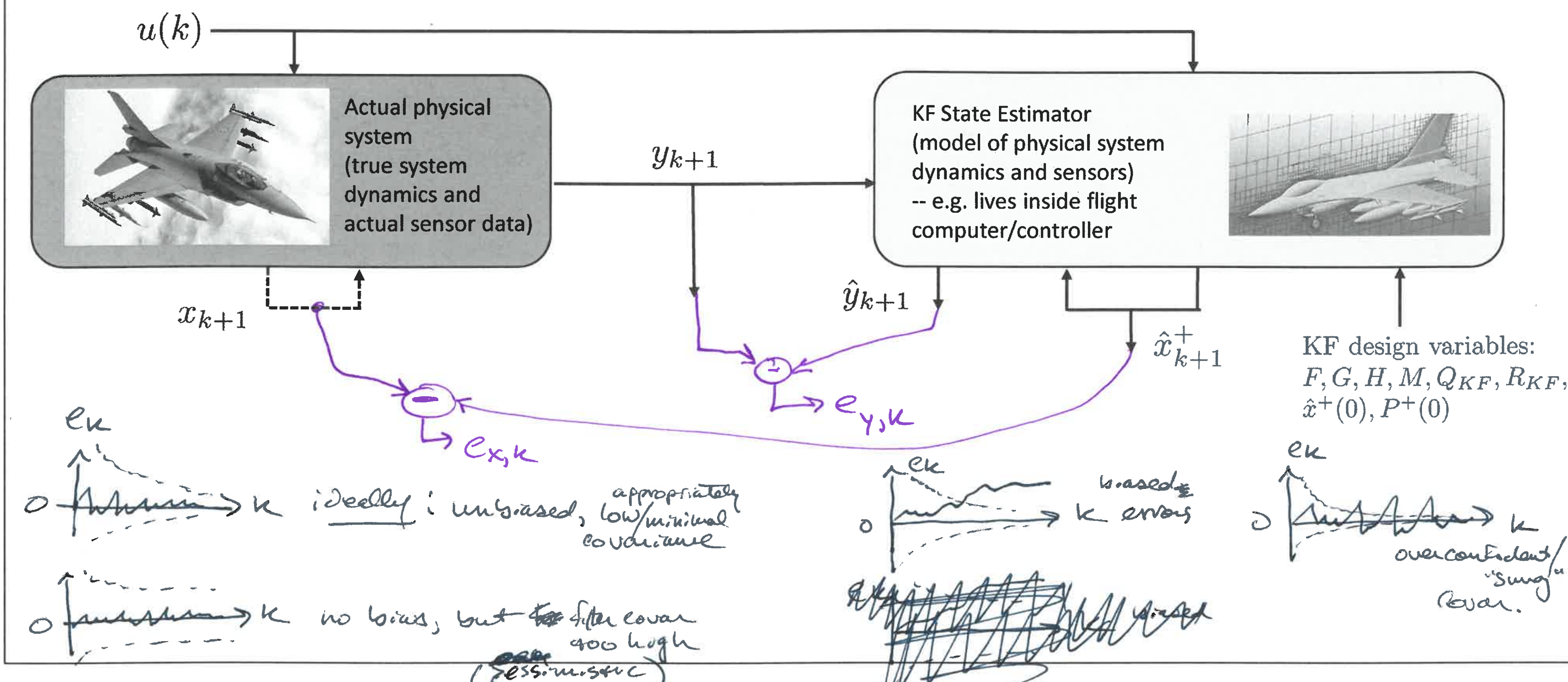
KF Consistency Analysis

- Like batch LLS and RLLS: estimates produced by the KF are random vectors
- For KF: this is due to uncertainties from process and measurement noise
- KF recursively assesses estimation uncertainty via error covariance matrix
- BUT: first need to set tuning parameters in Q_{KF} – generally not obvious!
- Other possible latent practical issues: approximate (F,G,H,M) model, unmodeled state dynamics, non-white noise, …
- So: how to know if KF estimates & covariances are correct for given system?

  i.e. do error statistics provided by KF reflect the actual error statistics?

Evaluating Errors in a KF State Estimator

- KF is itself a dynamical system (defined in software) that tries to track an actual physical system:



What Does "Minimal Error Estimation" Mean for the KF?

- In a perfect universe, we would like estimation errors to eventually vanish completely, i.e.

$$\text{if } e_{x,k} = x_k - \hat{x}_k^+, \text{ then } E[e_{x,k}] = 0 \text{ and } P_k^+ = E[e_{x,k} e_{x,k}^T] = 0 \text{ as } k \to \infty$$

(ie ideally: we would like perfect certainty in x_k as more measurements obtained)

(*)

- BUT, THIS DOES NOT HOLD AS EVIDENCED BY THE FACT THAT (IN MOST CASES) $P_\infty^- \neq 0$

(The Matrix Riccati & Algebraic Riccati Equations say so!!!)

- What is responsible for this?: Random Process Noise inputs incessantly disturb the true state x_k $\rightarrow \therefore e_{x,k} \not\to 0$ as $k \to \infty$! (generally speaking)

"Proper KF Error Characteristics": Dynamic Filter Consistency

- Because some finite/non-zero error will always exist, we instead say that the KF is "working properly" (for a given DT state space model and noise specs) if it satisfies the following **3 demands for <u>dynamic filter consistency</u>:**

1) Unbiasedness: $E[e_{x,k}] = 0$ for all k

2) Efficiency: $E[e_{x,k} e_{x,k}^T] = P_k^+$ (true state errors match filter covariance)

3) KF measurement residuals/innovations are a white Gaussian sequence:

$$e_{y,k} \sim \mathcal{N}(0, S_k), \quad E[e_{y,k} e_{y,j}^T] = S_k \cdot \delta(k, j)$$

where $e_{y,k} = y_k - \hat{y}_k = y_k - H\hat{x}_k^-$,

$$S_k = HP_k^- H^T + R$$

(S_k: "innovation covariance matrix" $\in \mathbb{R}^{p \times p}$)

Analyzing KF Estimation Errors and Measurement Innovations

- How do we analyze the two types of random error vectors in a KF?
  - State estimation errors (w.r.t. ground truth x_k): $e_{x,k} = x_k - \hat{x}_k^+ \in \mathbb{R}^n$ $\sim \mathcal{N}(0, P_k^+)$
  - Measurement innovations/residuals (w.r.t. observations y_k): $e_{y,k} = y_k - \hat{y}_k^- \in \mathbb{R}^p$ $\sim \mathcal{N}(0, S_k)$

- Simplest check is to look at the normalized magnitude of these vectors over time:

$$\epsilon_{x,k} = e_{x,k}^T (P_k^+)^{-1} e_{x,k} \quad \rightarrow \textbf{Normalized estimation error squared (NEES)} \text{ at time k}$$

$$\epsilon_{y,k} = e_{y,k}^T (S_k)^{-1} e_{y,k} \quad \rightarrow \textbf{Normalized innovation squared (NIS)} \text{ at time k}$$

$\rightarrow$ NEES and NIS are both **positive random scalars:** (weighted 2-norms)^2 of $e_{x,k}$ and $e_{y,k}$!

Key question: if dynamic consistency conditions (1)-(3) on previous slide hold, then what pdfs ought to describe our *expected* NEES & NIS outcomes?

ie what are $p(\epsilon_{x,k})$ & $p(\epsilon_{y,k})$?

Useful Fact #1: Squared 2-norms of Gaussian Random Vectors

- Suppose we are given some random vector $e \sim \mathcal{N}_e(0, P_e)$, $e \in \mathbb{R}^n$

→ let $ZZ^T = (P_e)^{-1}$, where $Z \in \mathbb{R}^{n\times n}$ is Cholesky factor of $(P_e)^{-1}$ [matrix square root]

→ Also define: $g = Z^T e \in \mathbb{R}^n$ (linear transformation of e via Z^T)

→ Now since $e \sim \mathcal{N}_e(0, P_e)$

& since $\mathcal{N}_e(0, P_e) \underset{\text{pdf def.}}{=} (\text{const.}) \cdot \exp\left(-\frac{1}{2} e^T (P_e)^{-1} e\right)$

$\underset{\text{(apply Chol. fact.)}}{=} (\text{const.}) \cdot \exp\left(-\frac{1}{2} e^T [ZZ^T] e\right)$ ~~$= (\text{const.}) \exp(-\frac{1}{2} g^T g)$~~

$\underset{\text{(apply def. of } g\text{)}}{=} (\text{const.}) \cdot \exp\left(-\frac{1}{2} g^T g\right) = \mathcal{N}_g(0, I)$ ~~(apply def. of g)~~ ~~$= \mathcal{N}_g(0, I)$~~

→ it follows that $g \sim \mathcal{N}_g(0, I)$ [std. multivariate normal pdf] ⟷ "decorrelated"/"whitened" Gaussian random vector

$\Rightarrow$ so $\epsilon = e^T (P_e)^{-1} e = e^T [ZZ^T] e = g^T g = \|g\|^2 = \sum_{i=1}^{n} g_i^2$ = Sum of squares of n iid Gaussian random scalar vars!

$g_i \sim \mathcal{N}(0,1)$ for $i = 1, \dots, n$

must have the same pdf!

$p(\epsilon) = p\left(\sum_{i=1}^{n} g_i^2\right)$

Example:
Correlated and Decorrelated Gaussian Random Vectors

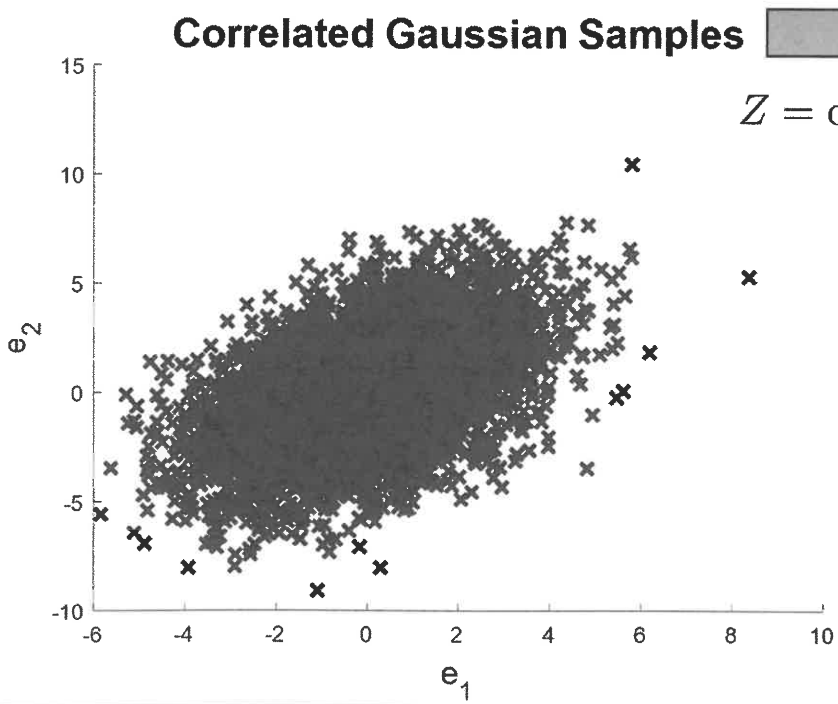
- Apply vector transform via Cholesky decomposition

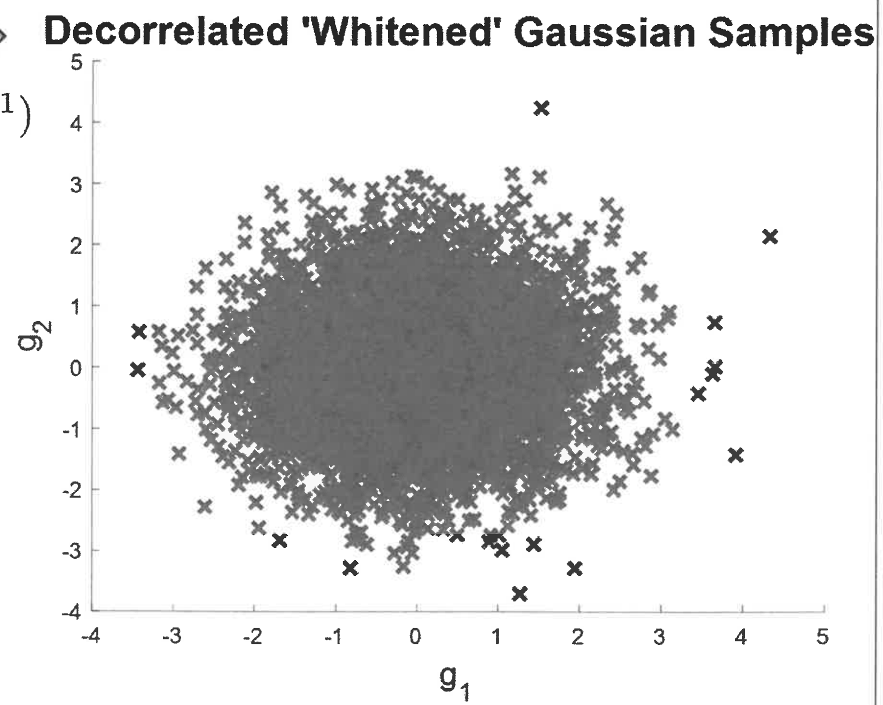
$$e \sim \mathcal{N}(0, P_e) \quad P_e = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix}$$

$$ZZ^T = P_e^{-1}$$

$$Z = \mathrm{chol}(P_e^{-1})$$

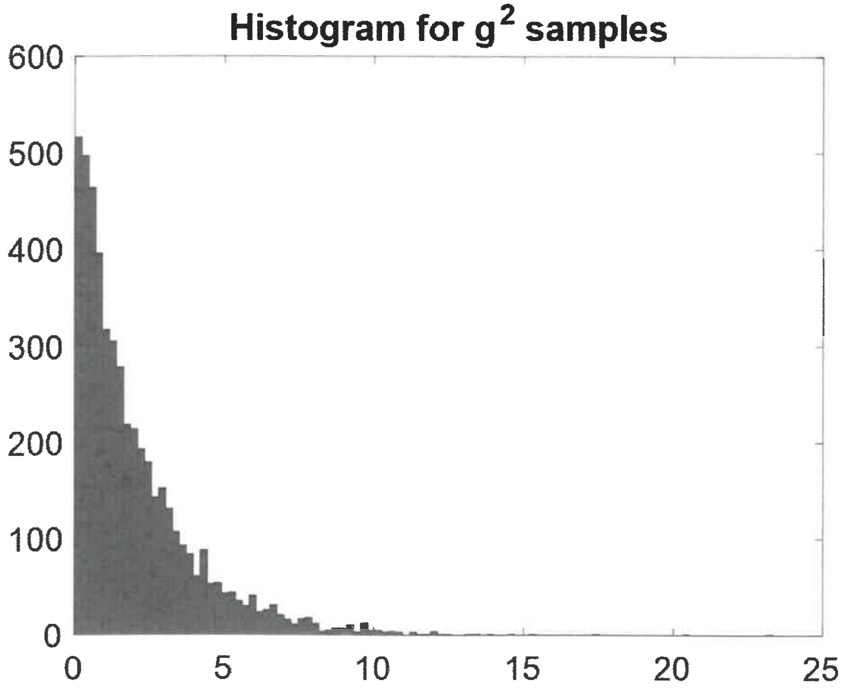
$$g = Z^T e, \text{ where } g \sim \mathcal{N}(0, I)$$

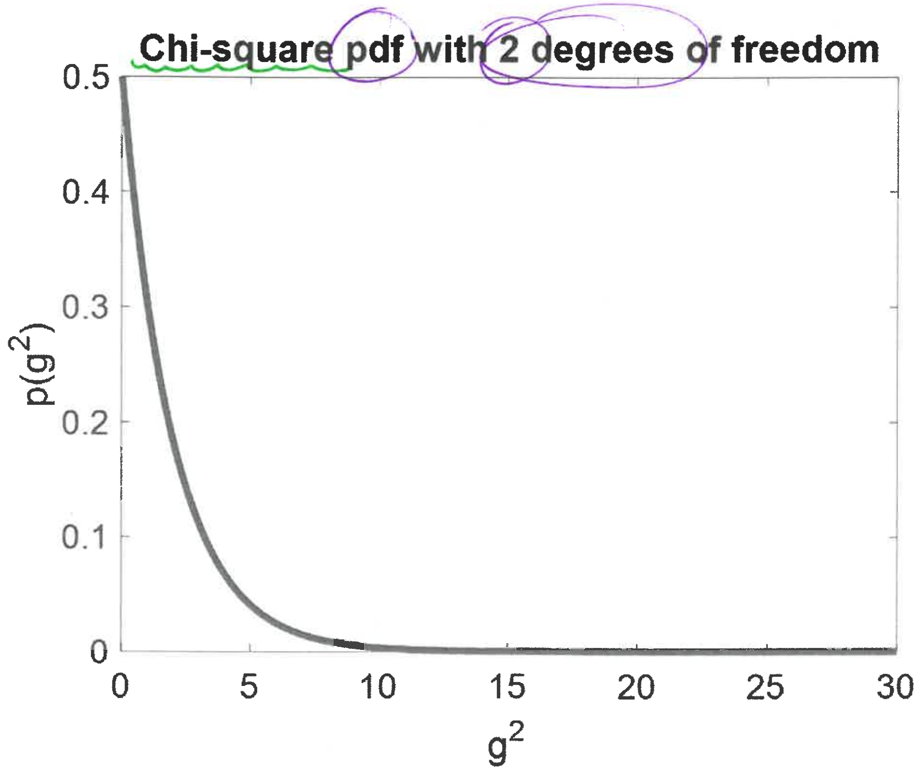




Example: Distribution of Gaussian RV Squared Magnitudes

- What does pdf look like for $g^T g = [\text{norm}(g)]^2 = g^2_1 + g^2_2$?





Useful Fact #2: The Chi-square Distribution

Suppose we have scalar i.i.d. random variables $g_1, ..., g_n$ where $g_i \sim \mathcal{N}(0,1)$ for $i = 1, ..., n$.

Define: random variable $q = \sum_{i=1}^{n} g_i^2 = \vec{g}^T \vec{g}$, where $\vec{g} = [g_1, \cdots, g_n]^T$ (note: $\vec{g} \sim \mathcal{N}(\vec{0}, I_{n\times n})$)

$\Rightarrow$ then the pdf $p(q)$ is a **chi-square** (χ^2) distribution with n degrees of freedom:

$$p(q) = \chi_n^2 = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} q^{\frac{n-2}{2}} \cdot \exp(-\frac{q}{2}), & \text{for } q \geq 0 \\ 0, & \text{for } q < 0 \end{cases}$$

where the 'gamma function' is

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$

$\Gamma(1) = 1$

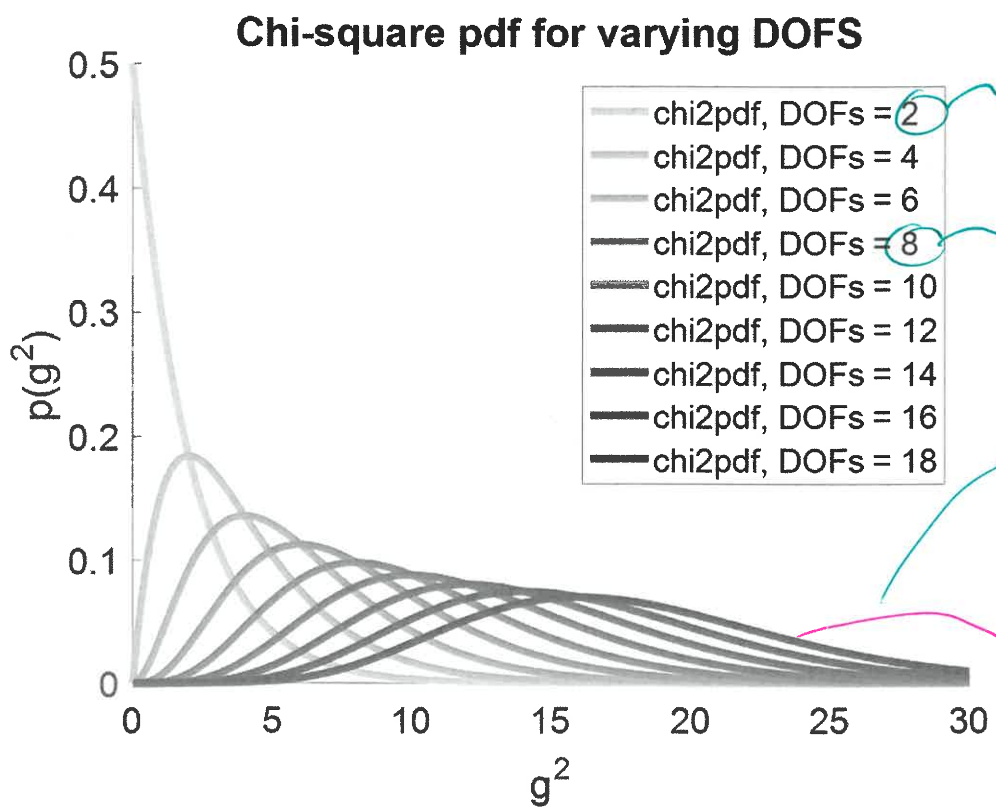
$\Gamma(m+1) = m \cdot \Gamma(m) = m!$ for integer m

* easy to show that: $E[q] = n$, and $\text{var}(q) = 2n$

* also can show: if $q_1 \sim \chi_{n_1}^2$ and $q_2 \sim \chi_{n_2}^2$, then $q_3 = q_1 + q_2 \Rightarrow q_3 \sim \chi_{n_1+n_2}^2$,
(i.e. $n_3 = n_1 + n_2$, so DOFs add!)

Example: Chi-square Distributions

- What does pdf look like for $[\text{norm}(g)]^2$ for different vector lengths n=length(e)?



$$g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} \sim \mathcal{N}(0, I_{2\times 2})$$

$$g = \begin{bmatrix} g_1 \\ \vdots \\ g_8 \end{bmatrix} \sim \mathcal{N}(0, I_{8\times 8})$$

$$p(\|g\|^2) = p(g^T g)$$
$$= p\left(\sum_{i=1}^{n} g_i^2\right) = p(\epsilon)$$

Note: Starts to look like Gaussian pdf for large n/DOFs! (by the CLT!)

Upshot: Theoretical KF NEES and NIS Error Distributions

- So, combining Facts #1 and #2, we deduce the following must be true:

If the KF works properly as per our DT state space model and noise specs
(i.e. if it meets the consistency criteria #1-#3 laid out earlier), then we must have:

$$\text{I. if } e_{x,k}(=x_k - \hat{x}_k^+) \sim \mathcal{N}(0, P_k^+) \text{ and } \epsilon_{x,k} = e_{x,k}^T (P_k^+)^{-1} e_{x,k} \text{ (NEES)}$$

$$\rightarrow \text{then } \epsilon_{x,k} \sim \chi_n^2 \ \forall k, \quad \text{where } E[\epsilon_{x,k}] = n, \ \text{var}(\epsilon_{x,k}) = 2n$$

$$\text{II. if } e_{y,k}(=y_k - \hat{y}_k^-) \sim \mathcal{N}(0, S_k) \text{ and } \epsilon_{y,k} = e_{y,k}^T (S_k)^{-1} e_{y,k} \text{ (NIS)}$$

$$\rightarrow \text{then } \epsilon_{y,k} \sim \chi_p^2 \ \forall k, \quad \text{where } E[\epsilon_{y,k}] = p, \ \text{var}(\epsilon_{y,k}) = 2p$$

We can use **"truth model testing" (TMT)** with NEES and use real/simulated sensor data with NIS to see if these pdfs actually show up!

→ if NOT, then we did something wrong!! (necessary but not sufficient conditions)