Technical Report PC-04-0220

Analysis of Emerson's Multiple Model Interpolation Estimation Algorithms: The MIMO Case

João P. Hespanha Dale E. Seborg

University of California, Santa Barbara

February 20, 2004

Analysis of Emerson's
Multiple Model Interpolation Estimation Algorithms:
The MIMO Case

Abstract

We extend to Multiple-Input/Multiple-Output (MIMO) processes the previous analysis of Emerson's Moving Multiple Model Interpolation (MMMI) algorithm for parameter estimation and compare it with standard least-squares estimators. We consider a MIMO ARX process model amenable to Model Predictive Control (MPC) sunthesis.

Our analysis proves that for MIMO ARX process models with affine unknown parameters, the only equilibrium point of Emerson's MMMI estimation algorithm is precisely the least-squares estimate of the parameters.

1 MIMO multiple-models interpolation

We consider a general MIMO ARX model for the process of the form

$$y(k) = \sum_{i=1}^{\tau_y} A_i(\theta_p) y(k-i) + \sum_{j=1}^{\tau_u} B_i(\theta_p) u(k-i) + n(k), \qquad \forall k \in \{1, 2, 3, \dots\}, \tag{1}$$

where $y(k) \in \mathbb{R}^{n_y}$ denotes the output, $u(k) \in \mathbb{R}^{n_u}$ the input, $n(k) \in \mathbb{R}^p$ output measurement noise, and $A_i(\theta_p)$, $B_i(\theta_p)$ matrices of coefficients that depends on an unknown parameter vector θ_p that belongs to a parameter set $\mathcal{P} \subset \mathbb{R}^n$. The process model (1), can be written in the *regression form*

$$y(k) = C(\theta_p)\varphi(k) + n(k), \qquad \forall k \in \{1, 2, 3, \dots\}, \tag{2}$$

where $\varphi(k)$ denotes the *regression matrix* defined by

$$\varphi(k) := \begin{bmatrix} y(k-1) \\ y(k-2) \\ \vdots \\ y(k-\tau_y) \\ u(k-1) \\ \vdots \\ u(k-\tau_u) \end{bmatrix} \in \mathbb{R}^{(n_y\tau_y + n_u\tau_u)\times 1}, \tag{3}$$

and

$$C(\theta_p) := \begin{bmatrix} A_1(\theta_p) & \cdots & A_{\tau_y}(\theta_p) & B_1(\theta_p) & \cdots & B_{\tau_u}(\theta_p) \end{bmatrix} \in \mathbb{R}^{n_y \times (n_y\tau_y + n_u\tau_u)}. \tag{4}$$

We assume that we have available a set of data

$$\Big\{ y(k), \varphi(k) : k = 1, 2, \dots, L \Big\} \tag{5}$$

collected over a time-window of length L. By "stacking" the outputs and regression vectors as follows

$$Y := \begin{bmatrix} y(1) & y(2) & \cdots & y(L) \end{bmatrix} \in \mathbb{R}^{n_y \times L}, \tag{6}$$
$$\Phi := \begin{bmatrix} \varphi(1) & \varphi(2) & \cdots & \varphi(L) \end{bmatrix} \in \mathbb{R}^{(n_y\tau_y + n_u\tau_u) \times L}, \tag{7}$$

we can write the process model as

$$Y = C(\theta_p)\Phi + N, \tag{8}$$

where

$$N := \begin{bmatrix} n(1) & n(2) & \cdots & n(L) \end{bmatrix} \in \mathbb{R}^{n_y \times L}, \tag{9}$$

is a matrix with measurement noise. The value of N is not available to estimate p.

The data-set is processed multiple times by examining its fits with respect to a finite bank of models that varies from iteration to iteration. We denote by

$$\mathcal{M}(i) := \{\theta_m^1(i), \theta_m^2(i), \dots, \theta_m^M(i)\} \subset \mathcal{P} \tag{10}$$

the values for the parameters for the bank of models used in the ith iteration. The estimate $Y_m^j(i)$ of Y based on the jth model during the ith iteration is defined by

$$Y_m^j(i) = C(\theta_m^j(i))\,\Phi, \qquad \forall j \in \{1, 2, \dots, M\}. \tag{11}$$

and the corresponding prediction error is given by

$$E_m^j(i) := Y_m^j(i) - Y, \qquad \forall j \in \{1, 2, \dots, M\}. \tag{12}$$

The *sum-of-squares error* $\mathrm{SSE}^j(i)$ for the jth model during the ith iteration is given by

$$\mathrm{SSE}^j(i) := \sum_{k=1}^{L} \|y_m^j(i)(k) - y(k)\|^2 = \mathrm{trace}\left[(Y_m^j(i) - Y)'(Y_m^j(i) - Y)\right], \tag{13}$$

and we define the corresponding *performance index* $J^j(i)$ by

$$J^j(i) := \frac{1}{\mathrm{SSE}^j(i)}. \tag{14}$$

Based on these definitions, we construct a *multiple model interpolation (MMI) estimator* by

$$\hat{\theta}_p(i) := \frac{\sum_{j=1}^{M} J^j(i)\theta_m^j(i)}{\sum_{j=1}^{M} J^j(i)}. \tag{15}$$

Example 1. For a one-step delay system with unknown gain $\theta_p \in [1, 10]$, we have

$$\varphi(k) := u(k-1), \qquad C(\theta_p) := \theta_p, \qquad \mathcal{P} := [1, 10] \tag{16}$$

leading to a model similar to the one considered in the August 8, 2003 report with $\tau = 1$:

$$y(k) = \theta_p u(k-1) + n(k), \qquad p \in [1, 10]. \tag{17}$$

Example 2. For a system with unknown gain $\theta_1 \in [1, 10]$ and unknown delay $\theta_2 \in \{1, 2, 3\}$, we would have

$$\varphi(k) := \begin{bmatrix} u(k-1) & u(k-2) & u(k-3) \end{bmatrix}', \quad C(\theta_1, \theta_2) := \begin{cases} \begin{bmatrix} \theta_1 & 0 & 0 \end{bmatrix}' & \theta_2 = 1 \\ \begin{bmatrix} 0 & \theta_1 & 0 \end{bmatrix}' & \theta_2 = 2 \\ \begin{bmatrix} 0 & 0 & \theta_1 \end{bmatrix}' & p_2 = 3 \end{cases}, \quad \mathcal{P} := [1, 10] \times \{1, 2, 3\} \tag{18}$$

leading to

$$y(k) = n(k) + \begin{cases} \theta_1 u(k-1) & \theta_2 = 1 \\ \theta_1 u(k-2) & \theta_2 = 2 \\ \theta_1 u(k-3) & \theta_2 = 3 \end{cases}, \qquad \theta_1 \in [1, 10]. \tag{19}$$

2 Moving multiple-models interpolation

The *moving multiple-models interpolation (MMMI) estimation algorithm* is defined as follows:

1. Set $i = 0$ (iteration index)
2. Set $\ell = 1$ (parameter index)
3. Compute the MMI estimate $\hat{\theta}_p(i)$ based on the family of model defined by the candidate parameters $\theta_m^j(i)$
4. Compute a new family of models by computing a new set of candidate parameters $\theta_m^j(i+1)$ *centered at the ℓth parameter* in $\hat{\theta}_p(i)$:

   $$\theta_m^j(i+1) = \hat{\theta}_p(i) + \Delta_\ell \big(j - \frac{M+1}{2}\big) e_\ell, \tag{20}$$

   where e_ℓ denotes the ℓth element of the canonical basis of $\mathbb{R}^n$
5. Increment i and ℓ (modulo n) and go to 3 until there is no significant change in $\hat{\theta}_p(i)$.

We are assuming here that the number of models M is odd and there is a constant spacing Δ_ℓ among the model values for the ℓth parameter.

According to the MMMI algorithm we obtain

$$\hat{\theta}_p(i+1) = \frac{\sum_{j=1}^{M} \theta_m^j(i+1) J^j(i+1)}{\sum_{j=1}^{M} J^j(i+1)} \tag{21}$$

$$= \frac{\sum_{j=1}^{M} \Big(\hat{\theta}_p(i) + \Delta_\ell \big(j - \frac{M+1}{2}\big) e_\ell\Big) J^j(i+1)}{\sum_{j=1}^{M} J^j(i+1)} \tag{22}$$

$$= \hat{\theta}_p(i) + \Delta_\ell \frac{\sum_{j=1}^{M} \big(j - \frac{M+1}{2}\big) J^j(i+1)}{\sum_{j=1}^{M} J^j(i+1)} e_\ell \tag{23}$$

where the performance indexes $J^j(i+1)$ are given by (14).

3 Analysis

We start with some preliminary results needed to analyze Emerson's MMMI algorithm.

3.1 Preliminaries

Lemma 1. *Given a positive semi-definite quadratic function $f : \mathbb{R}^n \to \mathbb{R}$, if there exists a vector $x_0 \in \mathbb{R}^n$, a basis $\{v_\ell : \ell = 1, 2, \ldots, n\}$ for $\mathbb{R}^n$, and positive scalars α_i, β_i, $i \in \{1, 2, \ldots, m\}$ such that*

$$\sum_{i=1}^{m} \alpha_i \big(f(x_0 + \beta_i v_\ell) - f(x_0 - \beta_i v_\ell) \big) = 0, \quad \forall \ell \in \{1, 2, \ldots, n\}, \tag{24}$$

then x_0 is a minimum of f.

Proof of Lemma 1. A generic positive semi-definite quadratic function is of the form

$$f(x) = (x - x^*)' Q (x - x^*) + c, \forall x \in \mathbb{R}^n, \tag{25}$$

where Q is a $n \times n$ positive semi-definite matrix, c a scalar, and x^* one of the minima of Q. Because of (24), we conclude that

$$\sum_{i=1}^{n} \alpha_i \big((x_0 + \beta_i v_\ell - x^*)' Q (x_0 + \beta_i v_\ell - x^*) - (x_0 - \beta_i v_\ell - x^*)' Q (x_0 - \beta_i v_\ell - x^*) \big) = 0, \quad \forall \ell \tag{26}$$

$$\Leftrightarrow \quad \sum_{i=1}^{n} \alpha_i \beta_i (x_0 - x^*)' Q v_\ell = \gamma (x_0 - x^*)' Q v_\ell = 0, \quad \forall \ell. \tag{27}$$

where $\gamma = \sum_{i=1}^{n} \alpha_i \beta_i > 0$ and the v_ℓ form a basis of $\mathbb{R}^n$, we conclude that

$$(x_0 - x^*)' Q = 0. \tag{28}$$

But then

$$f(x_0) = (x_0 - x^*)' Q (x_0 - x^*) + c = c = f(x^*), \tag{29}$$

which means that x_0 is also a minimum of f. ■

When the parametrization $C(\cdot)$ is **affine**[1], the sum-of-square errors for an arbitrary value $\theta \in \mathcal{P}$ of the parameter:

$$\text{SSE}(\theta) := \text{trace} \big[(C(\theta)\Phi - Y)'(C(\theta)\Phi - Y) \big], \tag{30}$$

is a quadratic function of θ. In this case, the following result is a consequence of Lemma 1:

Corollary 1. *When $C(\cdot)$ is affine, if there exists a vector $\theta_0 \in \mathbb{R}^n$, a basis $\{v_\ell : \ell = 1, 2, \ldots, n\}$ for $\mathbb{R}^n$, and positive scalars α_i, β_i, $i \in \{1, 2, \ldots, m\}$ such that*

$$\sum_{i=1}^{m} \alpha_i \big(\text{SSE}(\theta_0 + \beta_i v_\ell) - \text{SSE}(\theta_0 - \beta_i v_\ell) \big) = 0, \quad \forall \ell \in \{, 1, 2, \ldots, n\}, \tag{31}$$

then θ_0 is a least-squares estimate of θ.

[1] This implicitly assumes that $\mathcal{P}$ is the whole $\mathbb{R}^n$

3.2 Equilibrium

Assume that the MMMI algorithm converges to some value $\hat{\theta}_p$. Since $\hat{\theta}_p$ must be a fixed-point of (23) for every ℓ, we conclude that at equilibrium

$$\Delta_\ell \frac{\sum_{j=1}^{M} \big(j - \frac{M+1}{2}\big) J_\ell^j}{\sum_{j=1}^{M} J_\ell^j} e_\ell = 0 \quad \Leftrightarrow \quad \sum_{j=1}^{M} \big(j - \frac{M+1}{2}\big) J_\ell^j = 0, \qquad \forall \ell \in \{1, 2, \dots, n\}, \tag{32}$$

where J_ℓ^j denote the asymptotic value of $J^j(i)$ for the parameter index ℓ as $i \to \infty$, i.e.,

$$J_\ell^j = \lim_{i\to\infty} \frac{1}{\mathrm{SSE}^j(i)} = \lim_{i\to\infty} \frac{1}{\mathrm{SSE}\big(\theta_m^j(i)\big)} = \frac{1}{\mathrm{SSE}\big(\hat{\theta}_p + \Delta_\ell\big(j - \frac{M+1}{2}\big) e_\ell\big)}. \tag{33}$$

Separating the summation into terms with j smaller, equal, and larger than $\frac{M+1}{2}$, (32) can be re-written as

$$\sum_{j=1}^{\frac{M+1}{2}-1} \big(j - \frac{M+1}{2}\big) J_\ell^j + \big(\frac{M+1}{2} - \frac{M+1}{2}\big) J_\ell^j + \sum_{j=\frac{M+1}{2}+1}^{M} \big(j - \frac{M+1}{2}\big) J_\ell^j = 0, \tag{34}$$

$$\Leftrightarrow \quad \sum_{j=1}^{\frac{M+1}{2}-1} \big(j - \frac{M+1}{2}\big) J_\ell^j - \sum_{\bar{j}=1}^{\frac{M+1}{2}-1} \big(\frac{M+1}{2} - \bar{j}\big) J_\ell^{M+1-\bar{j}} = 0 \tag{35}$$

$$\Leftrightarrow \quad \sum_{j=1}^{\frac{M+1}{2}-1} \big(\frac{M+1}{2} - j\big)\big(J_\ell^j - J_\ell^{M+1-j}\big) = 0, \qquad \forall \ell. \tag{36}$$

Using (33), we further conclude that

$$\sum_{j=1}^{\frac{M+1}{2}-1} \big(\frac{M+1}{2} - j\big)\Big(\frac{1}{\mathrm{SSE}\big(\hat{\theta}_p + \Delta_\ell\big(j - \frac{M+1}{2}\big) e_\ell\big)} - \frac{1}{\mathrm{SSE}\big(\hat{\theta}_p - \Delta_\ell\big(j - \frac{M+1}{2}\big) e_\ell\big)}\Big) = 0, \qquad \forall \ell, \tag{37}$$

which can be written compactly as

$$\sum_{j=1}^{\frac{M+1}{2}-1} \alpha_j \Big(\mathrm{SSE}\big(\hat{\theta}_p + \beta_j\, e_\ell\big) - \mathrm{SSE}\big(\hat{\theta}_p - \beta_j\, e_\ell\big)\Big) = 0, \qquad \forall \ell. \tag{38}$$

where

$$\alpha_j := \frac{\frac{M+1}{2} - j}{\mathrm{SSE}\big(\hat{\theta}_p + \Delta_\ell\big(j - \frac{M+1}{2}\big) e_\ell\big)\ \mathrm{SSE}\big(\hat{\theta}_p - \Delta_\ell\big(j - \frac{M+1}{2}\big) e_\ell\big)} > 0, \quad \beta_j := \Delta_\ell\Big(\frac{M+1}{2} - j\Big) > 0. \tag{39}$$

Since $\{e_\ell : \ell = 1, 2, \dots, n\}$ is a basis of $\mathbb{R}^n$, the following result can be obtained from Corollary 1:

Theorem 1 (Equilibrium). *When $C(\cdot)$ is affine, any equilibrium point of the MMMI algorithm is a least-squares estimate of the parameter θ.*

4 Conclusions

The results in the report [1] were generalized for the MIMO case and with an arbitrary number of models. In particular, we showed that for MIMO ARX process models with affine unknown parameters, the only equilibrium point of Emerson's MMMI estimation algorithm is precisely the least-squares estimates of the parameters. The affine parameterization was used in the proofs but it does not seem necessary. In fact, it suffices that the sum-of-squares error function satisfy the condition that appears in Lemma 1 for the function f.

Future work include (i) the study of the convergence properties of the algorithm, (ii) the characterization of (non-affine) parameterizations for which the algorithm still converges to a least-squares estimate, and (iii) the characterization of (non-affine) parameterizations for which the algorithm converges to a value close to the least-squares estimate.

References

[1] J. P. Hespanha and D. E. Seborg. Analysis of Emerson's MMI estimation algorithm. Technical Report PC-03-0808, University of California, Santa Barbara, Aug. 2003.