Order Bases

George Labahn

Symbolic Computation Group
David R Cheriton School of Computer Science
University of Waterloo, Canada

MOCAA M^3 Workshop 2008

Outline

Outline

Purpose

We give a bit of background for the following problem :

Find all solutions of rational approximation problems

- Examples : Padé, Hermite-Padé, Simultaneous Padé, . . . ,
- Extensions : matrix versions, rational interpolation versions, . . .

Padé Approximation

Simplest case : Given $A(z)$ find $U(z)$, $V(z)$ so that

$$a_0 + a_1 z + a_2 z^2 + \cdots = \frac{U(z)}{V(z)} + O(z^N)$$

- Typically linearized as $A(z)V(z) - U(z) = O(z^N)$
- Usually some relationship between degrees (m, n) and order ($m + n + 1$).
- Sometimes coefficients a_i not scalars.
- Structure of solutions known in scalar case (Padé table)

More General : Hermite-Padé Approximation

Given power series $A_0(z), \ldots, A_m(z)$ and integers $n_0, \ldots, n_m$

Find $P_0(z), ..., P_m(z)$ with deg $P_i(z) \leq n_i$ and

- $A_0(z)P_0(z) + \cdots + A_m(z)P_m(z) \approx 0$

  that is:
- $A_0(z)P_0(z) + \cdots + A_m(z)P_m(z) = r_0 z^{N+1} + r_1 z^{N+2} + \cdots$

  with $N = n_0 + \cdots + n_m + m$.

Case $m = 1$ gives Padé approximation.

Examples

Given power series $y(z)$ find $P_0(z), P_1(z), P_2(z)$ such that

- $P_0(z)y(z) + P_1(z)y'(z) + P_2(z)y^{(2)}(z) \approx 0$
- $P_0(z) \quad + \quad P_1(z)y(z) \quad + \quad P_2(z)y^2(z) \approx 0$

(generalized rational reconstruction)

Simultaneous Padé Approximants

Given power series $A_1(z), \ldots, A_m(z)$ and integers $n_0, \ldots, n_m$

find $U_1(z), ..., U_m(z)$ and $V(z)$ with some degree bounds :

$$A_1(z) \approx \frac{U_1(z)}{V(z)}, \quad A_2(z) \approx \frac{U_2(z)}{V(z)}, \quad \ldots, \quad A_m(z) \approx \frac{U_1(z)}{V(z)}.$$

Outline

Rational Approximation Problems

Given matrix $\mathbf{A}(z)$, orders, $\vec{\sigma} = (\sigma_1, \ldots, \sigma_m)$ find all solutions of

$$\mathbf{A}(z) \cdot \mathbf{Q}(z) = z^{\vec{\sigma}} \mathbf{R}(z).$$

Often also have degree constraints $\vec{n} = (n_1, \ldots, n_m)$ and want

$$\deg Q_i(z) \leq n_i.$$

$\mathbf{R}(z)$ called residual.

Applications

Rational approximation problems appear in:

- Transcendence of *e* and other famous numbers
- Inversion formulae for structured matrices
- Linear diophantine equations (and hence to GCDs)
- Guessing recurrence formulae (e.g. Gfun)
- Reconstruction of power series to polynomial problems (e.g. DFactor)
- Matrix normal forms (Popov, etc)
- Fast polynomial matrix arithmetic
- ...

Example

Let

$$\mathbf{A}(z) = \begin{bmatrix} \frac{1}{2} + z^2 - z^4 & 1 + \sin(z^2)^4 & \frac{1}{\sqrt{1+z^2}} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

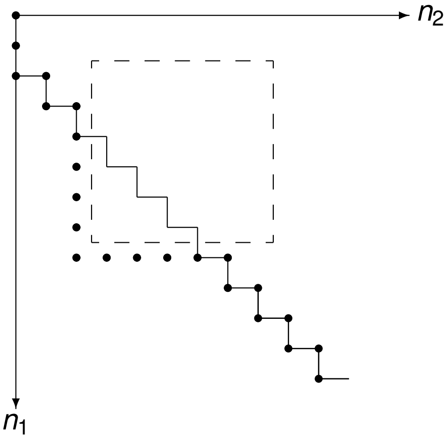
and $\vec{\sigma} = (z^8, 1, 1)$. Then a basis for all solutions given by

$$\mathbf{M}(z) = \begin{bmatrix} z^4 + \frac{11\,z^2}{2} & -\frac{10\,z^2}{19} + \frac{2}{19} & \frac{9\,z^2}{19} - \frac{11}{76} \\ -\frac{59\,z^2}{4} & z^2 - \frac{33}{19} & -\frac{5\,z^2}{4} + \frac{59}{152} \\ 12\,z^2 & \frac{32}{19} & z^2 - \frac{6}{19} \end{bmatrix}$$

with $\det \mathbf{M}(z) = z^8$. In this case the first 4 terms of the order residual $\mathbf{R}$ of $\mathbf{M}$ are given by

$$\mathbf{R}(z) = \begin{bmatrix} -\frac{19}{4} - \frac{367}{32}z^2 - \frac{189}{64}z^4 + O\left(z^6\right) & -\frac{97}{76} + \frac{89}{152}z^2 + O\left(z^4\right) & -\frac{13}{1216} - \frac{1093}{1216}z^2 + O\left(z^4\right) \\ -\frac{59\,z^2}{4} & -\frac{33}{19} + z^2 & \frac{59}{152} - \frac{5\,z^2}{4} \\ 12\,z^2 & \frac{32}{19} & -\frac{6}{19} + z^2 \end{bmatrix}.$$

Staircase path of computation in a Padé table:



Outline

Associated Linear System

$$A(z)V_n(z) - U_m(z) = z^{m+n+1}W(z)$$

$$(a_0 + a_1 z + \cdots)(v_0 + \cdots + v_n z^n) - (u_0 + \cdots + u_m z^m) = z^{m+n+1} w_0 + z^{m+n+2} w_1 + \cdots$$

Same as

$$\begin{bmatrix} a_{m-n+1} & \cdots & \cdots & a_{n-1} & a_n \\ a_{m-n+2} & \cdots & \cdots & a_n & a_{n+1} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ a_{n-1} & \cdots & \cdots & & a_{m+n-2} \\ a_n & \cdots & \cdots & a_{m+n-2} & a_{m+n-1} \end{bmatrix} \cdot \begin{bmatrix} v_n \\ v_{n-1} \\ \vdots \\ \vdots \\ v_2 \\ v_1 \end{bmatrix} = -v_0 \begin{bmatrix} a_{m+1} \\ a_{m+2} \\ \vdots \\ \vdots \\ a_{m+n-1} \\ a_{m+n} \end{bmatrix}$$

Similarly for a_i square matrices.
Similarly have structured linear system for other approx problems.
Nice when coefficient matrix is nonsingular.

Example : Hermite-Padé

$$a(x) \cdot p(x) + b(x) \cdot q(x) + c(x) \cdot r(x) = O(x^8)$$

with deg $p(x) \leq 2$, deg $q(x) \leq 3$, deg $r(x) \leq 1$

$$\left[\begin{array}{ccc|cccc|cc} a_0 & 0 & 0 & b_0 & 0 & 0 & 0 & c_0 & 0 \\ a_1 & a_0 & 0 & b_1 & b_0 & 0 & 0 & c_1 & c_0 \\ a_2 & a_1 & a_0 & b_2 & b_1 & b_0 & 0 & c_2 & c_1 \\ a_3 & a_2 & a_1 & b_3 & b_2 & b_1 & b_0 & c_3 & c_2 \\ a_4 & a_3 & a_2 & b_4 & b_3 & b_2 & b_1 & c_4 & c_3 \\ a_5 & a_4 & a_3 & b_5 & b_4 & b_3 & b_2 & c_5 & c_4 \\ a_6 & a_5 & a_4 & b_6 & b_5 & b_4 & b_3 & c_6 & c_5 \\ a_7 & a_6 & a_5 & b_7 & b_6 & b_5 & b_4 & c_7 & c_6 \end{array}\right] \cdot \left[\begin{array}{c} p_0 \\ p_1 \\ p_2 \\ \hline q_0 \\ q_1 \\ q_2 \\ q_3 \\ \hline r_0 \\ r_1 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right]$$

Note

- All Padé approximants known in scalar case
  - Padé table in scalar case has a type of block structure
- Padé approximants related to diophantine equations
  - There are algorithms corresponding to Euclidean algorithm
  - Fast way to compute Padé approximants in scalar case
- Nothing known about structure of matrix Padé case or Hermite-Padé case by 1990
- Use in inversion formulas for Hankel and Toeplitz matrices known in scalar and block cases.

Example : Inverses of Block Hankel Matrices

Inverse of block Hankel matrix, $H_{m,n}$, given by

$$\begin{bmatrix} v_{n-1} & \cdots & v_0 \\ \vdots & & \\ v_0 & & \end{bmatrix} \begin{bmatrix} q^*_{n-1} & \cdots & q^*_0 \\ & & \vdots \\ & \ddots & \\ & & q^*_{n-1} \end{bmatrix} - \begin{bmatrix} q_{n-2} & \cdots q_0 & 0 \\ \vdots & & \\ q_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} v^*_n & \cdots & v^*_1 \\ & \ddots & \vdots \\ & & v^*_n \end{bmatrix}$$

where

$$H_{m,n}[v_n, \cdots, v_0]^T = -[a_{m+1}, \cdots, a_{m+n+1}]^T, \quad H_{m,n}n[q_{n-1}, \cdots, q_0]^T = [0, \cdots, 0, 1]^T,$$

$$[v^*_n, \cdots, v^*_0]H_{m,n}n = -[a_{m+1}, \cdots, a_{m+n+1}], \quad [q^*_{n-1}, \cdots, q^*_0]H_{m,n}n = [0, \cdots, 0, 1].$$

Example : Inverses of Block Hankel Matrices

Inverse of block Hankel matrix, $H_{m,n}$, given by

$$\begin{bmatrix} v_{n-1} & \cdots & v_0 \\ \vdots & & \\ v_0 & & \end{bmatrix} \begin{bmatrix} q_{n-1}^* & \cdots & q_0^* \\ & & \vdots \\ & \ddots & \\ & & q_{n-1}^* \end{bmatrix} - \begin{bmatrix} q_{n-2} & \cdots q_0 & 0 \\ \vdots & & \\ q_0 & & \\ 0 & & \end{bmatrix} \begin{bmatrix} v_n^* & \cdots & v_1^* \\ & \ddots & \vdots \\ & & v_n^* \end{bmatrix}$$

where $Q_{n-1}(z)$, $V_n(z)$, $Q_{n-1}^*(z)$, $V_n^*(z)$

$$\begin{aligned} A(z)Q_{n-1}(z) - P_{m-1}(z) &= z^{m+n-1}R(z) \text{ with } R(0) = I \\ A(z)V_n(z) - U_m(z) &= z^{m+n+1}W(z) \text{ with } V_n(0) = I \\ Q_{n-1}^*(z)A(z) - P_{m-1}^*(z) &= z^{m+n-1}R^*(z) \text{ with } R^*(0) = I \\ V_n^*(z)A(z) - U_m^*(z) &= z^{m+n+1}W^*(z) \text{ with } V_n^*(0) = I \end{aligned}$$

Inversion Formulae : Fast, Stable, Numeric Algorithms

- Inversion formula determine condition number when properly scaled
- Allows one to build on stable subproblems
- Gives fast, provably numerically stable algorithms

Outline

People

- K. Mahler(1925-1969), J. Coates(1965) , J. Della Dora (1980),
  - (strong conditions always assumed)
- B. Beckermann and G. Labahn; A. Bultheel and M. van Barel
- B. Salvy and P. Zimmermann (gfun), (and H. Derksen) ; M. Rubey
- M. van Hoeij (use in differential factorization)
- G. Villard (matrix normal forms)
- P. Giorgi, C-P. Jeannerod, G. Villard (fast polynomial matrix arithmetic)
- H. Cheng; E. Schost et al;· · ·

Bernhard Beckermann

Order Bases

Idea : look at order condition independently of degree bounds,

$$R_\sigma = \{\mathbf{Q}(z) \in F^{(m)}[z] \ | \ \mathbf{A}(z) \cdot \mathbf{Q}(z) = O(z^\sigma)\}$$

Find basis of R_σ as a *module* over $F[z]$.

- Basis always has m elements
- Write as columns of an $m \times m$ matrix polynomial $\mathbf{M}(z)$.

Order Bases

Vector case ($s \times 1$). Convert to scalar series.

$$R_s = \{\mathbf{Q}(z) \in F^{(m)}[z] \mid \mathbf{A}(z) \cdot \mathbf{Q}(z^s) = O(z^\sigma)\}$$

Correct solutions handled via “s-trick.

$$A_1(z)Q_1(z^s) + \cdots + A_m(z)Q_m(z^s) = O(z^\sigma)$$

Conveniently converts vector problem to scalar problem.

Order Bases

Degree bounds? Given $\vec{n} = (n_1, \ldots, n_m)$:

Then

$$\mathbf{Q}(z) \;=\; \alpha_1(z)\mathbf{M}_1(z) + \cdots + \alpha_m(z)\mathbf{M}_m(z)$$

with

$$\deg \alpha_i(z) \leq \text{ defect } \mathbf{Q}(z) - \text{ defect } \mathbf{M}_i(z)$$

Here defect is a measure of the difference between degrees and bounds n_i.

Implies $\mathbf{M}(z)$ describes all solutions of $\mathbf{A}(z)\mathbf{Q}(z) = O(z^\sigma)$

Outline

Computation Complexity

Hermite-Padé approx. with degree bounds $(n_1, \ldots n_m)$

Set $N = n_1 + \cdots + n_m$. Then :

- Linear algebra : $O(N^3)$
- Sigma basis : $O(N \log^2 N)$ i.e. - $O(N^{1+\epsilon})$ in scalar case (BL - 1994)
- MBasis : $O(m^\omega N^{1+\epsilon})$ - in case of matrix input (GJV - 2003)
- Generating set : $O(m^\omega (N/m)^{1+\epsilon})$ (Stor. 2006)
- See following talk by Wei Zhou (2008)

Sigma Basis Algorithm [SIMAX - BL]

1. Start with Order basis $= \mathbf{I}$ and order $= 0$.
2. Of all the columns that need to have order increased:
   - pick one with minimal defect.
   - use to eliminate other columns needing order increase.
3. Multiply pivot column by z. Continue. Quadratic complexity.

Double order everytime : obtain superfast version.

Alternatively (GJV)

To get $(\sigma + 1)$-basis from a σ-basis do:

1. we compute the terms in z^σ in the residue $\mathbf{R}(z)$. This give us a matrix Δ
2. we compute a row echelon form of Δ
3. we apply some transformations according to the row echelon form. These transformations are of two types :
   - Either $\mathbf{M}_i$ is replaced by a linear combination of some $\mathbf{M}_j$
   - Or all the polynomials in $\mathbf{M}_i$ are multiplied by z

Example : calculation of a 1-basis

$$\mathbf{A}(z) = \begin{bmatrix} z+1 & z^2-1 & 5 \\ 2z-1 & 1 & z \end{bmatrix} \text{ and } \mathbf{M}(z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then: $\Delta = \begin{bmatrix} 1 & -1 & 5 \\ -1 & 1 & 0 \end{bmatrix} \xrightarrow{\text{row echelon form}} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

So we replace the 2^{nd} column by the 2^{nd} column minus the 1^{st} column, and we shift column 1 and 3 because they where used as pivots to obtain the 1-basis : $\mathbf{M}(z) = \begin{bmatrix} z & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & z \end{bmatrix}$

Outline

Fraction-free Computation

Fraction-free computation for given $\vec{n}$.

- set up linear system (structured Krylov matrix) for each order
- find so-called Cramer's solutions
- eliminate known divisors by a type of Sylvester's identity
- Order basis called Mahler system

Also version for Ore case. Also modular versions.

Goal: Try to find Cramer solutions

- e.g. Hermite-Padé problem

$$a(x) \cdot p(x) + b(x) \cdot q(x) + c(x) \cdot r(x) = O(x^6)$$

with deg $p(x) \leq 2$, deg $q(x) \leq 1$, deg $r(x) \leq 1$

$$\left[\begin{array}{ccc|cc|cc} a_0 & 0 & 0 & b_0 & 0 & c_0 & 0 \\ a_1 & a_0 & 0 & b_1 & b_0 & c_1 & c_0 \\ a_2 & a_1 & a_0 & b_2 & b_1 & c_2 & c_1 \\ a_3 & a_2 & a_1 & b_3 & b_2 & c_3 & c_2 \\ a_4 & a_3 & a_2 & b_4 & b_3 & c_4 & c_3 \\ a_5 & a_4 & a_3 & b_5 & b_4 & c_5 & c_4 \\ a_6 & a_5 & a_4 & b_6 & b_5 & c_6 & c_5 \end{array}\right] \cdot \left[\begin{array}{c} p_0 \\ p_1 \\ p_2 \\ \hline q_0 \\ q_1 \\ \hline r_0 \\ r_1 \end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ d \end{array}\right]$$

where d is determinant of coefficient matrix.

- Solution has determinant representation in nonsingular case:

$$\text{e.g. } p(z) = \det \left[\begin{array}{ccc|cccc|c} a_0 & 0 & 0 & b_0 & 0 & c_0 & 0 \\ a_1 & a_0 & 0 & b_1 & b_0 & c_1 & c_0 \\ a_2 & a_1 & a_0 & b_2 & b_1 & c_2 & c_1 \\ a_3 & a_2 & a_1 & b_3 & b_2 & c_3 & c_2 \\ a_4 & a_3 & a_2 & b_4 & b_3 & c_4 & c_3 \\ 1 & z & z^2 & 0 & 0 & 0 & 0 \end{array} \right]$$

- Unique in nonsingular case.
- Recursively build Cramer solutions from Cramer solutions of smaller problems along offdiagonal of associated table.

- Matrix $\mathbf{M}(z)$ of determinantal polynomials with degrees

$$\begin{bmatrix} n_1 & n_1 - 1 & \cdots & n_1 - 1 \\ n_2 - 1 & n_2 & \cdots & n_2 - 1 \\ \vdots & & \ddots & \vdots \\ n_m - 1 & \cdots & \cdots & n_m \end{bmatrix}$$

  and lcoeff of diagonal $=$ determinant of coeff matrix.
- Unique in nonsingular case
- Basic building block of recursions.
- Method 1: via modified Schur complements
  - nonsingular location to nonsingular location in table
  - similar to look ahead
- Method 2: via determinental identities
  - works in singular cases by computing at closest nonsingular locations (look around)

Outline

Shifted Popov Normal Forms

Shifted Popov form problem : gives

$$\mathbf{A}(z) \cdot \mathbf{U}(z) = \mathbf{P}(z)$$

Embed normal form problem inside part of a Mahler system for

$$[\mathbf{A}(z), -\mathbf{I}] \left[\begin{array}{cc} \mathbf{V}(z) & \mathbf{U}(z) \\ \mathbf{Q}(z) & \mathbf{P}(z) \end{array} \right] = [O(z^{\vec{n}}), 0].$$

Fraction-free computation of normal forms.

Outline

Currently we are working on:

- Fast algorithms in differential case
- Better fraction-free algorithms
- Use with differential-algebraic problems
  - Popov forms
  - Invariants and algorithms
- Alternate bases
- Multivariate Order Bases