Estimating parameters of dynamic errors-in-variables systems with polynomial nonlinearities

LEVENTE HUNYADI and ISTVÁN VAJK
Budapest University of Technology and Economics
Department of Automation and Applied Informatics
1111 Budapest, Goldmann György tér 3., HUNGARY
http://www.aut.bme.hu/portal/hunyadi

Abstract: An approach for identifying single-input single-output discrete-time dynamic nonlinear errors-in-variables systems is presented where the system model can be linearized such that it is expressed as a linear combination of polynomials of input and output observations. We assume white Gaussian noise on both input and output, characterized by a noise magnitude and a normalized noise covariance structure matrix, and employ a nonlinear extension of the generalized Koopmans–Levin method to estimate model parameters with an assumed noise structure and a subsequent covariance matching objective function minimization to estimate all noise parameters. The feasibility of the approach is demonstrated by Monte-Carlo simulations.

Key–Words: system identification; nonlinear systems; polynomial eigenvalue problem; covariance matching

1 Introduction

Errors-in-variables systems where all variables are equally observed with noise are of particular significance in applications where the quantitative description of the internal laws constituting the system is of interest rather than predicting future behavior. Such applications include computer vision, image reconstruction, speech and audio processing, signal processing, modal and spectral analysis, system identification, econometrics and time series analysis. In these applications, the task is to construct a best possible system model based on noisy observations. A general system model takes the form

\[ f(\theta, x^{(i)}_0) = 0 \]

where the vector \( x^{(i)}_0 \) denotes the true observations for \( i = 1, \ldots, N \), \( N \) is the number of observations, the vector \( \theta \) encapsulates the parameters of interest, and \( f \) represents some constraint between (for dynamic systems, past and present) observations. A usual assumption that is satisfied in most applications is that the constraint \( f \) is linear in \( \theta \), i.e.

\[ \theta^T g(x^{(i)}_0) = 0 \]

where \( g \) is a linearization of \( f \). Given that observations \( x^{(i)}_0 \) are not directly observable but contaminated with noise, the most common assumption being Gaussian noise, hence the actual observations \( x^{(i)} \) satisfy \( x^{(i)} = x^{(i)}_0 + z^{(i)} \), the objective is to derive estimates for \( \theta \) given the noisy observations \( x^{(i)} \).

For the special case of identifying linear errors-in-variables systems, where the constraint \( f \) is linear not only in \( \theta \) but also in observations \( x^{(i)}_0 \), a number of estimation schemes have been proposed. For dynamic single-input single-output (SISO) systems, where the system is described by the linear equation

\[ y_0[k] + a_1 y_0[k-1] + a_2 y_0[k-2] + \ldots + a_m y_0[k-m] = b_1 u_0[k-1] + b_2 u_0[k-2] + \ldots + b_n u_0[k-n] \]

but neither the noise-free true input \( u_0 \) nor the true output \( y_0 \) is observable but one is confined to their noise-contaminated variants \( u \) and \( y \), proposed methods include bias-compensating least squares, the Frisch scheme, instrumental variable, higher-order statistics, structured total least squares, frequency-domain and efficient maximum likelihood methods, see [2, 3, 7] and references within, or [6] for a comprehensive survey.

This paper deals with a nonlinear extension of the generalized Koopmans–Levin method to estimate model parameters of a dynamic system with given noise structure where the linearization \( g \) is a polynomial in terms of input and output observations, and a subsequent covariance matching objective function to
estimate noise covariance structure. The Koopmans–Levin method, proposed in [5], gives a non-iterative quick estimate of the model parameters of a linear system given a priori information on the noise structure. The original method was generalized in [8] to improve estimation accuracy at the cost of increased computational complexity, incorporating as special cases the original Koopmans–Levin method and the maximum likelihood method. In addition, a nonlinear extension to the original Koopmans method was proposed in [9] for static systems.

The Koopmans–Levin method and its generalization are briefly described in Section 2. Section 3 combines and extends the results of [8] and [9] to non-linear dynamic systems that comprise of polynomial nonlinearities yet are linear in model parameters. The outlined method assumes a preliminarily known noise structure, Section 4 extends the estimation method so that no such assumptions are required. In order to demonstrate the feasibility of the method, some simulation results are presented in Section 5 before the paper concludes with Section 6.

2 The generalized Koopmans–Levin method

Consider the linear SISO errors-in-variables system \( G(q^{-1}) \) described by the autoregressive moving average (ARMA) difference equation

\[
A(q^{-1})y_0[k] = B(q^{-1})u_0[k]
\]  

(1)

where \( q^{-1} \) denotes the backward shift operator such that \( q^{-1}u[k] = u[k-1] \) and

\[
A(q^{-1}) = a_0 + a_1 q^{-1} + \cdots + a_m q^{-m}
\]

\[
B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m}.
\]

Given the aforementioned system description, we may introduce the model parameter vector \( \theta \) and the extended regressor vector \( z \)

\[
\theta^T = \begin{bmatrix} a_0 & a_1 & \cdots & a_m & -b_0 & -b_1 & \cdots & -b_m \end{bmatrix}
\]

\[
z_k^T = \begin{bmatrix} y(k) & \cdots & y(k-m) & u(k) & \cdots & u(k-m) \end{bmatrix}
\]

in which the shorthand notation \( z(k) = z[k] \) has been used,\(^1\) such that \( \theta^T z(k) = 0 \) \( \forall k = m+1 \ldots N \) with \( m \) being the order of the model and \( z(k) = z(k) - z_0(k) \)

\(^1\)When not misleading, the compact notation \( z_k \) will also be used to denote the noise-contaminated \( z[k] = z(k) \).

is the noise contribution. Notice that the system is linear in components \( y^{(i)} \) and \( u^{(i)} \) as well as in model parameters \( a_i \) and \( b_i \). Furthermore, introduce the observation sample and noise covariance matrices as

\[
D = \mathbb{E} \left( z z^T \right) \approx \frac{1}{N} \sum_{k=1}^{N} z(k) [z(k)]^T
\]

\[
\mu C = \mathbb{E} \left( z z^T \right) \approx \frac{1}{N} \sum_{k=1}^{N} z(k) [z(k)]^T
\]

with \( \mu \) denoting noise magnitude and \( C \) representing a normalized noise covariance matrix, or noise structure, that expresses the relative distribution of noise between input and output.

The essence of the Koopmans–Levin method is that the (full-rank) sample covariance matrix \( D \) comprising of noisy observations can be decomposed into a (rank-deficient) noise-free component \( D_0 \) and a noise component \( C \):

\[
\theta^T D \theta = \theta^T D_0 \theta + \mu C \theta = \theta^T \mu C \theta
\]

\[
\text{in which } \theta^T D \theta = \theta^T \mathbb{E} \left( z z^T \right) \theta = \mathbb{E} (\theta^T z z^T \theta) \text{ and } \theta^T D_0 \theta = \theta^T \mathbb{E} (z_0 z_0^T) \theta = 0 \text{ so that finding } \theta \text{ entails minimizing the objective function}
\]

\[
J = \frac{1}{2} \theta^T D \theta
\]

which can be effectively tackled by solving the eigenvector decomposition problem

\[
(D - \mu C) \theta = 0
\]

or

\[
\det (D_0) = \det (D - \mu C) = 0
\]

so that the model parameter vector is found by solving a generalized eigenvector problem on the matrix pair \((D, C)\). The problem may alternatively be formulated using matrix notation where

\[
Z = \begin{bmatrix} y_1 & \cdots & y_m & u_1 & \cdots & u_m \\ y_2 & \cdots & y_{m+1} & u_2 & \cdots & u_{m+1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ y_N & \cdots & y_{N+m-1} & u_N & \cdots & u_{N+m-1} \end{bmatrix}
\]

\[
D = Z^T Z
\]

\[
\mu C = \text{diag} \left[ \sigma_y^2 \cdots \sigma_y^2 \sigma_u^2 \cdots \sigma_u^2 \right]
\]

with \( Z \) being an \((N - m + 1) \times 2m\) matrix and \( C = \mu C \sigma \otimes I_m \) denoting the noise structure such that the noise covariance matrix is known up to a multiplication by a scalar \( \mu \) representing the noise magnitude, i.e.

\[
\mu C \rho = \mu \begin{bmatrix} \sin^2 \rho & 0 \\ 0 & \cos^2 \rho \end{bmatrix} = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}
\]

(3)
in which we assume that the noise structure matrix $C_p$ is preliminarily known.

One way [8] to improve the robustness of the parameter estimation approach outlined above is by instead of (2) minimizing the objective function

$$J = \frac{1}{2 (q - m)} \text{trace} \left( (G_q^T C_q G_q)^{-1} G_q^T D_q G_q \right)$$

where $Z_q$ is an $(N - q + 1) \times 2q$ matrix obtained by augmenting $Z$ with $q - m$ columns of additional observations for both $y$ and $u$; $G_q$ is a $2q \times (q - m)$ matrix of model parameters such that $(Z_0)_q G_q = 0$; and $C_q = (\mu C_p) \otimes I_q$ is a diagonal covariance structure matrix of size $2q$, and $m + 1 \leq q \leq N$. In short, both the model parameter vector $\theta$ and the original observation matrix $Z$ have been extended from size $m$ to $q$.

The above problem can be reformulated as

$$J = \frac{1}{2 (q - m)} \text{trace} \left( (G_q^T C_q G_q)^{-1} G_q^T D_q G_q \right) \quad (4)$$

which can be gradually approximated with the iteration scheme

$$\frac{\text{trace} \left( ((G_q^T \theta_k) C_q G_q(\theta_k))^{-1} G_q^T (\theta) D_q G_q(\theta) \right)}{\text{trace} \left( ((G_q^T \theta_k) C_q G_q(\theta_k))^{-1} G_q^T (\theta) C_q G_q(\theta) \right)} \quad (5)$$

where $\theta_{k+1} = \arg \min_\theta (5)$, which yields the same extreme value upon convergence. To facilitate easier computation, the above scheme is equivalent to

$$\frac{\theta^T T^T \left( (G_q^T (\theta_k) C_q G_q(\theta_k))^{-1} \otimes D_q \right) T \theta}{\theta^T T^T \left( (G_q^T (\theta_k) C_q G_q(\theta_k))^{-1} \otimes C_q \right) T \theta} \quad (6)$$

where $\theta_{k+1} = \arg \min_\theta (6)$ and $T$ (a sparse matrix of zeros and ones) is chosen such that $\text{vec}(G_q) = T \theta$. In each iteration, minimization w.r.t. $\theta$ is attained by solving a generalized eigenvector decomposition problem on the matrix pair $(Q, R)$ with

$$Q = T^T \left( (G_q^T (\theta_k) C_q G_q(\theta_k))^{-1} \otimes D_q \right) T$$

$$R = T^T \left( (G_q^T (\theta_k) C_q G_q(\theta_k))^{-1} \otimes C_q \right) T$$

where $\theta$ is the eigenvector that belongs to the smallest eigenvalue $\mu$.

### 3 A nonlinear extension

In order to further generalize the Koopmans–Levin method to nonlinear systems, linear components $y_k$ and $u_k$ give way to the nonlinearities that occur in the model. For instance, for the nonlinearities $y_k, u_k, u_k^2$ and $y_k u_k$, the $(N - q + 1) \times nq$ matrix $Z_q$ takes the form

$$Z_q = \begin{bmatrix} y_1 & \cdots & y_q & u_1 & \cdots & u_q \\ y_2 & \cdots & y_{q+1} & u_2 & \cdots & u_{q+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_1^2 & \cdots & u_q^2 & y_1 u_1 & \cdots & y_q u_q \\ u_2^2 & \cdots & u_{q+1}^2 & y_2 u_2 & \cdots & y_{q+1} u_{q+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

and the general form for the matrix $G_q$ becomes

$$G_q = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \\ nq, q - m \end{bmatrix}$$

$n$ being the number of nonlinear components (in our case, $n = 4$) and $G_q$ encapsulating the parameters for the respective nonlinearity. Notice that the matrix product $Z_q G_q$ entails that the system is still linear in parameters. However, the covariance matrix structure $C_q$ is no longer a single diagonal matrix but is replaced by a matrix polynomial

$$C_q(\mu) = \mu C_q^{(1)} + \mu^2 C_q^{(2)} + \ldots + \mu^p C_q^{(p)}$$

in which $C_q^{(i)}$ is the $i$th coefficient of the matrix polynomial $C_q(\mu)$. One can use the following identities in deriving $C_q$:

$$E \left( x_k^p \right) = E \left( [x_0]_k + n_k \right)^p$$

$$E \left( n_k^{2i} \right) = (2k - 1)(2k - 3) \ldots 1 \sigma_{2i}^2$$

$$E \left( n_k^{2i-1} \right) = 0$$

$$E \left( x_k \right) \approx \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_k$$

$$E \left( x_k x_{k-\tau} \right) = E \left( x_k \right) E \left( x_{k-\tau} \right)$$

in which $n_k$ is a zero-mean $\sigma^2$-variance normally distributed random value at time instant $k$. [9] features a detailed description as well as an introductory example on how to derive these terms. Observe that the matrix entries usually depend not only on noise parameters $\sigma_u$ and $\sigma_y$ but also on (means of) the observations themselves.

The use of covariance matrix polynomials instead of regular covariance matrices necessitates some modifications to the objective function (4) as well as the iteration scheme (6). By including the noise magnitude within the trace operator in (4) yielding

$$\text{trace} \left( (G_q^T \mu C_q G_q)^{-1} G_q^T D_q G_q \right) ,$$
that it is apparent that the matrix product approaches the unit matrix should the best possible model parameters and noise covariance matrix be used. In this spirit, (4) can be reformulated as

$$J = \left(\text{trace}\left(\left(G_q^\top \mu C_q G_q\right)^{-1} G_q^\top D_q G_q\right) - (q - m)\right)^2$$

(7)

with $m$ being the the order of the dynamic model to estimate where the minimum of $J$ is attained when both model and noise magnitude estimates best match observations. Substituting the noise covariance polynomial $C_q(\mu)$ into (7), we get a parameter estimation scheme for nonlinear systems. Thus, we propose the following differentiable objective function:

$$J = \left(\text{trace}(\gamma^{-1} \delta) - (q - m)\right)^2$$

(8)

where

$$\delta = G_q^\top(\theta) D_q G_q(\theta)$$

$$\gamma = G_q^\top(\theta) C_q(\mu) G_q(\theta)$$

As the function (8) is differentiable, a direct search utilizing the Levenberg-Marquardt method yields model parameter and noise magnitude estimates. However, the Levenberg-Marquardt method finds only local minima, making the scheme sensitive to initial values.

Iterative schemes are more robust against local minima. Modifying (6) to incorporate the covariance polynomial $C_q(\mu)$ we get

$$\theta^T T^T \left(\left(G_q^\top(\theta_k) C_q(\mu_k) G_q(\theta_k)\right)^{-1} \otimes D_q\right) T \theta$$

(9)

where $\theta_{k+1} = \arg \min_{\theta, \mu} (9)$ with $\theta = [\theta \ \mu]$ needs the solution of a polynomial eigenvalue decomposition problem

$$\Psi(\mu) \theta = (Q - \mu R_1 - \mu^2 R_2 - \ldots - \mu^p R_p) \theta = 0$$

(10)

with

$$Q = T^T \left(\left(G_q^\top(\theta_k) C_q(\mu_k) G_q(\theta_k)\right)^{-1} \otimes D_q\right) T$$

$$R_i = T^T \left(\left(G_q^\top(\theta_k) C_q(\mu_k) G_q(\theta_k)\right)^{-1} \otimes C_q(i)\right) T$$

Further reduction in the computational space is possible if a priori knowledge of the equality of certain parameters is available. Introducing the structural constraint matrix $S$ with 0 and 1 entries such that $S \theta = \theta_p$, it is possible to augment the above equations to yield

$$Q = S^T T^T \left(\left(G_q^\top(\theta_k) C_q(\mu_k) G_q(\theta_k)\right)^{-1} \otimes D_q\right) TS$$

$$R_i = S^T T^T \left(\left(G_q^\top(\theta_k) C_q(\mu_k) G_q(\theta_k)\right)^{-1} \otimes C_q(i)\right) TS$$

and restrict the search for parameter estimates in $\theta_R$ with $\dim \theta_R < \dim \theta$.

One way to solve a polynomial eigenvalue decomposition problem is by linearization. As a result of linearization, the polynomial eigenvalue problem reduces to a generalized eigenvalue problem. In particular, (10), when subject to symmetry-preserving linearization $\psi(\mu)$ [1] becomes for even $p$

$$\text{diag} \left(\left[\begin{array}{cccc}0 & I & 0 & \ldots \end{array}\right], \left[\begin{array}{cccc}0 & I & R_1 & \ldots \end{array}\right], \ldots, \left[\begin{array}{cccc}0 & I & R_{p-1} & -R_p \end{array}\right]\right)$$

and for odd $p$

$$\text{diag} \left(\left[\begin{array}{cccc}Q & 0 & I & R_1 \end{array}\right], \left[\begin{array}{cccc}Q & 0 & I & R_3 \end{array}\right], \ldots, \left[\begin{array}{cccc}Q & 0 & I & -R_p \end{array}\right]\right)$$

As the linearized problem has eigenvectors of dimension $mp$ rather than $m$, the true polynomial eigenvector that belongs to the eigenvalue $\mu$ becomes the portion $v_i$ of the linearized eigenvector $\psi(\mu) x = 0$ that gives the smallest normalized residual, i.e.

$$v_i = \arg \min_{v_i} \frac{\sum_i |(\Psi(\mu) v_i)_i|}{\sum_i |v_i|}$$

With the iterative scheme (9) at hand, a few initial iterations can be used to seed the Levenberg-Marquardt search with appropriate initial values reducing the likelihood of (8) getting stuck in a local minimum.

4 Model and noise estimation

Contrary to the hidden assumption in the previous section, in a real-world scenario, the true noise structure $C_p$ (or equivalently, a noise direction $\rho$ that determines the ratio of input and output noise variances for a unit magnitude noise) is seldom at our disposal. As the final step of the parameter estimation method, we propose means to estimate $C_p$ for white noise.

One way to parametrize noise variances, as in (3), is by writing $\sigma_u^2 = \mu \cos^2 \rho$ and $\sigma_y^2 = \mu \sin^2 \rho$ such that $C_q = C_q(\mu, \rho)$. Let $\hat{\theta}$ denote (unit-normalized) estimates obtained with a particular assumption of $\rho$ using (9). Introduce the notations

$$\hat{\delta} = G_q^\top(\hat{\theta}) D_q G_q(\hat{\theta})$$

$$\hat{\gamma} = G_q^\top(\hat{\theta}) C_q(\hat{\mu}, \rho) G_q(\hat{\theta})$$
Varying $\rho$ in the range from $0$ to $\frac{\pi}{2}$, one can discover the “true” value by minimizing the loss function
\[ J(\rho) = \| \hat{\delta} - \hat{\gamma} \|_F \] (11)
where $\| \cdot \|_F$ denotes the Frobenius norm (a technique called covariance matching in [10]) or the “inverted” loss function
\[ J(\rho) = \text{trace}(\hat{\delta}^{-1}\hat{\gamma}) \] (12)
or the so-called Itakura–Saito divergence
\[ J(\rho) = \text{trace}(\gamma^{-1}\delta) - \log(\text{det}(\gamma\delta^{-1})) - n \] (13)
where $n$ is the dimension of the square matrices involved. The minimum value for $J$ in the above equations yields the optimal value for $\rho$.

As an alternative to the two-stage estimation procedure outlined above, a single-stage strategy might theoretically also be employed. Let
\[
\delta = G_q^\top(\theta)D_qG_q(\theta) \\
\gamma = G_q^\top(\theta)C_q(\mu, \rho)G_q(\theta)
\]
and introduce the objective functions
\[
J_1 = \left( \text{trace}(\gamma^{-1}\delta) - (q - m) \right)^2 \\
J_2 = \left( \text{trace}(\delta^{-1}\gamma) - (q - m) \right)^2 \\
J_3 = \| \delta - \gamma \|_F^2 = \text{trace}\left\{ (\delta - \gamma)^\top(\delta - \gamma) \right\}
\]
Notice that all are functions of $\theta$, $\mu$ as well as $\rho$ simultaneously. Due to their nonlinearity, they are likely to exhibit convergence to local minima if started with the wrong initial values.

5 Simulation results

Let us now draw our attention to an artificial yet relatively complex process described by the nonlinear relationship
\[
y_0[k] = p_1y_0[k - 1] + p_2y_0[k - 2] + p_3u_0[k - 1] + \\
+ p_4u_0^2[k - 1] + p_5y_0[k - 1]y_0[k - 2] \\
+ p_6u_0[k - 1]y_0[k - 1]
\]
comprising of both linear and polynomial terms as well as cross-correlating terms. The true parameter values are set to
\[
p_1 = 1.5 \quad p_2 = -0.7 \quad p_3 = 1 \quad p_4 = -0.3 \quad p_5 = -0.05 \quad p_6 = 0.1
\]
and a configuration of $N = 500$ samples, $q = 6$, $\sigma_u = 0.01$ and $\sigma_y = 0.01$ is set up. Next, a Monte-Carlo simulation of $M = 100$ runs of the outlined nonlinear extension to the generalized Koopmans–Levin method has been carried out. The mean values and variances of parameters thus obtained are shown in Table 1.

Figure 1 shows how to discover noise covariance structure (i.e. the noise “direction” $\rho$) by minimizing the distance or divergence metrics (11) and (12) over an interval to arrive at estimates for all parameters. In accordance, each independent simulation of the comprehensive results shown in Table 1 in which all model and noise parameters are estimated simultaneously entails a noise structure discovery step shown in Figure 1.

For the sake of comparison, an instrumental variable scheme, based on the bias-compensating least-squares technique for nonlinear polynomial systems (NBCLS) has been included in the table. The scheme minimizes the objective function [4]
\[ J = \| d_{IV} - c_{IV} - (D_{IV} - C_{IV})\hat{\theta} \| \] (14)
where
\[
\hat{\theta} = (D_{IV} - C_{IV})^{-1}(d_{IV} - c_{IV})
\]
with $D_{IV}$ and $d_{IV}$ being the (rectangular) covariance matrices of the regressor vector and the output vector, respectively, w.r.t. so-called instruments and $C_{IV}(\mu, \rho)$ and $c_{IV}(\mu, \rho)$ being the corresponding computed (rectangular) noise covariance matrices. Instruments include the regressor vector as well as past observations not in the system model. The notation $M^\dagger$ stands for the Moore-Penrose pseudoinverse $(M^\top M)^{-1}M^\top$.

6 Conclusion

After a brief description of the Koopmans–Levin method and its generalization for linear systems, we introduced a nonlinear extension by modifying the objective function and the iteration scheme of the linear generalization. As a result, the original dependency on the model parameters $\theta$ has been extended with dependency on the noise magnitude $\mu$. Consequently, the iterative problem is tackled efficiently by solving a polynomial rather than a generalized eigenvector decomposition problem. Next, an optimization scheme of minimizing an error term over an angle
Table 1: Simulation results with unknown noise structure and various methods to match covariance matrices (TS denotes time shift).

<table>
<thead>
<tr>
<th>nonlinearity</th>
<th>TS</th>
<th>NBCLS (14)</th>
<th>Frobenius norm (11)</th>
<th>“inverted” (12)</th>
<th>Itakura–Saito (13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{y}_0 )</td>
<td>1</td>
<td>-1.4996 ± 0.0045</td>
<td>-1.5002 ± 0.0039</td>
<td>-1.4997 ± 0.0044</td>
<td>-1.4998 ± 0.0036</td>
</tr>
<tr>
<td>( \hat{y}_0 )</td>
<td>2</td>
<td>0.6998 ± 0.0038</td>
<td>0.7002 ± 0.0032</td>
<td>0.6997 ± 0.0035</td>
<td>0.7000 ± 0.0028</td>
</tr>
<tr>
<td>( \hat{u}_0 )</td>
<td>1</td>
<td>-1.0015 ± 0.0096</td>
<td>-0.9998 ± 0.0182</td>
<td>-1.0021 ± 0.0221</td>
<td>-0.9980 ± 0.0198</td>
</tr>
<tr>
<td>( \hat{u}_0 )</td>
<td>2</td>
<td>0.3258 ± 0.3173</td>
<td>0.3101 ± 0.0953</td>
<td>0.2928 ± 0.1053</td>
<td>0.3004 ± 0.0898</td>
</tr>
<tr>
<td>( \hat{y}_{(k)} )</td>
<td>1</td>
<td>0.0500 ± 0.0040</td>
<td>0.0492 ± 0.0114</td>
<td>0.0496 ± 0.0104</td>
<td>0.0505 ± 0.0102</td>
</tr>
<tr>
<td>( \hat{u}_{(k)} )</td>
<td>2</td>
<td>-0.1058 ± 0.0477</td>
<td>-0.1048 ± 0.0391</td>
<td>-0.1017 ± 0.0309</td>
<td>-0.1010 ± 0.0338</td>
</tr>
<tr>
<td>( \sigma^2_y )</td>
<td>( 10^{-3} )</td>
<td>0.0984 ± 0.0153</td>
<td>0.1017 ± 0.0131</td>
<td>0.0999 ± 0.0099</td>
<td>0.1011 ± 0.0102</td>
</tr>
<tr>
<td>( \sigma^2_u )</td>
<td>( 10^{-3} )</td>
<td>0.1021 ± 0.0406</td>
<td>0.0915 ± 0.0138</td>
<td>0.0944 ± 0.0126</td>
<td>0.0911 ± 0.0102</td>
</tr>
</tbody>
</table>

Figure 1: Discovering noise direction by successive estimation over an angle range (continuous: Frobenius (11); dashed: Itakura–Saito (13) metric).

Acknowledgements

This work has been supported by the fund of the Hungarian Academy of Sciences for control research and the Hungarian National Research Fund (grant number T68370).

References

[1] Efstatios N. Antoniou and Stavros Vologiannidis. Linearizations of polynomial matrices with symme-


