Central extensions in closed-loop optimal experiment design

Roland Hildebrand[1] Michel Gevers[2] Gabriel Solari[3]

[1]WIAS Berlin

[2]Université Catholique de Louvain

[3]Tenaris Dalmine SpA

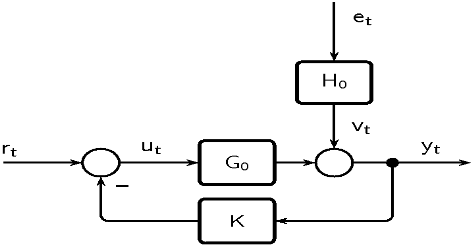
52nd CDC, Florence, December 12, 2013

Outline

Outline

Outline

System setup



identify a MIMO LTI system with a PE method in closed loop

$$y = G_0(q)u + H_0(q)e, \quad u = -K(q)y + r$$

r external input signal, K controller, u input, y output

both external input r and controller K are design variables

Assumptions

u, r are of dimension l_1, and e, y are of dimension l_2

- G_0, H_0 stable and H_0 inversely stable, e with power spectrum $\lambda_0 I$
- r quasi-stationary with power spectrum Φ_r
- $G_0(z), H_0(z)$ embedded in a model structure $G(z;\theta), H(z;\theta)$ with true parameter value θ_0, $G_0(z) = G(z;\theta_0)$, $H_0(z) = H(z;\theta_0)$
- asymptotic in the number of data parameter covariance formulas are assumed
- constraints and cost function depend on frequency weighted input and/or output with real-rational weightings

information matrix $\overline{M}$ is of this form

Model uncertainty

- parametric PE identification provides a parameter estimate $\hat{\theta}$ together with an ellipsoidal uncertainty region

$$E = \{\theta \,|\, (\theta - \hat{\theta})^T P^{-1} (\theta - \hat{\theta}) \leq \gamma\},$$

  P covariance matrix
- estimate $\hat{\theta}$ is applied as if it were the true parameter value θ_0
- $\hat{\theta}$ is distributed around θ_0, but covariance depends on experimental conditions
- distribution of the performance of the intended application depends on P and hence on r, K

Problem

How to optimally choose the design variables r, K in order to minimize a given criterion measuring the (expected) performance of the model in the application, while satisfying given constraints on the input and output?

Change of design variables

replace design variables r, K by the equivalent joint signal spectrum

$$\Phi_{\chi_0} = \begin{pmatrix} \Phi_u & \Phi_{ue} \\ \Phi_{ue}^* & \lambda_0 I \end{pmatrix}$$

Φ_u, Φ_{ue} are related to the design variables r, K by

$$\begin{aligned} \Phi_u(\omega) &= \lambda_0(I + KG_0)^{-1}KH_0H_0^*K^*(I + KG_0)^{-*} \\ &\quad +(I + KG_0)^{-1}\Phi_r(\omega)(I + KG_0)^{-*}, \\ \Phi_{ue}(\omega) &= -\lambda_0(I + KG_0)^{-1}KH_0, \end{aligned}$$

advantage: constraints and cost function usually become convex or even linear in Φ_{χ_0}

Moments

partial correlation approach:
replace infinite-dimensional design variables Φ_u, Φ_{ue} by finite-dimensional projection to the generalized (matrix-valued) moments

$$m_k = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{|d(e^{j\omega})|^2} \Phi_{\chi_0}(\omega) e^{jk\omega} \, d\omega = m_{-k}^T, \; k = 0, \dots, n$$

n and d chosen such that

- both cost function and constraints can be written as convex functions in the finite number of moments $m_0, \dots, m_n$
- the polynomial $d(z) = \sum_{l=0}^{m} d_l z^l$ has all roots outside of the closed unit disk
- d_l real and $d_0 \neq 0$, $d_m \neq 0$, $n \geq m$

Solution strategy

1. solve the optimization problem on the moments $m_0, \dots, m_n$
2. recover power spectrum Φ_{χ_0} producing these moments
3. construct external input r and the controller K from Φ_u, Φ_{ue}

set of moments which can be produced by a valid power spectrum Φ_{χ_0} is semi-definite representable [Hildebrand, Gevers, Solari 2010]
Carathéodory theorem not applicable because Φ_{χ_0} is structured (SE corner of Φ_{χ_0} is $\lambda_0 I$; NE corner is stable)

$$\Phi_r = (I + KG_0)(\Phi_u - \lambda_0^{-1}\Phi_{ue}\Phi_{ue}^*)(I + KG_0)^*,$$
$$K = -\Phi_{ue}(\lambda_0 H_0 + G_0\Phi_{ue})^{-1}.$$

Solution strategy

1. solve the optimization problem on the moments $m_0, \dots, m_n$
2. recover power spectrum Φ_{χ_0} producing these moments
3. construct external input r and the controller K from Φ_u, Φ_{ue}

set of moments which can be produced by a valid power spectrum Φ_{χ_0} is semi-definite representable [Hildebrand, Gevers, Solari 2010]
Carathéodory theorem not applicable because Φ_{χ_0} is structured (SE corner of Φ_{χ_0} is $\lambda_0 I$; NE corner is stable)

$$\Phi_r = (I + KG_0)(\Phi_u - \lambda_0^{-1}\Phi_{ue}\Phi_{ue}^*)(I + KG_0)^*,$$
$$K = -\Phi_{ue}(\lambda_0 H_0 + G_0\Phi_{ue})^{-1}.$$

Solution strategy

1. solve the optimization problem on the moments $m_0, \dots, m_n$
2. recover power spectrum Φ_{χ_0} producing these moments
3. construct external input r and the controller K from Φ_u, Φ_{ue}

set of moments which can be produced by a valid power spectrum Φ_{χ_0} is semi-definite representable [Hildebrand, Gevers, Solari 2010]
Carathéodory theorem not applicable because Φ_{χ_0} is structured (SE corner of Φ_{χ_0} is $\lambda_0 I$; NE corner is stable)

$$
\begin{aligned}
\Phi_r &= (I + KG_0)(\Phi_u - \lambda_0^{-1}\Phi_{ue}\Phi_{ue}^*)(I + KG_0)^*, \\
K &= -\Phi_{ue}(\lambda_0 H_0 + G_0\Phi_{ue})^{-1}.
\end{aligned}
$$

Solution strategy

1. solve the optimization problem on the moments $m_0, \dots, m_n$
2. recover power spectrum Φ_{χ_0} producing these moments
3. construct external input r and the controller K from Φ_u, Φ_{ue}

set of moments which can be produced by a valid power spectrum Φ_{χ_0} is semi-definite representable [Hildebrand, Gevers, Solari 2010]
Carathéodory theorem not applicable because Φ_{χ_0} is structured (SE corner of Φ_{χ_0} is $\lambda_0 I$; NE corner is stable)

$$\begin{aligned} \Phi_r &= (I + KG_0)(\Phi_u - \lambda_0^{-1}\Phi_{ue}\Phi_{ue}^*)(I + KG_0)^*, \\ K &= -\Phi_{ue}(\lambda_0 H_0 + G_0\Phi_{ue})^{-1}. \end{aligned}$$

Block-Toeplitz moment matrix

by the Carathéodory theorem, the block-Toeplitz matrix

$$T_n = \begin{pmatrix} m_0 & m_1^T & \ddots & m_{n-1}^T & m_n^T \\ m_1 & m_0 & \ddots & m_{n-2}^T & m_{n-1}^T \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ m_n & m_{n-1} & \ddots & m_1 & m_0 \end{pmatrix}$$

is positive semi-definite

Assumption: T_n is positive definite

Central extension

set [Delsarte, Genin, Kamp 1978]

$$\begin{aligned} U(z) &= \begin{pmatrix} z^n I & z^{n-1} I & \cdots & I \end{pmatrix}, \\ A(z) &= U_n(z) T_n^{-1} U_n^T(0), \\ \Phi(\omega) &= A(e^{j\omega})^{-*} A(0) A(e^{j\omega})^{-1} \end{aligned}$$

Φ is a rational matrix-valued function of order n

we have

$$m_k = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \Phi(\omega) e^{jk\omega} d\omega$$

for every $k = 0, \dots, n$

the moment sequence produced by Φ is called central extension of the finite sequence $m_0, \dots, m_n$

Main result

Theorem

Let $(m_0, \ldots, m_n)$ be a feasible finite moment sequence, and Φ be the spectrum generating the central extension of $(m_0, \ldots, m_n)$. Then the spectrum $\Phi_{\chi_0}(\omega) = \Phi(\omega)|d(e^{j\omega})|^2$ satisfies

- Φ_{χ_0} *rational of order n*
- $\Phi_{\chi_0}(-\omega) = \Phi_{\chi_0}(\omega)^T$
- Φ_{χ_0} *reproduces the moments $m_0, \ldots, m_n$*
- $\Phi_{\chi_0} = \begin{pmatrix} \Phi_u & \Phi_{ue} \\ \Phi_{ue}^* & \lambda_0 I \end{pmatrix}$ *with Φ_{ue} stable*

Φ_{χ_0} is explicitly given by the moments $m_0, \ldots, m_n$

Problem setup

consider an ARX model structure

$$G = \frac{\theta_1 z^{-1}}{1 + \theta_2 z^{-1}}, \qquad H = \frac{1}{1 + \theta_2 z^{-1}}$$

with true parameters θ_{10}, θ_{20}, $|\theta_{20}| < 1$

- output power constraint $\overline{E}y^2 \leq c$, $c > \lambda_0$
- maximize determinant of the information matrix $\overline{M}$ (D-optimality)

Cost and constraints as function of moments

set $n = m = 1$, $d(z) = 1 + \theta_{20} z$, then

$$\begin{aligned}
\overline{M}_{11} &= \lambda_0^{-1}((1+\theta_{20}^2)m_{0,11} + 2\theta_{20} m_{1,11}) \\
\overline{M}_{12} &= \lambda_0^{-1}(-\theta_{10} m_{1,11} - (1-\theta_{20}^2) m_{0,12} - \theta_{10}\theta_{20} m_{0,11}) \\
\overline{M}_{22} &= \lambda_0^{-1}(-2\theta_{10}\theta_{20} m_{0,12} + \frac{\lambda_0}{1-\theta_{20}^2} + \theta_{10}^2 m_{0,11}) \\
\overline{E}y^2 &= -2\theta_{10}\theta_{20} m_{0,12} + \frac{\lambda_0}{1-\theta_{20}^2} + \theta_{10}^2 m_{0,11}
\end{aligned}$$

Optimal moments

maximize $\det \overline{M}$ subject to $\overline{E}y^2 \leq c$

$$
\begin{aligned}
m_{0,12} &= \frac{\lambda_0\theta_{20}(2c-\lambda_0)}{\theta_{10}(1-\theta_{20}^2)(c+(c-\lambda_0)\theta_{20}^2)} \\
m_{0,22} &= \frac{\lambda_0}{1-\theta_{20}^2}, \quad m_{1,22} = -\frac{\lambda_0\theta_{20}}{1-\theta_{20}^2}, \quad m_{1,21} = -\theta_{20}m_{0,12} \\
m_{0,11} &= \frac{(c(1-\theta_{20}^2)+\lambda_0\theta_{20}^2)(c\theta_{20}^2+c-\lambda_0)}{\theta_{10}^2(1-\theta_{20}^2)(c+(c-\lambda_0)\theta_{20}^2)} \\
m_{1,11} &= -\frac{\lambda_0\theta_{20}(c\theta_{20}^2+c-\lambda_0)}{\theta_{10}^2(1-\theta_{20}^2)(c+(c-\lambda_0)\theta_{20}^2)} \\
m_{1,12} &= -\Delta^{-1}m_{0,12}\theta_{20}(\Delta+(c-\lambda_0)\lambda_0(1-\theta_{20}^2)^2)
\end{aligned}
$$

with $\Delta = c^2(1+\theta_{20}^2)^2 - c\lambda_0(2\theta_{20}^4+\theta_{20}^2+1)+\lambda_0^2\theta_{20}^4$

Explicit solution

the central extension of (m_0, m_1) yields

$$K = -\frac{\theta_{20}(2c-\lambda_0)(c\theta_{20}^2+c-\lambda_0)(1+\theta_{20}z^{-1})}{\theta_{10}(\Delta+\theta_{20}(2c(c-\lambda_0)(1+\theta_{20}^2)+\lambda_0^2\theta_{20}^2)z^{-1})},$$

$$\Phi_r = \frac{(c-\lambda_0)(c\theta_{20}^2+c-\lambda_0)(c+(c-\lambda_0)\theta_{20}^2)\Delta|e^{j\omega}+\theta_{20}|^2}{\theta_{10}^2|\Delta e^{j\omega}+\theta_{20}(2c(c-\lambda_0)(1+\theta_{20}^2)+\lambda_0^2\theta_{20}^2)|^2}.$$

Comparison with optimal open-loop experiment

- for $c < \lambda_0$ no experiment feasible
- for $\lambda_0 \leq c < \frac{\lambda_0}{(1-\theta_{20}^2)}$ only closed-loop experiments feasible
- for $\frac{\lambda_0}{(1-\theta_{20}^2)} \leq c < \frac{\lambda_0}{1-|\theta_{20}|}$ the optimal closed-loop experiment beats the optimal open-loop experiment
- for $\frac{\lambda_0}{1-|\theta_{20}|} \leq c$ both give the same information matrix

Simulation

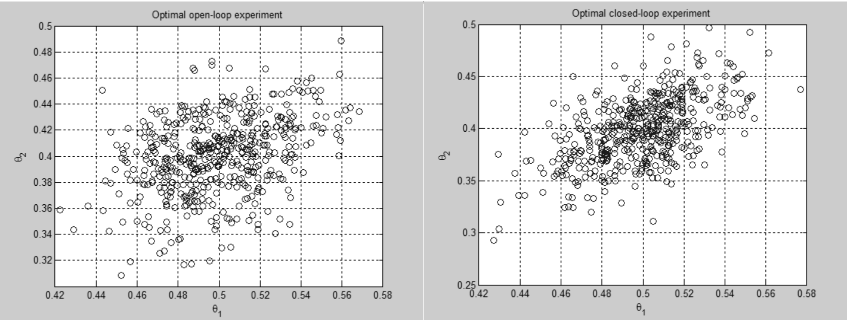
set $\lambda_0 = 1$, $c = 1.4$, $\theta_{10} = 0.5$, $\theta_{20} = 0.4$

- first identify in open-loop with white noise with variance $\sigma^2 = 1$
- from the identified parameters two experimental configurations are computed: the optimal open-loop input, and the optimal closed-loop input-controller pair
- an optimal open-loop and an optimal closed-loop experiment are performed and the parameter vector identified

500 runs, data length in each of the experiments is $N = 1000$

empirical covariance matrices have determinant:
$0.49736N^{-2}$ for open loop
$0.38796N^{-2}$ for closed-loop

Simulation cont'd



identified parameter vectors for optimal open-loop (left) and closed-loop (right) experiments

Thank you!