State Estimation ...

Lecture 16

... of Linear Dynamical Systems:

- Observability
- Luenberger Observer
- Discrete-Time Kalman Filter (KF)
  - Derivation of the KF, LS/Deterministic approach
  - ~~Review of Probabilistic concepts~~
  - ~~Probabilistic derivation of the KF~~
- Exercise: Estimation of Linear System
- Exercise: Constrainted KF

Linear Dynamical System

Lecture 16

- Consider a linear, (time-invariant,) deterministic dynamical system, represented in discrete time as:

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1},$$
$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k,$$

where $\mathbf{x} \in \mathbb{R}^n$ is the n-dimensional state vector, $\mathbf{u} \in \mathbb{R}^p$ is the p-dimensional input vector, $\mathbf{y} \in \mathbb{R}^m$ is the m-dimensional output vector. Furthermore , $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times p}$ and $\mathbf{H} \in \mathbb{R}^{m\times n}$.

Observability

Lecture 16

Estimation problem for a state-space (hidden states) model vs the input/output model. Is it even possible ?

- Observability is a measure of how well internal states of a system can be inferred from knowledge of its external outputs. (the mathematical dual of the controllability concept)
- A dynamical system designed to estimate the state of a system from measurements of the outputs is called a state observer.
- If the original system is not observable, we cannot design a state observer that can fulfill its purpose.

Observability of LTI Systems

Lecture 16

- The LTI system is observable if and only if, for any initial time k_0, any initial state $\mathbf{x}_{k_0}$ can be determined from the knowledge of output $\mathbf{y}_k$ and input $\mathbf{u}_k$, for $k_0 \leq k \leq$
- Equivalently, if the row rank of the observability matrix $\mathcal{O}$, defined as below, is equal to n (the state size), then the system is observable.

$$\mathcal{O} = \begin{bmatrix} \mathbf{H} & \mathbf{HA} & \mathbf{HA}^2 & \dots & \mathbf{HA}^{n-1} \end{bmatrix}^T$$

$$\text{rank}(\mathcal{O}) = n$$

Observability of LTV Systems

Lecture 16

For Linear Time Varying Systems,

- The LTV is completely observable if and only if for any initial time k_0, any initial state be $\mathbf{x}_{k_0}$ can be determined from the knowledge of output $\mathbf{y}_k$ and input $\mathbf{u}_k$ for $k_0 \leq k \leq k_N$, where k_N is some final finite time.
- Equivalently, if the matrix $\mathcal{O}_{k_0}$, defined below, has rank n (the state size) for any k_0, then the system is completely observable.

$$\mathcal{O}_{k_0} = \begin{bmatrix} \mathbf{H}_{k_0} & \mathbf{H}_{k_0+1}\mathbf{A}_{k_0} & \mathbf{H}_{k_0+2}\mathbf{A}_{k_0+1}\mathbf{A}_{k_0} & \dots & \mathbf{H}_{k_0+N-1}\mathbf{A}_{k_0+N-2}\dots\mathbf{A}_{k_0+1}\mathbf{A}_{k_0} \end{bmatrix}^T$$

$$\mathrm{rank}(\mathcal{O}_{k_0}) = n$$

- Various results when the time dependency has a certain structure

Luenberger Observer

Lecture 16

- A two step strategy: **Prediction** (of state and measurement) and **Error Correction** (based on actual measurements).
- Prediction is based on a model of the system, i.e. $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$, $\hat{\mathbf{H}}$, and a state estimate $\hat{\mathbf{x}}$.
- Correction is integrated using a feedback gain matrix $\mathbf{L} \in \mathbb{R}^{m \times n}$.
- Luenberger State Observer for deterministic discrete-time linear systems:

$$
\begin{aligned}
\hat{\mathbf{x}}_{k+1} &= \hat{\mathbf{A}}\hat{\mathbf{x}}_k + \hat{\mathbf{B}}\mathbf{u}_k + \mathbf{L}(\mathbf{y}_k - \hat{\mathbf{y}}_k), \\
\hat{\mathbf{y}}_k &= \hat{\mathbf{H}}\hat{\mathbf{x}}_k,
\end{aligned}
$$

Luenberger Observers

Lecture 16

- If the system is observable and if the model is correct (i.e. $\mathbf{A} = \hat{\mathbf{A}}, \mathbf{B} = \hat{\mathbf{B}}, \mathbf{H} = \hat{\mathbf{H}}$) then the convergence of the estimated state to the actual state can be guaranteed with a proper feedback gain $\mathbf{L}$.
- If the model is reasonably close to being correct (i.e. $\mathbf{A} \approx \hat{\mathbf{A}}, \mathbf{B} \approx \hat{\mathbf{B}}, \mathbf{H} \approx \hat{\mathbf{H}}$), the Luenberger observer can still work well.
- How to choose the feedback gain matrix $\mathbf{L}$?

Luenberger Observers

Lecture 16

- We next assume that the model is correct. Let's look at the dynamics of the state estimation error $\mathbf{e}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$.

$$
\begin{aligned}
\mathbf{e}_{k+1} &= \mathbf{A}\hat{\mathbf{x}}_k + \mathbf{B}\mathbf{u}_k + \mathbf{L}(\mathbf{H}\mathbf{x}_k - \mathbf{H}\hat{\mathbf{x}}_k) - \mathbf{A}\mathbf{x}_k - \mathbf{B}\mathbf{u}_k \\
&= (\mathbf{A}_k - \mathbf{L}\mathbf{H})\mathbf{e}_k
\end{aligned}
$$

- We want the dynamic $\mathbf{e}_{k+1} = (\mathbf{A} - \mathbf{LH})\mathbf{e}_k$ to be stable. This is a control problem. We want all the eigen values of the matrix $\mathbf{A} - \mathbf{LH}$ to be negative (or the real part of the eigen values to be negative).
- Examples: pole placement algorithms (especially suitable when $p = 1$, single input) or Liner Quadratic Regulators.

Kalman Filter

Lecture 16

- If the system model is not deterministic, that is we can better model the real-life situation considering multivariate random noise, process noise $\mathbf{w} \in \mathbb{R}^n$ and measurement (sensor) noise $\mathbf{v} \in \mathbb{R}^m$, making the state and the output essentially random variables,

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1} + \mathbf{G}\mathbf{w}_{k-1}$$
$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k,$$

- then the Kalman filter framework is most suitable. Also, about noise (...), $E[\mathbf{w}_k] = 0_n$, $E[\mathbf{v}_k] = 0_m$ and covariances $E[\mathbf{w}_k\mathbf{w}_k^T] = \mathbf{Q} \in \mathbb{R}^{n\times n}$, and $E[\mathbf{v}_k\mathbf{v}_k^T] = \mathbf{R} \in \mathbb{R}^{m\times m}$.

Kalman Filter
Lecture 16

- Kalman Filter Predict Step:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}$$
$$\mathbf{P}_{k|k-1} = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T$$

- Kalman Filter Update Step:

$$\hat{\mathbf{y}}_k = \mathbf{H}\hat{\mathbf{x}}_{k|k-1}; \quad \mathbf{P}_y = \mathbf{H}\mathbf{P}_{k|k-1}\mathbf{H}^T; \quad \mathbf{P}_{xy} = \mathbf{P}_{k|k-1}\mathbf{H}^T$$
$$\mathbf{K}_k = \mathbf{P}_{xy}(\mathbf{P}_y + \mathbf{R})^{-1}$$
$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_k)$$
$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{k|k-1}$$

- with some given initial conditions $\hat{\mathbf{x}}_0$, $\mathbf{P}_0$.

Kalman Filter: Derivations

Lecture 16

- In a “deterministic” framework, using a least-squares approach
- In a probabilistic framework imposing a minimum variance criteria (and that the result should be statistically unbiased)
- In a Bayesian probabilistic framework, a more general approach which leads also to the Particle Filter estimation algorithm
- . . .

Kalman Filter: LS Derivation

Lecture 16

- Consider a given series of measurements $\mathbf{y}_k$, $k \in \overline{1, N}$ where $\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{v}_k$, and $\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}$, all $\mathbf{u}_k$ known.
- We have essentially a single unknown, e.g. $\mathbf{x}_0$, or $\mathbf{x}_N$ or any $\mathbf{x}_k$.

$$
\begin{aligned}
\mathbf{x}_1 &= \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{u}_0 \\
\mathbf{x}_2 &= \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{u}_1 \\
&\vdots \\
\mathbf{x}_k &= \mathbf{A}^k\mathbf{x}_0 + \sum_{i=0}^{k-1} \mathbf{A}^{k-i-1}\mathbf{B}\mathbf{u}_i
\end{aligned}
\qquad
\begin{aligned}
\mathbf{x}_N &= \mathbf{A}\mathbf{x}_{N-1} + \mathbf{B}\mathbf{u}_{N-1} \\
\mathbf{x}_{N-1} &= \mathbf{A}^{-1}\mathbf{x}_N - \mathbf{A}^{-1}\mathbf{B}\mathbf{u}_{N-1} \\
&\vdots \\
\mathbf{x}_k &= \left(\mathbf{A}^{-1}\right)^{N-k}\mathbf{x}_N - \sum_{i=0}^{N-k} \left(\mathbf{A}^{-1}\right)^{N-k-i+1}\mathbf{B}\mathbf{u}_{N-i}
\end{aligned}
$$

Kalman Filter: LS Derivation

Lecture 16

- We'll try to optimize the following cost function:

$$J_N(\mathbf{x}_N) = \frac{1}{2}\sum_{k=1}^{N} \epsilon_k^T \mathbf{R}^{-1} \epsilon_k = \frac{1}{2}\sum_{k=1}^{N} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k)^T \mathbf{R}^{-1} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k) \in \mathbb{R}^+$$

$$= \frac{1}{2}\sum_{k=1}^{N} \left(\mathbf{y}_k - \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k} \mathbf{x}_N + \sum_{i=0}^{N-k} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k-i+1} \mathbf{B}\mathbf{u}_{N-i} \right)^T \mathbf{R}^{-1}$$

$$\left(\mathbf{y}_k - \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k} \mathbf{x}_N + \sum_{i=0}^{N-k} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k-i+1} \mathbf{B}\mathbf{u}_{N-i} \right)$$

- by setting the derivative to zero $\frac{\partial J_N(\mathbf{x}_N)}{\partial \mathbf{x}_N} = 0$.

Kalman Filter: LS Derivation
Lecture 16

- Why $\mathbf{R}^{-1}$? To scale the noises on each channel in the objective function. Imagine two measurement channels, each with a different level of noise σ_1 and σ_2. Therefore the appropriate residuals in the cost function are scaled such that they are all equivalent to a $\sigma = 1$, $J_N = \left(\frac{1}{\sigma_1}\epsilon_1\right)^2 + \left(\frac{1}{\sigma_2}\epsilon_2\right)^2$.
- Similary to LS in slide 7, using matrix calculus, we can show that

$$\frac{d}{d\mathbf{x}}(\mathbf{Ax}+\mathbf{b})^T\mathbf{R}^{-1}(\mathbf{Ax}+\mathbf{b}) = 2(\mathbf{Ax}+\mathbf{b})^T\mathbf{R}^{-1}\mathbf{A}$$

Kalman Filter: LS Derivation

Lecture 16

$$0 = \sum_{k=1}^{N} \left(\mathbf{y}_k - \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k} \mathbf{x}_N + \sum_{i=0}^{N-k} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k-i+1} \mathbf{B}\mathbf{u}_{N-i} \right)^T \mathbf{R}^{-1} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k}$$

$$\hat{x}_N = \Bigg(\underbrace{\sum_{k=1}^{N} \left[\mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k}\right]^T \mathbf{R}^{-1} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k}}_{\triangleq \mathbf{I}_N} \Bigg)^{-1} \cdot$$

$$\underbrace{\sum_{k=1}^{N} \left[\mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k}\right]^T \mathbf{R}^{-1} \left(\mathbf{y}_k + \sum_{i=0}^{N-k} \mathbf{H}\left(\mathbf{A}^{-1}\right)^{N-k-i+1} \mathbf{B}\mathbf{u}_{N-i} \right)}_{\triangleq \mathbf{M}_N}$$

Kalman Filter: LS Derivation
Lecture 16

- $\hat{\mathbf{x}}_N = \mathbf{I}_N^{-1}\mathbf{M}_{N} \ldots \quad \mathbf{M}_N = \ldots \hat{\mathbf{x}}_N$
- $\hat{\mathbf{x}}_{N+1} = \mathbf{I}_{N+1}^{-1} \ldots \mathbf{M}_{N+1}$
- $\mathbf{I}_{N+1} = \sum_{k=\ldots}^{N\ldots} \left[\mathbf{H}\left(\mathbf{A}^{-1}\right)^{\ldots-k+\ldots}\right]^T \mathbf{R}^{-1} \ldots \left(\mathbf{A}^{-1}\right)^{\ldots-k+1} \ldots$
  $\left(\mathbf{A}^{-1}\right)^T \mathbf{I}_N \left(\mathbf{A}^{-1}\right) + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$
- (whiteboard) $\mathbf{M}_{N+1} = \left(\mathbf{A}^{-1}\right)^T \mathbf{M}_N + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1} + \left(\mathbf{A}^{-1}\right)^T \mathbf{I}_N \left(\mathbf{A}^{-1}\right)\mathbf{B}\mathbf{u}_N =$
  $= \left(\mathbf{A}^{-1}\right)^T \mathbf{I}_N \hat{\mathbf{x}}_N + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1} + \left(\mathbf{A}^{-1}\right)^T \mathbf{I}_N \left(\mathbf{A}^{-1}\right)\mathbf{B}\mathbf{u}_N$

Kalman Filter: LS Derivation

Lecture 16

$$\hat{\mathbf{x}}_{N+1} = \left[\mathbf{A}^{-T}\mathbf{I}_N\mathbf{A}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\left[\underbrace{\mathbf{A}^{-T}\mathbf{I}_N\hat{\mathbf{x}}_N + \mathbf{A}^{-T}\mathbf{I}_N\mathbf{A}^{-1}\mathbf{B}\mathbf{u}_N}_{\mathbf{A}^{-T}\mathbf{I}_N\mathbf{A}^{-1}(\mathbf{A}\hat{\mathbf{x}}_N + \mathbf{B}\mathbf{u}_N)} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1}\right]$$

$$= \left[\mathbf{A}^{-T}\mathbf{I}_N\mathbf{A}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\left[\mathbf{A}^{-T}\mathbf{I}_N\mathbf{A}^{-1}\hat{\mathbf{x}}_{N+1|N} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1}\right]$$

$$= \left[\mathbf{P}_{N+1|N}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\left[\mathbf{P}_{N+1|N}^{-1}\hat{\mathbf{x}}_{N+1|N} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1}\right]$$

Where

$$\mathbf{P}_N = \mathbf{I}_N^{-1}; \qquad \mathbf{P}_{N+1|N} \triangleq \mathbf{A}\mathbf{I}_N^{-1}\mathbf{A}^T$$

$$\hat{\mathbf{x}}_{N+1|N} = \mathbf{A}\hat{\mathbf{x}}_N + \mathbf{B}\mathbf{u}_N$$

Kalman Filter: LS Derivation

Lecture 16

Matrix magic, again the Woodbury inversion lemma and also a related equality:

$$\left(\mathbf{P}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1} = \mathbf{P} - \mathbf{P}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}$$

$$\left(\mathbf{P}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{R}^{-1} = \mathbf{P}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}\right)^{-1}$$

Thus,

$$\hat{\mathbf{x}}_{N+1} = \left[\mathbf{P}^{-1}_{N+1|N} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\mathbf{P}^{-1}_{N+1|N}\hat{\mathbf{x}}_{N+1|N} \qquad \text{\#apply Woodbury}$$

$$+ \left[\mathbf{P}^{-1}_{N+1|N} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right]^{-1}\mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_{N+1} \qquad \text{\#apply related}$$

Kalman Filter: LS Derivation

Lecture 16

$$
\begin{aligned}
\hat{\mathbf{x}}_{N+1} &= \left[\mathbf{P}_{N+1|N} - \mathbf{P}_{N+1|N}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{N+1|N}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}_{N+1|N}\right]\mathbf{P}^{-1}_{N+1|N}\hat{\mathbf{x}}_{N+1|N} \\
&\quad + \mathbf{P}_{N+1|N}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{N+1|N}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{y}_{N+1} \\
&= \hat{\mathbf{x}}_{N+1|N} + \mathbf{K}_{N+1}\big(\mathbf{y}_{N+1} - \underbrace{\mathbf{H}\hat{\mathbf{x}}_{N+1|N}}_{\hat{y}}\big).
\end{aligned}
$$

where $\mathbf{K}_{N+1} = \mathbf{P}_{N+1|N}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{N+1|N}\mathbf{H}^T + \mathbf{R}\right)^{-1}$

Also,

$$
\begin{aligned}
\mathbf{P}_{N+1} = \mathbf{I}^{-1}_{N+1} &= \left(\mathbf{A}^{-T}\mathbf{P}_N^{-1}\mathbf{A}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}\right)^{-1} \qquad \text{\#apply Woodbury} \\
&= \mathbf{P}_{N+1|N} - \mathbf{P}_{N+1|N}\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_{N+1|N}\mathbf{H}^T + \mathbf{R}\right)^{-1}\mathbf{H}\mathbf{P}_{N+1|N} \\
&= \mathbf{P}_{N+1|N} - \mathbf{K}_{N+1}\mathbf{H}\mathbf{P}_{N+1|N} = \left(\mathbf{I} - \mathbf{K}_{N+1}\mathbf{H}\right)\mathbf{P}_{N+1|N}
\end{aligned}
$$

Kalman Filter: Joseph Form

Lecture 16

- A covariance update equation that is more numerically stable

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H})^T + \mathbf{K}_k\mathbf{R}\mathbf{K}_k^T$$

- Exercise: $\mathbf{P}_k = \mathrm{Cov}(\hat{\mathbf{x}}_k) = \mathrm{Cov}(\hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}\hat{\mathbf{x}}_{k|k-1}))$. Can you show the Joseph form?
- Exercise: Replace $\mathbf{K}_k$ into the Joseph formula to obtain the previous covariance update formula, i.e. $\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H})\mathbf{P}_{k|k-1}$;
- Exercise: Show also that this version is equivalent $\mathbf{P}_k = \mathbf{P}_{k|k-1} - \mathbf{K}(\mathbf{P}_y + \mathbf{R})\mathbf{K}^T$;
- Square-root Kalman filter (discussion)

Kalman Filter: Exercise

Lecture 16

- Consider the following discrete dynamical system,

$$\mathbf{x}_{k+1} = \begin{bmatrix} ? & .1 & .2 \\ .?5 & .8 & .1 \\ .?5 & .1 & .7 \end{bmatrix} \mathbf{x}_k + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \mathbf{w}_k; \quad \mathbf{y}_k = \begin{pmatrix} 1 & ? & 0 \\ 1 & ? & 1 \\ 0 & ? & 1 \end{pmatrix} \mathbf{x}_k + \mathbf{v}_k;$$

  where $\mathbf{x}_k, \mathbf{y}_k, \mathbf{v}_k \in \mathbb{R}^3$ and $\mathbf{w}_k \in \mathbb{R}^2$. Take the initial state of the system to be $\mathbf{x}_0 = \begin{bmatrix} 20.0 & 50.0 & 30.0 \end{bmatrix}^T$, and the noise covariances to be diagonal such that $\mathbf{Q} = \text{diag}(\begin{bmatrix} 1^2 & 1^2 \end{bmatrix})$ and $\mathbf{R} = (\text{diag}\begin{bmatrix} 5^2 & 5^2 & 5^2 \end{bmatrix})$.
- Check the observability of the system
- Implement a system simulator and a linear Kalman filter.

Kalman Filter: Exercise

Lecture 16

Consider the following properties of the real-system the above is a model of:

- The transition matrix is a left stochastic matrix (sum of all elements on the column is 1)
- The sum of all the state elements should be constant (and with these initial conditions, should be 100).
- State elements are all non-negative.

Implement this new information in the simulation/model (what is needed ?) and the evaluate if the KF estimation fulfills these (run a check on the estimate after each step).

Equality Constrained Kalman Filter

Lecture 16

- Consider the following generic equality constraint that we would like our system to uphold: $\mathbf{M}\mathbf{x} = \mathbf{b}$, $\mathbf{M} \in \mathbb{R}^{r \times n}$, and $\mathbf{b} \in \mathbb{R}^{r}$;
- [Gupta et al. 2007] present how we can correct both pairs $(\mathbf{x}_{k|k-1}, \mathbf{P}_{k|k-1})$ and $(\mathbf{x}_k, \mathbf{P}_k)$ in the following manner,
- Let $(\mathbf{x}, \mathbf{P} = \text{cov}(\mathbf{x}))$ be known quantities, to be modified into $(\mathbf{x}^C, \mathbf{P}^C)$ such that

$$\mathbf{x}^C = \arg\max_{\mathbf{z} \in \mathbb{R}^n} \left\{ (\mathbf{z} - \mathbf{x})^T \mathbf{W} (\mathbf{z} - \mathbf{x}) : \mathbf{M}\mathbf{z} = \mathbf{b} \right\}$$

- Then $\mathbf{x}^C = \mathbf{x} - \Lambda(\mathbf{M}\mathbf{x} - \mathbf{b})$ and $\mathbf{P}^C = (\mathbf{I}_n - \Lambda\mathbf{M})\mathbf{P}$; where $\Lambda = \mathbf{W}^{-1}\mathbf{M}^T \left(\mathbf{M}\mathbf{W}^{-1}\mathbf{M}^T\right)^{-1}$ and weight matrix $\mathbf{W} = \mathbf{P}^{-1}$ or $\mathbf{W} = \mathbf{I}_n$.

Equality Constrained Kalman Filter (cont)

Lecture 16

- Exercise: Add this to your algorithm for the previous problem and verify that the sum check is better.
- The more general case of inequality constraints does not have an analytic solution. A numerical solution can be calculated (so one or two for each KF step).

Well done!