Colorado Center for Astrodynamics Research

ASEN 5070

Statistical Orbit Determination I

Fall 2012

Professor Jeffrey S. Parker

Professor George H. Born

Lecture 18: CKF, Numerical Compensations

Announcements

- Homework 6 and 7 graded asap
- Homework 8 announced today.

- Exam 2 on the horizon. Actually, it's scheduled for 11/8! A week from Thursday.
- We still have quite a bit of material to cover.
- Exam 2 will cover:
  - Batch vs. CKF vs. EKF
  - Probability and statistics (good to keep this up!)
  - Observability
  - Numerical compensation techniques, such as the Joseph and Potter formulation.
  - No calculators should be necessary
  - Open Book, Open Notes

Quiz 13 Review

CCAR
Colorado Center for Astrodynamics Research

Question 1 (1 point)

Let's say we have the following scenario:

- One ground station tracking one satellite

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations.

In this scenario the ground station is perfectly known. So the state has 6 parameters, corresponding to the position and velocity of the satellite at a reference epoch. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable

Quiz 13 Review

CCAR
Colorado Center for Astrodynamics Research

Question 1 (1 point)

Let's say we have the following scenario:

- One ground station tracking one satellite
- Simple 2-body gravity model, no perturbations
- range and range-rate observations (lots of them). There are plenty of observations.

In this scenario the ground station is perfectly known. So the state has 6 parameters, corresponding to the position and velocity of the satellite at a reference epoch. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable

Quiz 13 Review

CCAR
Colorado Center for Astrodynamics Research

Question 2 (1 point)

Scenario 2:

- One ground station tracking one satellite

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations.

In this scenario the ground station is NOT well-known. So the state has 9 parameters, corresponding to the position and velocity of the satellite and the position of the ground station at a reference epoch. The a priori uncertainty is large. In theory the ground station's position could be off by many kilometers, even in altitude.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable
- No, only 6 parameters are observable

Quiz 13 Review

Question 2 (1 point)

Scenario 2:

- One ground station tracking one satellite

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations.

In this scenario the ground station is NOT well-known. So the state has 9 parameters, corresponding to the position and velocity of the satellite and the position of the ground station at a reference epoch. The a priori uncertainty is large. In theory the ground station's position could be off by many kilometers, even in altitude.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable
- No, only 6 parameters are observable

Only 46% of the class got this right

Quiz 13 Review

Question 3 (1 point)

Scenario 3:

- One on-orbit satellite tracking another satellite, in different orbits. No ground stations.

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations between the two satellites.

In this scenario the first satellite's orbit is perfectly well-known. So the state has 6 parameters, corresponding to the position and velocity of the unknown satellite orbit. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable

Quiz 13 Review

Question 3 (1 point)

Scenario 3:

- One on-orbit satellite tracking another satellite, in different orbits. No ground stations.

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations between the two satellites.

In this scenario the first satellite's orbit is perfectly well-known. So the state has 6 parameters, corresponding to the position and velocity of the unknown satellite orbit. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable

Quiz 13 Review

Question 4 (1 point)

Scenario 4:

- Two satellites, in different orbits that are both unknown. No ground stations.

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations between the two satellites.

In this scenario neither satellite's orbits are well-known. So the state has 12 parameters, corresponding to the position and velocity of both satellite's at a reference epoch. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable
- No, only 9 parameters are observable

Quiz 13 Review

Question 4 (1 point)

Scenario 4:

- Two satellites, in different orbits that are both unknown. No ground stations.

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations between the two satellites.

In this scenario neither satellite's orbits are well-known. So the state has 12 parameters, corresponding to the position and velocity of both satellite's at a reference epoch. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- Yes
- No, no parameters are observable
- No, only 3 parameters are observable
- No, only 9 parameters are observable

0% of the class got this right!

Quiz 13 Review

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

Scenario 4:

- Two satellites, in different orbits that are both unknown. No ground stations.

- Simple 2-body gravity model, no perturbations

- range and range-rate observations (lots of them). There are plenty of observations between the two satellites.

In this scenario neither satellite's orbits are well-known. So the state has 12 parameters, corresponding to the position and velocity of both satellite's at a reference epoch. The a priori uncertainty is large.

Question: Is this state fully observable? (Hopefully you can answer this by visualizing the scenario and not doing the math)

- ○ Yes
- ○ No, no parameters are observable
- ○ No, only 3 parameters are observable
- ○ No, only 9 parameters are observable

Observable quantities:

- Semimajor axis of both orbits
- Eccentricity of both orbits
- True anomaly of both satellites
- Inclination *difference* between the orbits
- Node *difference* between the orbits
- Arg Peri *difference* between the orbits

Quiz 14 Review

Question 1 (1 point)

What is a sign that the estimation error covariance matrix (the P matrix) has suffered from numerical computation errors?

- It is no longer a square matrix
- It is no longer a symmetric matrix
- It has become a diagonal matrix
- It has negative off-diagonal elements

Quiz 14 Review

Question 1 (1 point)

What is a sign that the estimation error covariance matrix (the P matrix) has suffered from numerical computation errors?

- ○ It is no longer a square matrix
- ○ It is no longer a symmetric matrix
- ○ It has become a diagonal matrix
- ○ It has negative off-diagonal elements

100% of the class got this right!

Quiz 14 Review

CCAR
Colorado Center for Astrodynamics Research

Information

Let's say we're processing observations using a Conventional Kalman Filter (the basic Sequential Processor) and we are set up to process an observation at time t_i. Let's say our state is 6 elements long (n=6) and we have one observation data type (p=1) at time t_i.

Question 2 (1 point)

What is the size of P, the estimation error covariance matrix?

- 1x1
- 1x6
- 6x1
- 6x6

Quiz 14 Review

CCAR
Colorado Center for
Astrodynamics Research

Information

Let's say we're processing observations using a Conventional Kalman Filter (the basic Sequential Processor) and we are set up to process an observation at time t_i. Let's say our state is 6 elements long (n=6) and we have one observation data type (p=1) at time t_i.

Question 2 (1 point)

What is the size of P, the estimation error covariance matrix?

- 1x1
- 1x6
- 6x1
- 6x6

Quiz 14 Review

Question 3 (1 point)

What is the size of R, the observation error covariance matrix?

- 1x1
- 1x6
- 6x1
- 6x6

Quiz 14 Review

CCAR
Colorado Center for Astrodynamics Research

Question 3 (1 point)

What is the size of R, the observation error covariance matrix?

- 1x1
- 1x6
- 6x1
- 6x6

Quiz 14 Review

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

Recall that the *a priori* state deviation vector (x-bar) and the *a priori* covariance (P-bar) are generated from all previous observations through t_i-1, and then they are mapped to the current time t_i. After processing the observation at time t_i, we see that the state estimate didn't change more than a tiny tiny amount from the *a priori* state ; the covariance also didn't change more than a tiny tiny amount.

What can we infer from this?

- We have far more certainty in the observation than we do with the estimated state.
- We have far more certainty in the estimated state than we do with the observation.

Quiz 14 Review

CCAR
Colorado Center for Astrodynamics Research

Question 4 (1 point)

Recall that the *a priori* state deviation vector (x-bar) and the *a priori* covariance (P-bar) are generated from all previous observations through t_i-1, and then they are mapped to the current time t_i. After processing the observation at time t_i, we see that the state estimate didn't change more than a tiny tiny amount from the *a priori* state ; the covariance also didn't change more than a tiny tiny amount.

What can we infer from this?

- We have far more certainty in the observation than we do with the estimated state.
- We have far more certainty in the estimated state than we do with the observation.

Quiz 14 Review

CCAR
Colorado Center for
Astrodynamics Research

Question 5 (1 point)

Mathematically, what does this mean? (where "this" is the Q&A of the previous question)

- R(t_i) >> Htilde(t_i)*Pbar(t_i)*Htilde(t_i)^T
- R(t_i) << Htilde(t_i)*Pbar(t_i)*Htilde(t_i)^T

Quiz 14 Review

Question 5 (1 point)

Mathematically, what does this mean? (where "this" is the Q&A of the previous question)

- R(t_i) >> Htilde(t_i)*Pbar(t_i)*Htilde(t_i)^T
- R(t_i) << Htilde(t_i)*Pbar(t_i)*Htilde(t_i)^T

Probably the hardest question on this quiz and 78% got it right.

$$K_k = \bar{P}_k \tilde{H}_k^T \left[\tilde{H}_k \bar{P}_k \tilde{H}_k^T + R_k \right]^{-1} \qquad \hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + K_k \left[\mathbf{y}_k - \tilde{H}_k \bar{\mathbf{x}}_k \right]$$

Assignment #8

- Homework 8
- Due in 9 days

Beginning with equations (4.7.20) and the a priori information given in the text, write a Matlab program to compute the following :

1a. The exact value of P2 using equation (4.7.24).

1b. P2 using a conventional Kalman algorithm

1c. P2 using the Joseph algorithm

1d. P2 using the Potter algorithm (equation (5.7.17))

1e. P2 using the batch algorithm

2. Plot the trace of the exact value of P2 minus the trace of P2 vs ε for each of the following

a. conventional Kalman

b. Joseph algorithm

c. Potter algorithm

d. batch algorithm

plot for $\varepsilon = 1\text{x}10^{-6}, 1\text{x}10^{-7}, \ldots\ldots, 1\text{x}10^{-15}$.

3. Compare and contrast the behavior and stability of the various filters.

4. Complete this problem

CKF vs. Batch

CCAR
Colorado Center for
Astrodynamics Research

$$\left[\bar{x}_k\right]_{n\times 1},\ \left[\Phi\right]_{n\times n},\ \left[\bar{P}_k\right]_{n\times n},\ \left[y_k\right]_{p\times 1},\ \left[K_k\right]_{n\times p},\ \left[P_k\right]_{n\times n}$$

The Kalman (sequential) filter differs from the batch filter as follows:

1. It uses $\tilde{H}$ in place of H.
2. It uses $\Phi\left(t_k, t_{k-1}\right)$ in place of $\Phi\left(t_k, t_0\right)$

   This means that $\dot{\Phi} = A\Phi$ is reinitialized at each observation time.

Or if we don't reinitialize, then

$$\begin{aligned}\Phi\left(t_k, t_{k-1}\right) &= \Phi\left(t_k, t_0\right)\Phi\left(t_0, t_{k-1}\right)\\ &= \Phi\left(t_k, t_0\right)\Phi^{-1}\left(t_{k-1}, t_0\right)\\ &= \Phi\left(t_k, t_{k-1}\right)\end{aligned}$$

i.e., we must invert $\Phi\left(t_{k-1}, t_0\right)$ at each stage

Processing Observations One at a Time

Given a $p \times 1$ vector of observations at t_k we may process these as p scalar observations by performing the measurement update p times and skipping the time update.

Algorithm

1. Do the time update at t_k

$$\bar{P}_k = \Phi(t_k, t_{k-1}) P_{k-1} \Phi^T(t_k, t_{k-1})$$

$$\bar{x}_k = \Phi(t_k, t_{k-1}) \hat{x}_{k-1}$$

2. Measurement update

$$\vec{Y}_k = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix}_k, \quad G(\mathrm{X}_k^*) = \begin{bmatrix} G_1(\mathrm{X}_K^*) \\ G_2(\mathrm{X}_K^*) \\ \vdots \\ G_p(\mathrm{X}_K^*) \end{bmatrix}_k, \quad y_k = \vec{Y}_k - G(\mathrm{X}_k^*) = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}_k$$

Processing Observations One at a Time

$$\tilde{H}_k = \begin{bmatrix} \tilde{H}_1 \\ \tilde{H}_2 \\ \vdots \\ \tilde{H}_p \end{bmatrix}_{\substack{k \\ p\times n}} = \begin{bmatrix} \partial G_1\left(\mathrm{X}_K^*\right) / \partial \mathrm{X}_K \\ \partial G_2\left(\mathrm{X}_K^*\right) / \partial \mathrm{X}_K \\ \vdots \\ \partial G_p\left(\mathrm{X}_K^*\right) / \partial \mathrm{X}_K \end{bmatrix}, \quad R_k = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_p^2 \end{bmatrix}_{p\times p}$$

2a. Process the 1st element of Y_k. Compute $y_1 = \left(Y_1 - G_1\left(\mathrm{X}_k^*\right)\right)$,

$$\tilde{H}_1 = \frac{\partial G_1\left(\mathrm{X}_k^*\right)}{\partial \mathrm{X}_k^*}, \quad K_1 = \bar{P}_k \tilde{H}_1 \left(\tilde{H}_1 \bar{P}_k \tilde{H}_1^T + \sigma_1^2\right)^{-1}$$

$$\hat{x}_{k_1} = K_1\left(y_1 - \tilde{H}_1 \bar{x}_k\right) \qquad \text{where} \qquad \left[K_1\right]_{n\times1}, \left[y_1\right]_{n\times1}, \left[\tilde{H}_1\right]_{1\times n}, \left[\bar{x}_k\right]_{n\times1}$$

$$P_{k_1} = \left(\mathrm{I} - K_1 \tilde{H}_1\right)\bar{P}_k$$

Processing Observations One at a Time

2. Do not do a time update but do a measurement update by processing the 2nd element of Y_k (i.e. $\hat{x}_{k_1}$ and P_{k_1} are not mapped to t_{k+1})

Compute $y_2, \tilde{H}_2,\quad K_2 = P_{k_1}\tilde{H}_2\left(\tilde{H}_2 P_{k_1}\tilde{H}_2^T + \sigma_2^2\right)^{-1}$

$$\hat{x}_{k_2} = K_2\left(y_2 - \tilde{H}_2\hat{x}_{k_1}\right)$$

$$P_{k_2} = \left(\mathrm{I} - K_2\tilde{H}_2\right)P_{k_1}$$

Here $\bar{x}_{k2} = \hat{x}_{k1}$

$\bar{P}_{k2} = P_{k1}$

etc.

Process pth element of Y_k. Compute $y_p, \tilde{H}_p, K_p = P_{k_{p-1}}\tilde{H}_p\left(\tilde{H}_p P_{k_{p-1}}\tilde{H}_p^T + \sigma_p^2\right)^{-1}$

$$\hat{x}_{k_p} = K_p\left(y_p - \tilde{H}_p\hat{x}_{k_{p-1}}\right) \qquad P_{k_p} = \left(\mathrm{I} - K_p\tilde{H}_p\right)P_{k_{p-1}}$$

Processing Observations One at a Time

CCAR
Colorado Center for Astrodynamics Research

let $$\hat{x}_k = \hat{x}_{k_p}, \qquad P_k = P_{k_p}$$

Time update to t_{k+1} and repeat this procedure

Note: R_k must be a diagonal matrix. Why?

If not we would apply a "Whitening Transformation" as described in the next slides.

Whitening Transformation

CCAR
Colorado Center for Astrodynamics Research

$$y = \tilde{H}x + \varepsilon \quad (1)$$

$$E[\varepsilon] = 0 \qquad E\left[\varepsilon\varepsilon^T\right] = R$$

Where R is any sort of observation error covariance

Whitening Transformation

CCAR
Colorado Center for
Astrodynamics Research

$$y = \tilde{H}x + \varepsilon \qquad (1)$$

$$E\left[\varepsilon\right] = 0 \qquad E\left[\varepsilon\varepsilon^T\right] = R$$

Where R is any sort of observation error covariance

Factor $R = SS^T$ where S is the square root of R

multiply Eq. (1) by S^{-1}

$$S^{-1}y = S^{-1}\tilde{H}x + S^{-1}\varepsilon$$

S is chosen to be upper triangular

Whitening Transformation

$$y = \tilde{H}x + \varepsilon \quad (1)$$

$$E\left[\varepsilon\right] = 0 \qquad E\left[\varepsilon\varepsilon^{T}\right] = R$$

Where R is any sort of observation error covariance

Factor $R = SS^{T}$ where S is the square root of R

multiply Eq. (1) by S^{-1}

$$S^{-1}y = S^{-1}\tilde{H}x + S^{-1}\varepsilon$$

S is chosen to be upper triangular

let

$$y' \equiv S^{-1}y, \qquad \tilde{H}' \equiv S^{-1}\tilde{H}, \qquad \varepsilon' \equiv S^{-1}\varepsilon$$

$$y' \equiv \tilde{H}'x + \varepsilon'$$

Whitening Transformation

Now $$E\left[\varepsilon'\right] = S^{-1}E\left[\varepsilon\right] = 0$$

$$E\left[\varepsilon'\varepsilon'^{T}\right] = S^{-1}E\left[\varepsilon'\varepsilon'^{T}\right]S^{T} = S^{-1}RS^{-T}$$

$$= S^{-1}SS^{T}S^{-T} = \mathrm{I},\ \varepsilon' \sim N\left[0,\mathrm{I}\right]$$

Hence, the new observation $y' = S^{-1}y$ has an error with zero mean and unit variance. We would now process the new observation y' and use $\tilde{H}'$.

Cholesky Decomposition

- Notice the $R = SS^T$ decomposition in the whitening algorithm.
  - S is a "square root" of R.
- Several ways of computing S.
- A good way is the Cholesky Decomposition (Section 5.2).

- Major benefit of S vs. R: if the condition number of R is large, say 16, the condition number of S is half as large, say 8.
- It's much more accurate to invert S than R.

Cholesky Decomposition

- Let *M* be a symmetric positive definite matrix, and let *R* be an upper triangular matrix computed such that

$$R^T R = M$$

- By inspection,

$$r_{11} = \sqrt{M_{11}}$$
$$r_{1k} = \left(M_{1k}\right)/r_{11}$$
$$r_{22} = \sqrt{\left(M_{22} - r_{12}^2\right)}$$
$$r_{2k} = \left(M_{2k} - r_{12}r_{1k}\right)/r_{22}$$

Etc. See 5.2.1

Cholesky Decomposition

- *M*:

```
>> disp(M)
    3.2443    2.6518    2.3555    1.4758    2.1490    2.6638
    2.6518    2.8509    2.2282    1.7568    2.1782    2.1157
    2.3555    2.2282    2.6323    1.4811    1.6396    2.6423
    1.4758    1.7568    1.4811    1.9426    1.0814    1.2941
    2.1490    2.1782    1.6396    1.0814    1.7627    1.5862
    2.6638    2.1157    2.6423    1.2941    1.5862    3.3699
```

$$R^T R = M$$

- Use Matlab's "chol" function to compute *R*:

```
>> R = chol(M)
R =
    1.8012    1.4722    1.3078    0.8193    1.1931    1.4789
         0    0.8267    0.3664    0.6660    0.5101   -0.0745
         0         0    0.8876    0.1866   -0.1212    0.8286
         0         0         0    0.8905   -0.2394   -0.0255
         0         0         0         0    0.0840   -0.5475
         0         0         0         0         0    0.4361

>> R'*R
ans =
    3.2443    2.6518    2.3555    1.4758    2.1490    2.6638
    2.6518    2.8509    2.2282    1.7568    2.1782    2.1157
    2.3555    2.2282    2.6323    1.4811    1.6396    2.6423
    1.4758    1.7568    1.4811    1.9426    1.0814    1.2941
    2.1490    2.1782    1.6396    1.0814    1.7627    1.5862
    2.6638    2.1157    2.6423    1.2941    1.5862    3.3699
```

Cholesky Decomposition

CCAR
Colorado Center for Astrodynamics Research

- Consider the case where we have the least squares condition

$$M\hat{\mathbf{x}} = N$$

- If *M* is poorly conditioned, then inverting *M* will not produce good results.
- If we decompose *M* first, then the accuracy of the solution will improve.

$$R^T R\hat{\mathbf{x}} = N$$

- Let $$\mathbf{z} = R\hat{\mathbf{x}}$$
- Then $$R^T\mathbf{z} = N$$

- Easy to forward-solve for *z* and then we can backward-solve for *x*-hat.

Cholesky Decomposition

- Notice that the Cholesky Decomposition includes square roots. These are very expensive operations!

$$r_{11} = \sqrt{M_{11}}$$
$$r_{22} = \sqrt{(M_{22} - r_{12}^2)}$$

- A square root free algorithm is given in 5.2.2.
- Idea:

$$R^T R = M$$
$$UDU^T = M$$

Questions?

CCAR
Colorado Center for Astrodynamics Research

- Processing an observation vector one element at a time.
- Whitening
- Cholesky

- Quick Break

- Next topic: Joseph, Potter, …

Kalman Filter History

- Least squares estimation began with Gauss
- 1963: Kalman's sequential approach
  - Introduced minimum variance
  - Introduced process noise
  - Permitted covariance analyses without data
- Schmidt proposed a linearization method that would work for OD problems
  - Supposed that linearizing around the best estimate trajectory is better than linearizing around the nominal trajectory
- 1970: Extended Kalman Filter
- Gradually, researchers identified problems.
  - (a) Divergence due to the use of incorrect *a priori* statistics and unmodeled parameters.
  - (b) Divergence due to the presence of nonlinearities.
  - (c) Divergence due to the effects of computer round-off.

Kalman Filter History

- Numerical issues cause the covariance matrix to lose their symmetry and nonnegativity
- Possible corrections:
  - (a) Compute only the upper (or lower) triangular entries and force symmetry
  - (b) Compute the entire matrix and then average the upper and lower fields
  - (c) Periodically test and reset the matrix
  - (d) Replace the optimal Kalman measurement update by other expressions (Joseph, Potter, etc)
  - (e) Use larger process noise and measurement noise covariances.

Kalman Filter History

- Potter is credited with introducing square root factorization.
  - Worked for the Apollo missions!
- 1968: Andrews extended Potter's algorithms to include process noise and correlated measurements.
- 1965 – 1969: Development of the Householder transformation
  - Worked for Mariner 9 in 1971!
- 1969: Dyer-McReynolds filter added additional process noise effects.
  - Worked for Mariner 10 in 1973 for Venus and Mercury!

Numerical Issues

- Batch:
  - Large matrix inversion w/ potential poorly-conditioned matrix.
- Sequential:
  - Lots and lots of matrix inversions w/ potentially poorly-conditioned matrices.
- EKF:
  - Divergence: If the observation data noise is too large.
  - Saturation: If covariance gets too small, new data stops influencing the solution.

Joseph Formulation

- Replace

$$P_k = \left(I - K_k \tilde{H}_k^T\right) \bar{P}_k$$

with

$$P_k = \left(I - K_k \tilde{H}_k^T\right) \bar{P}_k \left(I - K_k \tilde{H}_k^T\right)^T + K_k R_k K_k^T$$

- This formulation will always retain a symmetric matrix, but it may still lose positive definiteness.

Need to show that

$$\left(\mathrm{I}-KH\right)\bar{P}\left(\mathrm{I}-KH\right)^T+KRK=\left(\mathrm{I}-KH\right)\bar{P}$$

First derive the Joseph formulation

$$\hat{x}=\bar{x}+K\left(y-H\bar{x}\right)$$

subst. $y=Hx+\varepsilon$

$$\hat{x}=\bar{x}+K\left(Hx+\varepsilon-H\bar{x}\right)$$

$$\hat{x}=\bar{x}+KH\left(x-\bar{x}\right)+K\varepsilon$$

subtract x from both sides and rearrange

$$\hat{x}-x=\left(\bar{x}-x\right)-KH\left(\bar{x}-x\right)+K\varepsilon$$

$$=\left(\mathrm{I}-KH\right)\left(\bar{x}-x\right)+K\varepsilon$$

Equivalence of Joseph and Conventional Formulation of the Measurement Update of P

form $P \equiv E(\hat{x}-x)(\hat{x}-x)^T$ $\quad\quad$ $\hat{x}-x=(\mathrm{I}-KH)(\bar{x}-x)+K\varepsilon$

$$\boxed{P=(\mathrm{I}-KH)\bar{P}(\mathrm{I}-KH)^T+KRK^T}$$

where

$$\bar{P}=E(\bar{x}-x)(\bar{x}-x)^T, \quad R=E\left[\varepsilon\varepsilon^T\right]$$

$E\left[(\bar{x}-x)\varepsilon^T\right]=0,$ since $\bar{x}(t_i)$ is independent of $\varepsilon(t_i)$

Equivalence of Joseph and Conventional Formulation of the Measurement Update of P

Next, show that this is the same as the conventional Kalman

$$\left(\mathrm{I}-KH\right)\bar{P}\left(\mathrm{I}-KH\right)^{T}+KRK^{T}=\left(\mathrm{I}-KH\right)\bar{P}-\left(\mathrm{I}-KH\right)\bar{P}H^{T}K^{T}+KRK^{T}$$

$$=\left(\mathrm{I}-KH\right)\bar{P}-\bar{P}H^{T}K^{T}+KH\bar{P}H^{T}K^{T}+KRK^{T}$$

$$=\left(\mathrm{I}-KH\right)\bar{P}-\bar{P}H^{T}K^{T}+K\left(H\bar{P}H^{T}+R\right)K^{T}$$

Equivalence of Joseph and Conventional Formulation of the Measurement Update of P

Next, show that this is the same as the conventional Kalman

$$(\mathrm{I}-KH)\bar{P}(\mathrm{I}-KH)^T + KRK^T = (\mathrm{I}-KH)\bar{P} - (\mathrm{I}-KH)\bar{P}H^T K^T + KRK^T$$

$$= (\mathrm{I}-KH)\bar{P} - \bar{P}H^T K^T + KH\bar{P}H^T K^T + KRK^T$$

$$= (\mathrm{I}-KH)\bar{P} - \bar{P}H^T K^T + K\left(H\bar{P}H^T + R\right)K^T$$

subst. for $K = \bar{P}H^T\left(H\bar{P}H^T + R\right)^{-1}$

$$= (\mathrm{I}-KH)\bar{P} - \bar{P}H^T K^T + \bar{P}H^T\left(H\bar{P}H^T + R\right)^{-1}\left(H\bar{P}H^T + R\right)K^T$$

$$= (\mathrm{I}-KH)\bar{P} - \bar{P}H^T K^T + \bar{P}H^T K^T$$

$$= (\mathrm{I}-KH)\bar{P}$$

- Question on Joseph?
- Next: Demo!

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

CCAR
Colorado Center for Astrodynamics Research

The following example problem from Bierman (1977) illustrates the numerical characteristics of several algorithms we have studied to date. Consider the problem of estimating x_1 and x_2 from scalar measurements z_1 and z_2

$$z_1 = x_1 + \varepsilon x_2 + v_1$$
$$z_2 = x_1 + x_2 + v_2$$

Where v_1 and v_2 are uncorrelated zero mean random variables with unit variance. If they did not have the above traits we could perform a whitening transformation so that they do. In matrix form

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} 1 & \varepsilon \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \qquad (4.7.20)$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

CCAR
Colorado Center for Astrodynamics Research

The *a priori* covariance associated with our knowledge of $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ is assumed to be

$$P_0 = \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \sigma^2 I$$

Where $\sigma = 1/\varepsilon$ and $0 < \varepsilon \leq 1$. The quantity ε is assumed to be small enough such that computer round-off produces

$$1 + \varepsilon^2 = 1 \qquad (4.7.21)$$

This estimation problem is well posed. The observation z_1 provides an estimate of x_1 which, when coupled with z_2 should accurately determine x_2. However, when the various data processing algorithms are applied several diverse results are obtained.

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

Let the gain and estimation error covariance associated with z_1 be denoted as k_1 and P_1 respectively. Similarly the measurement z_2 is associated with k_2 and P_2 respectively. Note that this is a linear parameter (constant) estimation problem.

Hence,

$$\Phi\left(t, t_k\right) = I$$

$$\tilde{H} = H = \begin{bmatrix} 1 & \varepsilon \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

We will process the observation one at a time. Hence,

$$\bar{P}_k = P_{k-1}$$

$$P_k = \left(I - k_k h_k\right) P_{k-1}$$

$$k_k = P_{k-1} h_k^T \left[h_k P_{k-1} h_k^T + 1\right]^{-1} \qquad k = 1, 2$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

CCAR
Colorado Center for Astrodynamics Research

Note: Since we will process the observations one at a time, the matrix inversion in previous equation is a scalar inversion. The exact solution is

$$k_1 = P_0 h_1^T \left[h_1 P_0 h^T + 1\right]^{-1} = \sigma^2 \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \left[\begin{bmatrix} 1 & \varepsilon \end{bmatrix} \sigma^2 \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} + 1 \right]^{-1}$$

$$k_1 = \frac{1}{1+2\varepsilon^2} \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \qquad (4.7.22)$$

where $\sigma^2 \varepsilon^2 = 1$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

CCAR
Colorado Center for Astrodynamics Research

The estimation error covariance associated with processing the first data point is

$$P_1 = \left(I - k_1 h_1\right) P_0$$

$$P_1 = \left(I - \frac{1}{1+2\varepsilon^2} \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \begin{bmatrix} 1 & \varepsilon \end{bmatrix} \right) \sigma^2 I$$

$$P_1 = \frac{1}{\alpha} \begin{Bmatrix} 2 & -\sigma \\ -\sigma & \sigma^2 + 1 \end{Bmatrix} \qquad (4.7.23)$$

where $\alpha = 1 + 2\varepsilon^2$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

CCAR
Colorado Center for Astrodynamics Research

Processing the second observation:

$$k_2 = P_1 h_2^T \left(h_2 P_1 h_2^T + 1 \right)^{-1}$$

$$k_2 = \frac{1}{\alpha} \begin{Bmatrix} 2 & -\sigma \\ -\sigma & \sigma^2 + 1 \end{Bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left[\begin{pmatrix} 1 & 1 \end{pmatrix} \frac{1}{\alpha} \begin{bmatrix} 2 & -\sigma \\ -\sigma & \sigma^2 + 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1 \right]^{-1}$$

$$= \frac{1}{\Delta} \begin{bmatrix} -\varepsilon \left(1 - 2\varepsilon \right) \\ 1 - \varepsilon + \varepsilon^2 \end{bmatrix} \quad \text{where} \quad \Delta = 1 - 2\varepsilon + 2\varepsilon^2 \left(2 + \varepsilon^2 \right)$$

The exact solution for P_2 is given by:

$$P_2 = \left(I - k_2 h_2 \right) P_1 = \frac{1}{\Delta} \begin{bmatrix} 1 + 2\varepsilon^2 & -\left(1 + \varepsilon \right) \\ -\left(1 + \varepsilon \right) & 2 + \varepsilon^2 \end{bmatrix} \qquad (4.7.24)$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

The conventional Kalman filter yields (using $1+\varepsilon^2=1$)

$$k_1 = \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix}$$

$$P_1 = \left[I - \begin{bmatrix} 1 \\ \varepsilon \end{bmatrix} \begin{pmatrix} 1 & \varepsilon \end{pmatrix} \right] \sigma^2 I$$

$$P_1 = \begin{bmatrix} 0 & -\sigma \\ -\sigma & \sigma^2 \end{bmatrix} \tag{4.7.25}$$

Note that P_1 is no longer positive definite. The diagonals of a PD matrix must be positive. However P_1^{-1} does exist since $|P_1| \neq 0$.

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

Compute P_2 using the Conventional Kalman

$$k_2 = P_1 h_2^T \left(h_2 P_1 h_2^T + 1\right)^{-1} = \frac{1}{1-2\varepsilon}\begin{bmatrix} -\varepsilon \\ 1-\varepsilon \end{bmatrix}$$

$$P_2 = \left[I - k_2 h_2\right] P_1 = \frac{1}{1-2\varepsilon}\begin{bmatrix} -1 & +1 \\ +1 & -1 \end{bmatrix} \qquad (4.7.26)$$

Now P_2 is not positive definite ($|P_2| = 0$) nor does it have positive terms on the diagonal. In fact, the conventional Kalman filter has failed for these observations.

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

The Joseph formulation yields

$$P_1 = \begin{bmatrix} 2 & -\sigma \\ -\sigma & \sigma^2 \end{bmatrix}$$

$$P_2 = \left(I - k_2 h_2\right) P_1 \left(I - k_2 h_2\right)^T + k_2 k_2^T$$

$$P_2 = \begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+\varepsilon \end{bmatrix}$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

The Batch Processor yields

$$P_1 = \left[\sum_{i=1}^{2} h_i^T h_i + P_0^{-1}\right]^{-1} = \left[\begin{bmatrix}1\\ \varepsilon\end{bmatrix}\begin{bmatrix}1 & \varepsilon\end{bmatrix} + \begin{bmatrix}1\\ 1\end{bmatrix}\begin{bmatrix}1 & 1\end{bmatrix} + \varepsilon^2 I\right]$$

$$= \begin{bmatrix}2+\cancel{\varepsilon^2} & 1+\varepsilon \\ 1+\varepsilon & 2\cancel{\varepsilon^2}+1\end{bmatrix}^{-1}$$

P_2 for the batch

$$P_2 = \begin{bmatrix}1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+4\varepsilon\end{bmatrix}$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

P_2 for the batch

$$P_2 = \begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+4\varepsilon \end{bmatrix}$$

To order ε, the exact solution for P_2 is

$$P_2 = (1+2\varepsilon)\begin{bmatrix} 1 & -(1+\varepsilon) \\ -(1+\varepsilon) & 2 \end{bmatrix} = \begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+4\varepsilon \end{bmatrix}$$

Example Illustrating Numerical Instability of Sequential (Kalman) Filter (see 4.7.1)

Summary of P_2 Results

Exact to order ε

$$\begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+4\varepsilon \end{bmatrix}$$

Conventional Kalman

$$\frac{1}{1-2\varepsilon}\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

Joseph

$$\begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+\varepsilon \end{bmatrix}$$

Batch

$$\begin{bmatrix} 1+2\varepsilon & -(1+3\varepsilon) \\ -(1+3\varepsilon) & 2+4\varepsilon \end{bmatrix}$$

The End

- Homework 6 and 7 graded asap
- Homework 8 announced today.

- Exam 2 on the horizon. Actually, it's scheduled for 11/8! A week from Thursday.
- We still have quite a bit of material to cover.
- Exam 2 will cover:
  - Batch vs. CKF vs. EKF
  - Probability and statistics (good to keep this up!)
  - Observability
  - Numerical compensation techniques, such as the Joseph and Potter formulation.
  - No calculators should be necessary
  - Open Book, Open Notes