Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

6.242, Fall 2004: MODEL REDUCTION *

Model reduction via moments matching[1]

This lecture investigates the use of interpolation, or *moments matching*, for model reduction.

7.1 Mathematics of moments matching

This section contains basic definitions and abstract algebraic results associated with moments matching.

7.1.1 Moments of analytical functions

Recall that a complex-valued function $f: \Omega \mapsto \mathbf{C}$ defined on an open subset Ω of the complex plane is called *analytical* if it can be represented by the expansion

$$f(s) = \sum_{i=0}^{\infty} f_i (s - s_0)^i, \tag{7.1}$$

exponentially converging in a neigborhood of every point $s_0 \in \Omega$. The number

$$f_i = M_i = M_i(s_0) = M_i^f(s_0)$$

is called the i-th moment of f at s_0. It is easy to see that

$$M_i^f(s_0) = (i!)^{-1} f^{(i)}(s_0)$$

[1]Version of October 13, 2004

is a scaled i-th derivative of f at s_0.

Rational functions are analytical on the complement of the set of their poles. The i-th moment of $f(s) = C(sI - A)^{-1}B$ at $s = s_0$ equals $C(s_0I - A)^{-i-1}B$. Elementary functions, such as $f(s) = \exp(s)$, $f(s) = \sin(s)$, $f(s) = \log(s)$ etc., are analytical on the open sets on which they can be defined as continuous functions (so that $f(s) = \sqrt{s(s-1)}$ cannot be made analytical on $\mathbf{C}$, but, with a right definition, is analytical on $\mathbf{C}$ without the real axis interval $[0, 1]$. Compositions of analytical functions are analytical. Not every "simple" continuous function is analytical: for example, $f(s) = \mathrm{Re}(s)$ is not.

7.1.2 Moments matching as a model reduction algorithm

One common formulation of the moments matching problem can be introduced as follows: given a positive integer n, a sequence of complex numbers $(s_k)_{k=1}^d$, a sequence of positive integers $(m_k)_{k=1}^d$, and a function f which is analytical in a neighorhood of points $(s_k)_{k=1}^d$, find a strictly proper real rational function $\hat{f}(s) = p(s)/q(s)$ of degree n with no poles at $(s_k)_{k=1}^d$, such that

$$\hat{f}^{(i)}(s_k) = f^{(i)}(s_k) \quad \text{for} \quad 1 \le k \le d, \quad 0 \le i < m_k. \tag{7.2}$$

Since a strictly proper transfer function of order n is defined by $2n$ independent real parameters, it will be natural to assume that

$$2n = \sum_{k=1}^{d} m_k, \tag{7.3}$$

to make the number of parameters equal to the number of equations.

When f is a transfers function of a large (or infinite) order system, the solution $\hat{f}$ of the moments matching problem, with $s_k = \pm j\omega_k$ chosen on the imaginary axis, is frequently used as a reduced order model of f. This leads to computationally inexpensive algorithms which provide high accuracy in the frequency regions which are located near the matching points ω_k. On the other hand, the approximation quality tends to be poor away from ω_k. Worse, $\hat{f}$ can be unstable when f is stable.

In studying moments matching as a method od model reduction, this lecture will concentrate on the following questions.

(a) Which conditions guarantee existence, uniqueness and continuous dependence of $\hat{f}$ on f?

(b) What are numerically robust ways of calculating $\hat{f}$?

(c) Which general accuracy guarantees can be proven for $\hat{f}$ as an approximation of f?

Question (a) can be answered in algebraic terms. A projection method based on Theorem 4.2 (Lecture 4) delivers numerically robust calculations for $\hat{f}$ which avoids direct calculation of the moments of f (those moments can be very large while the final answer may be scaled well). Most results related to question (c) will be negaite examples.

7.1.3 Moments matching problem as a system of linear equations

Assume for simplicity that $m_k = 1$ for all k (while $s_k \neq s_i$ for $k \neq i$), and hence $d = 2n$. Then the moments matching condition can be written simply as

$$\frac{p(s_k)}{q(s_k)} = f(s_k),$$

or, equivalently (since $q(s_k) \neq 0$),

$$p(s_k) = f(s_k)q(s_k).$$

The last equation is linear with respect to the coefficients of p, q. Representing p, q as

$$p(s) = \sum_{i=0}^{n-1} p_i s^i, \quad q(s) = s^n + \sum_{i=0}^{n-1} q_i s^i,$$

the equations can be written as $az = b$, where

$$a = \begin{bmatrix} 1 & s_1 & \dots & s_1^{n-1} & h_1 & h_1 s_1 & \dots & h_1 s_1^{n-1} \\ 1 & s_2 & \dots & s_2^{n-1} & h_2 & h_2 s_2 & \dots & h_2 s_2^{n-1} \\ & & & \vdots & & & & \vdots \\ 1 & s_d & \dots & s_d^{n-1} & h_d & h_d s_d & \dots & h_d s_d^{n-1} \end{bmatrix}, \quad b = -\begin{bmatrix} h_1 s_1^n \\ h_2 s_2^n \\ \vdots \\ h_d s_d^n \end{bmatrix},$$

$h_k = -f(s_k)$.

Thus, the original problem is reduced to solving a linear system with $2n$ variables and $2n$ unknowns. After a solution z of $az = b$ is found, it can be converted into the polynomials p and q. However, these p, q will solve the original matching problem only if $q(s_k) \neq 0$ for all k, which is not quaranteed automatically under the assumptions made.

For example, when $n = 1$, $d = 2$, $s_1 = 0$, $s_2 = 1$, $f(s_1) = 0$, $f(s_2) = 1$, the resulting linear equations take the form

$$p_0 = 0, \quad p_0 = q_0 + 1,$$

which has a unique solution $p_0 = 0$, $q_0 = -1$. This solution of the system of linear equations does not lead, however, to a solution of the moments matching problem, since the resulting polynomial $q(s) = s - 1$ is zero at $s = s_2 = 1$.

A similar reduction to a system of linear equations can be made when $m_k > 1$ for some k.

7.1.4 Moments matching with a fixed denominator

To understand the original moments matching problem better, it is beneficial to consider a different setup, in which a polynomial $\tilde{q}$, such that $\deg(\tilde{q}) = 2n$ and $\tilde{q}(s_k) \neq 0$ for all $k = 1, \dots, d$, is given, and one has to find a polynomial $\tilde{p}$ of degree not larger than $2n-1$ such that the first m_k moments of $\tilde{p}(s)/\tilde{q}(s)$ match the first m_k moments of $f(s)$ at $s = s_k$ for all $k = 1, \dots, d$. As follows from Theorem 7.1 below, such moments matching problem *always* has a unique solution $\tilde{p}$.

Theorem 7.1 *Let $s_1, \dots, s_d$ be* different *complex numbers. Let $m_1, \dots, m_d$ be positive integers. Let $f_{k,i}$, where the integer indexes k, i satisfy $1 \leq k \leq d$ and $0 \leq i < m_k$, be given complex numbers. Let $\tilde{q}$ be a polynomial of degree*

$$N = \sum_{k=1}^{d} m_k$$

such that $\tilde{q}(s_k) \neq 0$ for all $k = 1, \dots, d$. Then there exists a unique polynomial $\tilde{p}$ of degree $N - 1$ such that the i-th moment of $\tilde{f}(s) = \tilde{p}(s)/\tilde{q}(s)$ at $s = s_k$ equals $f_{k,i}$ for $1 \leq k \leq d$ and $0 \leq i < m_k$; Moreover, if, in addition, $\tilde{q}$ has real coefficients and for every $k \in \{1, \dots, d\}$ there exists $r \in \{1, \dots, d\}$ such that $s_k = \bar{s}_r$, $m_k = m_r$, and $f_{k,i} = \bar{f}_{r,i}$ for $0 \leq i < m_k$, then the polynomial $\tilde{p}$ has real coefficients.

Proof Consider the function H which maps the complex N-vector of coefficients of $\tilde{p}$ into the N-vector of the moments of $\tilde{f} = \tilde{p}/\tilde{q}$ at points s_k (a total of m_k moments at each s_k). This is a linear map $H: \mathbf{C}^N \mapsto \mathbf{C}^N$. Note that if all the moments are equal to zero then $\tilde{p}$ has a root of multiplicity m_k at each s_k, i.e. a total of N roots, which implies that $\tilde{p} = 0$, since the degree of $\tilde{p}$ is less than N. Hence H is a one-to-one map, and a (complex) solution of the moments matching problem exists and is unique. To show that, under the additional assumptions, $\tilde{p}$ has real coefficients, note that

$$\frac{p(s)}{q(s)} = O((s - s_k)^{m_k}) + \sum_{i=0}^{m_k - 1} f_{ki}(s - s_k)^i$$

implies

$$\frac{p^{\nabla}(s)}{q^{\nabla}(s)} = O((s-\bar{s}_k)^{m_k}) + \sum_{i=0}^{m_k-1} \bar{f}_{ki}(s-\bar{s}_k)^i,$$

where α^{∇} denotes the polynomial with the coefficients which are complex conjugates of the coefficients of polynomial α. Since $\tilde{q}$ has real coefficients, this means that the polynomial $\tilde{p}^{\nabla}$ is also a solution of the same moments matching problem. Since the solution is unique, $\tilde{p}$ must have real coefficients. ■

7.1.5 Homogeneous moments matching

For the original moments matching problem described in subsection 7.1.2, select a real polynomial $\tilde{q}$ of degree $2n$ such that $\tilde{q}(s_k) \neq 0$ for all $k = 1, \dots, d$, and find the polynomial $\tilde{p}$ which solves the fixed denominator moments matching problem as in Theorem 7.1 with f_{ki} being the i-th moment of f at s_k. Note that if transfer function f has real coefficient, so does the resulting polynomial $\tilde{q}$.

The first m_k moments of f and $\tilde{f}$ at s_k are identical. Hence, a strictly proper rational function $\hat{f} = p/q$ of degree n, such that $q(s_k) \neq 0$ for all k, solves the original moments matching problem if and only if the polynomial $\delta = p\tilde{q} - q\tilde{p}$ is divisible by

$$\theta = \theta(s) = \Pi_{k=1}^{d}(s-s_k)^{m_k}.$$

The task of finding polynomials polynomials p_0, q_0 such that

$$q_0 \not\equiv 0, \quad \deg(q_0) \leq n, \quad \deg(p_0) < n, \quad p_0\tilde{q} - q_0\tilde{p} \mathrel{\vdots} \theta, \tag{7.4}$$

where $\alpha \mathrel{\vdots} \beta$ means that division of polynomial α by polynomial β yields a zero reminder, is of an independent interest, and will be called the auxiliary *homogeneous* moments matching problem associated with the original setup.

It turns out that the homogeneous moments matching problem always has a unique solution.

Theorem 7.2 *Let θ be a polynomial of degree $2n$. Let $\tilde{q}$ be a polynomial which has no common roots with θ. Then the homogeneous moments matching problem (7.4) has a solution p_0, q_0. If θ, $\tilde{p}$ and $\tilde{q}$ have real coefficients then p_0, q_0 can also be chosen to be real. Moreover, if p_0^*, q_0^* is another such solution then*

$$f_0(s) = \frac{p_0(s)}{q_0(s)} = f_0^*(s) = \frac{p_0^*(s)}{q_0^*(s)}$$

for almost all s.

Proof Consider the linear map H from the $(2n+1)$-vector of coefficients of polynomials p, q, where $\deg(p) < n$ and $\deg(q) \leq n$, into the $(2n)$-vector of the coefficients of the reminder of the division of $p\tilde{q} - q\tilde{p}$ by θ. This is a linear map $H: \mathbf{C}^{2n+1} \mapsto \mathbf{C}^{2n}$, and hence $Hz = 0$ for some $z = (p_0, q_0) \neq 0$. Note that $q_0 \equiv 0$ would imply that $p_0\tilde{q}$ is divisible by θ, which is only possible when $q_0 \equiv 0$. Hence $q_0 \not\equiv 0$, which proves the *existence* of the desired p_0, q_0.

To show existence of a *real* solution, note that, when $\tilde{p}$ and $\tilde{q}$ are real, the polynomials defined by the real and imaginary parts of p_0, q_0 satisfy the conditions from (7.4), except, possibly, the first one. Since either real or imaginary part of q_0 is not identically zero, existence of a non-zero *real* solution of (7.4) follows.

Finally, if p_0^*, q_0^* is such solution then

$$\tilde{q}(p_0 q_0^* - q_0 p_0^*) \,\vdots\, \theta,$$

which implies $\delta = p_0 q_0^* - q_0 p_0^* = 0$ since $\deg(\delta) < 2n = \deg(\theta)$ and θ and $\tilde{q}$ have no common roots. ■

7.1.6 Existence and uniqueness of a moments matching solution

Existence and uniqueness of a solution in the original moments matching problem from subsection 7.1.2 can be established in terms of a solution $\tilde{p}, \tilde{q}$ of the homogeneous moments matching problem.

Theorem 7.3 *Let p_0, q_0 be the polynomials defined in Theorem 7.2.*

(a) If there exists $k \in \{1, \ldots, d\}$ such that s_k is a root of both q_0 and p_0 of multiplicity at least $r > 0$, but the multiplicity of s_k as a root of $\tilde{q}p_0 - \tilde{p}q_0$ is less that $r + m_k$, then the original moments matching problem has no solution;

(b) If $\deg(p_0) \geq \deg(q_0)$ then the original moments matching problem has no solution;

(c) If neither (a) nor (b) take place then the original moments matching problem has a solution. All such solutions are related to solutions of the system of $2n$ linear equations with $2n$ variables $x_1, \ldots, x_{2n}$, resulting from the polynomial relation

$$\tilde{q}p - \tilde{p}q \,\vdots\, \theta, \tag{7.5}$$

where

$$q(s) = s^n + \sum_{k=1}^{n} x_k s^{k-1}, \quad p(s) = \sum_{k=n+1}^{2n} x_k s^{n+1-k}.$$

The linear system of equations has a unique solution if and only if $\deg(q_0) = n$ *and* p_0, q_0 *have no common roots.*

Proof Let h, h_0 be the greatest common divisors of the pairs (p, q) and (p_0, q_0) respectively, normalized in such a way that the highest powers of s enter q/h and q_0/h_0 with coefficient 1. Then, since $p/q = p_0/q_0$ almost everywhere, $p/h = p_0/h_0$ and $q/h = q_0/h_0$.

To prove (a), note that $\deg(p) < \deg(q)$ implies $\deg(q_0) > \deg(p_0)$.

To prove (b), note that

$$(\tilde{q}p - \tilde{p}q)h_0 = (\tilde{q}p_0 - \tilde{p}q_0)h.$$

Since s_k is not a zero of h (because it is not a zero of q), the multiplicity of s_k as a root of $\tilde{q}p_0 - \tilde{p}q_0$ is at least $r + m_k$.

To prove the existence in (c), let $r = n - \deg(q_0/h_0)$, take a real number $\sigma > 0$ which does not belong to the set $\{-s_1, \dots, -s_d\}$, and define p, q by

$$p(s) = (s + \sigma)^r p_0(s)/h_0(s), \quad q(s) = (s + \sigma)^r (q_0(s)/h_0(s)).$$

Since there is a continuum of possible σ, the solution is not unique when $r > 0$. ■

7.2 Numerical algorithms for moments matching

This section describes a projection-based approach to numerically robust calculation of solutions of moments matching problems.

7.2.1 An example

While Theorem 7.3 provides important fundamental insight into the inner workings of moments matching, solving the linear system of $2n$ equations with $2n$ variables from (c) is usually not a viable option. The following example is aimed at demonstrating this.

Consider the task of finding an n-th strictly proper transfer function

$$\hat{f}(s) = \frac{p_0 + p_1 s + \dots + p_{n-1}s^{n-1}}{q_0 + q_1 s + \dots + q_{n-1}s^{n-1} + s^n}$$

which matches the first $2n$ moments of $f(s) = (1 + s)^N$ at $s = 0$, where $N \gg 2n$. The equations resulting from (7.5) will have the form $ax = b$, where

$$a = \begin{bmatrix} 1 & 0 & 0 & & 0 & -1 & 0 & 0 & & 0 \\ h_1 & 1 & 0 & & 0 & 0 & -1 & 0 & & 0 \\ h_2 & h_1 & 1 & & 0 & 0 & 0 & -1 & & 0 \\ & & & \ddots & & & & & \ddots & \\ h_{n-1} & h_{n-2} & h_{n-3} & & 1 & 0 & 0 & 0 & & -1 \\ h_n & h_{n-1} & h_{n-2} & & h_1 & 0 & 0 & 0 & & 0 \\ h_{n+1} & h_n & h_{n-1} & & h_2 & 0 & 0 & 0 & & 0 \\ h_{n+2} & h_{n+1} & h_n & & h_3 & 0 & 0 & 0 & & 0 \\ & & & \ddots & & & & & \ddots & \\ h_{2n-1} & h_{2n-2} & h_{2n-3} & & h_n & 0 & 0 & 0 & & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

where h_k are the binomial coefficients

$$(1+s)^N = 1 + h_1 s + h_2 s^2 + \dots .$$

According to Theorem 7.3, matrix a is invertible when $N \geq 2n-1$. However, the largest singular number of a is at least as large as h_{2n-1}, and the smallest singular number of a is not larger than 1. Hence the conditioning number of a (a numerical measure of how close a is to being singular) is at least

$$h_{2n-1} = \frac{N(N-1)(N-2)\cdot\dots\cdot(N-2n+2)}{1\cdot 2\cdot 3\cdot\dots\cdot(2n-1)},$$

which is very large when $N \gg n \gg 1$.

7.2.2 Krylov subspaces and the Arnoldi method

Theorem 4.2 from Lecture 4 can be used to obtain the solution of the moments matching problem more efficiently.

Let

$$f(s) = C(sI_N - A)^{-1}B$$

be a state space model of the original transfer function, where A is an N-by-N matrix. According to Theorem 4.2, if matrices U, V of dimensions n-by-N and N-by-n respectively are such that for every $k \in \{1, \dots, d\}$ the vectors $(s_k I - A)^{-i}B$ are linear combinations of the columns of V for $i = 1, \dots, i_k$, the row vectors $C(s_k I - A)^{-r}$ are linear combinations of the rows of U for $r = 1, \dots, r_k$, where $i_k + r_k \geq m_k$ and $UV = I_n$, then the transfer function

$$\hat{f}(s) = CV(sI_n - UAV)^{-1}UB$$

solves the moments matching problem.

Here the subspace of N-dimensional column vectors spanned by $(s_kI - A)^{-i}B$ with $k = 1, \ldots, d$, $i = 1, \ldots, i_k$, and the subspace of N-dimensional row vectors spanned by $C(s_kI - A)^{-r}$ with $k = 1, \ldots, d$, $r = 1, \ldots, r_k$, are called the *Krylov subspaces*. The moments matching problem can be solved. When the pair (A, B) is controllable, the pair (C, A) is observable, and

$$\sum_{k=1}^{d} i_k = \sum_{k=1}^{d} r_k = n \leq N,$$

the Krylov subspaces have dimension n. Hence, solving the moments matching problem reduces to finding matrices U_0, V_0 columns of which form bases in the Krylov subspaces, verifying that U_0V_0 is invertible, and then forming U, V by re-normalizing U and V, as in

$$U = (U_0V_0)^{-1}U_0, \quad V_0 = V.$$

In practice, forming the column vectors $(s_kI - A)^{-i}B$ with large i explicitly leads to poor conditioning of U_0V_0. Better results are achieved by applying the *Arnoldi method*, based on forming a recurrent sequence of vectors $B_1, B_2, \ldots,$ where $B_1 = (s_kI - A)^{-1}B$, and B_{i+1} is the normalized orthogonal complement of $(s_kI - A)^{-1}B_i$ to $B_1, \ldots, B_i$.