An Algebraic Method for System Reduction of Stationary Gaussian Systems

Jan H. van Schuppen (CWI)
with Dorina Jibetean (CWI)

26 January 2004, Rome
CC Meeting

Outline

- Problem formulation. Gaussian systems.
- System identification - System reduction
  - Infimization of divergence rate.
- Global optimization of rational functions.
- Procedure for infimization of divergence.
- Example
- Concluding remarks.

Problem formulation

Motivation System identification leads to infimization problem. In general, local minima occur because function is not convex!

Problem Determine global optimum for approximation problem of system identification.

Approach

- System reduction of Gaussian systems with divergence rate criterion.
- Apply algebraic methods for rational functions to determine the global minimum.

Preliminaries

Def. Gaussian system, Kalman canonical form,

$$\begin{aligned} x(t+1) &= Ax(t) + Bv(t), \quad x(t_0) = x_0, \\ y(t) &= Cx(t) + Dv(t), \quad v(t) \in G(0, V), \\ & m = p, \quad \text{rank}(D) = p, \quad V = I, \\ & \text{spec}(A), \text{spec}(A - BD^{-1}C) \subset \mathbb{C}^-, \\ & Q = AQA^T + BVB^T, \quad G = AQC^T + BVD^T, \\ & (A, B) \text{ reachable pair}, \ (A, C), \ (A, G) \text{ observable pairs}, \\ & SGSP_{min}(p, n, p) = \{(A, B, C, D) \text{ as above}\}, \\ & q \in QD \subset \mathbb{R}^{n_Q} \mapsto (A(q), B(q), C(q), D(q)) \in SGSP_{min}, \\ & \text{rational map. Canonical form: } p = 1 \text{ global}, \ p > 1 \text{ generic.} \end{aligned}$$

Def. Divergence or Kullback-Leibler pseudo-distance

$$
\begin{aligned}
D(P_1 \| P_2) &= E_Q[r_1 \ln(\frac{r_1}{r_2}) I_{(r_2>0)}] \\
&= \int_\Omega r_1(\omega) \ln(\frac{r_1(\omega)}{r_2(\omega)}) I_{(r_2(\omega)>0)} Q(d\omega), \\
& \quad P_1 \ll Q,\ r_1 = dP_1/dQ,\ \ P_2 \ll Q,\ r_2 = dP_2/dQ.
\end{aligned}
$$

Divergence rate for stochastic process $y : \Omega \times T \to \mathbb{R}^p$,

$$
D_r(P_1 \| P_2) = \lim_{t\to\infty} \frac{1}{2t+1} D(P_1|_{[-t,+t]} \| P_2|_{[-t,+t]}).
$$

Remark

Divergence equals expected value of natural logarithm of likelihood function.

System identification - System reduction - Infimization divergence rate

Procedure for system identification

1. From signals to system: Determine from a finite time series a high-order Gaussian system.
2. System reduction: Determine from a high-order Gaussian system a low-order Gaussian system.

Approximation criterion is divergence rate $D_r(P_1 \| P_2)$.

Problem Optimal system reduction

$$\inf_{q \in QD} D_r(P_1 \| P_2(q)), \quad q^* = \arg \min_{q \in QD} D_r(P_1 \| P_2(q)). \tag{1}$$

Determine global minimum q^*.

Procedure for computation of divergence rate

Notation

System 1 $\quad n_1 \in \mathbb{N}, (A_1, B_1, C_1, D_1) \in SGSP_{min}(p, n_1, p)$,

high-order system,

System 2 $\quad n_2 \in \mathbb{N}, (A_2, B_2, C_2, D_2) \in SGSP_{min}(p, n_2, p)$,

low-order approximant.

System 3 $\quad$ defined as inverse of System 2, $n_3 = n_2$,

$$\begin{aligned} & (A_3, B_3, C_3, D_3) \\ = \quad & (A_2 - B_2 D_2^{-1} C_2, B_2 D_2^{-1}, -D_2^{-1} C_2, D_2^{-1}) \\ & \in SGSP_{min}(p, n_3, p). \end{aligned}$$

$D_r(P_1 \| P_2)$ is computed using the series interconnection of System 3 and System 1.

Procedure for computation of divergence rate (continued)

1. Construct (A_4, B_4, C_4, D_4) according to the formulas $n_4 = n_1 + n_3$,

$$(A_4, B_4, C_4, D_4) = \left(\begin{pmatrix} A_1 & 0 \\ B_3C_1 & A_3 \end{pmatrix}, \begin{pmatrix} B_1 \\ B_3D_1 \end{pmatrix}, \begin{pmatrix} D_3C_1 & C_3 \end{pmatrix}, D_3D_1 \right)$$

2. Solve the discrete-time Lyapunov equation for the matrix $Q_4 \in \mathbb{R}^{n_4 \times n_4}$,

$$Q_4 = A_4Q_4A_4^T + B_4B_4^T; \quad Q_4 = \begin{pmatrix} Q_1 & Q_2 \\ Q_2^T & Q_3 \end{pmatrix}, \tag{2}$$

where $Q_1 \in \mathbb{R}^{n_1 \times n_1}$, $Q_2 \in \mathbb{R}^{n_1 \times n_2}$, $Q_3 \in \mathbb{R}^{n_2 \times n_2}$.

3. Calculate

$$\begin{aligned} f_c(q) &= D_r(P_1 \| P_2(q)) & (3) \\ &= \frac{1}{2}\text{trace}(C_4Q_4C_4^T + D_4D_4^T - I) - \frac{1}{2}\ln\det(D_4D_4^T). \end{aligned}$$

Theorem Algebraic method

Consider $\inf_{q \in QD} f_c(q)$.

(a) The optimum C_3^* is

$$C_3^* \quad = \quad -D_3 C_1 Q_2 Q_3^{-1}. \tag{4}$$

(b) The optimum D_3^* satisfies ${D_3^*}^T D_3^* = M^{-1}$,
where $\quad M = C_1 \left(Q_1 - Q_2 Q_3^{-1} Q_2^T \right) C_1^T + D_1 D_1^T.$

(c) The criterion simplifies to

$$f_c(q) = -\frac{1}{2} \ln \det (D_1^T M^{-1} D_1) = \frac{1}{2} \ln \det M - \ln \det D_1.$$

$\det(M)$ is a rational function w.r.t. entries of the (A_3, B_3).
Thus one obtains an infimization of a rational function with constraints encoded in QD.

Global optimization of rational functions

$$\text{Find} \quad \inf_{(x_1,\ldots,x_k)\in S\subseteq\mathbb{R}^k} \frac{p(x_1,\ldots,x_k)}{q(x_1,\ldots,x_k)},$$

where p, q are multivariate polynomials without common factor.

Note NP-hard problem.

Algebraic methods: see

(D. Jibetean (2003, Ph.D. thesis); Research advisor B. Hanzon).

(a) **Eigenvalue method**

$\Longrightarrow$ Global minimal value and a set of global minimizers (at least one point in each connected component).

(b) **Linear Matrix Inequalities method**

$\Longrightarrow$ Global lower bound; combined with local search leads to global minimum.

LMI method. Main results.

Assumption: $S \subset \mathbb{R}^k$ is an open, semi-algebraic set (partial closure).

Theorem Let $p(x),\ q(x)$ relatively prime.

$$\exists x_1, x_2 \in S,\ q(x_1) > 0,\ q(x_2) < 0 \implies \inf_{x \in S \subseteq \mathbb{R}^k} \frac{p(x)}{q(x)} = -\infty.$$

Remark Converse *not* true.

Assume w.l.g. $q(x) \geq 0,\ \forall x \in S$.

Optimal system reduction is equivalent to:

Problem

$$\begin{aligned} &\sup \quad \alpha \in \mathbb{R} \\ &\text{s.t.} \ \ p(x) - \alpha q(x) \geq 0, \quad \forall x \in S. \end{aligned} \tag{5}$$

Approach for (**??**): LMI relaxation and constraints encoded in S.

Transformation

$$\begin{aligned} p(x) - \alpha q(x) &= z^T N(\lambda, \alpha) z, \\ z^T &= \begin{pmatrix} 1 & x_1 & \dots & x_k & x_1^2 & x_1 x_2 & x_2^2 & \dots & x_k^d \end{pmatrix}, \\ N(\lambda, \alpha) &= Q_0 + \sum_{i=1}^{s} Q_i \lambda_i + Q_{s+1} \alpha. \end{aligned}$$

Gram matrix $N(\lambda, \alpha)$ is affine, symmetric, $\alpha \in \mathbb{R}$, $\lambda \in \mathbb{R}^s$;

$$\exists \lambda^* \in \mathbb{R}^s, \quad N(\lambda^*, \alpha) \succeq 0 \Longrightarrow p(x) - \alpha q(x) \geq 0, \ \forall x \in \mathbf{R}^k.$$

Generalized Gram matrix $N_S(\alpha, \lambda)$

$$\exists \lambda^* \in \mathbb{R}^s, \quad N_S(\lambda^*, \alpha) \succeq 0 \Longrightarrow p(x) - \alpha q(x) \geq 0, \ \forall x \in S \subseteq \mathbf{R}^k.$$

Theorem The convex problem

$$\begin{aligned} &\sup \quad \alpha \in \mathbb{R} \\ &\text{s.t.} \;\; N_S(\lambda, \alpha) \succeq 0, \end{aligned} \tag{6}$$

is an LMI relaxation of (**??**), specifically,

$$\begin{aligned} &\sup \quad \alpha && \leq \quad \inf_{(x_1,\ldots,x_k)\in S\subseteq\mathbb{R}^k} \frac{p(x_1,\ldots,x_k)}{q(x_1,\ldots,x_k)} \\ &\text{s.t.} \;\; N_S(\lambda, \alpha) \succeq 0. \end{aligned}$$

Is the lower bound tight?

$$N_S(\lambda^*, \alpha^*) \Longrightarrow Y^* = Y(\lambda^*, \alpha^*).$$

If $\operatorname{rank}(Y^*) = 1$, $Y^* = z^* {z^*}^T$, $z^* \mapsto (x_1^*, \ldots, x_k^*)$ global optimum then the bound is tight.

Procedure for infimization of divergence rate

(1) Select a canonical form (A_3, B_3, C_3, D_3), with C_3, D_3 fully parametrized, independent from A_3, B_3. For example, the control canonical form.

(2) Solve symbolically for Q_1, Q_2, Q_3 $(Q_1 \to Q_2 \to Q_3)$

$$\begin{pmatrix} Q_1 & Q_2 \\ {Q_2}^T & Q_3 \end{pmatrix}$$

$$= \begin{pmatrix} A_1 & 0 \\ B_3 C_1 & A_3 \end{pmatrix} \begin{pmatrix} Q_1 & Q_2 \\ {Q_2}^T & Q_3 \end{pmatrix} \begin{pmatrix} A_1 & 0 \\ B_3 C_1 & A_3 \end{pmatrix}^T +$$

$$+ \begin{pmatrix} B_1 \\ B_3 D_1 \end{pmatrix} \begin{pmatrix} B_1 \\ B_3 D_1 \end{pmatrix}^T .$$

(3) Calculate $\det M$, a rational function, where

$$M(q) = C_1 \left(Q_1 - Q_2 Q_3{}^{-1} Q_2{}^T\right) C_1^T + D_1 D_1^T .$$

(4) Compute (see page 12)

$$\inf_{q \in QD} \det M(q).$$

If, moreover, the infimum is attained, i.e. the global minimum exists, then determine its location $q^* \in QD$ as well.

(5) Evaluate System 3 at the optimal value and compute

$$q^* \mapsto (A_3^*, B_3^*, C_3^*, D_3^*).$$

(6) Compute the approximant System 2

$$(A_2^*, B_2^*, C_2^*, D_2^*).$$

Example 1

Consider a Gaussian system of order 2 in the control canonical form

$$\begin{aligned} A_1 &= \begin{pmatrix} -0.4 & -0.32 \\ 1 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ C_1 &= \begin{pmatrix} 0 & -0.28 \end{pmatrix}, \quad D_1 = \begin{pmatrix} 1 \end{pmatrix}. \end{aligned}$$

System 3 is parametrized by

$$(a_3, 1, c_3, d_3) \in SGSP_{min}(1,1,1),$$

that is, $d_3 > 0,\ |a_3| < 1,\ |a_3 - c_3 d_3^{-1}| < 1,\ c_3 \neq 0.$

Example 1 (Continued)

The criterion becomes

$$f_c = -\frac{1}{2} \ln \left(\frac{-34\left(25 + 10\, a_3 + 8\, {a_3}^2\right)\left(56907\, {a_3}^2 - 230375 - 79900\, a_3\right)}{(731\, {a_3}^2 + 1801\, a_3 + 19500)\,(391\, {a_3}^2 + 7039\, a_3 + 11000)} \right)$$

In the stability region we find two local minima

$$\begin{aligned} f_c((0.6353, 1, 0.1059, 0.9631)) &= 0.0376, \\ f_c((-0.7835, 1, -0.1269, 0.9693)) &= 0.0312. \end{aligned}$$

The second point is global minimum.

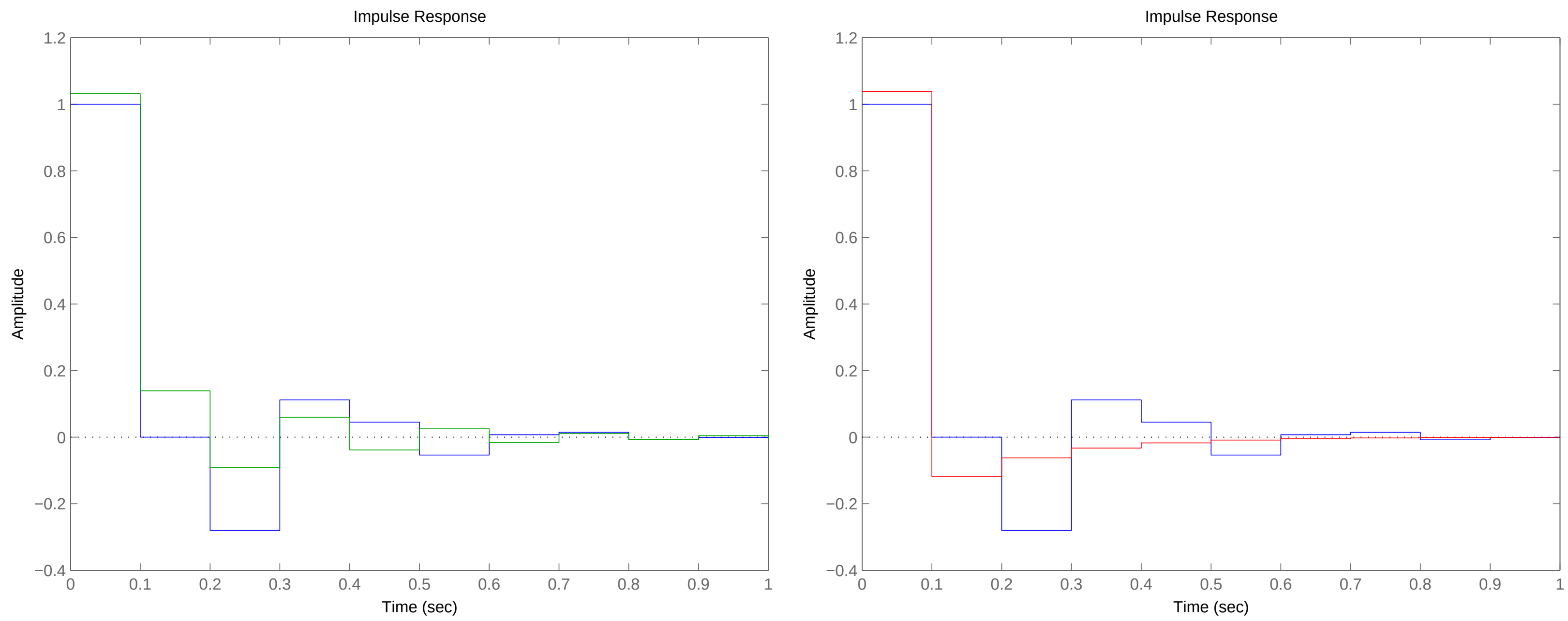
Compute the approximants (System 2)

$$(a_2^*, b_2^*, c_2^*, d_2^*) = (0.5253, 1.0383, -0.1142, 1.0383),$$

respectively $(-0.6525, 1.0317, 0.1351, 1.0317)$.

Example 1 (Continued)

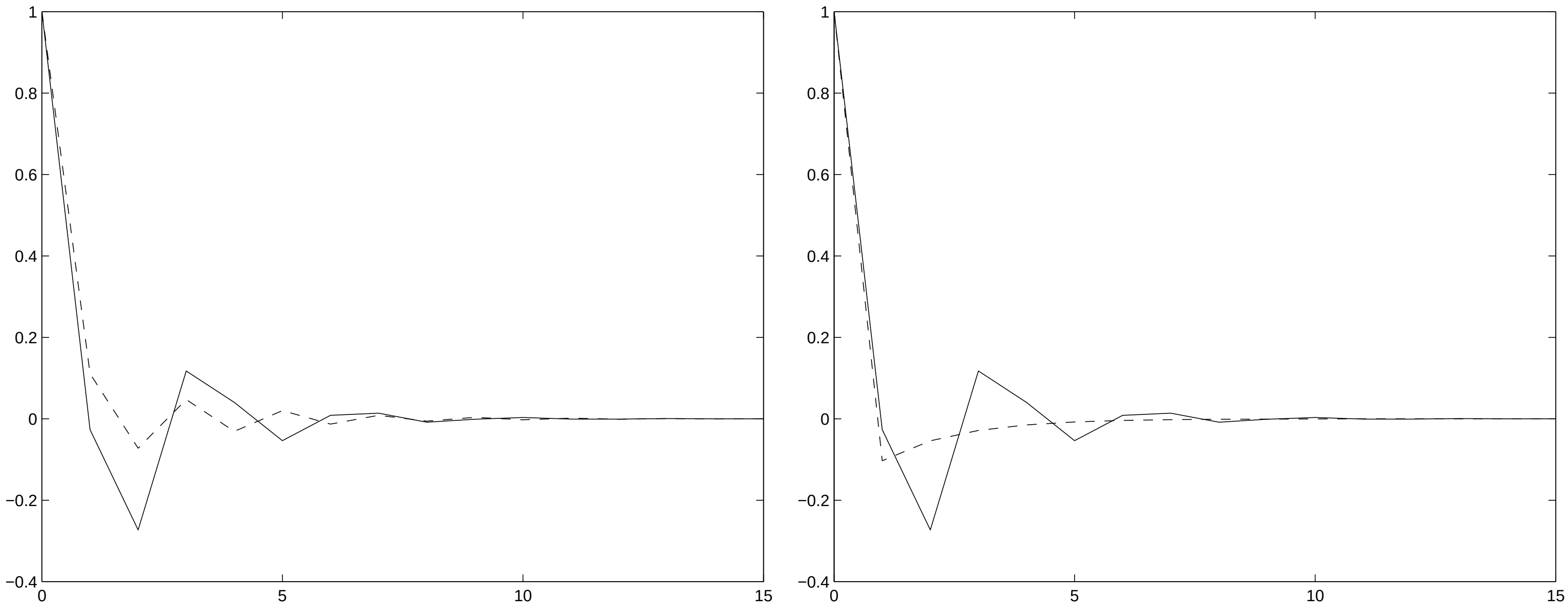
The impulse response of the original system, global approximant, and local approximant are plotted.



Example 1 (Continued)

The covariance function of the original system,

global approximant, and local approximant are plotted.



Example 2

System reduction from $n_1 = 3$ to $n_2 = 2$.

Consider Gaussian system with

$$\begin{aligned} A_1 &= \begin{pmatrix} -1/4 & 1/2 & 1/3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \\ C_1 &= \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}, \quad D_1 = 2, \\ \mathrm{spec}(A_1) &= \{0.8322, \; -0.5411 + / - i0.3281\}. \end{aligned}$$

Approximant System 3 in control canonical form with parameters,
α_1, α_2, γ_1, γ_2, δ.

Example 2

Optimize analytically with respect to γ_1, γ_2, and δ.

Criterion becomes,

$$\sup_{\alpha_1, \alpha_2 \in A_D} \frac{5640\alpha_1^3 + 85896\alpha_1^2 + ...}{376\alpha_2^2 - 618\alpha_2^2 + \dots},$$

$$A_d = \left\{ \begin{array}{l} (\alpha_1, \alpha_2) \in A_D | 1 + \alpha_2 \geq 0, \\ \alpha_1 - \alpha_2 + 1 \geq 0, \ - \alpha_1 - \alpha_2 + 1 \geq 0 \end{array} \right) .$$

Optimization of rational function.

Gröbner basis, produces at most 100 local minima. Not all computed.

Bounding technique using SOSTOOLS (Pablo Parrilo) yields a global approximant.

Example 2

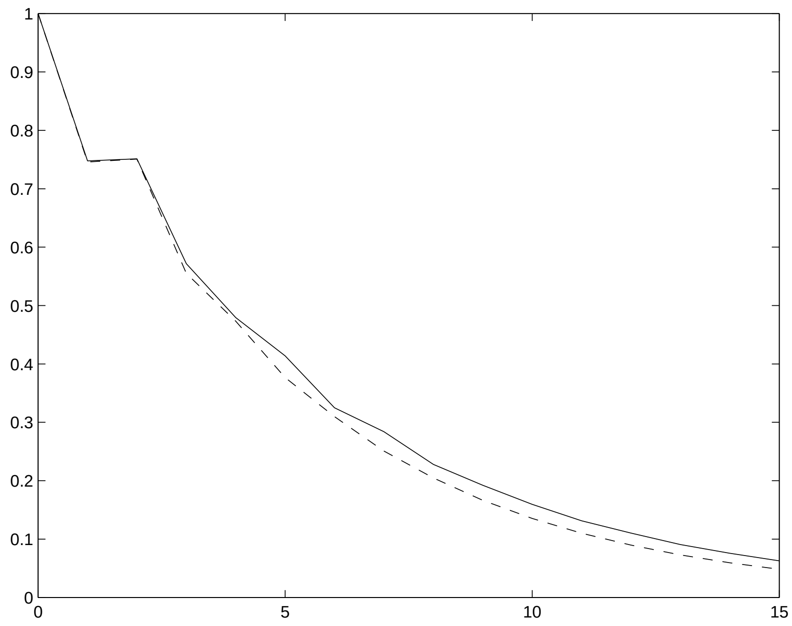
Global approximant,

$$
\begin{aligned}
A_2 &= \begin{pmatrix} 0.4252 & 0.3162 \\ 1 & 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 2.0071 \\ 0 \end{pmatrix}, \\
C_2 &= \begin{pmatrix} 0.5033 & 0.6739 \end{pmatrix}, \; D_2 = 2.0071, \\
\mathrm{spec}(A_2) &= \{0.8138, \; -0.3886\}, \\
f_c(q^*) &= 0.0036.
\end{aligned}
$$

Case $n_2 \geq 3$ requires investigation of parameter constraints.

Example 2 (Continued)

The covariance function of the original system (continuous line), best approximant (dashed line) are plotted.



Concluding remarks

- System reduction by algebraic method.
- Procedure algebraic infimization of divergence rate.
- Examples
- Implications for practice of system identification.

Further research

- Class of Gaussian systems of order $n_2 \geq 3$.
- Algebraic theory and algorithms, in particular, how to better handle parametrizations and constraints.

The End!