Some remarks on the Vector Fitting iteration

Wouter Hendrickx[1], Dirk Deschrijver, and Tom Dhaene

University of Antwerp, Department of Mathematics and Computer Science, Middelheimlaan 1, B-2020 Antwerpen, Belgium, wouter.hendrickx@ua.ac.be

Summary. *Vector Fitting* (VF) is an iterative technique to construct rational approximations based on multiple frequency domain samples, introduced by Gustavsen and Semlyen [1, 3]. VF is nowadays widely investigated and used in the *Power Systems* and *Microwave Engineering* communities. Numerical experiments show that VF has favorable convergence properties. However, so far, no theoretical proof for its convergence, or conditions to guarantee convergence, have been published. This paper gives a description of a general iterative Least-Squares framework for rational approximation and shows that VF fits into this framework.

Key words: Vector Fitting, Rational interpolation, System Identification, Least-Squares

1 Introduction

In *System Theory*, it is common practice to approximate the frequency domain response of a *Linear Time-Invariant system* (LTI system) by a rational pole-zero function. Finding such an approximation is inherently a difficult problem due to the non-linearity of the approximant. To remove the non-linearity, the denominator is often fixed at some well-chosen polynomial or the system is linearized in some way. Of course, this can degrade the quality of the approximant, or can even make accurate approximation impossible.

VF consists of an iterative pole relocation scheme. In each iteration step a linear Least-Squares (LS) problem is solved, to come up with more accurate approximations of numerator and denominator. New estimates of the poles are based on the approximations of the previous iteration.

In this contribution we position the VF technique in a broader LS rational approximation framework. This way, we want to facilitate further exploration of the theoretical properties of the VF technique. Furthermore, we offer some insight into the initial choice of pole locations of the VF algorithm. For com-

pleteness, we note that the iteration treated in this paper is related to the Sanathanan and Koerner iteration [2].

2 An iterative scheme for solving rational LS problems

Suppose we're trying to approximate a function H by a model of the form

$$\tilde{H}(s) = \frac{\sum_{i=1}^{N} \alpha_i f_i(s)}{\sum_{j=1}^{D} \beta_j g_j(s)} =: \frac{p(s,\alpha)}{q(s,\beta)} \tag{1}$$

where the f_i and g_j are fixed basis functions for the nominator and denominator respectively. Furthermore, α_i and β_j are unknown coefficients. To resolve the ambiguity in the definition, it's possible to choose $\alpha_N = 1$ for example. The p and q serve as an abbreviation, α and β are shorthands for the tuples $(\alpha_1, \ldots, \alpha_N)$ and $(\beta_1, \ldots, \beta_D)$ respectively.

Now suppose we have sampled H at certain points $(s_k)_{k=1}^n$. Our goal is to approximate H by a function of the form $\tilde{H}$ in a LS sense:

$$\text{argmin}_{\alpha,\beta} \sum_{k=1}^{n} \left| H(s_k) - \tilde{H}(s_k) \right|^2 \tag{2}$$

The problem with this formulation is that both numerator and denominator contain unknown variables α_i and β_j, so basic techniques for solving LS problems do not apply.

It is tempting to rewrite the LS problem as

$$\text{argmin}_{\alpha,\beta} \sum_{k=1}^{n} \left| \sum_{i=1}^{N} \alpha_i f_i(s_k) - H(s_k) \sum_{j=1}^{D} \beta_j g_j(s_k) \right|^2 \tag{3}$$

which is a simple linear LS problem of the form $\text{argmin}_x \|Ax - b\|_{l^2}$. Unfortunately this formulation is not equivalent with problem (2). Rewriting (2) gives:

$$\text{argmin}_{\alpha,\beta} \sum_{k=1}^{n} \frac{1}{|q(s_k,\beta)|^2} \; |p(s_k,\alpha) - H(s_k)\, q(s_k,\beta)|^2 \tag{4}$$

which resembles (3), except for the weighting factor $\dfrac{1}{|q(s_k,\beta)|^2}$.

The following iterative scheme can be applied: Start by setting $|q(s,\beta^{(0)})| = 1$. Calculate the sequences $\alpha^{(t)}$ and $\beta^{(t)}$ by iteratively solving

$$\text{argmin}_{\alpha^{(t)},\beta^{(t)}} \sum_{k=1}^{n} \frac{1}{\left|q(s_k,\beta^{(t-1)})\right|^2} \; \left|p(s_k,\alpha^{(t)}) - H(s_k)\, q(s_k,\beta^{(t)})\right|^2 \tag{5}$$

(which is a basic LS problem in $\alpha^{(t)}$ and $\beta^{(t)}$) for $t = 1, 2, \ldots$ Note that the weighting factor is approximated by the denominator from the last iteration.

3 The Vector Fitting methodology

In this section we will repeat the classical formulation of the VF methodology. Suppose we want to approximate the function $f : \mathbb{C} \to \mathbb{C}$ by a rational function and that f is known at a fixed set of sample points $(s_k)_{k=1}^n$. Now take an arbitrary function $\sigma : \mathbb{C} \to \mathbb{C}$ and assume that both $\sigma(s)f(s)$ and $\sigma(s)$ can be approximated by rational functions using *the same set of poles* $(\overline{a}_i)_{i=1}^D$ (and linear and constant terms). Formally, we have:

$$\begin{pmatrix} f(s)\,\sigma(s) \\ \\ \sigma(s) \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{D} \dfrac{c_i}{s-\overline{a}_i} + e + h\,s \\ \\ \sum_{i=1}^{D} \dfrac{\tilde{c}_i}{s-\overline{a}_i} + 1 \end{pmatrix} \tag{6}$$

We can now multiply the second row by $f(s)$ and evaluate the system in each of the samples s_k. If we assume the poles of $\overline{a}_i$ are fixed beforehand, we get a system of linear equations in the unknowns $(c_i, \tilde{c}_i, e, h)$ by equating the first row with the second row. This system is overdetermined if a lot of samples are available. In that case it can be solved using classical LS techniques.

We now proceed by writing both $\sigma(s)\,f(s)$ and $\sigma(s)$ in function of their zeros and poles:

$$\begin{pmatrix} f(s)\,\sigma(s) \\ \sigma(s) \end{pmatrix} = \begin{pmatrix} \dfrac{\prod_{i=1}^{D+1}(s-z_i)}{\prod_{i=1}^{D}(s-\overline{a}_i)} \\ \\ \dfrac{\prod_{j=1}^{D}(s-\tilde{z}_j)}{\prod_{i=1}^{D}(s-\overline{a}_i)} \end{pmatrix} \tag{7}$$

Dividing the first row by the second, we get an approximation for f of the form:

$$f(s) = \frac{\prod_{i=1}^{D+1}(s-z_i)}{\prod_{j=1}^{D}(s-\tilde{z}_j)} \tag{8}$$

Note that the zeroes of σ became the poles of our approximation.

The above procedure can be applied in an iterative fashion: the poles found in the last iteration can be inserted in equation (7) as *guesses* for the actual poles $(\overline{a}_i)_{i=1}^D$. Eventually, we want this procedure to converge. By this we mean that the guessed poles $\overline{a}_i$ become close enough to the real poles of f. In that case σ will be approximately 1 and we have found an approximation for f.

One problem that remains is the choice of the initial pole locations. This problem will be addressed in the section 5.

4 How VF fits in

Fix D (the degree of the denominator), fix $(a_i)_{i=1}^D$ (the *starting poles*) and set $N = D + 2$. Now choose the following basis for the first iteration:

$$f_i(s) = g_i(s) = \frac{1}{s - a_i} \quad \text{for } i = 1, \ldots, D$$
$$f_{D+1}(s) = g_{D+1}(s) = 1 \quad f_{D+2}(s) = s \tag{9}$$

First note the following:

$$\text{span}\,\langle f_1, \ldots, f_{D+2} \rangle = \frac{\mathbb{C}_{D+1}[s]}{\prod_{i=1}^{D}(s - a_i)} \tag{10}$$

$$\text{span}\,\langle g_1, \ldots, g_{D+1} \rangle = \frac{\mathbb{C}_{D}[s]}{\prod_{i=1}^{D}(s - a_i)} \tag{11}$$

where $\mathbb{C}_k[s]$ denotes the polynomials in s of degree equal or less than k. If a polynomial $p(s) \in \mathbb{C}_k[s]$, it's always possible to factor p completely, i.e. $p(s) = \prod_{i=1}^{k}(s - x_i)$ for certain $x_i \in \mathbb{C}$.

Now using the basis functions specified in (9) we can proceed with the iterative method proposed in section 2. The first iteration produces two sets of coefficients $\alpha_i^{(1)}$ and $\beta_i^{(1)}$. Using (10) and (11) we can rewrite p and q as

$$p(s, \alpha^{(1)}) = \frac{\prod_{i=1}^{D+1}(s - z_i^{(p,1)})}{\prod_{i=1}^{D}(s - a_i)} \qquad q(s, \beta^{(1)}) = \frac{\prod_{j=1}^{D}(s - z_j^{(q,1)})}{\prod_{i=1}^{D}(s - a_i)} \tag{12}$$

The second iteration (using $q(s, \beta^{(1)})$ as a weighting factor, as in section 2) produces $p(s, \alpha^{(2)})$ and $q(s, \beta^{(2)})$. (12) can be applied to these new functions (just replace the 1's by 2's). At first sight this does not really resemble the vector fitting methodology. Rewriting the defining equation (5) of the iterative scheme shows the following:

$$\sum_{k=1}^{n} \left| \frac{\prod_{i=1}^{D}(s_k - a_i)}{\prod_{j=1}^{D}(s_k - z_j^{(q,1)})} \right|^2 \left| \frac{\prod_{i=1}^{D+1}(s_k - z_i^{(p,2)})}{\prod_{i=1}^{D}(s_k - a_i)} - H(s_k) \frac{\prod_{j=1}^{D}(s_k - z_j^{(q,2)})}{\prod_{i=1}^{D}(s_k - a_i)} \right|^2 \tag{13}$$

which simplifies to

$$\sum_{k=1}^{n} \left| \frac{\prod_{i=1}^{D+1}(s_k - z_i^{(p,2)})}{\prod_{j=1}^{D}(s_k - z_j^{(q,1)})} - H(s_k) \frac{\prod_{j=1}^{D}(s_k - z_j^{(q,2)})}{\prod_{j=1}^{D}(s_k - z_j^{(q,1)})} \right|^2 \tag{14}$$

Using (10) and (11) with a_i replaced by $z_j^{(q,1)}$, we see that the LS problem we solve in the second iteration is exactly that solved in the vector fitting technique:

$$\sum_{k=1}^{n} \left| \sum_{i=1}^{D} \frac{c_i}{(s_k - z_i^{(q,1)})} + e + h\, s_k - H(s_k) \sum_{j=1}^{D} \frac{d_j}{(s_k - z_j^{(q,1)})} - H(s_k)\, \gamma \right|^2$$

where γ is chosen 1.

5 Initial pole placement

In order to get a system of linear equations that's not too badly conditioned, it's important to choose the initial poles at *good locations*. As all samples lie on the complex axis, choosing poles too far to the left in the complex plane makes the real part of the poles dominate the matrix entries

$$\frac{1}{s_k - \overline{a}_i} = \frac{1}{j(\omega_k - \mathcal{I}\overline{a}_i) - \mathcal{R}\overline{a}_i} \approx \frac{-1}{\mathcal{R}\overline{a}_k}$$

which renders all entries equally small.

Ideally, one would like to put the initial poles close to some of the sample points. Doing so makes some of the elements in the linear systems matrix significantly larger than all the other elements. This improves conditioning of the system. Of course the limiting case, where we let the poles coincide with some sample points, produces a matrix with some elements infinitely large and all others zero. In that case all information would be lost.

Therefore, we suggest to place poles on a line, parallel and close to the imaginary axis in order to get good conditioning. Originally, pole placement on a line through the origin was suggested [1]. To our experience this gives similar results.

The VF methodology also introduces the flipping of the poles around the imaginary axis between each two iterations in order to obtain a model which has all its poles in the left half-plane. In the context of system identification, this means that the modeled system is stable. Flipping a pole to the left half plane is equivalent to multiplying the approximant by the all-pass function

$$F(s) = \frac{s - p}{s - (-\mathcal{R}p + \mathcal{I}p)} \qquad |F(j\omega)|^2 = \frac{|\omega - \mathcal{I}p|^2 + |\mathcal{R}p|^2}{|\omega - \mathcal{I}p|^2 + |-\mathcal{R}p|^2} = 1$$

where p is a pole. This means that the amplitude of the system remains the same, only the phase changes.

References

1. Bjorn Gustavsen and Adam Semlyen. Rational approximation of frequency domain responses by Vector Fitting. *IEEE transactions on power delivery*, 14(3):1052–1061, 1999.
2. C. K. Sanathanan and J. Koerner. Transfer function synthesis as a ratio of two complex polynomials. *IEEE trans. Automat. Contr.*, 8:56–58, 1963.
3. Adam Semlyen and Bjorn Gustavsen. Vector Fitting by pole relocation for the state equation approximation of nonrational transfer matrices. *Circuits and systems: Analog and Digital signal processing*, 19(6):549–566, 2000.