Stochastic Adaptive Control (02421)

Lecture 4

Tobias K. S. Ritschel

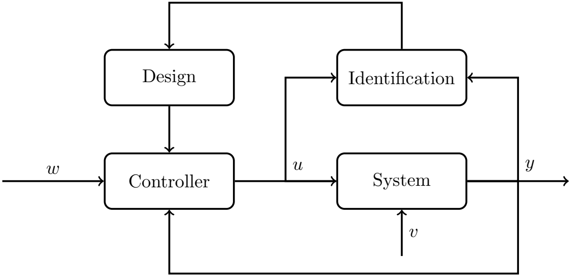
Section for Dynamical Systems, DTU Compute

DTU Compute

Department of Applied Mathematics and Computer Science

Lecture Plan

1. Systems theory
2. Stochastics
3. State estimation - Kalman filter 1
4. **State estimation - Kalman filter 2**
5. Optimal control 1 - internal models
6. External models
7. Prediction + optimal control 1 - external models
8. Optimal control 2 - external models
9. System identification 1
10. System identification 2
11. System identification 3 + model validation
12. Adaptive control 1
13. Adaptive control 2



Today's Agenda

- Follow-up from last lecture
- Kalman assumptions
- Noise correlation
- Kalman errors and properties
- Examples

Follow-up From Last Time: Question 1

Find the stationary distribution of

$$x_{k+1} = \begin{bmatrix} 2/5 & 0 \\ -3/5 & 1/5 \end{bmatrix} x_k + v_k, \qquad v_k \in N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \right), \tag{1}$$

$$y_k = \begin{bmatrix} 1 & 1 \end{bmatrix} x_k + e_k, \qquad e_k \in N(0, 2). \tag{2}$$

The system is linear and the noise is Gaussian. Therefore the distributions are Gaussian:

$$E\{x_{k+1}\} = AE\{x_k\} + E\{v_k\}, \tag{3}$$

$$Var\{x_{k+1}\} = AVar\{x_k\}A^T + Var\{v_k\}, \tag{4}$$

$$E\{y_k\} = CE\{x_k\} + E\{e_k\}, \tag{5}$$

$$Var\{y_k\} = CVar\{x_k\}C^T + Var\{e_k\}. \tag{6}$$

The stationary distributions are

$$x_\infty \in N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.1905 & -0.3106 \\ -0.3106 & 2.6074 \end{bmatrix} \right), \tag{7}$$

$$y_\infty \in N\left(0, 5.1768\right). \tag{8}$$

Follow-up from last time

Questions?

Kalman Filter: The Standard System

The theory behind the Kalman filter is based on the assumption of a linear system with both white process noise and white measurement noise

$$x_{t+1} = Ax_t + Bu_t + v_t \quad v_t \in N(0, R_1), \tag{9}$$

$$y_t = Cx_t + Du_t + e_t \quad e_t \in N(0, R_2), \tag{10}$$

$$CoV(v_t, e_t) = 0, \quad e_t, v_t \text{ white} \perp x_s \quad s \leq t. \tag{11}$$

Furthermore, the noises are assumed to be independent of the state history and each other. In summary,

1. $x_0 \in N(\hat{x}_0, P_0)$
2. $v_t \in N(0, R_1)$, white
3. $e_t \in N(0, R_2)$, white
4. $CoV(v_t, e_t) = 0$
5. $v_t, e_t \perp x_s, \quad s \leq t$

Kalman Filter: Deviations from the Standard Assumptions

We will now consider systems where one of these assumptions do not apply:

1. Non-zero mean process disturbances
2. Non-zero mean output disturbances
3. Colored (non-white) process noise
4. Colored (non-white) output noise
5. Noise correlated with the state
6. Correlated noises

Uncertain Offset in the Process

Consider the system that contains a stochastic process offset,

$$x_{t+1} = Ax_t + Bu_t + Gd_t + v_t, \tag{12}$$

$$y_t = Cx_t + Du_t + e_t, \tag{13}$$

where process offset is given by

$$d_{t+1} = d_t + w_t. \tag{14}$$

We can obtain the standard system description by:

$$\begin{bmatrix} x \\ d \end{bmatrix}_{t+1} = \begin{bmatrix} A & G \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_t + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v \\ w \end{bmatrix}_t, \tag{15}$$

$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_t + Du_t + e_t. \tag{16}$$

Uncertain Offset in the Output

Consider the system that contains a stochastic measurement offset,

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{17}$$

$$y_t = Cx_t + Du_t + Hd_t + e_t, \tag{18}$$

where output offset is given by

$$d_{t+1} = d_t + w_t. \tag{19}$$

We can obtain the standard system description by:

$$\begin{bmatrix} x \\ d \end{bmatrix}_{t+1} = \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_t + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v \\ w \end{bmatrix}_t, \tag{20}$$

$$y_t = \begin{bmatrix} C & H \end{bmatrix} \begin{bmatrix} x \\ d \end{bmatrix}_t + Du_t + e_t. \tag{21}$$

The Case of Coloured Noise

In Kalman filters, we assume that the noise is white.

White noise (discrete-time):
In discrete time, a white noise signal ϵ_t has zero mean, finite variance, and it is time-wise independent: $\epsilon_t \perp \epsilon_s$ for $s \neq t$.

If the noise w_t of a system is colored (non-white), it can be described as a system of white noises (η_t, ξ_t):

$$z_{t+1} = A_w z_t + \eta_t, \tag{22}$$

$$w_t = C_w z_t + \xi_t. \tag{23}$$

The Case of Coloured Process Noise

Consider a system with colored process noise v_t:

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{24}$$

$$y_t = Cx_t + Du_t + e_t. \tag{25}$$

We can obtain a standard system description by:

$$\begin{bmatrix} x \\ z \end{bmatrix}_{t+1} = \begin{bmatrix} A & C_w \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_t + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} \xi \\ \eta \end{bmatrix}_t, \tag{26}$$

$$y_t = \begin{bmatrix} C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_t + Du_t + e_t. \tag{27}$$

The Case of Coloured Output Noise

Let us consider a system with colored measurement noise e_t:

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{28}$$
$$y_t = Cx_t + Du_t + e_t \tag{29}$$

We can then obtain a standard system description by:

$$\begin{bmatrix} x \\ z \end{bmatrix}_{t+1} = \begin{bmatrix} A & 0 \\ 0 & A_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_t + \begin{bmatrix} B \\ 0 \end{bmatrix} u_t + \begin{bmatrix} v \\ \eta \end{bmatrix}_t, \tag{30}$$
$$y_t = \begin{bmatrix} C & C_w \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix}_t + Du_t + \xi_t. \tag{31}$$

Can you think of a disadvantage of augmenting the states?
How can you mitigate it?

Think about it for yourself for one minute and
then discuss with the person next to you for one minute.

Correlation with State History

Consider the system

$$x_{t+1} = Ax_t + Bu_t + Gw_t, \tag{32}$$
$$y_t = Cx_t + Du_t + \eta_t, \tag{33}$$

where the noises w_t and η_t are correlated with the states:

$$w_t = Hx_t + v_t, \tag{34}$$
$$\eta_t = Fx_t + e_t, \tag{35}$$
$$v_t \in N(0, R_1), \quad x_t \perp v_t, \tag{36}$$
$$e_t \in N(0, R_2), \quad x_t \perp e_t. \tag{37}$$

Then we can obtain an uncorrelated description:

$$x_{t+1} = (A + GH)x_t + Bu_t + Gv_t, \tag{38}$$
$$y_t = (C + F)x_t + Du_t + e_t. \tag{39}$$

Ordinary Kalman Filter with Correlated Noise

Consider a system with correlated process and measurement noise:

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \in N(0, R_1), \tag{40}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \in N(0, R_2), \tag{41}$$

$$CoV(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s \quad s \leq t. \tag{42}$$

For this system, the ordinary Kalman filter is

Data Update:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{43}$$

$$P_{t|t} = (1 - \kappa_t C)P_{t|t-1}, \tag{44}$$

$$\kappa_t = P_{t|t-1}C^T(CP_{t|t-1}C^T + R_2)^{-1}. \tag{45}$$

Time Update:

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t} + Bu_t + M(y_t - C\hat{x}_{t|t} - Du_t), \tag{46}$$

$$= (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t, \tag{47}$$

$$P_{t+1|t} = (A - MC)P_{t|t}(A - MC)^T + R_1 - MR_{12}, \quad M = R_{12}R_2^{-1}. \tag{48}$$

Predictive Kalman Filter with Correlated Noise

Consider a system with correlated process and measurement noise:

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \in N(0, R_1), \tag{49}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \in N(0, R_2), \tag{50}$$

$$CoV(v_t, e_t) = R_{12}, \quad e_t, v_t \text{ white} \perp x_s \quad s \leq t. \tag{51}$$

For this system, the predictive Kalman filter is

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{52}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{53}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1}. \tag{54}$$

Despite the difference in the two filters, the gains are still related:

$$K_t = (A - MC)\kappa_t + M. \tag{55}$$

Correlated noise - Derivation (Ordinary)

Consider the system with a slight change:

$$
\begin{aligned}
x_{t+1} &= Ax_t + Bu_t + v_t + M(y_t - y_t) && (56)\\
&= Ax_t + Bu_t + v_t + M(y_t - Cx_t - Du_t - e_t) && (57)\\
&= (A - MC)x_t + (B - MD)u_t + My_t + \tilde{v}_t, \tilde{v}_t = v_t - Me_t). && (58)
\end{aligned}
$$

Define the covariance of $\tilde{v}_t$ and e_t to be zero:

$$
\begin{aligned}
\tilde{R}_{12} &= E\{\tilde{v}_t e_t^T\} = E\{(v_t - Me_t)e_t^T\} && (59)\\
&= R_{12} - MR_2 = 0 \Rightarrow M = R_{12}R_2^{-1}, && (60)\\
\tilde{R}_1 &= E\{\tilde{v}_t \tilde{v}_t^T\} = R_1 - R_{12}R_2^{-1}R_{12}^T. && (61)
\end{aligned}
$$

The correlated version of the ordinary Kalman filter is then defined by the original data update but with the new time update:

$$
\begin{aligned}
\hat{x}_{t+1|t} &= (A - MC)\hat{x}_{t|t} + (B - MD)u_t + My_t && (62)\\
P_{t+1|t} &= (A - MC)P_{t|t}(A - MC)^T + \tilde{R}_1 && (63)
\end{aligned}
$$

Correlated Noise - Derivation (Predicted)

Consider the conditional distribution of the future state and current measurement based on past measurements:

$$\begin{bmatrix} x_{t+1} \\ y_t \end{bmatrix} |Y_{t-1} \in E \left\{ \begin{bmatrix} A\hat{x}_{t|t-1} + Bu_t \\ C\hat{x}_{t|t-1} + Du_t \end{bmatrix}, \begin{bmatrix} AP_{t|t-1}A^T + R_1 & AP_{t|t-1}C^T + R_{12} \\ CP_{t|t-1}A^T + R_{12}^T & CP_{t|t-1}C^T + R_2 \end{bmatrix} \right\}. \tag{64}$$

Considering $x_{t+1}|Y_t \in N(\hat{x}_{t+1|t}, P_{t+1|t})$, and using the projection theorem, we get

$$\hat{x}_{t+1|t} = A\hat{x}_{t|t-1} + Bu_t + K_t(y_t - C\hat{x}_{t|t-1} - Du_t), \tag{65}$$

$$P_{t+1|t} = AP_{t|t-1}A^T + R_1 - K_t(AP_{t|t-1}C^T + R_{12})^T, \tag{66}$$

$$K_t = (AP_{t|t-1}C^T + R_{12})(CP_{t|t-1}C^T + R_2)^{-1}. \tag{67}$$

Correlated Noise - A Special Case

In the case of perfect correlation between the process noise and measurement noise,

$$v_t = Ge_t, \tag{68}$$

$$R_1 = GR_2G^T, \tag{69}$$

$$R_{12} = GR_2, \tag{70}$$

and the stationary predictive Kalman filter is characterized by

$$P_\infty = 0, \quad K_\infty = G. \tag{71}$$

Proof

If $P_{t|t-1} = 0$, the predictive variance and gain become

$$P_{t+1|t} = R_1 - K_t R_{12}^T, \tag{72}$$
$$K_t = R_{12} R_2^{-1}. \tag{73}$$

We substitute the correlation:

$$P_{t+1|t} = G R_2 G^T - K_t (G R_2)^T = 0, \tag{74}$$
$$K_t = G R_2 R_2^{-1} = G, \tag{75}$$

This proves that $P_\infty = 0$ and $K_\infty = G$ is a solution.

Then, if $(A - GC,\, R_1)$ is reachable and $R_2 \succ 0$, we know that it is the only solution (from the Riccatti equation and (63)).

Kalman Errors

For a given system,

$$x_{t+1} = Ax_t + Bu_t + v_t, \quad v_t \in N(0, R_1), \tag{76}$$

$$y_t = Cx_t + Du_t + e_t, \quad e_t \in N(0, R_2), \tag{77}$$

$$e_t, v_t \text{ white} \perp x_s \quad s \le t, \tag{78}$$

the properties of the Kalman errors are well-known provided that the model is correct:

$$\tilde{x}_{t|t} = x_t - \hat{x}_{t|t} \in N(0, P_{t|t}), \tag{79}$$

$$\tilde{x}_{t|t-1} = x_t - \hat{x}_{t|t-1} \in N(0, P_{t|t-1}), \tag{80}$$

$$\epsilon_t = y_t - C\hat{x}_{t|t-1} \in N(0, CP_{t|t-1}C^T + R_2), \text{white.} \tag{81}$$

State Estimation and Its Errors

The estimation errors (Ee) are

$$\tilde{x}_{t|t} = x_t - \hat{x}_{t|t}, \qquad \text{(Ordinary Ee)} \tag{82}$$

$$\tilde{x}_{t|t-1} = x_t - \hat{x}_{t|t-1}, \qquad \text{(Predictive Ee)} \tag{83}$$

and have the stochastic properties

$$\tilde{x}_{t|t} \in N(0, P_{t|t}), \tag{84}$$

$$\tilde{x}_{t|t-1} \in N(0, P_{t|t-1}). \tag{85}$$

The process description of the estimation errors are

$$\tilde{x}_{t+1|t+1} = (I - \kappa_{t+1}C)(A\tilde{x}_{t|t} + v_t) - \kappa_{t+1}e_{t+1}, \tag{86}$$

$$\tilde{x}_{t+1|t} = (A - K_tC)\tilde{x}_{t|t-1} - K_te_t + v_t, \quad K_t = A\kappa_t. \tag{87}$$

The Prediction Error (the Innovation)

The prediction error ϵ_t (also called the innovation) is

$$\epsilon_t = y_t - C\hat{x}_{t|t-1}. \tag{88}$$

The prediction error is white and has the following stochastic properties, when the system has deterministic parameters (correct model):

$$\epsilon_t \in N(0, CP_{t|t-1}C^T + R_2), \text{white}, \tag{89}$$

$$\epsilon_s \perp \epsilon_t, \quad s \neq t. \tag{90}$$

Therefore, the prediction error can be used for

1. Model validation (i.e., validating A, B, $\ldots$).
2. System representation.
3. Fault detection.

Example of Prediction Error

Consider the system

$$x_{t+1} = 0.5x_t + v_t, \qquad v_t \in N(0, 0.1), \tag{91}$$
$$y_t = x_t + e_t, \qquad e_t \in N(0, 0.5). \tag{92}$$

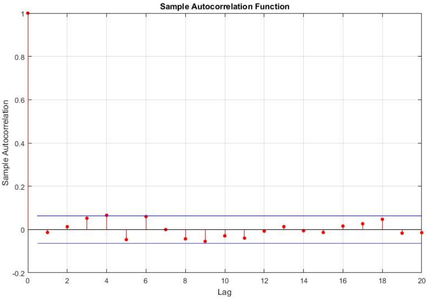
Then, the prediction error should be

$$\epsilon_t \in N(0, 0.625). \tag{93}$$

Based on several simulations,

$$E\{\epsilon_t\} = -0.0047, \tag{94}$$
$$V\{\epsilon_t\} = 0.6218, \tag{95}$$

and the autocorrelation shows that it is white.



Prediction Error: Innovation Form

The prediction error can be used to rewrite the system Consider

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{96}$$
$$y_t = Cx_t + Du_t + e_t, \tag{97}$$

in the innovation form based on predictive Kalman filter,

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K\epsilon_t, \tag{98}$$
$$y_t = C\bar{x}_t + Du_t + \epsilon_t, \tag{99}$$

where K_t is the Kalman gain:

$$K_t = (AP_tC^T + R_{12})(CP_tC^T + R_2)^{-1}. \tag{100}$$

Notice that the innovation form is exactly the case of perfectly correlated process and measurement noise.

Innovation Form: Internal to External (Stationary Case)

If we assume stationary conditions, then, by using the innovation form, we can transform an internal model to an external model.

If we consider the system

$$x_{t+1} = Ax_t + Bu_t + v_t, \tag{101}$$

$$y_t = Cx_t + Du_t + e_t, \tag{102}$$

we can obtain the stationary innovation form

$$\bar{x}_{t+1} = A\bar{x}_t + Bu_t + K\epsilon_t, \tag{103}$$

$$y_t = C\bar{x}_t + Du_t + \epsilon_t, \tag{104}$$

which allows us to define the external model as

$$y_t = H_u(q)u_t + H_\epsilon(q)\epsilon_t \tag{105}$$

$$H_u(q) = C[qI - A]^{-1}B + D \tag{106}$$

$$H_\epsilon(q) = C[qI - A]^{-1}K + 1 \tag{107}$$

$$\tag{108}$$

Trick - Hemes Inversion Lemma

A trick, when computing the variance matrix, is the Hemes inversion lemma:

$$[A + BC^{-1}D^T]^{-1} = A^{-1} - A^{-1}B(C + D^T A^{-1}B)^{-1}D^T A^{-1}. \tag{109}$$

It allows us to rewrite the variance equation

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}C^T(CP_{t|t-1}C^T + R_2)^{-1}CP_{t|t-1}, \tag{110}$$

into its inverse form with corresponding Kalman gain:

$$P_{t|t}^{-1} = P_{t|t-1}^{-1} + C^T R_2^{-1} C, \tag{111}$$

$$\kappa_t = P_{t|t} C^T R_2^{-1}. \tag{112}$$

Consider the scalar case. Which is bigger; $P_{t|t}$ or $P_{t|t-1}$?
Is that what you would expect?

Think about it for yourself for one minute and
then discuss with the person next to you for one minute.

Example: Estimation of Constant State

We want to determine a scalar constant. This can be described as a state estimation problem where

$$x_{t+1} = x_t, \tag{113}$$

$$y_t = x_t + e_t, \quad e_t \in N(0, r_2). \tag{114}$$

If we define $q_t = p_t^{-1}$, we get the inverse form as:

$$\kappa_{t+1} = \frac{p_t}{p_t + r_2} = \frac{1}{1 + r_2 q_t}, \tag{115}$$

$$p_{t+1} = (1 - \kappa_{t+1})p_t, \quad q_{t+1} = q_t + \frac{1}{r_2} = q_0 + t\frac{1}{r_2}, \tag{116}$$

$$\hat{x}_{t+1} = \hat{x}_t + \kappa_{t+1}(y_{t+1} - \hat{x}_t). \tag{117}$$

We can the see that, if $q_0 = 0$ ($p_0 = \infty$), the estimate becomes

$$\hat{x}_{t+1} = \hat{x}_t + \frac{1}{1+t}(y_{t+1} - \hat{x}_t) = \frac{1}{t+1}\sum_{i=1}^{t+1} y_i. \tag{118}$$

The Prediction Problem

Consider the below system with measurement y_t and the desired output z_t.

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + v_t, & v_t &\in N(0, R_v), & (119)\\ y_t &= Cx_t + Du_t + e_t, & e_t &\in N(0, R_e), & (120)\\ z_t &= C_z x_t + D_z u_t + w_t, & w_t &\in N(0, R_w). & (121) \end{aligned}$$

From the Kalman filter, we know that $x_t|Y_t \in N(\hat{x}_{t|t}, P_{t|t})$. If we assume that the future control inputs are known, we can predict the future state estimates by

$$\hat{x}_{t+1|s} = A\hat{x}_{t|s} + Bu_t, \quad t \geq s, \tag{122}$$

$$P_{t+1|s} = AP_{t|s}A^T + R_1. \tag{123}$$

Considering the desired output, we can also derive the conditional predictions of the outputs:

$$z_t|Y_s \in N(C_z x_{t|s} + D_z u_t, C_z P_{t|s} C_z^T + R_w). \tag{124}$$

The Prediction Problem

If we define $Z_{s:t}$ as all future desired outputs in the interval [s:t], and $U_{s:t}$ to be the corresponding control inputs; then the future outputs are given by

$$Z_{s:t} = \begin{bmatrix} z_s \\ \vdots \\ z_t \end{bmatrix} \quad U_{s:t} = \begin{bmatrix} u_s \\ \vdots \\ u_t \end{bmatrix} \quad V_{s:t} = \begin{bmatrix} v_s \\ \vdots \\ v_t \end{bmatrix} \quad W_{s:t} = \begin{bmatrix} w_s \\ \vdots \\ w_t \end{bmatrix}, \tag{125}$$

$$Z_{s:t} = W_o(t-s)x_s + \Pi_{u,t-s}U_{s:t} + \Pi_{v,t-s}V_{s:t} + W_{s:t}. \tag{126}$$

The conditional output $Z_{s:t}|Y_s \in N(\hat{Z}_{s:t|s}, \Sigma_{s:t|s})$ then provides us with a prediction estimate:

$$\hat{Z}_{s:t|s} = W_o(t-s)\hat{x}_{s|s} + \Pi_{u,t-s}U_{s:t}, \tag{127}$$

$$\begin{aligned} \Sigma_{s:t|s} &= W_o(t-s)P_{s|s}W_o^T(t-s) + \Pi_{v,t-s}\tilde{R}_v\Pi_{v,t-s}^T + \tilde{R}_w \\ &+ \Pi_{v,t-s}\tilde{R}_{v,w} + \tilde{R}_{v,w}^T\Pi_{v,t-s}^T, \end{aligned} \tag{128}$$

$$\tilde{R}_x = \text{blkdiag}(R_x, ..., R_x) \in \mathbb{R}^{(t-s)n_x \times (t-s)n_x}, \tag{129}$$

$$\tilde{R}_{x,y} = \text{blkdiag}(R_{x,y}, ..., R_{x,y}) \in \mathbb{R}^{(t-s)n_x \times (t-s)n_y}. \tag{130}$$

Nonlinear Kalman Filters

In nonlinear Kalman filters, we consider the estimation of nonlinear systems:

$$x_{t+1} = f(x_t, u_t, v_t), \quad v_t \in N(0, R_v), \tag{131}$$
$$y_t = g(x_t, u_t, e_t), \quad e_t \in N(0, R_e). \tag{132}$$

We linearize the right-hand side functions:

$$f(x_t, u_t, v_t) \simeq d + Ax_t + Bu_t + Gv_t, \tag{133}$$
$$g(x_t, u_t, v_t) \simeq \delta + Cx_t + Du_t + He_t. \tag{134}$$

The system matrices are the Jacobian matrices:

$$A = \left.\frac{\partial f}{\partial x}\right|_{x_t^*, u_t^*, v_t^*} \qquad B = \left.\frac{\partial f}{\partial u}\right|_{x_t^*, u_t^*, v_t^*} \qquad G = \left.\frac{\partial f}{\partial e}\right|_{x_t^*, u_t^*, v_t^*} \tag{135}$$
$$C = \left.\frac{\partial g}{\partial x}\right|_{x_t^*, u_t^*, e_t^*} \qquad D = \left.\frac{\partial g}{\partial u}\right|_{x_t^*, u_t^*, e_t^*} \qquad H = \left.\frac{\partial g}{\partial e}\right|_{x_t^*, u_t^*, e_t^*} \tag{136}$$
$$R_1 = GR_vG^T, \qquad R_2 = HR_eH^T. \tag{137}$$

Nonlinear Kalman Filters

The Kalman filter for the nonlinear case, can then be formulated based on the approximation in (133) and (134).

Data Update:

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + \kappa_t(y_t - \hat{g}(\hat{x}_{t|t-1}, u_t)), \tag{138}$$

$$\kappa_t = P_{t|t-1}C^T(CP_{t|t-1}C^T + R_2)^{-1}, \tag{139}$$

$$P_{t|t} = [I - \kappa_t C]P_{t|t-1}A. \tag{140}$$

Time Update:

$$\hat{x}_{t+1|t} = \hat{f}(\hat{x}_{t|t}, u_t), \tag{141}$$

$$P_{t+1|t} = AP_{t+1|t}A^T + R_1. \tag{142}$$

Nonlinear Kalman filters are, in general, not optimal.

When is it *not* appropriate to use the extended Kalman filter, i.e., for which types of systems or system properties?

Think about it for yourself for one minute and then discuss with the person next to you for two minutes.

Questions

Questions?