Katholieke Universiteit Leuven

Departement Elektrotechniek ESAT-SISTA/TR 1994-11

Frequency weighted H_2 and Hilbert-Schmidt-Hankel model reduction. [1] [2]

Geert Schelfhout and Bart De Moor[3]

Published in the Proceedings of the 33rd IEEE Conference on Decision and Control, Dec. 14-16, Lake Buena Vista, Florida, pp. 3215-3216

[1]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/SISTA/schelfhout/reports/h2red.ps.Z*

[2]The following text presents research results obtained within the framework of the Belgian programme on interuniversity attraction poles (IUAP-17 and IUAP-50) initiated by the Belgian State - Prime Minister's Office - Science Policy Programming. The scientific responsibility is assumed by its authors.

[3]ESAT - Katholieke Universiteit Leuven,
Kardinaal Mercierlaan 94, 3001 Leuven (Heverlee), Belgium,
tel. 32/16/32.18.08, fax 32/16/32.19.86,
geert.schelfhout@esat.kuleuven.ac.be *bart.demoor@esat.kuleuven.ac.be*.

Frequency weighted H_2 and Hilbert-Schmidt-Hankel model reduction.

Geert Schelfhout*† Bart De Moor‡

ESAT–SISTA

Katholieke Universiteit Leuven

Kard. Mercierlaan 94

B-3001 Leuven

BELGIUM

Tel.: 00-32-16-32 18 08 FAX: 00-32-16-32 19 86

e-mail : geert.schelfhout@esat.kuleuven.ac.be bart.demoor@esat.kuleuven.ac.be

February 27, 1995

Keywords: H_2 norm, frequency weighting, inverse iteration

33rd CDC

Abstract

We consider the problem of approximating a discrete time SISO system of order n by one of order at most r ($r < n$), minimizing the frequency weighted H_2 norm of the modeling error.

Using Lagrange multipliers, we derive a set of necessary conditions, which are solved iteratively in a way inspired by the inverse power iteration algorithm for singular value decomposition.

*GS is a research assistant with the NFWO (Belgian National Fund for Scientific Research)

†The following text presents research results obtained within the framework of the Belgian programme on Interuniversity Attraction Poles (IUAP-17 and IUAP-50) initiated by the Belgian State - Prime Minister's Office - Science Policy Programming. The scientific responsibility is assumed by its authors.

‡BDM is a senior research associate with the NFWO

Furthermore, we consider the problem of optimal model reduction in the (frequency weighted) Hilbert-Schmidt-Hankel norm, and propose an algorithm similar to the one for the H_2 case.

1 Introduction and notation

The H_2 norm is a natural criterion for model reduction, and the H_2 optimal model reduction problem has received attention for some decades. Most approaches compute the gradient of the quadratic criterion w.r.t. the system parameters and use optimization techniques to find a local minimum. However, computing such a gradient is generally costly. We present a novel non-gradient algorithm.

The importance of frequency weighting for the applicability of a model reduction technique to controller design has been stressed by various authors [AM89, DS81]. Nevertheless, frequency weighted H_2 optimal model reduction algorithms have been dealt with much more rarely [ND76, H90]. We have included frequency weighting with a stable and minimum phase weighting function.

Some notation: (let $x(z) \triangleq \sum_{-\infty}^{\infty} x_i z^{-i}$)
$[x(z)]_i \triangleq x_i$: i-th Markov parameter of $x(z)$
$[x(z)]_+ \triangleq \sum_{i=1}^{\infty} x_i z^{-i}$: strictly proper part of $x(z)$
$\|x(z)\|_2 \triangleq \sqrt{\sum_{i=-\infty}^{\infty} h_i^2}$
$L_2 \triangleq \{x(z) | \; \|x(z)\|_2 < \infty\}$
$L_2^+ \triangleq \{x(z) \in L_2 | \; x(z) = [x(z)]_+\}$ (the set of strictly proper series $\in L_2$)
$L_2^{0+} \triangleq \{x(z) \in L_2 | \; z^{-1}x(z) \in L_2^+\}$
RL_2, RL_2^+, RL_2^{0+} : L_2, resp. L_2^+, L_2^{0+} restricted to (series converging to) rational functions

2 Problem statement and necessary conditions

Consider a discrete-time stable LTI system $h(z)$ of order n and a stable, invertible and minimum-phase weighting function $w(z)$. We want a system $\hat{h}(z) = \sum_{i=1}^{\infty} \hat{h}_i z^{-i}$ such that $\frac{1}{2}\sum_{i=1}^{\infty}([w \star h]_i - [w \star \hat{h}]_i)^2 = \frac{1}{4j\pi} \oint_T z^{-1} \left(h(z^{-1}) - \hat{h}(z^{-1})\right) w(z^{-1}) w(z) \left(h(z) - \hat{h}(z)\right) dz$ is minimized (T is the unit circle in the complex plane), and $\hat{h}(z)$ has order less or equal to r ($r < n$). For simplicity, we consider only systems $h(z)$ and $w(z)$ with distinct poles (and zeroes for $w(z)$) in this paper, even though the method can be extended to systems with multiple poles and zeroes.

The constraint on the order of $\hat{h}(z)$ is equivalent to the rank-deficiency of the $\infty \times (r+1)$ Hankel matrix $\hat{H}^\infty$ constructed from the Markov parameters $\hat{h}_i$, which is expressed through the existence of a nonzero null vector, i.e. $\exists y : \|y\| = 1$ and $\hat{H}^\infty y = 0$.

Using Lagrange multipliers l_i and λ, we may formulate (using somewhat informal notation) the constrained minimization problem as

$$\mathrm{ext}_{\hat{h},y,l,\lambda} \frac{1}{2} \sum_{i=1}^{\infty} ([w \star h]_i - [w \star \hat{h}]_i)^2 + l^T \hat{H}^\infty y + \lambda (y^T y - 1)$$

Defining $y(z) \triangleq \sum_{i=0}^{r} y_i z^{-i}$ and $l(z) \triangleq \sum_{i=1}^{\infty} l_i z^{-i}$, taking the derivatives of the Lagrangian w.r.t. $\hat{h}, l, y$ yields the following necessary conditions:

$\left[w(z^{-1})w(z)(h(z) - \hat{h}(z))\right]_+ = l(z)y(z)$, $\quad \left[\hat{h}(z)y(z^{-1})\right]_+ = 0$, $\quad$ and $\left[\hat{h}(z)l(z^{-1})\right]_{0..r} = 0$.

Using Lemma 1, we find the first condition to be equivalent to:

$$h(z) - \hat{h}(z) = w^{-1}(z) \left[w^{-1}(z^{-1}) l(z) y(z)\right]_+ \tag{1}$$

Now we can eliminate $\hat{h}(z)$, and find necessary conditions on $y(z)$ and $l(z)$:

$$\begin{aligned} \left[h(z)y(z^{-1})\right]_+ &= \left[w^{-1}(z) \left[w^{-1}(z^{-1}) l(z) y(z)\right]_+ y(z^{-1})\right]_+ \\ \left[h(z)l(z^{-1})\right]_{0..r} &= \left[w^{-1}(z) \left[w^{-1}(z^{-1}) l(z) y(z)\right]_+ l(z^{-1})\right]_{0..r} \end{aligned} \tag{2}$$

3 Solution scheme

In [D93a], a time-domain formulation for the unweighted H_2 approximation problem is dealt with, and an iterative scheme is proposed there inspired by the inverse power iteration for the singular value problem. Note that a solution of the necessary conditions need only be a stationary point; the number of solutions may grow combinatorially with the problem size. Therefore at least some heuristics will be necessary to pick a good candidate for the global minimum; these heuristics are exactly provided by this algorithm. We refer to [D93a] for the details.

By analogy, we propose an iteration to find a good solution to the necessary conditions (2). Each iteration step, the following equations are solved for $l^{[k]}(z)$ and $y^{[k]}(z)$:

$$\left[h(z)y^{[k]}(z^{-1})\right]_+ = \left[w^{-1}(z) \left[w^{-1}(z^{-1}) l^{[k]}(z) y^{[k-1]}(z)\right]_+ y^{[k-1]}(z^{-1})\right]_+ \tag{3}$$

$$\left[h(z)l^{[k]}(z^{-1})\right]_{0..r} = \left[w^{-1}(z) \left[w^{-1}(z^{-1}) l^{[k-1]}(z) y^{[k-1]}(z)\right]_+ l^{[k-1]}(z^{-1})\right]_{0..r} \tag{4}$$

We will show that $l^{[k]}(z)$ is rational, and derive its form; for this purpose, we introduce some more definitions:

$h(z) \triangleq \frac{b(z)}{a(z)}$
$w(z) \triangleq \frac{n_w(z)}{d_w(z)}$
$\hat{a}^{[k]}(z) \triangleq y^{[k-1]}(z^{-1})$ (i.e. a polynomial in z of degree r)
$\alpha_i \triangleq$ roots of $a(z)$ (= poles of $h(z)$)
$\hat{\alpha}_i^k \triangleq$ roots of $\hat{a}^{[k]}(z)$
$\alpha_{wi} \triangleq$ roots of $d_w(z)$ (= poles of $w(z)$)
$\beta_{wi} \triangleq$ roots of $n_w(z)$ (= zeroes of $w(z)$)
$a_i^k \triangleq$ roots of $a(z)\hat{a}^{[k]}(z)d_w(z)$
$\xi^{[k]}(z) \triangleq \left[w^{-1}(z^{-1})l^{[k]}(z)y^{[k-1]}(z)\right]_+$

As $\left[\frac{d_w(z)}{n_w(z)}\xi^{[k]}(z)\hat{a}^{[k-1]}(z)\right]_+ = \left[w^{-1}(z)\xi^{[k]}(z)y^{[k-1]}(z^{-1})\right]_+ = \left[h(z)y^{[k]}(z^{-1})\right]_+ = \left[\frac{b(z)}{a(z)}\hat{a}^{[k]}(z)\right]_+$, assuming that $\hat{a}^{[k-1]}(z)$ is Hurwitz (i.e. has only roots in the unit disk), we find

$$\xi^{[k]}(z) = \frac{n_w(z)\phi^{[k]}(z)}{a(z)\hat{a}^{[k-1]}(z)d_w(z)}, \text{ where } \deg \phi^{[k]}(z) = n + r - 1$$

since $\xi^{[k]}(z)$ is causal and can have no other poles than those of $h(z)$ and those that are canceled by $d_w(z)$ and $\hat{a}^{[k-1]}(z)$, and moreover the zeros of $\xi^{[k]}(z)$ must cancel the poles β_{wi}. The assumption that $\hat{a}^{[k-1]}(z)$ is Hurwitz can be violated, but will only occur in practice during the first few iterations and will not influence the ultimate convergence of the algorithm.
From the definition of $\xi^{[k]}(z) \triangleq \left[w^{-1}(z^{-1})l^{[k]}(z)z^{-r}\hat{a}_r^{[k-1]}(z)\right]_+$, one can find an expression for $l^{[k]}(z)$:

$$l^{[k]}(z) = \frac{z^r f^{[k]}(z)}{a(z)\hat{a}^{[k-1]}(z)d_w(z)}, \qquad \deg f^{[k]}(z) = n + w - 1$$

as $l^{[k]}(z)$ must cancel the r-multiple pole at $z = 0$ and have the same poles as $\xi^{[k]}(z)$.

Therefore we have the partial fraction forms $l^{[k]}(z) \triangleq \sum_{i=1}^{n} \frac{C_{\alpha i}^k}{z-\alpha_i} + \sum_{i=1}^{r} \frac{C_{\hat{\alpha} i}^k}{z-\hat{\alpha}_i^{[k-1]}} + \sum_{i=1}^{w} \frac{C_{wi}^k}{z-\alpha_{wi}}$ and $\xi^{[k]}(z) \triangleq \sum_{i=1}^{n} \frac{X_{\alpha i}^k}{z-\alpha_i} + \sum_{i=1}^{r} \frac{X_{\hat{\alpha} i}^k}{z-\hat{\alpha}_i^{[k-1]}} + \sum_{i=1}^{w} \frac{X_{wi}^k}{z-\alpha_{wi}}$. Interpolation in the poles in the expression defining $\xi^{[k]}(z)$ will yield a linear relation between each residue X_i^k and C_i^k, so that e.g. the C_i^k can be eliminated and the only unknowns remaining are the $n + r + w$ residues X_i^k and the $r+1$ coefficients of $\hat{a}^{[k]}(z)$. Interpolation in the poles α_i of $\left[h(z)\hat{a}^{[k]}(z)\right]_+ = \left[w^{-1}(z)\hat{a}^{[k-1]}(z)\xi^{[k]}(z)\right]_+$ yields n conditions on $\hat{a}_i^{[k]}$ and X_i^k, of which $r+1$ may be used to eliminate the $\hat{a}_i^{[k]}$. Expressing that $\xi^{[k]}(\beta_{wi}) = 0$ and that $l^{[k]}(z)$ has an r-multiple zero at $z = 0$ yields w conditions resp. r conditions on X_i^k, and finally $\left[h(z)l^{[k]}(z^{-1})\right]_s = \left[\left[w^{-1}(z^{-1})l^{[k-1]}(z)y^{[k-1]}(z)\right]_+ l^{[k-1]}(z^{-1})\right]_s$, $s = 0 \ldots r$, gives $r+1$ conditions on X_i^k, so there is a $(n+w+r)$-by-$(n+w+r)$ linear system to be solved each iteration.

The convergence of the algorithm is typically linear, as one could expect from the similarity with inverse power iteration (see Fig. 1 for typical convergence behavior of $\|y^{[k]} - y\|$). Once solutions $y(z)$ and $l(z)$ to the necessary conditions have been found, $\hat{h}(z)$ is computed using (1).

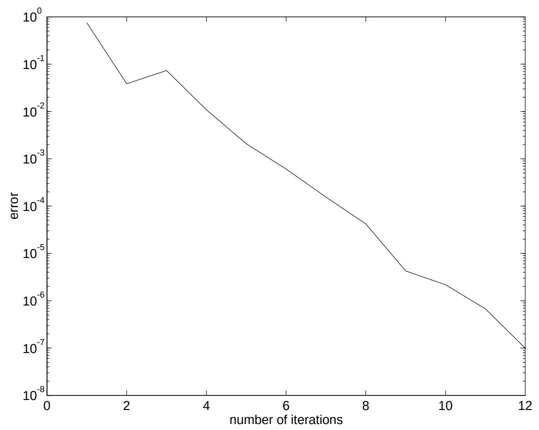

Figure 1: Convergence of $\|\hat{a}_k - \hat{a}\|$

4 Frequency weighted Hilbert-Schmidt-Hankel norm model reduction

The Hilbert-Schmidt-Hankel (HSH) norm is defined as

$$\|h(z)\|^2_{HSH} \triangleq \sum_i^{\infty} i h_i^2 = -\frac{1}{2\pi j}\oint_T h(z^{-1})\frac{d}{dz}h(z)dz$$

This norm has some interesting properties, e.g. it can be shown to be equivalent to the H_∞ norm for strictly proper systems (see [H92]).

Using similar manipulations as in the first section, we derive the following necessary conditions for the optimal HSH norm model reduction problem:

$$\begin{aligned}
\left[h(z)y(z^{-1})\right]_+ &= -\left[w^{-1}(z)y(z^{-1})\int z^{-1}\left[w^{-1}(z^{-1})l(z)y(z)\right]_+ dz\right]_+ \\
\left[h(z)l(z^{-1})\right]_{0\ldots r} &= -\left[w^{-1}(z)l(z^{-1})\int z^{-1}\left[w^{-1}(z^{-1})l(z)y(z)\right]_+ dz\right]_{0\ldots r}
\end{aligned}$$

where the integral is defined in such a way that it has no constant term, i.e. $\forall x(z) \in L_2^+$, $\int z^{-1}x(z)dz \triangleq \sum_{i=1}^{\infty} \frac{x_i}{i} z^{-i}$.

Analogously, we propose to solve the following equations for $y^{[k]}(z)$ and $l^{[k]}(z)$ each iteration:

$$\begin{aligned}
\left[h(z)y^{[k]}(z^{-1})\right]_+ &= -\left[w^{-1}(z)y^{[k-1]}(z^{-1})\int z^{-1}\left[w^{-1}(z^{-1})l^{[k]}(z)y^{[k-1]}(z)\right]_+ dz\right]_+ \\
\left[h(z)l^{[k]}(z^{-1})\right]_{0\ldots r} &= -\left[w^{-1}(z)l^{[k-1]}(z^{-1})\int z^{-1}\left[w^{-1}(z^{-1})l^{[k-1]}(z)y^{[k-1]}(z)\right]_+ dz\right]_{0\ldots r}
\end{aligned}$$

Defining $\xi^{[k]}(z) \triangleq - \int z^{-1} \left[w^{-1}(z^{-1})\hat{a}^{[k-1]}(z^{-1})l^{[k]}(z)\right]_+ dz$, then again

$$\xi^{[k]}(z) = \frac{n_w(z)\phi^{[k]}(z)}{a(z)\hat{a}^{[k-1]}(z)d_w(z)} = \sum_{i=1}^{n+r+w} \frac{X_i^k}{z - a_i^{k-1}}, \ \deg \phi^{[k]}(z) = n + r - 1$$

Using the same arguments as before, an expression for $l^{[k]}(z)$ is derived from:

$$\left[w^{-1}(z^{-1})\hat{a}^{[k-1]}(z^{-1})l^{[k]}(z)\right]_+ = -z\frac{d}{dz}\xi^{[k]}(z) = \sum_i \frac{zX_i}{(z-a_i^{k-1})^2} = \sum_i \frac{a_i^{k-1}X_i}{(z-a_i^{k-1})^2} + \sum_i \frac{X_i}{z-a_i^{k-1}}$$

$$\Rightarrow l^{[k]}(z) = \frac{z^r\psi^{[k](z)}}{(a(z)\hat{a}^{[k-1]}(z)d_w(z))^2} = \sum_i \frac{\gamma_i^k X_i^k}{(z-a_i^{k-1})^2} + \sum_i \frac{\delta_i^k X_i^k}{z-a_i^{k-1}}$$

The coefficients γ_i^k and δ_i^k are determined from interpolation in the poles a_i^{k-1}. Again, after elimination of the coefficients $\hat{a}_i^{[k]}$, $n+r+w$ linear conditions on the $n+r+w$ residues X_i^k are found. Only the computation of $-\left[w^{-1}(z)l^{[k-1]}(z^{-1}) \int z^{-1} \left[w^{-1}(z^{-1})l^{[k-1]}(z)y^{[k-1]}(z)\right]_+ dz\right]_{0\ldots r}$ is slightly more involved here, for which one needs lemma 2.

A Appendix : lemmas

Lemma 1

$$x(z) \in L_2^+,\ w(z),\ w^{-1}(z) \in RL_2^{0+} \ \Rightarrow \ \left[w^{-1}(z^{-1})\left[w(z^{-1})x(z)\right]_+\right]_+ = x(z),\ |z| > 1$$

Lemma 2

$$\left[w(z)l(z^{-1})\log|1 - az^{-1}|\right]_s = \int z^{-1}\left[z^{-s}w(z^{-1})l(z)\right]_+ dz|_{z=a^{-1}}$$

References

[AM89] ANDERSON, B.D.O., and J.B. MOORE, 1989, Optimal Control - Linear Quadratic Methods (Chapter 10). *Prentice-Hall*

[DVS93] DE MOOR, B., P. VAN OVERSCHEE, and G. SCHELFHOUT, 1993, Rational H_2 model reduction for SISO systems. *Proc. IFAC 1993, Sydney,* **II**, 227-230

[D93a] DE MOOR, B., 1993a, Structured total least squares and L_2 approximation problems. *Special Issue of Linear Algebra and its Applications on Numerical Linear Algebra in Systems & Control*, **188-189**, 163-207

[D93b] DE MOOR, B., 1993b, Total least squares for affinely structured matrices and the noisy realization problem. *To appear in IEEE Trans. Sign. Proc.*

[DS81] DOYLE, J.C., and G. STEIN, 1981, Multivariable feedback design: Concepts for a classical/modern synthesis. *IEEE Tr AC*, **26**, 4-16

[H90] HALEVI, Y., 1990, Frequency weighted model reduction via optimal projection. *Proc. 29th CDC, Honolulu, Hawaii*, 2906-2911

[H92] HANZON, B., 1992, The area enclosed by the oriented Nyquist diagram and the Hilbert-Schmidt-Hankel norm of a linear system. *IEEE Trans. Aut. Control*, **37**, 835-839

[ND76] NOLDUS, E., and M. DECOSTER, 1976, Reduction of linear systems by weighted frequency response matching. *Regelungstechnik*, 122-128