Multi-model identification and the separability index

R. Nikoukhah* F. Delebecque* S. L. Campbell† Kirk Horton†

Keywords: Failure detection, Identification, Auxiliary signal design

Abstract

Under the assumption that one of two given models is the real underlying model of the system, a valid auxiliary signal is defined as an input signal that allows the selection of the correct model. Under the assumption that the noise energy is bounded, the separability index is defined as the energy of the proper auxiliary signal of least energy. A constructive method for the computation of this index is presented.

1 Introduction

The design of auxiliary signals has been a major issue in system identification; it has been studied in the framework of multi-model identification, and in particular failure detection, in [9, 3, 4]; see also [7]. These works consider linear statistical models. The work presented here is different in that uncertainties are modeled as bounded energy signals and zero-error identification is envisaged. A similar approach has been undertaken in [6] for the study of robust failure detection problem.

In particular, we consider the problem of model selection. We suppose that two models are available for the system, one of which is correct. System outputs are measured, and the inputs (auxiliary signal) can be chosen.

Let v denote an auxiliary signal and let $\mathcal{A}^0(v)$ represent the set of possible outputs y associated with this input, if Model 0 were the correct model; similarly, let $\mathcal{A}^1(v)$ represent the set of outputs under the assumption that Model 1 is the correct model. Then clearly for perfect identification we need that

$$\mathcal{A}^0(v) \cap \mathcal{A}^1(v) = \emptyset. \tag{1.1}$$

The minimum energy required by v to impose (1.1) is a measure of how distinct the two models are, and how easy it is to distinguish them apart. We call it the separability index. In this paper, we propose a constructive method for the computation of this index and the corresponding auxiliary signal.

*INRIA, Rocquencourt, BP 105, 78153 Le Chesnay cedex, France

†North Carolina State University, Raleigh, NC 27695-8205. Research supported in part by the National Science Foundation under DMS-9714811, DMS-9802259, and INT-9605114.

2 Problem formulation

2.1 System models

The true system model is supposed to be one of the following two models

$$\dot{x}_i = A_i x_i + B_i v + M_i \nu_i, \tag{2.1}$$
$$y = C_i x_i + D_i v + N_i \nu_i \tag{2.2}$$

for $i = 0$ and 1. v is the auxiliary signal, y is the output and ν_i's represent perturbations, noises and unmeasured inputs. A_i, B_i, C_i, M_i, N_i are matrices of appropriate dimensions. N_i's have full row rank. The states of the two models x_0 and x_1 need not have the same dimensions. The same is true for ν_0 and ν_1.

The two models are supposed to be stable and satisfy:

$$\begin{pmatrix} sI - A_i & M_i \\ C_i & N_i \end{pmatrix} \quad \text{has full row rank } \forall s \text{ on the } j\omega\text{-axis.} \tag{2.3}$$

2.2 Proper auxiliary signal

We say that the L_2 vector function v is a *proper auxiliary signal* if its application implies that we are able to always distinguish the two candidate models based on observation y.

Definition 2.1 *Suppose*

$$||\nu_i||^2 = \int_0^\infty ||\nu_i(t)||^2 dt < 1, \quad i = 0, 1. \tag{2.4}$$

Then, we say that the auxiliary signal v is not proper if there exist x_0, x_1, ν_0, ν_1, and y satisfying (2.1), (2.2) and (2.4) both for $i = 0$ and $i = 1$. The auxiliary signal v is called proper otherwise.

Definition 2.2 *Let V denote the set of proper auxiliary signals v. Then,*

$$\gamma^* = \frac{1}{\inf_{v \subset V} ||v||} \tag{2.5}$$

where $||v||$ denotes the L_2-norm of v, is called the separability index associated with (2.1)-(2.2).

Here, we are interested in minimum energy proper auxiliary signals. Let us introduce an auxiliary cost function.

Definition 2.3 *The function $J_v(\beta)$ is the auxiliary cost function associated with problem (2.1)-(2.2) if*

$$J_v(\beta) = \inf \int_0^\infty \beta ||\nu_0(t)||^2 + (1-\beta)||\nu_1(t)||^2 dt \quad (2.6)$$

for $0 \le \beta \le 1$ where the infimum is taken over x_i, ν_i and y subject to (2.1)-(2.2), $i = 0, 1$.

Note that ν_0, ν_1 in (2.6) need not satisfy (2.4).

Lemma 2.1 *For all v, for $0 \le \beta \le 1$, $J_v(\beta)$ is defined and has the following properties:*

1. *it is zero for $\beta = 0$ and $\beta = 1$,*
2. *it is quadratic in v, i.e., for all scalar c, $J_{cv}(\beta) = |c|^2 J_v(\beta)$.*
3. *it is a continuous function of β,*
4. *it is a strictly concave function of β if the set of proper auxiliary signals is not empty, otherwise it is identically zero.*

Theorem 2.1 *Let*

$$J^*(\beta) = \sup_{v \neq 0} \frac{J_v(\beta)}{\int_0^\infty ||v(t)||^2 dt} \quad (2.7)$$

then

$$\gamma^{*2} = \max_{0 \le \beta \le 1} J^*(\beta). \quad (2.8)$$

The proof is based on the fact that for all $0 \le \beta \le 1$, $J_v(\beta) < 1$ if and only if v is not proper.

Note that the larger γ^* is, the easier it is to separate the two models. And when $\gamma^* = 0$, then the two model are indistinguishable no matter what the input v is. So, γ^* can be considered as the deterministic counterpart of the Kullback distance [5] used in some stochastic formulations.

2.3 Problem simplification

As we have seen in the previous section, the first optimization problem to solve is the following:

$$J_v(\beta) = \inf \int_0^\infty \beta ||\nu_0(t)||^2 + (1-\beta)||\nu_1(t)||^2 dt \quad (2.9)$$

subject to

$$\begin{aligned} \dot{x}_0 &= A_0 x_0 + B_0 v + M_0 \nu_0 && (2.10)\\ y &= C_0 x_0 + N_0 \nu_0 && (2.11)\\ \dot{x}_1 &= A_1 x_1 + B_1 v + M_1 \nu_1 && (2.12)\\ y &= C_1 x_1 + N_1 \nu_1 && (2.13)\\ & && (2.14) \end{aligned}$$

This problem can be expressed as follows:

$$J_v(\beta) = \inf \int_0^\infty \nu(t)^T V_\beta \nu(t) dt \quad (2.15)$$

subject to

$$E\dot{\xi} = F\xi + G\nu + Hv, \quad (2.16)$$

where

$$E = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & I \end{pmatrix}, \; F = \begin{pmatrix} A_0 & 0 & 0 \\ 0 & A_1 & 0 \\ C_0 & 0 & 0 \\ 0 & C_1 & 0 \end{pmatrix}, \quad (2.17)$$

$$G = \begin{pmatrix} M_0 & 0 \\ 0 & M_1 \\ N_0 & 0 \\ 0 & N_1 \end{pmatrix}, \; H = \begin{pmatrix} B_0 \\ B_1 \\ 0 \\ 0 \end{pmatrix}, \; V_\beta = \begin{pmatrix} \beta I & 0 \\ 0 & (1-\beta) I \end{pmatrix}, \quad (2.18)$$

and

$$\nu = \begin{pmatrix} \nu_0 \\ \nu_1 \end{pmatrix}, \; \xi = \begin{pmatrix} x_0 \\ x_1 \\ y \end{pmatrix}. \quad (2.19)$$

Constraints (2.16) can be simplified without affecting the solution of the optimization problem (2.15). For example, clearly y can be removed from the constraints by replacing (2.11) and (2.13) with their difference. These simplifications allow us to reduce the size of (2.16).

This type of simplification can be done systematically as described in the following Lemma.

Lemma 2.2 *There exist a full row rank matrix S and a full column rank matrix T such that the set of ν's satisfying*

$$SET\dot{\tilde{\xi}} = SFT\tilde{\xi} + SG\nu + SHv, \quad (2.20)$$

is identical to that of (2.16), for all v, where

- *SET has full column rank and*
- *$(sSET - SFT)$ has full column rank $\forall s$.*

Proof Let us first put the pencil $\{E, F\}$ in Kronecker form. As shown in [8], there exist orthogonal matrices Q and Z such that

$$Q(sE - F)Z = \begin{pmatrix} sE_\epsilon - F_\epsilon & * & * & * \\ 0 & sE_\infty - F_\infty & * & * \\ 0 & 0 & sE_f - F_f & * \\ 0 & 0 & 0 & sE_\eta - F_\eta \end{pmatrix} \quad (2.21)$$

where the the eigenmodes of the square pencils $\{E_f, F_f\}$ and $\{E_\infty, F_\infty\}$ are respectively the finite and infinite eigenmodes of $\{E, F\}$, $sE_\epsilon - F_\epsilon$ and $sE_\eta - F_\eta$ are respectively full row rank and full column rank, for all s, and E_ϵ and E_η are respectively full row rank and full column rank. Let

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = Z^T \xi. \tag{2.22}$$

Then (2.16) can be expressed as follows

$$\begin{pmatrix} E_\epsilon & * & * & * \\ 0 & E_\infty & * & * \\ 0 & 0 & E_f & * \\ 0 & 0 & 0 & E_\eta \end{pmatrix} \frac{d}{dt} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} - \begin{pmatrix} F_\epsilon & * & * & * \\ 0 & F_\infty & * & * \\ 0 & 0 & F_f & * \\ 0 & 0 & 0 & F_\eta \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{pmatrix} = \begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}. \tag{2.23}$$

It is straightforward to verify that there always exist ξ_1, ξ_2 and ξ_3 such that (2.23) is satisfied. This means that the top 3 equations in (2.23) do not impose any constraint on ν and v. We can thus take

$$S = \begin{pmatrix} 0 & 0 & 0 & I \end{pmatrix} Q, \quad T = Z \begin{pmatrix} 0 & 0 & 0 & I \end{pmatrix}^T, \tag{2.24}$$

and of course $\tilde{\xi} = \xi_4$. ■

It is straightforward to show that Lemma 2.2 can still be used even if we generalize the two models (2.1)-(2.2) by allowing for measured inputs (inputs which are measured online similar to the output y) and unknown disturbances.

3 Construction of the separability index

3.1 Characterization of $J^*(\beta)$

We have seen that the separability index γ^* can be constructed from $J^*(\beta)$ where

$$J^*(\beta) = \sup_{v \neq 0} \frac{\inf_\nu \int_0^\infty \nu(t)^T V_\beta \nu(t) dt}{\int_0^\infty \|v(t)\|^2 dt} \tag{3.1}$$

subject to constraint

$$E\dot{\xi} = F\xi + G\nu + Hv \tag{3.2}$$

The matrix V_β is positive-definite and, thanks to Lemma 2.2, we can assume that

$$E \text{ has full column rank} \tag{3.3}$$
$$\begin{pmatrix} sE - F \end{pmatrix} \text{ has full column rank}, \forall s \tag{3.4}$$
$$\begin{pmatrix} sE - F & G \end{pmatrix} \text{ has full row rank } \forall s = j\omega \tag{3.5}$$
$$\begin{pmatrix} E & G \end{pmatrix} \text{ has full row rank.} \tag{3.6}$$

Assumption (3.5) follows from (2.3), and (3.6) the full rankedness of N_i's.

The matrix pencil

$$\{\mathcal{E}, \mathcal{F}\} = \left\{ \begin{pmatrix} E & 0 \\ 0 & E^T \end{pmatrix}, \begin{pmatrix} F & GV_\beta^{-1}G^T - HH^T/\gamma^2 \\ 0 & -F^T \end{pmatrix} \right\} \tag{3.7}$$

plays a key role in the computation of $J^*(\beta)$.

Lemma 3.1 *For large enough γ, the pencil $\{\mathcal{E}, \mathcal{F}\}$*

- *is regular,*
- *has index 1,*
- *has no (finite) eigenmodes on the $j\omega$-axis.*

Proof It suffices to prove the three properties for the case $\gamma = \infty$. Without any loss of generality, we can suppose that

$$E = \begin{pmatrix} I \\ 0 \end{pmatrix} \quad , \quad F = \begin{pmatrix} A \\ C \end{pmatrix}, \tag{3.8}$$

$$GV_\beta^{-1}G^T = \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix} \quad , \quad H = \begin{pmatrix} H_1 \\ H_2 \end{pmatrix}. \tag{3.9}$$

This can be done by left multiplication of $\{\mathcal{E}, \mathcal{F}\}$ by an invertible matrix and a trivial coordinate change. Thanks to assumptions (3.3)-(3.6), it is easy to see that the pair (C, A) is observable, the pair (A, Q) is controllable on the $j\omega$-axis and R is positive definite.

The resulting pencil, with $\gamma = \infty$, can be expressed as follows

$$\{\mathcal{E}', \mathcal{F}'\} = \left\{ \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & I & 0 \end{pmatrix}, \begin{pmatrix} A & Q & 0 \\ C & 0 & R \\ 0 & -A^T & -C^T \end{pmatrix} \right\}. \tag{3.10}$$

One step application of the shuffle algorithm yields the E-matrix

$$\begin{pmatrix} I & 0 & 0 \\ C & 0 & R \\ 0 & I & 0 \end{pmatrix}$$

which clearly is invertible. Thus $\{\mathcal{E}, \mathcal{F}\}$ is regular has index 1.

By left and right multiplication by invertible matrices, we can transform the pencil $\{\mathcal{E}', \mathcal{F}'\}$ as follows

$$\{\mathcal{E}'', \mathcal{F}''\} = \left\{ \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} A & Q & 0 \\ C^T R^{-1} C & -A^T & 0 \\ 0 & 0 & R \end{pmatrix} \right\}. \tag{3.11}$$

Clearly the finite eigenmodes of $\{\mathcal{E}'', \mathcal{F}''\}$ are the eigenvalues of

$$\mathcal{H} = \begin{pmatrix} A & Q \\ C^T R^{-1} C & -A^T \end{pmatrix}.$$

But $\mathcal{H}$ is the Hamiltonian associated with a standard LQ problem, which thanks to observability of (C, A) and controllability of (A, Q) on the $j\omega$-axis, has no eigenvalue on the unit circle. ■

Theorem 3.1 *Let Γ denote the set of all γ satisfying the three conditions of Lemma 3.1. Then,*

$$J^*(\beta) = \inf_{\gamma \in \Gamma} \gamma^2. \tag{3.12}$$

Proof Let us suppose, without any loss of generality, that the system matrices are as in (3.8)-(3.9). Then the problem (2.15) can be expressed as follows

$$J_v(\beta) = \inf_{\mu} \int_0^\infty \mu^T Q \mu + (C\xi + H_2 v)^T R^{-1} (C\xi + H_2 v) dt \tag{3.13}$$

subject to constraints

$$\dot{\xi} = A\xi + Q\mu + H_1 v \tag{3.14}$$

This problem is a standard L_∞ problem. The solution to this problem is given by the unique L_2 solution to

$$\frac{d}{dt}\begin{pmatrix} \xi \\ \mu \end{pmatrix} = \begin{pmatrix} A & Q \\ C^T R^{-1} C & -A^T \end{pmatrix} \begin{pmatrix} \xi \\ \mu \end{pmatrix} + \begin{pmatrix} H_1 \\ C^T R^{-1} H_2 \end{pmatrix} v \tag{3.15}$$

So, if consider the system

$$\mathcal{S} : \left\{ \begin{array}{rcl} \dot{x} & = & \mathcal{A}x + \mathcal{B}v \\ z & = & \mathcal{C}x + \mathcal{D}v \end{array} \right. \tag{3.16}$$

where

$$\mathcal{A} = \begin{pmatrix} A & Q \\ C^T R^{-1} C & -A^T \end{pmatrix}, \mathcal{B} = \begin{pmatrix} H_1 \\ C^T R^{-1} H_2 \end{pmatrix},$$
$$\mathcal{C} = \begin{pmatrix} CR^{-\frac{1}{2}} & 0 \\ 0 & Q^{\frac{1}{2}} \end{pmatrix}, \mathcal{D} = \begin{pmatrix} H_2 R^{-\frac{1}{2}} \\ 0 \end{pmatrix} \tag{3.17}$$

we get that

$$J^*(\beta) = \sup_{v \neq 0} \frac{\int_0^\infty ||z(t)||^2 dt}{\int_0^\infty ||v(t)||^2 dt} = ||\mathcal{S}||_\infty^2. \tag{3.18}$$

But then $||\mathcal{S}||_\infty$ can be constructed from the Hamiltonian matrix

$$\mathcal{K}(\gamma) = \begin{pmatrix} \mathcal{A} - \mathcal{B}S^{-1}\mathcal{D}^T\mathcal{C} & -\gamma\mathcal{B}S^{-1}\mathcal{B}^T \\ \gamma\mathcal{C}^T T^{-1}\mathcal{C} & -\mathcal{A}^T + \mathcal{C}^T\mathcal{D}S^{-1}\mathcal{B}^T \end{pmatrix} \tag{3.19}$$

where

$$S = (\mathcal{D}^T\mathcal{D} - \gamma^2 I), \quad T = (\mathcal{D}\mathcal{D}^T - \gamma^2 I). \tag{3.20}$$

It is shown in particular in [1] that $\gamma > ||\mathcal{S}||_\infty$ if and only if $\gamma > \bar{\sigma}(\mathcal{D})$ and $\mathcal{K}(\gamma)$ has no eigenvalue on the $j\omega$-axis.

It is straightforward, but tedious, to show that the finite eigenvalues of $\{\mathcal{E}, \mathcal{F}\}$ are the union of the set of eigenvalues of $\mathcal{K}(\gamma)$ and that of $\mathcal{H}$. But, as we have seen before, $\mathcal{H}$ has no eigenvalues on the $j\omega$-axis. The proof of the theorem then follows from this, plus the fact that $\{\mathcal{E}, \mathcal{F}\}$ has index 1 if and only if $\gamma \neq \bar{\sigma}(\mathcal{D})$. ■

Lemma 3.2 *The pencil $\{\mathcal{E}, \mathcal{F}\}$ is regular and has index 1 if and only if*

$$\det(E^\perp (GV_\beta^{-1}G^T - \frac{1}{\gamma^2}HH^T){E^\perp}^T) \neq 0 \tag{3.21}$$

where $E^\perp$ is any highest rank left annihilator of E.

Theorem 3.2 *Let γ_l denote the largest value of γ for which (3.21) does not hold. If $\gamma \geq \gamma_l$ then*

- *if the pencil $\{\mathcal{E}, \mathcal{F}\}$ has no eigenmodes on the $j\omega$-axis, then $\gamma^2 > J^*(\beta)$,*
- *if the pencil $\{\mathcal{E}, \mathcal{F}\}$ has an eigenmode on the $j\omega$-axis, then $\gamma^2 \leq J^*(\beta)$.*

3.2 Computation of the separability index

The results of the previous section can be used to construct an algorithm based on the bisection method for the computation of $J^*(\beta)$, for any given β. Note that γ_l is a lower bound for $J^*(\beta)$ so Theorem 3.2 gives us the necessary test for implementing the algorithm.

$J^*(\beta)$ can then be used to compute the separability index, as seen previously:

$$\gamma^* = \sqrt{\max_{0 \leq \beta \leq 1} J^*(\beta)}. \tag{3.22}$$

Even though $J^*(\beta)$ is not concave, it has nice properties making the optimization problem (3.22) numerically tractable. In particular, thanks to Lemma 2.1, we can show that $J^*(\beta)$ is a sup over concave functions each of which is zero at $\beta = 0$ and $\beta = 1$. Using this fact, it is easy to show the following result.

Lemma 3.3 *Consider two scalars β_1 and β_2 satisfying $0 \leq \beta_1 < \beta_2 \leq 1$. Then*

$$\max_{\beta_1 \leq \beta \leq \beta_2} J^*(\beta) \leq \frac{J^*(\beta_1)J^*(\beta_2)}{J^*(\beta_1)(1-\beta_2) + J^*(\beta_2)\beta_1}. \tag{3.23}$$

The proof follows a straightforward geometric argument and is illustrated in Figure 3.2.

Now consider the following simple optimization strategy for estimating γ^* which consists of taking the square-root of the maximum of $J^*(\beta)$ for $n-1$ regularly spaced values of β over $[0, 1]$:

$$\hat{\gamma}^2 = \max_{k=1,..,n-1} J^*(k/n). \tag{3.24}$$

Then, thanks to Lemma 3.3, it is straightforward to show that

$$\frac{{\gamma^*}^2 - \hat{\gamma}^2}{\hat{\gamma}^2} \leq \frac{1}{n-1}. \tag{3.25}$$

This shows that we are not dealing with a difficult optimization problem. We can of course use more sophisticated algorithms to estimate γ^* and even couple the β and γ iterations.

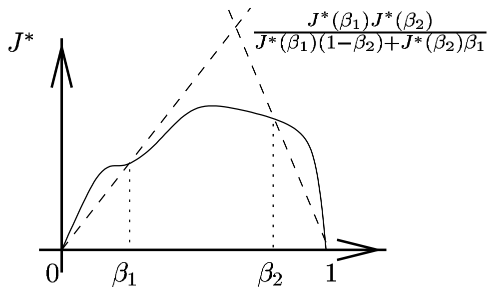

Figure 3.1: Since $J^*(\beta)$ is the sup over concave functions going through the points $(0,0)$ and $(1,0)$, it remains necessarily below the two dashed lines inside $[\beta_1, \beta_2]$ (and above outside).

References

[1] Bruinsma, N. A., Steinbuch, M., Fast computation of H_∞-norm of transfer function, *Systems & Control Letters*, 14, 1990, pp. 287-293.

[2] Campbell, S. L. and Meyer, C.D. Jr., *Generalized Inverses of Linear Transformations*, Dover, 1991.

[3] Kerestecioğlu, F., Change detection and input design in dynamical systems, *Research Studies Press*, Taunton, U.K., 1993.

[4] Kerestecioğlu, F. and Zarrop, M. B., Input design for detection of abrupt changes in dynamical systems, *Int. J. Control*, vol. 59, no. 4, 1994, pp. 1063-1084.

[5] Kullback, S., *Information Theory and Statistics*, John Wiley & Sons, 1971.

[6] Nikoukhah, R., Campbell, S. L. and Delebecque, F., Detection signal design for failure detection: a robust approach, *Internat. J. Adaptive Control and Signal Processing*, to appear.

[7] Uosaki, K, Tanaka, I., and Sugiyama, H., Optimal input design for autoregressive model discrimination with constrained output variance, *IEEE Trans. Auto. Contr.*, vol. AC-29, no. 4, 1984, pp. 348-350.

[8] Van Dooren, P., The computation of Kronecker's Canonical Form of a singular pencil, *Linear Algebra and its Applications*, vol. 27, pp. 103-140, 1979.

[9] Zhang, X. J., *Auxiliary signal design in fault detection and diagnosis*, Springer-Verlag, Heidelberg, 1989.