Stochastic Adaptive Control (02421)

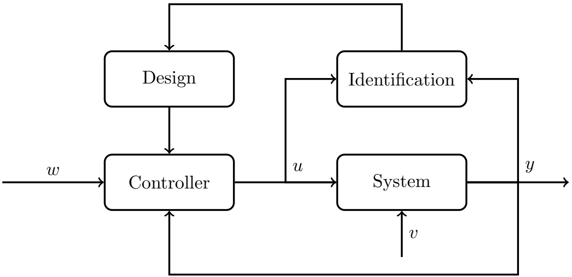
Lecture 6

Tobias K. S. Ritschel

Section for Dynamical Systems, DTU Compute

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
- Model structures
- Spectral signals
- Properties of external systems

Follow-up from last time: General Pole placement

Last time, we saw how we could change the poles of a system on controller canonical form, by a control law:

$$u = -Lx \quad L = \begin{bmatrix} \alpha_1 - a_1 & \dots & \alpha_n - a_n \end{bmatrix} \tag{1}$$

where α and a are the coefficients of the A polynomial of the desired and current system respectively.

We will now consider the general linear system:

$$x_{k+1} = Ax_k + Bu_k \tag{2}$$

and remember the controllability matrix:

$$W_c = \begin{bmatrix} B & AB & A^2B & \dots A^nB \end{bmatrix} \tag{3}$$

Follow-up from last time: General Pole placement

We use a similarity transformation to transform the original system to controller canonical form. The transformed state variables are $\bar{x} = Tx$ and

$$\begin{aligned} \bar{x}_{k+1} &= Tx_{k+1} && (4) \\ &= TAx_k + TBu_k && (5) \\ &= TAT^{-1}\bar{x}_k + TBu_k && (6) \\ &= \bar{A}\bar{x}_k + \bar{B}u_k. && (7) \end{aligned}$$

Next, we compute the feedback matrix such that

$$u_k = -L_{cc}\bar{x}_k. \quad (8)$$

Finally, the feedback law for the original system is

$$\begin{aligned} u_k &= -L_{cc}Tx_k && (9) \\ &= -Lx_k, && (10) \end{aligned}$$

where

$$L = L_{cc}T. \quad (11)$$

Follow-up from last time: General Pole placement

1. Choose the desired poles (eigenvalues of the system matrix), $\{\lambda_{d,i}\}_{i=1}^n$, and compute the actual poles, $\{\lambda_i\}_{i=1}^n$.
2. Compute the polynomial coefficients of the desired, $\{\alpha_i\}_{i=1}^n$, and actual, $\{a_i\}_{i=1}^n$, polynomial:
$$A_d(q) = \prod_{i=1}^n (q - \lambda_{d,i}), \qquad A(q) = \prod_{i=1}^n (q - \lambda_i). \tag{12}$$
In our case, $A_d(q) = q^2 - (\lambda_{d,1} + \lambda_{d,2})q + \lambda_{d,1}\lambda_{d,2}$ and similarly for $A(q)$.
3. Compute the feedback matrix, L_{cc}, for the system transformed to controller canonical form:
$$L_{cc} = \begin{bmatrix} \alpha_1 - a_1 & \dots & \alpha_n - a_n. \end{bmatrix} \tag{13}$$
4. Compute the similarity transformation matrix by
$$T = W_{c,cc} W_c^{-1}, \tag{14}$$
where $W_{c,cc}$ is the controllability matrix of the controller canonical form.
5. Compute the feedback matrix for the original system:
$$L = L_{cc} T. \tag{15}$$

Follow-up from last time

Questions?

Follow-up from last time

The pole placement approach is relatively simple. Which control aspects do we not address with this approach?

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

The Moving-Average (MA) Process

The MA(q) process is defined according to:

The process $\{y_t\}$ given by

$$y_t = \varepsilon_t + \sum_{k=1}^{n} c_k \varepsilon_{t-k}, \quad c_0 = 1 \tag{16}$$

where $\{\varepsilon_k\}$ represents a white-noise process (i.e. independent and Gaussian with variance σ_ε^2), is called a *moving average* process of order n.

The Moving-Average (MA) Process

Let q denote the shift-operator defined according to

$$q^{-1}y_t = y_{t-1}, \tag{17}$$

then the MA(q) process can defined using the compact notation

$$y_t = \theta(q)\varepsilon_t, \tag{18}$$

where the shift-polynomial θ is defined according to

$$\theta(q) = 1 + \sum_{k=1}^{n} \theta_k q^{-k} \tag{19}$$

The corresponding transfer function is given as $\theta(z)$.

$$\theta(z) = \frac{z^n + \sum_{k=1}^{n} \theta_k z^{n-k}}{z^n} \tag{20}$$

The Moving-Average (MA) Process

Finite-order MA processes have the following properties:

- They are always stationary.
- Invertible if the zeros $\theta(z) = 0$ lie within the unit circle.

Remember that a process is invertible if *the innovations can be represented as a function of past observations*.

The Moving-Average (MA) Process

The auto-covariance function of an MA(q) process is given by

$$\gamma(k) = \begin{cases} \sigma_\varepsilon^2 \Big(c_k + c_1 c_{k+1} + \cdots + c_{q-k} c_q \Big), & |k| = 0 ... q \\ 0, & |k| > 0 ... q, \end{cases} \tag{21}$$

and in particular, the (always stationary!) variance is given by

$$\sigma_y^2 = \gamma(0) = \sigma_\varepsilon^2 \left(1 + \sum_{k=1}^{q} c_k^2 \right). \tag{22}$$

The spectral density of the MA(q) process is given by

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \theta \left(e^{i\omega} \right) \theta \left(e^{-i\omega} \right) = \frac{\sigma_\varepsilon^2}{2\pi} \left| 1 + \sum_{k=1}^{q} c_k e^{-ik\omega} \right|^2, \quad \omega \in [-\pi, \pi]. \tag{23}$$

The Auto-Regressive (AR) Process

The AR(m) process is defined according to:

The process $\{y_t\}$ given by

$$y_t + \sum_{k=1}^{m} a_k y_{t-k} = \varepsilon_t, \quad a_0 = 1 \tag{24}$$

where $\{\varepsilon_k\}$ represents a white-noise process, is called an *auto-regressive* process of order m.

The Auto-Regressive (AR) Process

The AR(m) process can defined using the compact notation

$$A(q)y_t = \varepsilon_t, \tag{25}$$

where the shift-polynomial $A(q)$ is defined according to

$$A(q) = 1 + \sum_{k=1}^{m} a_k q^{-k}. \tag{26}$$

The corresponding transfer function is given as $\dfrac{1}{A(z)}$.

The term *auto-regressive* is framed based on the fact that y_t can be viewed as a regression on past values

$$y_t = \varepsilon_t - \sum_{k=1}^{m} a_k y_{t-k}. \tag{27}$$

The Auto-Regressive (AR) Process

Finite-order AR processes have the following properties:

- They are always invertible.
- Stationary if the roots $A(z) = 0$ lie within the unit circle.

The equation

$$A(z) = 0 \tag{28}$$

is also called the characteristic equation.

The Auto-Regressive (AR) Process

The auto-covariance function of an AR(m) process satisfies the linear difference equations given by

$$\gamma(k) + \sum_{j=1}^{m} a_j \gamma(k-j) = 0, \quad k > 0, \tag{29}$$

with initial conditions given by

$$\gamma(0) + \sum_{j=1}^{m} a_j \gamma(j) = \sigma_\varepsilon^2, \tag{30}$$

where we remember the symmetry of auto-covariance functions $\gamma(k) = \gamma(-k)$. The spectrum of the process is given by

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{1}{\left|1 + \sum_{k=1}^{m} a_k e^{-ik\omega}\right|^2} \tag{31}$$

The ARMA Process

The ARMA(m,n) process is defined according to:

The process $\{y_t\}$ given by

$$y_t + \sum_{k=1}^{m} a_k y_{t-k} = \varepsilon_t + \sum_{k=1}^{n} c_k \varepsilon_{t-k}, \tag{32}$$

where $\{\varepsilon_k\}$ represents a white-noise process, is called an auto-regressive, moving-average process of order (m,n).

The ARMA Process

The ARMA(m,n) process can defined using the compact notation

$$A(q)y_t = C(q)\varepsilon_t, \tag{33}$$

where the shift-polynomials $A(q)$ and $C(q)$ are defined according to

$$A(q) = 1 + \sum_{k=1}^{m} a_k q^{-k} \quad \text{and} \quad C(q) = 1 + \sum_{k=1}^{n} c_k q^{-k} \tag{34}$$

The corresponding transfer function is given as $\dfrac{C(q)}{A(q)}$.

Advanced External Model structures

The ARMAX structure:

$$A(q^{-1})y_t = B(q^{-1})u_t + C(q^{-1})e_t \tag{35}$$

The Box-Jenkins Structure:

$$y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t \tag{36}$$

The L-Structure:

$$A(q^{-1})y_t = \frac{B(q^{-1})}{F(q^{-1})}u_t + \frac{C(q^{-1})}{D(q^{-1})}e_t \tag{37}$$

Covariance functions - reminder

Auto covariance function:

$$r_x(s,t) = cov(x_s, x_t) = E(x_s x_t^T) - E(x_s)E(x_t^T) \tag{38}$$

cross covariance function:

$$r_{xy}(s,t) = cov(x_s, y_t) = E(x_s y_t^T) - E(x_s)E(y_t^T) \tag{39}$$

Rules and Notation:

$$r_x(k) = r_x(t+k,t) \quad r_{xy}(k) = r_{xy}(t+k,t) \tag{40}$$

$$r_x(k) = r_x^T(-k) \quad r_{xy}(k) = r_{yx}^T(-k) \tag{41}$$

$$z_t = x_t + y_t: \qquad r_z(k) = r_x(k) + r_y(k) + r_{xy}(k) + r_{xy}^T(-k) \tag{42}$$

$$r_{zx}(k) = r_x(k) + r_{xy}^T(-k) \tag{43}$$

$$z_t = Ax_t: \qquad r_z(k) = Ar_x(k)A^T \quad r_{zx}(k) = Ar_x(k) \tag{44}$$

Variance and Spectral properties

Let us now continue with the ARMA model:

$$A(q^{-1})y_t = C(q^{-1})e_t \Leftrightarrow y_t = \sum_{i=0}^{\infty} h_i q^{-i} e_t, \quad e_t \sim N(0, \sigma_e^2) \tag{45}$$

The cross covariance of the ARMA model is given by:

$$A(q^{-1})r_{ye}(k) = C(q^{-1})\delta_k \sigma_e^2, \quad \delta_k = \begin{cases} 1 & k = 0 \\ 0 & else \end{cases} \tag{46}$$

$$r_{ye}(k) = h_k \sigma_e^2 \tag{47}$$

The auto covariance, can be obtained from the Yule-Walker equation:

$$A(q^{-1})r_y(k) = C(q^{-1})r_{ey}(k) \tag{48}$$

$$r_y(k) = \sigma_e^2 h_k \star h_{-k} \tag{49}$$

Spectrum and Spectral density

While covariance is a time-domain characteristic, a corresponding characteristic in the frequency-domain is the spectrum

$$\Psi_x(z) = \mathcal{Z}_b\{r_x(k)\} = \sum_{k=-\infty}^{\infty} r_x(k) z^{-k} \tag{50}$$

$$\Psi_{xy}(z) = \mathcal{Z}_b\{r_{xy}(k)\} = \sum_{k=-\infty}^{\infty} r_{xy}(k) z^{-k} \tag{51}$$

A subset of the spectrum is the spectral density $z = e^{jw}$ of the unit circle:

$$\phi_x(\omega) = \Psi_x(e^{j\omega}) = \mathcal{F}(r_x(k)), \quad \omega \in [-\pi, \pi] \tag{52}$$

$$r_x(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_x(\omega) e^{j\omega k} d\omega \tag{53}$$

$\mathcal{F}$ and $\mathcal{Z}_b$ indicating Fourier transform and bilateral Z-transform, respectively.

Spectrum and Spectral density

Considering an ARMA model with the transfer function:

$$H(z) = \frac{b_0 + b_1 z^{-1} + \ldots + b_{n_b} z^{-n_b}}{1 + a_1 z^{-1} + \ldots + a_{n_a} z^{-n_a}} \tag{54}$$

Then its spectrum is given by:

$$\Psi(z) = H(z)H(z^{-1}) = \frac{\bar{b}_0 + \sum_{i=1}^{n_b} \bar{b}_i (z^i + z^{-i})}{\bar{a}_0 + \sum_{i=1}^{n_a} \bar{a}_i (z^i + z^{-i})} \tag{55}$$

$$\bar{a}_i = \sum_{j=i}^{n_a} a_j a_{j-i}, \quad \bar{b}_i = \sum_{j=i}^{n_b} b_j b_{j-i} \tag{56}$$

The spectrum density is then given by

$$\phi(w) = \Psi(e^{jw}) = \frac{\bar{b}_0 + \sum_{i=1}^{n_b} 2\bar{b}_i \cos(iw)}{\bar{a}_0 + \sum_{i=1}^{n_a} 2\bar{a}_i \cos(iw)} \tag{57}$$

Spectrum and Spectral density

Continuing with the ARMA model:

$$y_t = H(z)e_t \tag{58}$$

Then its spectrums is given by:

$$\Psi_y(z) = H(z)H(z^{-1})\sigma_e^2 \qquad \Psi_{ye}(z) = H(z)\sigma_e^2 \tag{59}$$

With the spectrum density given by

$$\phi_y(z) = H(e^{jw})H(e^{-jw})\sigma_e^2 \qquad \phi_{ye}(z) = H(e^{-jw})\sigma_e^2 \tag{60}$$

Spectral Factorization

Let us consider the opposite situation, Spectral factorization.
Consider a stationary process given by its spectral density $\phi(\omega) \geq 0$ and rational, then there exists a $H(z)$ with only zeroes/poles inside the stability area, such that:

$$\phi(\omega) = H(e^{j\omega})H(e^{-j\omega})\sigma^2 \tag{61}$$

The representation theorem: Given a weak stationary stochastic process with rational spectral density $\phi(\omega) \geq 0$, this can be represented by:

$$y_t = H(q)e_t, \quad e_t \text{ (white)} \tag{62}$$

where $H(q)$ and its inverse is asymptotically stable, and the spectral density of y_t is $\phi(\omega)$

Spectral Factorization

If we have a polynomial $\Psi(z)$ with $\Psi(e^{-jw}) \geq 0 \in \mathbb{R}$:

$$\Psi(z) = r_n z^{-n} + r_{n-1} z^{-(n-1)} + \ldots + r_{n-1} z^{n-1} + r_n z^n \tag{63}$$

then there exist a polynomial $P(z)$, such that:

$$\Psi(z) = P(z^{-1})P(z) \tag{64}$$

$$P(z^{-1}) = p_0 + p_1 z^{-1} + \ldots + p_n z^{-n} \tag{65}$$

with all zeros lying within the unit circle.
The spectrum of $H(z)$ can be considered a ratio of spectra:

$$\Psi_H(z) = H(z)H(z^{-1}) = \frac{C(z)}{A(z)} \frac{C(z^{-1})}{A(z^{-1})} = \frac{C(z)C(z^{-1})}{A(z)A(z^{-1})} = \frac{\Psi_C(z)}{\Psi_A(z)} \tag{66}$$

Spectral Factorization

Using a correction polynomial, $X(z)$, we can compute the factorized polynomial in iterative approach:

1. $P_i(z^{-1})X_i(z) + P_i(z)X_i(z^{-1}) = 2\Psi(z)$
2. $P_{i+1}(z^{-1}) = \frac{1}{2}(P_i(z^{-1}) + X_i(z^{-1}))$

with each correction been computed from:

$$\begin{bmatrix} p_n & 0 & \dots & 0 \\ p_{n-1} & p_n & \dots & 0 \\ \vdots & \vdots & & \vdots \\ p_0 & p_1 & \dots & p_n \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 0 & \dots & 0 & p_0 \\ 0 & \dots & p_0 & p_1 \\ \vdots & & \dots & \dots \\ p_0 & \dots & p_{n-1} & p_n \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} = 2 \begin{bmatrix} r_n \\ r_{n-1} \\ \vdots \\ r_0 \end{bmatrix} \tag{67}$$

Stochastic systems on external form

Consider the system

$$y_t = H_u(q)u_t + H_d(q)v_t, \quad v_t \sim N(0, \sigma^2) \text{ (white)} \tag{68}$$

The stochastic description then becomes

$$E\{y_t\} = m_t = H_u(q)u_t \tag{69}$$

$$A(q^{-1})r_{yv}(k) = C(q^{-1})r_v(k) \tag{70}$$

$$A(q^{-1})r_y(k) = C(q^{-1})r_{vy}(k) \tag{71}$$

$$r_{vy}(k) = r_{yv}^T(-k) \tag{72}$$

In the case of non-white disturbance, a rational assumption, makes the substitution of $v_t = H_n e_t$ sufficient. where H_n and its inverse are asymp. stable.

Stochastic systems on external form

Consider the asymp. stable system

$$y_t = H_u(q)u_t + H_d(q)v_t, \quad v_t \sim F(\mu_v, \sigma_v^2) \tag{73}$$

If v_t is weakly stationary process, then y_t is also a weakly stationary process given by

$$E\{y_t\} = \mu_{y,t} = H_u(1)u_0 + H_d(1)\mu_v \tag{74}$$

$$A(q^{-1})r_{yv}(k) = C(q^{-1})r_v(k) \tag{75}$$

$$A(q^{-1})r_y(k) = C(q^{-1})r_{vy}(k) \tag{76}$$

$$r_{vy}(k) = r_{yv}^T(-k) \tag{77}$$

If v_t is Gaussian, y_t is strongly stationary process.
In the frequency-domain we have:

$$\Psi_y(z) = H_d(z)\Psi_v(z)H_d^T(z^{-1}) \tag{78}$$

$$\Psi_{yv}(z) = H_d(z)\Psi_v(z) \tag{79}$$

System gains

Consider the system on both internal and external form:

$$x_{t+1} = Ax_t + Be_t \tag{80}$$

$$y_t = Cx_t + De_t = (C(qI - A)^{-1}B + D)e_t = H(q)e_t \tag{81}$$

Any system then have a DC-Gain:

$$K_{dc} = \frac{y_\infty}{e_\infty} = H(1) = C(I - A)^{-1}B + D \tag{82}$$

Similarly for a stochastic process, systems also have an AC-Gain or Variance-Gain:

$$K_{ac} = \frac{\sigma_y^2}{\sigma_e^2} \tag{83}$$

if $e_i \sim N(0, \sigma_e^2)$. The relations can also be expressed as:

$$P_x = AP_xA^T + B\sigma_e^2B^T \qquad \sigma_y^2 = \int_{-\pi}^{\pi} H(e^{jw})H(e^{-jw})dw\,\sigma_e^2 \tag{84}$$

$$\sigma_y^2 = CP_xC^T + D\sigma_e^2D^T \tag{85}$$

System gains - Variance

If we consider the external description of the variance

$$\sigma_y^2 = \int_{-\pi}^{\pi} H(e^{jw})H(e^{-jw})dw\,\sigma_e^2, \quad H(z) = \frac{B(z)}{A(z)} \tag{86}$$

The variance of a nth order system can be computed by

$$\sigma_y^2 = \frac{1}{a_0}\sum_{i=0}^{n} b_i^i \beta_i \tag{87}$$

where the parameters are given by

$$a_i^{k-1} = a_i^k - \alpha_k a_{k-i}^k, \qquad \alpha_k = a_k^k / a_0^k, \qquad a_i^n = a_i, \tag{88}$$

$$b_i^{k-1} = b_i^k - \beta_k b_{k-i}^k, \qquad \beta_k = b_k^k / a_0^k, \qquad b_i^n = b_i \tag{89}$$

System gains - Variance

If we consider the system

$$y_t = \frac{1 + 0.25q^{-1}}{1 + 0.5q^{-1}} e_t, \quad e_t \sim N(0, 1), \tag{90}$$

then we have $n = 1$:

$$\sigma_y^2 = \frac{1}{a_0}(b_0^0 \beta_0 + b_1 \beta_1) \tag{91}$$

where the parameters are given by

$$\alpha_1 = a_1/a_0 = 0.5, \qquad \beta_1 = b_1/a_0 = 0.25 \tag{92}$$

$$a_0^0 = a_0 - \alpha_1 a_1 = 0.75, \qquad b_0^0 = b_0 - \beta_1 b_1 = 0.875 \tag{93}$$

$$\beta_0 = b_0^0/a_0^0 = 1.1667 \tag{94}$$

with the variance being:

$$\sigma_y^2 = 1.0833 \tag{95}$$

System Gains - multiple inputs and multiple outputs

Let us consider the MIMO system:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} H_{11}(w) & H_{12}(w) \\ H_{21}(w) & H_{22}(w) \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \tag{96}$$

The DC-gains are then given by

$$K_{dc} = \begin{bmatrix} H_{11}(1) & H_{12}(1) \\ H_{21}(1) & H_{22}(1) \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_\infty \begin{bmatrix} e_1 & e_2 \end{bmatrix}_\infty \left(\begin{bmatrix} e_1 & e_2 \end{bmatrix}_\infty \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}_\infty \right)^{-1} \tag{97}$$

System Gains - multiple inputs and multiple outputs

Let us consider the MIMO system:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} H_{11}(w) & H_{12}(w) \\ H_{21}(w) & H_{22}(w) \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \tag{98}$$

For the variance or AC-Gain, we will consider independent noises e_i. Under this assumption, the AC-gain is given from:

$$P_y = \sum_{i=1}^{2} K_{ac,i} \sigma_{in,i}^2 \tag{99}$$

Where $K_{ac,i}$ is defined as:

$$K_{ac,i} = \begin{bmatrix} \int_{-\pi}^{\pi} H_{1i}(e^{jw}) H_{1i}(e^{-jw}) dw & \int_{-\pi}^{\pi} H_{1i}(e^{jw}) H_{2i}(e^{-jw}) dw \\ \int_{-\pi}^{\pi} H_{2i}(e^{jw}) H_{1i}(e^{-jw}) dw & \int_{-\pi}^{\pi} H_{2i}(e^{jw}) H_{2i}(e^{-jw}) dw \end{bmatrix} \tag{100}$$

$$P_{x,i} = A_i P_{x,i} A_i^T + B_i \sigma_{e_i}^2 B_i^T \tag{101}$$

$$P_y^2 = \sum_{i=1}^{2} C_i P_{x,i} C_i^T + D_i \sigma_{e_i}^2 D_i^T \tag{102}$$

Questions

Questions?

Examples

Today's Matlab example topics:

- Spectrum/Spectral density: back and forth
- Spectral factorization
- Addition of Spectra
- Plotting Spectra
- Matlab functions