REF. NO. KF #8

1

STOCHASTIC ESTIMATION

The Continuous Time Kalman-Bucy Filter

2

MOTIVATION

- In many continuous time decision and control problems accurate estimates of the system state variables are needed continuously in time to generate real time decision and feedback control strategies.

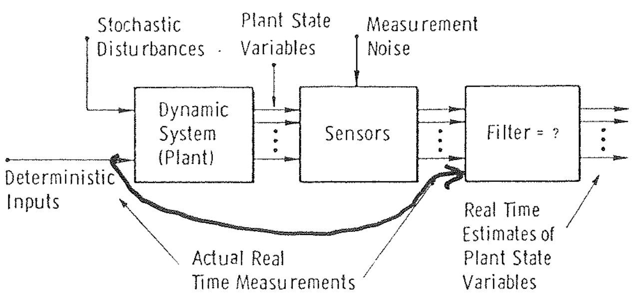
3

- Basic problem:

Design a physical system or data-processing algorithm that generates "good" on-line estimates of the plant state variables based upon unreliable sensor measurements.

4

VISUALIZATION



5

MATHEMATICAL MODELLING

- State Dynamics

$$\dot{\underline{x}}(t) = \underline{A}(t)\underline{x}(t) + \underline{B}(t)\underline{u}(t) + \underline{L}(t)\underline{\xi}(t) \qquad (1)$$

- Measurement Equation

$$\underline{z}(t) = \underline{C}(t)\underline{x}(t) + \underline{\theta}(t) \qquad (2)$$

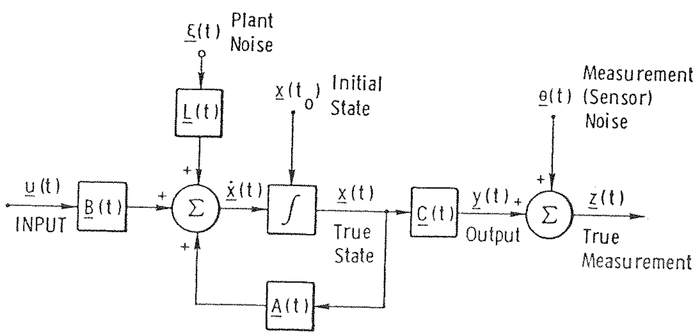
6

- Variable Dimensions

$$\left.\begin{array}{ll} \underline{x}(t),\ \dot{\underline{x}}(t) \in R_n & \underline{\xi}(t) \in R_p \\ \underline{u}(t) \in R_m & \underline{\theta}(t) \in R_r \\ \underline{y}(t) \in R_r & \underline{z}(t) \in R_r \end{array}\right\} \qquad (3)$$

7

VISUALIZATION



8

ASSUMPTIONS

- Deterministic Quantities (Assumed known exactly for all $t \geq t_0$

System matrix: $\underline{A}(t)$ $(n \times n)$

Control gain matrix: $\underline{B}(t)$ $(n \times m)$

Plant noise gain matrix: $\underline{L}(t)$ $(n \times p)$

Measurement gain matrix: $\underline{C}(t)$ $(r \times n)$

Control input vector: $\underline{u}(t)$ $(m \times 1)$

9

- STOCHASTIC QUANTITIES

Initial state: $\underline{x}(t_0)$

Plant noise: $\underline{\xi}(t)$ } continuous-time white-noise

Measurement noise: $\underline{\theta}(t)$

10

STATISTICAL INFORMATION

Initial State Uncertainty

- $\underline{x}(t_0)$ modelled as random vector

$$E\{\underline{x}(t_0)\} = \underline{\bar{x}}_0 = \text{initial mean state} \quad (4)$$

$$\text{cov}[\underline{x}(t_0);\underline{x}(t_0)] = \underline{\Sigma}_0 = \text{initial state covariance} \quad (5)$$

$$\underline{\Sigma}_0 = \underline{\Sigma}_0' \geq \underline{0} \quad (6)$$

11

- These quantities are used to model the fact that the initial state (initial conditions) are not precisely known
- $\underline{\bar{x}}_0$ tells mathematics best "guess" on value of initial state
- $\underline{\Sigma}_0$ tells mathematics how much to "believe" $\underline{\bar{x}}_0$ (via specification of standard deviations, etc)

12

PLANT NOISE $\underline{\xi}(t)$ ← Continuous-time

- $\underline{\xi}(t)$ is a white noise stochastic process

$$E\{\underline{\xi}(t)\} = \underline{0} \text{ for all } t \quad (7)$$

$$\text{cov}[\underline{\xi}(t);\underline{\xi}(\tau)] = \underline{\Xi}(t)\,\delta(t-\tau) \quad (8)$$

- $\underline{\Xi}(t)$ called plant noise intensity matrix

$$\underline{\Xi}(t) = \underline{\Xi}'(t) \geq \underline{0} \quad (p\times p \text{ matrix})$$

13

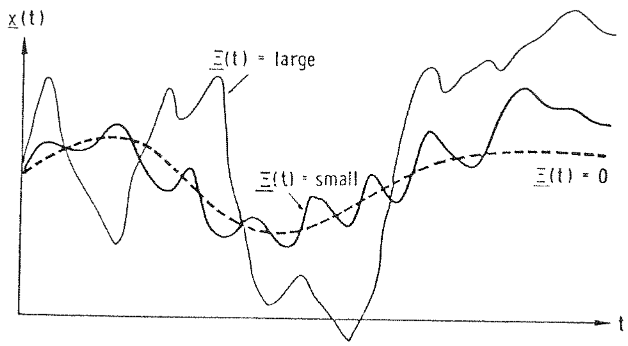
Plant noise $\underline{\xi}(t)$ is used to model

- actuator errors
- external disturbances
- modelling errors in $\underline{A}(t)$, $\underline{B}(t)$, $\underline{L}(t)$

that cause "wiggles" in state $\underline{x}(t)$

- The "larger" $\underline{\Xi}(t)$, the greater the plant uncertainty, the "more random" the state $\underline{x}(t)$

14



15

MEASUREMENT NOISE $\underline{\theta}(t)$ ← continuous-time

- $\underline{\theta}(t)$ is a white noise stochastic process

$$E\{\underline{\theta}(t)\} = \underline{0} \text{ for all } t \qquad (10)$$

$$\text{cov}\left[\underline{\theta}(t);\underline{\theta}(\tau)\right] = \underline{\Theta}(t)\delta(t-\tau) \qquad (11)$$

- $\underline{\Theta}(t)$ is called the measurement intensity matrix $(r \times r)$

$$\underline{\Theta}(t) = \underline{\Theta}'(t) > \underline{0} \qquad (12)$$

⟹ every measurement contains white noise

⟹ $\underline{\Theta}^{-1}(t)$ exists (13)

16

- Measurement noise $\underline{\theta}(t)$ is used to model
  - actual sensor innacuracies
  - modelling errors in $\underline{C}(t)$
- The "larger" $\underline{\Theta}(t)$, the "noisier" the measurements, the "more high frequency wiggles" in $\underline{z}(t)$

17

- ADDITIONAL ASSUMPTIONS

$\underline{x}(t_0)$, $\underline{\xi}(t)$, $\underline{\theta}(\tau)$ are independent for all t_0, t, τ

$\Rightarrow$

$$\left.\begin{aligned} \text{cov}\left[\underline{x}(t_0);\ \underline{\xi}(t)\right] &= \underline{0}\ \forall\ t_0, t \\ \text{cov}\left[\underline{x}(t_0);\ \underline{\theta}(t)\right] &= \underline{0}\ \forall\ t_0, t \\ \text{cov}\left[\underline{\xi}(t);\ \underline{\theta}(\tau)\right] &= \underline{0}\ \forall\ t, \tau \end{aligned}\right\} \qquad (14)$$

18

- This assumption implies that different, and unrelated, physical phenomena give rise to
  - initial state uncertainty
  - plant disturbances
  - sensor inaccuracies

19

DEFINITION OF FILTERING PROBLEM

- Given
  - past measurements time functions

$$Z(t) \triangleq \left\{\underline{z}(\tau);\ t_0 \leq \tau \leq t\right\} \qquad (15)$$

  - past inputs time functions

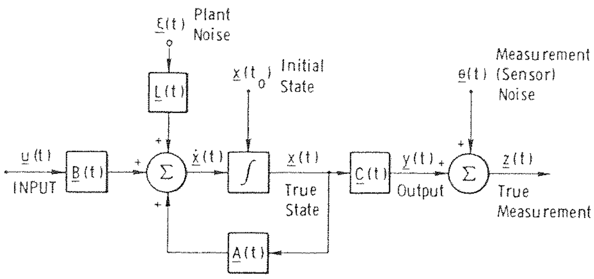
$$U(t) \triangleq \left\{\underline{u}(t);\ t_0 \leq \tau \leq t\right\} \qquad (16)$$

- Find

a vector $\hat{\underline{x}}(t) \in R_n$ which is a "good" estimate of the actual state vector $\underline{x}(t) \in R_n$

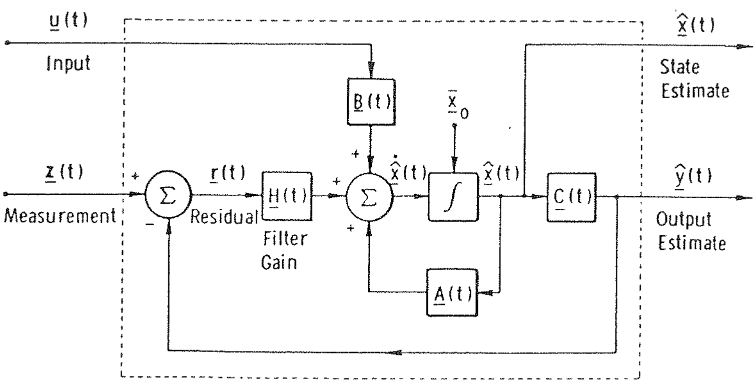
20

VISUALIZATION



21

THE KALMAN-BUCY FILTER



22

MATHEMATICAL SPECIFICATION OF KALMAN-BUCY FILTER

OFF-LINE CALCULATIONS

- Determine nxn matrix $\underline{\Sigma}(t)$ by numerical integration (forward in time) of matrix Riccati equation

$$\frac{d}{dt}\underline{\Sigma}(t) = \underline{A}(t)\,\underline{\Sigma}(t) + \underline{\Sigma}(t)\,\underline{A}'(t) + \underline{L}(t)\,\underline{\Xi}(t)\,\underline{L}'(t) - \underline{\Sigma}(t)\,\underline{C}'(t)\,\underline{\Theta}^{-1}(t)\,\underline{C}(t)\,\underline{\Sigma}(t); \qquad (17)$$

$$\underline{\Sigma}(t_0) = \underline{\Sigma}_0 \qquad (18)$$

- Compute the n x r filter gain matrix $\underline{H}(t)$

$$\underline{H}(t) = \underline{\Sigma}(t)\,\underline{C}'(t)\,\underline{\Theta}^{-1}(t) \qquad (19)$$

23

• ON LINE CALCULATIONS

Construct an (analog or digital) simulation that accepts as "inputs"

- .the actual applied input, $\underline{u}(t)$
- .the actual measurement, $\underline{z}(t)$

and generates the state estimate $\hat{\underline{x}}(t)$ by

$$\boxed{\frac{d}{dt}\hat{\underline{x}}(t) = \underline{A}(t)\hat{\underline{x}}(t) + \underline{B}(t)\underline{u}(t) + \underline{H}(t)\left[\underline{z}(t) - \underline{C}(t)\hat{\underline{x}}(t)\right]} \qquad (20)$$

$$\hat{\underline{x}}(t_0) = \bar{\underline{x}}_0 \qquad (21)$$

Recall eq. (19)
$\underline{H}(t) = \underline{\Sigma}(t)\underline{C}(t)\underline{\Theta}^{-1}(t)$

Residual $\underline{r}(t)$:
$\underline{r}(t) \triangleq \underline{z}(t) - \underline{C}(t)\hat{\underline{x}}(t)$

- $E\{\underline{r}(t)\} = \underline{0}$
- $\operatorname{cov}[\underline{r}(t); \underline{r}(\tau)] = E\{\underline{r}(t)\underline{r}'(\tau)] = \underline{\Theta}(t)\,\delta(t-\tau)$
- Residual is cont. time white noise process

24

OPTIMALITY CONSIDERATIONS

1. Assumption

$$\boxed{\underline{x}(t_0),\ \underline{\xi}(t),\ \underline{\theta}(\tau) \text{ are all gaussian}}$$

• Then KBF generates conditional mean

$$\hat{\underline{x}}(t) = E\{\underline{x}(t)/Z(t), U(t)\} = \hat{\underline{x}}(t/t) \qquad (22)$$

$$\underline{\Sigma}(t) = \operatorname{cov}\left[\underline{x}(t); \underline{x}(t)/Z(t), U(t)\right] = \underline{\Sigma}(t/t) \qquad (23)$$

$$\boxed{p\left(\underline{x}(t)/Z(t), U(t)\right) \text{ is gaussian}} \qquad (24)$$

25

2. Assumption

(a) The gaussian assumption is made

(b) For any estimate $\hat{\underline{x}}(t)$ of the state $\underline{x}(t)$, given past measurement data Z(t) and past input data U(t), one measures the performance by the "least squares" criterion.

$$J = E\left\{\|\underline{x}(t) - \hat{\underline{x}}(t)\|^2 / Z(t), U(t)\right\} \qquad (25)$$

$$= \int_{-\infty}^{+\infty}\!\!\cdots\!\int \|\underline{x}(t) - \hat{\underline{x}}(t)\|^2\, p\Big(\underbrace{\underline{x}(t)/Z(t), U(t)}_{\text{Gaussian}}\Big)\, d\underline{x}(t)$$

Then: The KBF estimate $\hat{\underline{x}}(t)$ is optimal, in the sense that it minimizes J

Additional proofs in next lecture