Ch. 13
Kalman Filters

Introduction

In 1960, Rudolf Kalman developed a way to solve some of the practical difficulties that arise when trying to apply Weiner filters.

There are D-T and C-T versions of the Kalman Filter… we will only consider the D-T version.

The Kalman filter is widely used in:

- Control Systems
- Navigation Systems
- Tracking Systems

KF initially arose in the field of control systems – in order to make a system do what you want, you must know what it is doing now

It is less widely used in signal processing applications

The Three Keys to Leading to the Kalman Filter

Wiener Filter: LMMSE of a Signal (i.e., a Varying Parameter)

Sequential LMMSE: Sequentially Estimate a *Fixed* Parameter

State-Space Models: Dynamical Models for Varying Parameters

Kalman Filter: Sequential LMMSE Estimation for a time-varying parameter vector – but the time variation is constrained to follow a "state-space" dynamical model.

Aside: There are many ways to mathematically model dynamical systems…

- Differential/Difference Equations
- Convolution Integral/Summation
- Transfer Function via Laplace/Z transforms
- State-Space Model

13.3 State-Variable Dynamical Models

System *State*: the collection of variables needed to know how to determine how the system will "exist" at some future time (in the absence of an input)…

For an RLC circuit… you need to know all of its current capacitor voltages and all of its current inductor currents

Motivational Example: *Constant* Velocity Aircraft in 2-D

$$\mathbf{s}(t) = \begin{bmatrix} r_x(t) \\ r_y(t) \\ v_x(t) \\ v_y(t) \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \text{A/C positions (m)} \\ \left.\vphantom{\begin{matrix} a \\ b \end{matrix}}\right\} \text{A/C velocities (m/s)} \end{matrix}$$

For the constant velocity model we would constrain $v_x(t)$ & $v_y(t)$ to be constants V_x & V_y.

If we know $\mathbf{s}(t_o)$ and there is no input we know how the A/C behaves for all future times: $r_x(t_o + \tau) = V_x\tau + r_x(t_o)$

$$\boldsymbol{r_x(t_o + \tau) = r_x(t_o) + V_x\tau}$$
$$\boldsymbol{r_y(t_o + \tau) = r_y(t_o) + V_y\tau}$$

D-T State Model for Constant Velocity A/C

Because measurements are often taken at discrete times… we often need D-T models for what are otherwise C-T systems

(This is the same as using a difference equation to approximate a differential equation)

If every increment of n corresponds to a duration of Δ sec and there is no driving force then we can write a D-T State Model as:

$$\boxed{\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1]}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

State Transition Matrix

$$r_x[n] = r_x[n\text{-}1] + v_x[n\text{-}1]\Delta$$

$$r_y[n] = r_y[n\text{-}1] + v_y[n\text{-}1]\Delta$$

$$v_x[n] = v_x[n\text{-}1]\Delta$$

$$v_y[n] = v_y[n\text{-}1]\Delta$$

We can include the effect of a vector input:

$$\boxed{\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1] + \mathbf{B}\mathbf{u}[n]}$$

Input could be deterministic and/or random.
Matrix **B** combines inputs & distributes them to states.

Thm 13.1 Vector Gauss-Markov Model

Don't confuse with the G-M Thm. of Ch. 6

This theorem characterizes the probability model for a specific state-space model with Gaussian Inputs

Linear State Model: $\mathbf{s}[n] = \mathbf{A}\mathbf{s}[n-1] + \mathbf{B}\mathbf{u}[n]$ $\quad n \geq 0$

$p\times 1$ $\quad$ $p\times p$ known $\quad$ $p\times r$ known $\quad$ $r\times 1$

$\mathbf{s}[n]$: "state vector" is a vector Gauss-Markov <u>process</u>

$\mathbf{A}$: "state transition matrix"; assumed $|\lambda_i| < 1$ for stability

eigenvalues

$\mathbf{B}$: "input matrix"

$\mathbf{u}[n]$: "driving noise" is vector WGN w/ zero mean

$\mathbf{s}[-1]$: "initial state" ~ $N(\boldsymbol{\mu}_s, \mathbf{C}_s)$ and independent of $\mathbf{u}[n]$

$$\mathbf{u}[n] \sim N(\mathbf{0}, \mathbf{Q})$$

$$E\{\mathbf{u}[n]\,\mathbf{u}^T[m]\} = \mathbf{0}, \quad n \neq m$$

Theorem:

- $\mathbf{s}[n]$ for $n \geq 0$ is Gaussian with the following characteristics…
- Mean of state vector is $\boxed{E\left\{\mathbf{s}[n]\right\} = \mathbf{A}^{n+1}\boldsymbol{\mu}_{\mathbf{s}}}$

diverges if e-values have $|\lambda_i| \geq 1$

- Covariance between state vectors at m and n is

$$\text{for} \quad m \geq n: \quad \mathbf{C}_{\mathbf{s}}[m,n] = E\left\{[\mathbf{s}[m] - E\{\mathbf{s}[m]\}][\mathbf{s}[n] - E\{\mathbf{s}[n]\}]^T\right\}$$

$$= \mathbf{A}^{m+1}\mathbf{C}_{\mathbf{s}}\left(\mathbf{A}^{n+1}\right)^T + \sum_{k=m-n}^{m} \mathbf{A}^k \mathbf{B}\mathbf{Q}\mathbf{B}^T \left(\mathbf{A}^{n-m+k}\right)^T$$

$$\text{for} \quad m < n: \quad \mathbf{C}_{\mathbf{s}}[m,n] = \mathbf{C}_{\mathbf{s}}^T[n,m]$$

State Process is <u>Not</u> WSS!

- Covariance Matrix: $\mathbf{C}[n] = \mathbf{C}_{\mathbf{s}}[n,n]$ (this is just notation)
- Propagation of Mean & Covariance:

$$E\left\{\mathbf{s}[n]\right\} = \mathbf{A}E\left\{\mathbf{s}[n-1]\right\}$$

$$\mathbf{C}[n] = \mathbf{A}\mathbf{C}[n-1]\mathbf{A}^T + \mathbf{B}\mathbf{Q}\mathbf{B}^T$$

Proof: (only for the scalar case: $p = 1$)

differs a bit from (13.1) etc.

For the scalar case the model is: $s[n] = a\, s[n\text{-}1] + b\, u[n] \qquad n \geq 0$

Now we can just iterate this model and surmise its general form:

$$\boxed{s[0] = as[-1] + bu[0]}$$

$$s[1] = as[0] + bu[1]$$
$$= a^2 s[-1] + abu[0] + bu[1]$$

$$s[2] = as[1] + bu[2]$$
$$= a^3 s[-1] + a^2 bu[0] + abu[1] + bu[2]$$

⋮

$$s[n] = \underbrace{a^{n+1} s[-1]}_{\text{z.i. response … exponential}} + \underbrace{\sum_{k=0}^{n} a^k bu[n-k]}_{\text{z.s. response … convolution}}$$

Now easy to find the mean:

$$E\{s[n]\} = a^{n+1} \underbrace{E\{s[-1]\}}_{=\mu_s} + \sum_{k=0}^{n} a^k b \underbrace{E\{u[n-k]\}}_{=0}$$

$$= a^{n+1} \mu_s \qquad \text{… as claimed!}$$

Covariance between $s[m]$ and $s[n]$ is:

$$C_s[m,n] = E\left\{[s[m]-a^{m+1}\mu_s][s[n]-a^{n+1}\mu_s]^T\right\}$$

$$= E\left\{\left(a^{m+1}[s[m]-\mu_s]+\sum_{k=0}^{m}a^k bu[m-k]\right)\times\right.$$

Must use <u>different</u> dummy variables!!

Cross-terms will be zero… why?

$$\left.\left(a^{n+1}[s[n]-\mu_s]+\sum_{l=0}^{n}a^l bu[n-l]\right)\right\}$$

$$= a^{m+1}a^{n+1}\sigma_s^2 + \sum_{k=0}^{m}\sum_{l=0}^{m}a^k b\underbrace{E\{u[m-k]u[n-l]\}}_{=\sigma_u^2\delta[l-(n-m+k)]}ba^l$$

For $m \geq n$: $C_s[m,n] = a^{m+1}a^{n+1}\sigma_s^2 + \sum_{k=m-n}^{m}a^k b\sigma_u^2 ba^{n-m+k}$

For $m < n$: $C_s[m,n] = C_s[n,m]$

For mean & cov. propagation: from $s[n] = a\,s[n-1] + b\,u[n]$

$$\underbrace{E\{s[n]\} = aE\{s[n-1]\}}_{\text{propagates as in theorem}} + b\underbrace{E\{u[n]\}}_{=0}$$

$$\begin{aligned}
\operatorname{var}\{s[n]\} &= E\left\{(s[n] - E\{s[n]\})^2\right\} \\
&= E\left\{(as[n-1] + bu[n] - aE\{s[n-1]\})^2\right\} \\
&= a\underbrace{E\left\{(s[n-1] - E\{s[n-1]\})^2\right\}}_{=\operatorname{var}\{s[n-1]\}}a + b\underbrace{E\left\{u^2[n]\right\}}_{=\sigma_u^2}b
\end{aligned}$$

… which propagates as in theorem **< End of Proof >**

<u>So we now have:</u>

- **Random Dynamical Model (A State Model)**
- **Statistical Characterization of it**

Random Model for "Constant" Velocity A/C

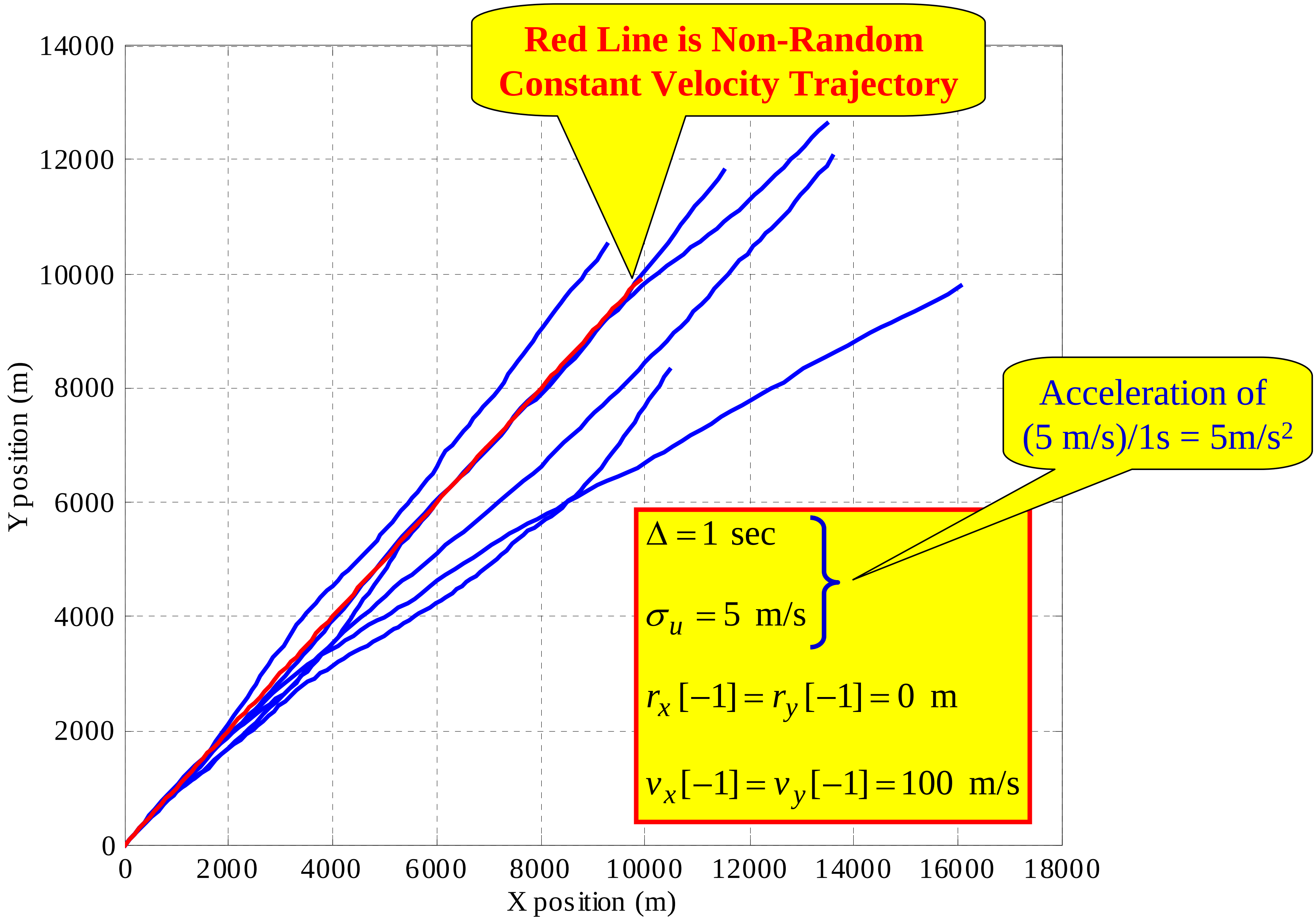
$$\mathbf{s}[n] = \begin{bmatrix} r_x[n] \\ r_y[n] \\ v_x[n] \\ v_y[n] \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_x[n-1] \\ r_y[n-1] \\ v_x[n-1] \\ v_y[n-1] \end{bmatrix}}_{\text{Deterministic Propagation of Constant-Velocity}} + \begin{bmatrix} 0 \\ 0 \\ u_x[n] \\ u_y[n] \end{bmatrix}$$

Deterministic Propagation of Constant-Velocity

Random Perturbation of Constant Velocities

$$\operatorname{cov}\{\mathbf{u}[n]\} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ 0 & 0 & 0 & \sigma_u^2 \end{bmatrix}$$

Ex. Set of "Constant-Velocity" A/C Trajectories



Observation Model

So… we have a random state-variable model for the dynamics of the "signal" (… the "signal" is often some true A/C trajectory)

We need to have some observations (i.e., measurements) of the "signal"

- In Navigation Systems… inertial sensors make noisy measurements at intervals of time

- In Tracking Systems… sensing systems make noisy measurements (e.g., range and angles) at intervals of time

Linear Observation Model: $\boxed{\mathbf{x}[n] = \mathbf{H}[n]\mathbf{s}[n] + \mathbf{w}[n]}$

Measured "observation" vector at each time

Observation Matrix …can change w/ time

State Vector Process being observed

Vector Noise Process

allows multiple measurements at each time

The Estimation Problem

Observe a Sequence of Observation Vectors $\{\mathbf{x}[0], \mathbf{x}[1], \ldots \mathbf{x}[n]\}$

Compute an $\underbrace{\text{Estimate of the State Vector } \mathbf{s}[n]}_{\hat{\mathbf{s}}[n|n]}$

using observation up to n

estimate state at n

Notation: $\hat{\mathbf{s}}[n|m]$ = Estimate of $\mathbf{s}[n]$ using $\{\mathbf{x}[0], \mathbf{x}[1], \ldots \mathbf{x}[m]\}$

Want Recursive Solution:

Given: $\hat{\mathbf{s}}[n|n]$ and a new observation vector $\mathbf{x}[n+1]$

Find: $\hat{\mathbf{s}}[n+1|n+1]$

Three Cases of Interest:

- **Scalar State – Scalar Observation**
- **Vector State – Scalar Observation**
- **Vector State – Vector Observation**