Stochastic Adaptive Control (02421)

Lecture 3

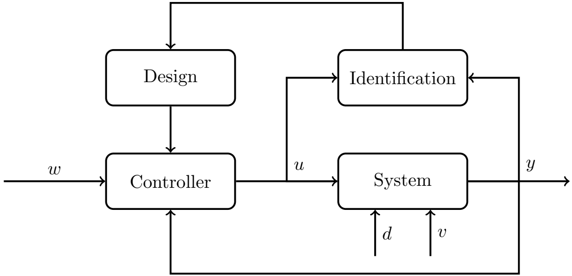
Tobias K. S. Ritschel

Assistant Professor

Section for Dynamical Systems, DTU Compute

Lecture Plan

1. ARX models
2. ARX estimation
3. ARMAX estimation
4. ARMAX identification
5. ARMAX control
6. ARMAX adaptive control
7. Systems and control theory
8. Stochastic systems + SS estimation
9. SS estimation + control
10. SS control
11. To be decided . . .
12. SS nonlinear estimation
13. SS nonlinear control



Today's Agenda

- ARMAX estimation
- Recursive estimation
- Time-varying parameters

ARMAX estimation

ARMAX least-squares estimation (extended least-squares)

ARMAX model

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d, \qquad e_t \sim N(0,\sigma^2) \tag{1}$$

Recall the normalization: $a_0 = c_0 = 1$

Formulate in terms of regressors

$$y_t = \phi_t^T \theta + e_t \tag{2}$$

Regressors and parameters

$$\phi_t^T = \left[-y_{t-1}, \ldots, -y_{t-n_a}, u_{t-k}, \ldots, u_{t-k-n_b}, e_{t-1}, \ldots, e_{t-n_c}, 1\right], \tag{3}$$

$$\theta = \left[a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b}, c_1, \ldots, c_{n_c}, d\right] \tag{4}$$

ARMAX least-squares estimation – Approximate the noise

Problem: $e_{t-1}, \ldots, e_{t-n_c}$ are not known

Solution: Approximate them by the residuals, $\epsilon_{t-1}, \ldots, \epsilon_{t-n_c}$

$$\epsilon_t = y_t - \phi_t^T \theta, \qquad t = 1, \ldots, N \tag{5}$$

Initialize with zeros

$$\epsilon_t = 0, \qquad t = 0, -1, -2, \ldots \tag{6}$$

ARMAX least-squares estimation – Iterative solution

Iterative solution

$$\hat{\theta}_{n+1} = (\Phi_{t,n}^T \Phi_{t,n})^{-1} \Phi_{t,n}^T Y_t, \qquad n = 0, 1, \ldots \tag{7}$$

Simple stopping criteria (e.g., $\rho = 10^{-6}$)

$$\max \left\{ \frac{|\hat{\theta}_{i,n+1} - \hat{\theta}_{i,n}|}{|\hat{\theta}_{i,n}|} \right\}_{i=1}^{n_\theta} \leq \rho, \qquad \max \left\{ \frac{|\hat{\theta}_{i,n+1} - \hat{\theta}_{i,n}|}{1 + |\hat{\theta}_{i,n}|} \right\}_{i=1}^{n_\theta} \leq \rho \tag{8}$$

ARMAX least-squares estimation – Algorithm

Step 1. Select an initial guess, $\hat{\theta}_0$, of the parameters. Consider setting many of the parameters to zero.

Step 2. Evaluate the residuals, ϵ_t, in (5) for $t = 1, \ldots, N$. Use, e.g., the initial condition (6).

Step 3. Update the parameter estimate, i.e., compute $\hat{\theta}_{n+1}$ from $\hat{\theta}_n$ using (7).

Step 4. Check whether the convergence criterion, e.g., one of those in (8), is satisfied. If so, the estimate is $\hat{\theta}_{n+1}$. If not, go to Step 2.

ARMAX maximum likelihood estimation

Maximum likelihood estimation problem (equivalent formulations)

$$\max_{\theta} p(Y_N|\theta), \qquad \min_{\theta} -\ln p(Y_N|\theta) \tag{9}$$

Write out the probability density in terms of conditional densities

$$p(Y_N|\theta) = p(y_N|Y_{N-1},\theta)p(Y_{N-1}|\theta) = \prod_{t=1}^{N} p(y_t|Y_{t-1},\theta) \tag{10}$$

ARMAX maximum likelihood estimation – conditional distribution

ARMAX model

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d, \qquad e_t \sim N(0, \sigma^2) \tag{11}$$

If $Y_{t-1} = \{\ldots, y_1, y_2, \ldots, y_{t-2}, y_{t-1}\}$ is known, y_t is normally distributed

$$p(y_t|Y_{t-1}, \theta) = \frac{1}{\sqrt{2\pi\sigma^2_{t|t-1}}} \exp\left(-\frac{1}{2}\left(\frac{y_t - \hat{y}_{t|t-1}}{\sigma_{t|t-1}}\right)^2\right) \tag{12}$$

Conditional mean: Take expectations on both sides

$$\mathbb{E}\left[A(q^{-1})y_t|Y_{t-1}, \theta\right] = \mathbb{E}\left[q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d|Y_{t-1}, \theta\right] \tag{13}$$

$$A(q^{-1})\mathbb{E}[y_t|Y_{t-1}, \theta] = q^{-k}B(q^{-1})u_t + d \tag{14}$$

$$A(q^{-1})\hat{y}_{t|t-1} = q^{-k}B(q^{-1})u_t + d \tag{15}$$

Note that $\hat{y}_{t-j|t-1} = y_{t-j}$ for $j \geq 1$

ARMAX maximum likelihood estimation – conditional distribution

Conditional variance

$$\text{Cov}\left[A(q^{-1})y_t, y_t|Y_{t-1},\theta\right] = \text{Cov}\left[q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d, y_t|Y_{t-1},\theta\right] \tag{16}$$

$$\begin{aligned}
\text{Cov}[y_t, y_t|Y_{t-1},\theta] &= \text{Cov}\left[C(q^{-1})e_t, y_t|\theta\right] && (17)\\
&= \text{Cov}\left[\sum_{j=0}^{n_c} c_j e_{t-j}, y_t|\theta\right] && (18)\\
&= \sum_{j=0}^{n_c} c_j \,\text{Cov}[e_{t-j}, y_t|\theta] && (19)\\
\sigma^2_{t|t-1} &= \sum_{j=0}^{n_c} c_j \sigma^2_{ye,t,t-j|t-1} && (20)
\end{aligned}$$

ARMAX maximum likelihood estimation – conditional distribution

Conditional cross-covariance between y_t and e_{t-i}

$$\mathrm{Cov}\left[A(q^{-1})y_t, e_{t-i}|Y_{t-1},\theta\right] = \mathrm{Cov}\left[q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d, e_{t-i}|Y_{t-1},\theta\right] \tag{21}$$

$$\mathrm{Cov}[y_t, e_{t-i}|Y_{t-1},\theta] = \mathrm{Cov}\left[C(q^{-1})e_t, e_{t-i}|\theta\right] \tag{22}$$

$$= \mathrm{Cov}\left[\sum_{j=0}^{n_c} c_j e_{t-j}, e_{t-i}|\theta\right] \tag{23}$$

$$= \sum_{j=0}^{n_c} c_j \,\mathrm{Cov}[e_{t-j}, e_{t-i}] \tag{24}$$

$$\sigma^2_{ye,t,t-i|t-1} = c_i\sigma^2 \tag{25}$$

Conditional covariance (from previous slide)

$$\sigma^2_{t|t-1} = \sum_{j=0}^{n_c} c_j \sigma^2_{ye,t,t-j|t-1} = \sigma^2 \sum_{j=0}^{n_c} c_j^2 = \bar{\sigma}^2 \tag{26}$$

ARMAX maximum likelihood estimation – conditional distribution

Summary of conditional distribution

$$p(y_t|Y_{t-1},\theta) = \frac{1}{\sqrt{2\pi\bar{\sigma}^2}} \exp\left(-\frac{1}{2}\left(\frac{y_t - \hat{y}_{t|t-1}}{\bar{\sigma}}\right)^2\right), \tag{27}$$

$$A(q^{-1})\hat{y}_{t|t-1} = q^{-k}B(q^{-1})u_t + d, \tag{28}$$

$$\bar{\sigma}^2 = \sigma^2 \sum_{j=0}^{n_c} c_j^2 \tag{29}$$

Negative logarithm of density

$$-\ln p(y_t|Y_{t-1},\theta) = \frac{1}{2}\ln 2\pi + \frac{1}{2}\ln \bar{\sigma}^2 + \frac{1}{2}\left(\frac{y_t - \hat{y}_{t|t-1}}{\bar{\sigma}}\right)^2 \tag{30}$$

ARMAX maximum likelihood estimation – Objective function

Maximum likelihood estimate

$$\min_{\theta} \phi = -\ln p(Y_N|\theta) = -\sum_{t=1}^{N} \ln p(y_t|Y_{t-1}, \theta) \tag{31}$$

Insert conditional density

$$\phi = \sum_{t=1}^{N} \left(\frac{1}{2} \ln 2\pi + \frac{1}{2} \ln \bar{\sigma}^2 + \frac{1}{2} \left(\frac{y_t - \hat{y}_{t|t-1}}{\bar{\sigma}} \right)^2 \right) \tag{32}$$

$$= \frac{N}{2} \ln 2\pi + \frac{N}{2} \ln \bar{\sigma}^2 + \frac{1}{2\bar{\sigma}^2} \sum_{t=1}^{N} (y_t - \hat{y}_{t|t-1})^2 \tag{33}$$

Conclusion: The same as least-squares estimation (for the coefficients of A and B and the disturbance, d)

ARMAX maximum likelihood estimation

Express conditional mean in terms of regressors

$$\hat{y}_{t|t-1} = \phi_t^T \theta, \tag{34}$$

$$\phi_t^T = \left[-y_{t-1}, \dots, -y_{t-n_a}, u_{t-k}, \dots, u_{t-k-n_b}, 1\right], \tag{35}$$

$$\theta^T = \left[a_1, \dots, a_{n_a}, b_0, \dots, b_{n_b}, d\right] \tag{36}$$

Maximum likelihood problem

$$\min_\theta \frac{1}{2}(Y - \hat{Y})^T (Y - \hat{Y}), \qquad \hat{Y} = \Phi\theta \tag{37}$$

Maximum likelihood estimator

$$\hat{\theta} = (\Phi^T \Phi)^{-1} \Phi^T Y \tag{38}$$

ARMAX maximum likelihood estimation – Normalized variance

Objective function

$$\phi = \frac{N}{2}\ln 2\pi + \frac{N}{2}\ln\bar{\sigma}^2 + \frac{1}{2\bar{\sigma}^2}\sum_{t=1}^{N}(y_t - \hat{y}_{t|t-1})^2 \tag{39}$$

Differentiate objective function

$$\frac{\partial\phi}{\partial\bar{\sigma}^2} = \frac{N}{2\bar{\sigma}^2} - \frac{1}{2\bar{\sigma}^4}\sum_{t=1}^{N}(y_t - \hat{y}_{t|t-1})^2 = 0 \tag{40}$$

$$N\bar{\sigma}^2 = \sum_{t=1}^{N}(y_t - \hat{y}_{t|t-1})^2 \tag{41}$$

Maximum likelihood estimate of normalized variance

$$\hat{\bar{\sigma}}^2 = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_{t|t-1})^2 \tag{42}$$

Recursive estimation

RELS - Recursive Extended Least Squares

ARMAX model

$$A(q^{-1})y_t = q^{-k}B(q^{-1})u_t + C(q^{-1})e_t + d, \tag{43}$$

$$y_t = \phi_t^T\theta + e_t, \tag{44}$$

$$\phi_t^T = \left[-y_{t-1}, \dots, -y_{t-n_a}, u_{t-k}, \dots, u_{t-k-n_b}, e_{t-1}, \dots, e_{t-n_c}, 1\right] \tag{45}$$

$$\theta^T = \left[a_1, \dots, a_{n_a}, b_0, \dots, b_{n_b}, c_1, \dots, c_{n_c}, d\right] \tag{46}$$

$$\tag{47}$$

Least squares estimator (approximate e_i by ϵ_i in ϕ)

$$\epsilon_t = y_t - \phi_t^T\hat{\theta}_{t-1}, \tag{48}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_t\phi_t\epsilon_t, \tag{49}$$

$$P_t^{-1} = P_{t-1}^{-1} + \phi_t\phi_t^T \tag{50}$$

Time-varying parameters

Time-varying estimation - first example

ARX model

$$A(q^{-1})y_t = B_t(q^{-1})u_t + e_t, \tag{51}$$

$$b_{1,t} = b_{1,0} + b_{1,1}t \tag{52}$$

Treat time-varying coefficient as two coefficients with their own inputs

$$y_t = \phi^T\theta + e_t \tag{53}$$

$$\theta^T = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n_a} & b_{1,0} & b_{1,1} & b_2 & \cdots & b_{n_b} \end{bmatrix} \tag{54}$$

$$\phi^T = \begin{bmatrix} -y_{t-1} & -y_{t-2} & \cdots & -y_{t-n_a} & u_{t-1} & tu_{t-1} & u_{t-2} & \cdots & u_{t-n_b} \end{bmatrix} \tag{55}$$

Time-varying estimation

For deterministic time-varying systems, rearrange the parameters

$$y_t = \phi_t^T \theta_t + e_t \tag{56}$$

$$\theta_t = \alpha + f(t)\beta = \begin{bmatrix} I & f(t) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{57}$$

$$y_t = \begin{bmatrix} \phi_t^T & \phi_t^T f(t) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + e_t \tag{58}$$

For piece-wise linear parameters, rearrange the parameters

$$y_t = \phi_t^T \theta_t + e_t, \tag{59}$$

$$\theta_t = \alpha_i + (t - T_i)\beta_i, \quad T_i \le t \le T_{i+1}, \tag{60}$$

$$y_t = \begin{bmatrix} \phi_t^T & \phi_t^T (t - T_i) \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} + e_t \tag{61}$$

Time-varying systems

System with general time-varying parameters

$$\theta_{t+1} = f(t, \theta_t, v_t) \tag{62}$$

- The methods discussed so far cannot estimate the time-varying dynamics and were not designed to do it
- In practice, the problem is that the correction factor diminishes over time

$$P_t \to 0 \tag{63}$$

Time-varying systems - Forgetting methods: Resetting

Reset the covariance after some time, t_i

$$P_{t_i} = P_i > P_{t_i-1}, \quad \hat{\theta}_{t_i} = \hat{\theta}_{t_i-1} \tag{64}$$

The appropriate restarting time depends on the application

For instance, restart at fixed intervals

$$t_i = Ni \tag{65}$$

This can be useful for periodic systems

Time-varying systems - Forgetting methods: Constant Gain

Another method: Keep the correction term large

For instance, keep the correction term κ constant

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \kappa\epsilon_t \tag{66}$$

Alternatively, keep the variance constant

$$P_t = P \tag{67}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \kappa\epsilon_t \tag{68}$$

$$\kappa_t = \frac{P\phi_t}{1 + \phi_t^T P \phi_t} \tag{69}$$

Time-varying systems - Forgetting methods: Exponential Forgetfulness

Another approach: Forget a little bit all the time (exponential forgetfulness)

$$J_t = \frac{1}{2}\sum_{i=1}^{t} \lambda^{t-i}\epsilon_i^2 = \lambda J_{t-1} + \frac{1}{2}\epsilon_t^2 \tag{70}$$

The recursion is similar to the previous methods

$$\epsilon_t = y_t - \phi_t^T \hat{\theta}_{t-1}, \tag{71}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + P_t \phi_t \epsilon_t, \tag{72}$$

$$P_t^{-1} = \lambda P_{t-1}^{-1} + \phi_t \phi_t^T \tag{73}$$

The forgetting factor λ can be expressed in terms of a horizon, N_∞

$$\lambda = 1 - \frac{1}{N_\infty} \tag{74}$$

Time-varying systems - Fortescue's Method

Improve with a time-varying forgetting factor depending on the prediction error, ϵ_t

$$\lambda_t = 1 - \frac{1}{N_0}\frac{\epsilon_t^2}{\sigma^2 s_t} \tag{75}$$

N_0 is the approx. horizon over which the parameter is roughly constant

Recursion

$$\epsilon_t = y_t - \phi_t^T \hat{\theta}_{t-1} \tag{76}$$

$$s_t = 1 + \phi_t^T P_{t-1} \phi_t \tag{77}$$

$$K_t = \frac{P_{t-1}\phi_t}{\lambda_t + s_t} \tag{78}$$

$$\hat{\theta}_t = \hat{\theta}_{t-1} + K_t \epsilon_t \tag{79}$$

$$P_t = (I - K_t \phi_t^T) P_{t-1} \frac{1}{\lambda_t} \tag{80}$$

Time-varying systems - Fortescue's Method

If the variance is unknown, we can introduce an estimate

$$\lambda_t = 1 - \frac{1}{N_0}\frac{\epsilon_t^2}{r_t s_t} \tag{81}$$

$$r_t = r_{t-1} + \frac{1}{t}\left(\frac{\epsilon_t^2}{s_t} - r_{t-1}\right), \qquad r_0 = \epsilon_0^2 \tag{82}$$

Time-varying systems - Model Estimators

Introduce model of parameters

$$\theta_{t+1} = \theta_t + v_t, \qquad v_t \sim N(0, R_1\sigma^2) \tag{83}$$

$$y_t = \phi_t^T \theta_t + e_t, \qquad e_t \sim N(0, \sigma^2) \tag{84}$$

Estimate parameters using the Kalman filter

Data update

$$\hat{\theta}_{t|t} = \hat{\theta}_{t|t-1} + P_{t|t-1}\phi_t(y_t - \phi_t^T \hat{\theta}_{t|t-1}) \tag{85}$$

$$P_{t|t}^{-1} = P_{t|t-1}^{-1} + \phi_t\phi_t^T \tag{86}$$

Time update

$$\hat{\theta}_{t+1|t} = \hat{\theta}_{t|t} \tag{87}$$

$$P_{t+1|t} = P_{t|t} + R_1 \tag{88}$$

Questions?