Lecture 3

Order Bases

The Sigma Basis Algorithm

George Labahn

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

AEC Summer School : July 31, 2019

Outline

1. Power Hermite-Padé Approximants

2. Order and Defect

3. Sigma Basis Algorithm

4. Complexity

Preamble

In lecture 3 we take a different approach to compuing our equations of the form

$$a_1(z)p_1(z) + \cdots + a_m(z)p_m(z) = O(z^\sigma)$$

In this case we separate satisfying the order condition with also trying the satisfy the degree bounds. We also show how these same methods can be used to solve the case where the $a_i(z)$ are vectors of power series. We then present the sigma-basis algorithm, a contructive procedure for determining an order basis (also sometimes a sigma basis or a minimal approximant basis). We give both a simple algorithm quadratic in the order and a recursive algorithm which computes with quasi-linear complexity.
This work was done jointly with Bernhard Beckermann.

Recall from last day

- Extended Euclidean Algorithm vs Padé Approximation

Recall from last day

- Extended Euclidean Algorithm vs Padé Approximation
  - Implies superfast algorithm for Padé
  - Implies fraction-free algorithm for Padé
  - Implies superfast Hankel inversion

Recall from last day

- Extended Euclidean Algorithm vs Padé Approximation
  - Implies superfast algorithm for Padé
  - Implies fraction-free algorithm for Padé
  - Implies superfast Hankel inversion
- Padé problem $\equiv$ solving structured linear system
  - e.g. describes recursive computation

Recall from last day

- Extended Euclidean Algorithm vs Padé Approximation
  - Implies superfast algorithm for Padé
  - Implies fraction-free algorithm for Padé
  - Implies superfast Hankel inversion
- Padé problem $\equiv$ solving structured linear system
  - e.g. describes recursive computation
- Still not so precise for Hermite-Padé, matrix problems, etc

Goal of Today's Lecture

- Describe *all* well-known approx problems uniformly.
  - Include both scalar and matrix versions

Goal of Today's Lecture

- Describe *all* well-known approx problems uniformly.
  - Include both scalar and matrix versions
- Describe *all* solutions to such problems.

Goal of Today's Lecture

- Describe *all* well-known approx problems uniformly.
  - Include both scalar and matrix versions
- Describe *all* solutions to such problems.
- Provide algorithm to efficiently compute such solutions .
  - Uniformize to model Hermite-Padé approximants
  - Model algorithm on Hermite-Padé computation

Problem and Techniques

Recall our problem:

- Given $\mathbf{G}(z) \in \mathbb{K}^{s\times m}[[z]]$, some degree constraints and order σ find solutions to $\mathbf{G}(z) \cdot \mathbf{P}(z) = O(z^{\sigma})$ satisfying degree constraints.

We make use of two insights

Problem and Techniques

Recall our problem:

- Given $\mathbf{G}(z) \in \mathbb{K}^{s\times m}[[z]]$, some degree constraints and order σ find solutions to $\mathbf{G}(z)\cdot\mathbf{P}(z) = O(z^{\sigma})$ satisfying degree constraints.

We make use of two insights

Insight I:

- Treat order part and degree parts separately.

Problem and Techniques

Recall our problem:

- Given $\mathbf{G}(z) \in \mathbb{K}^{s\times m}[[z]]$, some degree constraints and order σ find solutions to $\mathbf{G}(z)\cdot\mathbf{P}(z) = O(z^{\sigma})$ satisfying degree constraints.

We make use of two insights

Insight I:

- Treat order part and degree parts separately.

Insight II:

- Make things look like Hermite-Padé problem
- Solve Hermite-Padé problem.
- Show similar techniques work in more general case.

Power Hermite-Padé Approximants

Vector Hermite Padé Approximants

Deal with vector problem by converting to scalar problem.

$$\mathbf{G}(z) \cdot \mathbf{P}(z) = z^{\vec{v}}\mathbf{R}(z)$$

converted to scalar problem via

$$\mathbf{A}(z) = [1, z, \ldots, z^{s-1}]\mathbf{G}(z^s)$$

Order problem now given by

$$\mathbf{A}(z) \cdot \mathbf{P}(z^s) = z^{\sigma}\mathbf{S}(z).$$

Example

Consider 2×2 model : $\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) \\ f_1(z) & g_1(z) \end{bmatrix}$

Example

Consider 2×2 model : $\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) \\ f_1(z) & g_1(z) \end{bmatrix}$

Form $\quad \mathbf{A}(z) = [f_0(z^2) + zf_1(z^2), g_0(z^2) + zg_1(z^2)]$

Example

Consider 2×2 model : $\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) \\ f_1(z) & g_1(z) \end{bmatrix}$

Form $\quad \mathbf{A}(z) = [f_0(z^2) + zf_1(z^2), g_0(z^2) + zg_1(z^2)]$

Then $\quad \mathbf{A}(z)\mathbf{P}(z^2) = z^{\sigma}\mathbf{S}(z)$ with $\mathbf{P} = [p_1, p_2]^T$ gives:

$$\begin{aligned} \mathbf{A}(z)\mathbf{P}(z^2) &= f_0(z^2)p_1(z^2) + zf_1(z^2)p_1(z^2) + g_0(z^2)p_2(z^2) + zg_1(z^2)p_2(z^2) \\ &= z^{\sigma}\mathbf{S}(z) \end{aligned}$$

Example

Consider 2×2 model : $\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) \\ f_1(z) & g_1(z) \end{bmatrix}$

Form $\mathbf{A}(z) = [f_0(z^2) + z f_1(z^2), g_0(z^2) + z g_1(z^2)]$

Then $\mathbf{A}(z)\mathbf{P}(z^2) = z^\sigma \mathbf{S}(z)$ with $\mathbf{P} = [p_1, p_2]^T$ gives:

$$\begin{aligned} \mathbf{A}(z)\mathbf{P}(z^2) &= f_0(z^2)p_1(z^2) + z f_1(z^2)p_1(z^2) + g_0(z^2)p_2(z^2) + z g_1(z^2)p_2(z^2) \\ &= z^\sigma \mathbf{S}(z) \end{aligned}$$

Same as

$$\begin{aligned} f_0(z)p_1(z) + g_0(z)p_2(z) &= z^\tau r_0(z) \\ f_1(z)p_1(z) + g_1(z)p_2(z) &= z^\tau r_1(z) \end{aligned}$$

Example

Consider 2×2 model : $\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) \\ f_1(z) & g_1(z) \end{bmatrix}$

Form $\quad \mathbf{A}(z) = [f_0(z^2) + zf_1(z^2), g_0(z^2) + zg_1(z^2)]$

Then $\quad \mathbf{A}(z)\mathbf{P}(z^2) = z^\sigma \mathbf{S}(z)$ with $\mathbf{P} = [p_1, p_2]^T$ gives:

$$\begin{aligned} \mathbf{A}(z)\mathbf{P}(z^2) &= f_0(z^2)p_1(z^2) + zf_1(z^2)p_1(z^2) + g_0(z^2)p_2(z^2) + zg_1(z^2)p_2(z^2) \\ &= z^\sigma \mathbf{S}(z) \end{aligned}$$

Same as

$$\begin{aligned} f_0(z)p_1(z) + g_0(z)p_2(z) &= z^\tau r_0(z) \\ f_1(z)p_1(z) + g_1(z)p_2(z) &= z^\tau r_1(z) \end{aligned}$$

Thus $\quad \mathbf{A}(z)\mathbf{P}(z^2) = z^\sigma \mathbf{S}(z)$ same as $\quad \mathbf{G}(z)\mathbf{P}(z) = z^\tau \mathbf{R}(z)$

Another Example

$$\mathbf{G}(z) = \begin{bmatrix} f_0(z) & g_0(z) & h_0(z) \\ f_1(z) & g_1(z) & h_1(z) \\ f_2(z) & g_2(z) & h_2(z) \end{bmatrix}$$

$\mathbf{G}(z)\mathbf{P}(z) = z^{\tau}\mathbf{R}(z)$ gives

$$\begin{aligned} \mathbf{A}(z) &= [f_0(z^3) + zf_1(z^3) + z^2 f_2(z^3), \\ &\qquad g_0(z^3) + zg_1(z^3) + z^2 g_2(z^3), \\ &\qquad\qquad h_0(z^3) + zh_1(z^3) + z^2 h_2(z^3)] \end{aligned}$$

$\mathbf{A}(z)\mathbf{P}(z^3) = z^{\sigma}\mathbf{S}(z)$ with $\mathbf{P} = [p_1, p_2, p_3]^T$ then gives

$$\begin{aligned} f_0(z)p_1(z) + g_0(z)p_2(z) + h_0(z)p_3(z) &= z^{\tau} r_0(z) \\ f_1(z)p_1(z) + g_1(z)p_2(z) + h_1(z)p_3(z) &= z^{\tau} r_1(z) \\ f_2(z)p_1(z) + g_2(z)p_2(z) + h_2(z)p_3(z) &= z^{\tau} r_2(z) \end{aligned}$$

Power Hermite-Padé

Let $\sigma \geq 0$, $s > 0$, $n_1, \ldots, n_m$ be integers, $n_l \geq -1$

Definition (Power Hermite Padé approximant)

$\mathbf{P} = (p_1, \ldots, p_m)$ of PHPA of type $(\mathbf{n}, \sigma, s)$ consists of scalar polynomials p_l having degrees bounded by the n_l with

$$\mathbf{A}(z) \cdot \mathbf{P}(z^s) = a_1(z)p_1(z^s) + \ldots + a_m(z)p_m(z^s) = c_\sigma z^\sigma + \cdots$$

Generalizes

- Hermite-Padé, Simultaneous-Padé, etc
- vector and matrix versions of the above

Order and Defect

Order and Defect

Given : $\mathbf{A} = (a_1, \ldots, a_m) \in \mathbb{K}^m[[z]]$ with $a_m(0) \neq 0$.

Definition (Order)

The *order* of a $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$

$$\textit{ord}\,\mathbf{P} := \sup\{\sigma \in \mathbb{N}_0 : \mathbf{A}(z) \cdot \mathbf{P}(z^s) = z^\sigma \cdot R(z) \text{ with } R \in \mathbb{K}[[z]]\}.$$

Order and Defect

Given : $\mathbf{A} = (a_1, \ldots, a_m) \in \mathbb{K}^m[[z]]$ with $a_m(0) \neq 0$.

Definition (Order)

The *order* of a $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$

$$\textit{ord}\,\mathbf{P} := \sup\{\sigma \in \mathbb{N}_0 : \mathbf{A}(z) \cdot \mathbf{P}(z^s) = z^\sigma \cdot R(z) \text{ with } R \in \mathbb{K}[[z]]\}.$$

Definition (Defect)

The *defect* of $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$ (w.r.t. $\mathbf{n} = (n_1, \ldots, n_m)$):

$$\textit{dct}\,\mathbf{P} := \min_l \{n_l + 1 - \textit{deg}\, P_l\}$$

Order and Defect

Given : $\mathbf{A} = (a_1, \ldots, a_m) \in \mathbb{K}^m[[z]]$ with $a_m(0) \neq 0$.

Definition (Order)

The *order* of a $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$

$$\mathit{ord}\, \mathbf{P} := \sup\{\sigma \in \mathbb{N}_0 : \mathbf{A}(z) \cdot \mathbf{P}(z^s) = z^\sigma \cdot R(z) \text{ with } R \in \mathbb{K}[[z]]\}.$$

Definition (Defect)

The *defect* of $\mathbf{P} = (P_1, \ldots, P_m) \in \mathbb{K}^m[z]$ (w.r.t. $\mathbf{n} = (n_1, \ldots, n_m)$):

$$\mathit{dct}\, \mathbf{P} := \min_l \{n_l + 1 - \mathit{deg}\, P_l\}$$

- defect is a measure of closeness to set of degree bounds

Defect, Order (cont.)

- *dct* $\mathbf{P} > 0$ same as *deg* $P_\ell \leq n_\ell$ for all ℓ.
- Hermite-Padé problem is same as : find $\mathbf{P}$ in

$$\mathcal{L}^\sigma = \{\mathbf{P} \in \mathbb{K}^m[z] : \textit{dct}\, \mathbf{P} > 0, \textit{ord}\, \mathbf{P} \geq \sigma\}$$

where $\sigma = n_1 + \cdots + n_m + m - 1$.

- In general look at

$$\mathcal{L}^\sigma = \{\mathbf{P} \in \mathbb{K}^m[z] : \textit{dct}\, \mathbf{P} > 0, \textit{ord}\, \mathbf{P} \geq \sigma\}$$

for arbitary σ.

Sigma Basis Algorithm

Recall : Order Bases

Let $\sigma \in \mathbb{N}_0$.

Definition (Order Bases)

The system $\mathbf{P}_1, \ldots, \mathbf{P}_m \in \mathbb{K}^m[z]$ is called an *order-basis* if:

(a) *ord* $\mathbf{P}_\ell \geq \sigma$ for all ℓ,

(b) For each $\mathbf{F} \in \mathcal{L}^\sigma$ there exists one and only one tuple of polynomials $(\alpha_1, \ldots, \alpha_m)$, *deg* $\alpha_l <$ *dct* $\mathbf{P}_l$ such that

$$\mathbf{F} = \alpha_\mathbf{1} \cdot \mathbf{P_1} + \ldots + \alpha_\mathbf{m} \cdot \mathbf{P_m}.$$

Recall : Order Bases

Let $\sigma \in \mathbb{N}_0$.

Definition (Order Bases)

The system $\mathbf{P}_1, \ldots, \mathbf{P}_m \in \mathbb{K}^m[z]$ is called an *order-basis* if:

(a) *ord* $\mathbf{P}_\ell \geq \sigma$ for all ℓ,

(b) For each $\mathbf{F} \in \mathcal{L}^\sigma$ there exists one and only one tuple of polynomials $(\alpha_1, \ldots, \alpha_m)$, *deg* $\alpha_l <$ *dct* $\mathbf{P}_l$ such that

$$\mathbf{F} = \alpha_\mathbf{1} \cdot \mathbf{P_1} + \ldots + \alpha_\mathbf{m} \cdot \mathbf{P_m}.$$

Also called Sigma Bases or Minimal Approximant Bases

Sigma Bases (cont.)

An order basis $\mathbf{P}_1, \ldots, \mathbf{P}_m$ is linearly independent with respect to polynomial coefficients - i.e. a basis of the order module.

Moreover, in terms of vector spaces, we have

$$\mathcal{L}^{\sigma} = \text{span}\{z^j \cdot \mathbf{P}_l : 1 \leq l \leq m, 0 \leq j < \textit{dct}\, \mathbf{P}_l\}$$

$$\text{dim}\mathcal{L}^{\sigma} = \max\{\textit{dct}\, \mathbf{P}_1, 0\} + \ldots + \max\{\textit{dct}\, \mathbf{P}_m, 0\}.$$

The Sigma Basis Algorithm

Constructive process:

- Start with $\sigma = 0$. $\mathbf{P} = I_m$. $\mathbf{AP} = z^{\sigma}\mathbf{R}$.
- At each step do:
  - If each $\mathbf{R}_{\ell}(0) = 0$ then increment σ
  - If there $\mathbf{R}_{\ell}(0) \neq 0$ then find pivot and eliminate rest of initial values
  - Pivot should be first column having maximal defect.
  - Multiply pivot column by z.

The Algorithm

INPUT: $m \geq 2$, $s \in \mathbb{N}$, $\mathbf{A} = (a_1, \dots a_m)^T$, multi-index $\mathbf{n} = (n_1, \dots, n_m)$

INITIALIZATION: Let $\sigma = 0$, $d_{l,0} = n_l$, $\mathbf{P}_{l,0} = (0, .., 0, 1, 0, .., 0)$

RECURSIVE STEP: For $\sigma = 0, 1, 2, \dots$:

Let $l = 1, \dots, m$: $c_{l,\sigma} = z^{-\sigma} \cdot \mathbf{P}_{l,\sigma}(z^s) \cdot \mathbf{A}(z) \mid_{z=0}$ and $\Lambda_\sigma = \{l : c_{l,\sigma} \neq 0\}$

CASE $\Lambda_\sigma = \{\}$, then for $l = 1, \dots, m$: $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}$, $d_{l,\sigma+1} = d_{l,\sigma}$

CASE $\Lambda_\sigma \neq \{\}$, then let $\pi = \pi_\sigma \in \Lambda_\sigma$ be defined by

$d_{\pi,\sigma} = \max\{d_{l,\sigma} : l \in \Lambda_\sigma\}$

and compute for $l = 1, \dots, m$:

$l \in \Lambda_\sigma, l \neq \pi$: $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma} - \frac{c_{l,\sigma}}{c_{\pi,\sigma}} \cdot \mathbf{P}_{\pi,\sigma}$, $d_{l,\sigma+1} = d_{l,\sigma}$

$l \notin \Lambda_\sigma$: $\mathbf{P}_{l,\sigma+1} = \mathbf{P}_{l,\sigma}$, $d_{l,\sigma+1} = d_{l,\sigma}$

$l = \pi$: $\mathbf{P}_{\pi,\sigma+1} = z \cdot \mathbf{P}_{\pi,\sigma}$, $d_{\pi,\sigma+1} = d_{\pi,\sigma} - 1$

OUTPUT: σ-bases $\mathbf{P}_{1,\sigma}, \dots, \mathbf{P}_{m,\sigma}$ with *dct* $\mathbf{P}_{l,\sigma} = d_{l,\sigma} + 1$, $1 \leq l \leq m$.

Complexity

Complexity of Sigma Bases Algorithm

(Complexity : I) *For computing vector HPA's of order* $\sigma = 0, 1, \ldots, \|\mathbf{n}\|$ *:*

$$4(m-s)\cdot\|\mathbf{n}\|^2 + \mathcal{O}(m^2\cdot\|\mathbf{n}\|)$$

roughly half additions and half multiplications plus $\mathcal{O}(m\cdot\|\mathbf{n}\|)$ *divisions.*

(Complexity : II) *For the case* $\mathbf{n} = (n, \ldots, n)$*, we obtain the sharper bound*

$$(1-\frac{s}{m})\cdot(2m - \text{card}\, L)\cdot\|\mathbf{n}\|^2 + \mathcal{O}(m^2\cdot\|\mathbf{n}\|)$$

where $L = \{l : a_l(z) = z^j$ *with a* $j \in \mathbb{N}_0\}$*.*

Faster Sigma Bases Algorithm

Recursive Computation : *Suppose* $0 \leq \rho \leq \sigma$ *and*

$$(\mathbf{P}^{(1)}, \mathbf{d}^{(1)}) \longleftarrow \textit{FPHPS}(\mathbf{A}, \rho, \mathbf{n})$$

Let $\mathbf{A}^{(1)}(z) := z^{-\rho} \cdot \mathbf{P}^{(1)}(z^s) \cdot \mathbf{A}(z)$. *Compute*

$$(\mathbf{P}^{(2)}, \mathbf{d}^{(2)}) \longleftarrow \textit{FPHPS}(\mathbf{A}^{(1)}, \sigma - \rho, \mathbf{d}^{(1)}).$$

Then $(\mathbf{P}^{(3)}, \mathbf{d}^{(3)}) \longleftarrow \textit{FPHPS}(\mathbf{A}, \sigma, \mathbf{n})$ *where*

$$\mathbf{P}^{(3)} = \mathbf{P}^{(2)} \cdot \mathbf{P}^{(1)} \quad \textit{and} \quad \mathbf{d}^{(3)} = \mathbf{d}^{(2)}.$$

Complexity of Superfast Method

(Complexity) *The superfast algorithm for computing vector HPA's of order* σ *has a complexity of at most*

$$\frac{3}{2} \cdot (m+s) \cdot m \cdot \sigma \cdot \log^2 \sigma + \mathcal{O}(\sigma \cdot \log \sigma)$$

roughly half multipications as additions.