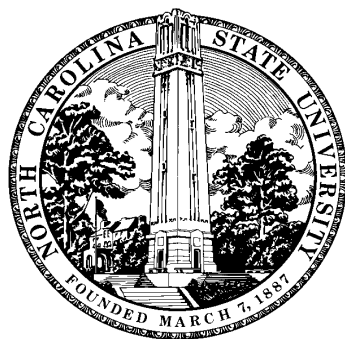


*Approximate Greatest Common Divisors of
Several Multivariate Polynomials
with Linearly Constrained Coefficients
and Singular Polynomials*

Erich Kaltofen

North Carolina State University

google->kaltofen



Joint work with Zhengfeng Yang and Lihong Zhi

Approximate GCD Problem

Given polynomials $f_1, \dots, f_s \in F[y_1, y_2, \dots, y_r] \setminus \{0\}$, where F is \mathbb{R} or \mathbb{C} ; let $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ for all i with $1 \leq i \leq s$.

We wish to compute $\Delta f_1, \dots, \Delta f_s \in F'[y_1, y_2, \dots, y_r]$, where F' is \mathbb{R} or \mathbb{C} , such that $\text{tdeg}(\Delta f_1) \leq d_1, \dots, \text{tdeg}(\Delta f_s) \leq d_s$,

- $\text{tdeg}(\gcd(f_1 + \Delta f_1, \dots, f_s + \Delta f_s)) \geq k$,
- $\|\Delta f_1\|_2^2 + \dots + \|\Delta f_s\|_2^2$ is minimized.

Approximate GCD Problem

Given polynomials $f_1, \dots, f_s \in F[y_1, y_2, \dots, y_r] \setminus \{0\}$, where F is \mathbb{R} or \mathbb{C} ; let $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ for all i with $1 \leq i \leq s$.

We wish to compute $\Delta f_1, \dots, \Delta f_s \in F'[y_1, y_2, \dots, y_r]$, where F' is \mathbb{R} or \mathbb{C} , such that $\text{tdeg}(\Delta f_1) \leq d_1, \dots, \text{tdeg}(\Delta f_s) \leq d_s$,

- $\text{tdeg}(\gcd(f_1 + \Delta f_1, \dots, f_s + \Delta f_s)) \geq k$,
- $\|\Delta f_1\|_2^2 + \dots + \|\Delta f_s\|_2^2$ is minimized.

Problem depends on choice of norm $\|\cdot\|$, and notion of degree.
We use **2-norm** on the coefficient vector, and **total degree**.

Previous Work on Approximate GCD

- Modified Euclidean algorithm for polynomials with floating point coefficients
[Dunaway '74, Schönhage '85, Sasaki and Noda '89 & '91, Ochi et al. '91, Hribernic and Stetter '97, Beckermann and Labahn '98, Sasaki and Sasaki '01, Sanuki '05]
- Nearby roots matching, resultant-based algorithms, QR factorization, Hensel lifting strategy [Pan '01, Emiris et al. '96, Rupprecht '99, Zhi and Noda '00, Corless et al. '04]
- Least squares and SVD-based total least squares methods
[Corless et al. '95, Chin et al. '98, Karmarkar and Lakshamn '96 & '98, Zeng '03 & '04, Gao et al. '04, Diaz-Toca and Gonzalez-Vega '02 & '06]
- Structure preserving total least squares algorithms
[Chu et al. '03, Li et al. '05, Kaltofen et al. '05, Botting et al. '05, Markovsky and Huffel '05]

Globally Nearest Complex/Real Solution

Theorem 1. Let $f_1, \dots, f_s \in F[y_1, \dots, y_r] \setminus \{0\}$, where F is \mathbb{C} or \mathbb{R} , $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ with $1 \leq i \leq s$. There exist $\hat{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\hat{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\hat{f}_1, \dots, \hat{f}_s)) \geq k,$$

Globally Nearest Complex/Real Solution

Theorem 1. Let $f_1, \dots, f_s \in F[y_1, \dots, y_r] \setminus \{0\}$, where F is \mathbb{C} or \mathbb{R} , $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ with $1 \leq i \leq s$. There exist $\hat{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\hat{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\hat{f}_1, \dots, \hat{f}_s)) \geq k,$$

such that for all $\bar{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\bar{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\bar{f}_1, \dots, \bar{f}_s)) \geq k,$$

Globally Nearest Complex/Real Solution

Theorem 1. Let $f_1, \dots, f_s \in F[y_1, \dots, y_r] \setminus \{0\}$, where F is \mathbb{C} or \mathbb{R} , $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ with $1 \leq i \leq s$. There exist $\hat{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\hat{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\hat{f}_1, \dots, \hat{f}_s)) \geq k,$$

such that for all $\bar{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\bar{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\bar{f}_1, \dots, \bar{f}_s)) \geq k,$$

we have

$$\|\hat{f}_1 - f_1\|_2^2 + \dots + \|\hat{f}_s - f_s\|_2^2 \leq \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2.$$

Globally Nearest Complex/Real Solution

Theorem 1. Let $f_1, \dots, f_s \in F[y_1, \dots, y_r] \setminus \{0\}$, where F is \mathbb{C} or \mathbb{R} , $d_i = \text{tdeg}(f_i)$ and $k \leq d_i$ with $1 \leq i \leq s$. There exist $\hat{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\hat{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\hat{f}_1, \dots, \hat{f}_s)) \geq k,$$

such that for all $\bar{f}_i \in F[y_1, y_2, \dots, y_r]$, $1 \leq i \leq s$ with

$$\text{tdeg}(\bar{f}_i) \leq d_i, \quad \text{and} \quad \text{tdeg}(\text{gcd}(\bar{f}_1, \dots, \bar{f}_s)) \geq k,$$

we have

$$\|\hat{f}_1 - f_1\|_2^2 + \dots + \|\hat{f}_s - f_s\|_2^2 \leq \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2.$$

Remark: Theorem is false if $\text{tdeg gcd} = k$.

Multi-polynomial Generalized Sylvester Matrix

Denote the coefficient matrix of the system

$$u_1 \neq 0, \quad \forall i, 2 \leq i \leq s: u_i f_1 + u_1 f_i = 0, \quad \text{tdeg}(u_i) \leq d_i - k$$

as $S_k(f_1, \dots, f_s) =$

$$\begin{bmatrix} C^{[d_2-k]}(f_1) & \mathbf{0} & \dots & \mathbf{0} & C^{[d_1-k]}(f_2) \\ \mathbf{0} & C^{[d_3-k]}(f_1) & & \mathbf{0} & C^{[d_1-k]}(f_3) \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & C^{[d_s-k]}(f_1) & C^{[d_1-k]}(f_s) \end{bmatrix}$$

Convolution Matrix

The convolution matrix $\mathbf{C}^{[l]}(f)$ produces the coefficient vector of $u \cdot f$ as $\mathbf{C}^{[l]}(f) \cdot \vec{u}$, where $l = \text{tdeg}(u)$. For instance,

$$\mathbf{C}^{[2]}(a_2y^2 + a_1y + a_0) \cdot \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} a_2 & 0 & 0 \\ a_1 & a_2 & 0 \\ a_0 & a_1 & a_2 \\ 0 & a_0 & a_1 \\ 0 & 0 & a_0 \end{bmatrix} \cdot \begin{bmatrix} b_2 \\ b_1 \\ b_0 \end{bmatrix}.$$

In the univariate case, the matrix is of Toeplitz form. In the multivariate case, the dimensions of $\mathbf{C}^{[l]}(f)$ with $\text{tdeg}(f) = m$ are $\binom{l+m+r}{r} \times \binom{l+r}{r}$.

The Minimal Perturbation Problem

Lemma 1. $\text{tdeg}(\gcd(f_1, \dots, f_s)) \geq k$ iff $S_k(f_1, \dots, f_s)$ has rank deficiency at least one.

The Minimal Perturbation Problem

Lemma 1. $\text{tdeg}(\gcd(f_1, \dots, f_s)) \geq k$ iff $S_k(f_1, \dots, f_s)$ has rank deficiency at least one.

Our problem can be transformed into:

$$\begin{aligned} & \min_{\text{tdeg}(\gcd(\bar{f}_1, \dots, \bar{f}_s)) \geq k} \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2 \\ & \iff \min_{\dim \text{Nullspace}(\bar{S}_k) \geq 1} \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2 \end{aligned}$$

where \bar{S}_k is the k -th Sylvester matrix generated by $\bar{f}_1, \dots, \bar{f}_s$ with $\text{tdeg} \bar{f}_i \leq d_i, 1 \leq i \leq s$.

The Minimal Perturbation Problem

Lemma 1. $\text{tdeg}(\gcd(f_1, \dots, f_s)) \geq k$ iff $S_k(f_1, \dots, f_s)$ has rank deficiency at least one.

Our problem can be transformed into:

$$\begin{aligned} \min_{\text{tdeg}(\gcd(\bar{f}_1, \dots, \bar{f}_s)) \geq k} \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2 \\ \iff \min_{\dim \text{Nullspace}(\bar{S}_k) \geq 1} \|\bar{f}_1 - f_1\|_2^2 + \dots + \|\bar{f}_s - f_s\|_2^2 \end{aligned}$$

where \bar{S}_k is the k -th Sylvester matrix generated by $\bar{f}_1, \dots, \bar{f}_s$ with $\text{tdeg} \bar{f}_i \leq d_i, 1 \leq i \leq s$.

Nearest singular matrix + Sylvester structure

\implies Approximate GCD

Structure Preserving Low Rank Approximation

Let $S(\zeta) = [A_1(\zeta) \mid b(\zeta) \mid A_2(\zeta)]$ and let $A(\zeta) = [A_1(\zeta) \mid A_2(\zeta)]$, where ζ contains the coefficients of f_1, \dots, f_s . We solve the structure-preserving total least norm problem

$$\min_{\Delta \mathbf{c} \in \mathbb{R}^v} \|\Delta \mathbf{c}\| \text{ or } \min_{\Delta \mathbf{c} \in \mathbb{C}^v} \|\Delta \mathbf{c}\| \quad \text{with} \quad A(\mathbf{c} + \Delta \mathbf{c})\mathbf{x} = b(\mathbf{c} + \Delta \mathbf{c})$$

for some vector \mathbf{x} . Here \mathbf{c} is fixed to the initial coefficient vector.

Structure Preserving Low Rank Approximation

Let $S(\zeta) = [A_1(\zeta) \mid b(\zeta) \mid A_2(\zeta)]$ and let $A(\zeta) = [A_1(\zeta) \mid A_2(\zeta)]$, where ζ contains the coefficients of f_1, \dots, f_s . We solve the structure-preserving total least norm problem

$$\min_{\Delta \mathbf{c} \in \mathbb{R}^v} \|\Delta \mathbf{c}\| \text{ or } \min_{\Delta \mathbf{c} \in \mathbb{C}^v} \|\Delta \mathbf{c}\| \quad \text{with} \quad A(\mathbf{c} + \Delta \mathbf{c})\mathbf{x} = b(\mathbf{c} + \Delta \mathbf{c})$$

for some vector \mathbf{x} . Here \mathbf{c} is fixed to the initial coefficient vector.

Remark: We choose the column $b(\zeta)$ corresponding to the absolutely largest component in the first singular vector of $S(\zeta)$.

The STLN Algorithm

The STLN Algorithm first initializes \mathbf{x} as a solution to

$$A(\mathbf{c})\mathbf{x} \approx b(\mathbf{c})$$

for the input parameters \mathbf{c} and $\Delta\mathbf{c} = \mathbf{z} = 0$, then updates $\mathbf{x} + \Delta\mathbf{x}$ and $\mathbf{z} + \Delta\mathbf{z}$ satisfy

$$\min \|\mathbf{z} + \Delta\mathbf{z}\| \quad \text{with} \quad A(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z})(\mathbf{x} + \Delta\mathbf{x}) \approx b(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z}).$$

References: Rosen et al. '96, Park et al. '99

New Initialization of \mathbf{x} and $\Delta\mathbf{c}$

Suppose $H(\xi)\zeta = S(\zeta)\xi$ and \mathbf{v} is the first singular vector of $S(\mathbf{c})$ and

$$\Delta\mathbf{c} = \mathbf{z} = -H(\mathbf{v})^{Tr} (H(\mathbf{v})H(\mathbf{v})^{Tr})^{-1} S(\mathbf{c})\mathbf{v}.$$

We have $-S(\mathbf{z})\mathbf{v} = -H(\mathbf{v})\mathbf{z} = S(\mathbf{c})\mathbf{v}$, hence $S(\mathbf{c} + \mathbf{z})\mathbf{v} = 0$.

New Initialization of \mathbf{x} and $\Delta\mathbf{c}$

Suppose $H(\xi)\zeta = S(\zeta)\xi$ and \mathbf{v} is the first singular vector of $S(\mathbf{c})$ and

$$\Delta\mathbf{c} = \mathbf{z} = -H(\mathbf{v})^{Tr} (H(\mathbf{v})H(\mathbf{v})^{Tr})^{-1} S(\mathbf{c})\mathbf{v}.$$

We have $-S(\mathbf{z})\mathbf{v} = -H(\mathbf{v})\mathbf{z} = S(\mathbf{c})\mathbf{v}$, hence $S(\mathbf{c} + \mathbf{z})\mathbf{v} = 0$.

We initialize \mathbf{x} as

$$\mathbf{x} = \left[-\frac{\mathbf{v}[1]}{\mathbf{v}[t]}, \dots, -\frac{\mathbf{v}[t-1]}{\mathbf{v}[t]}, -\frac{\mathbf{v}[t+1]}{\mathbf{v}[t]}, \dots \right]^{Tr},$$

where $\mathbf{v}[t]$ is the absolutely largest component. We have

$$A(\mathbf{c} + \mathbf{z})\mathbf{x} = b(\mathbf{c} + \mathbf{z}).$$

Reference: Lemmerling'99

Penalty Method

Let the residue

$$\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) = b(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z}) - A(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z})(\mathbf{x} + \Delta\mathbf{x})$$

Penalty Method

Let the residue

$$\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) = b(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z}) - A(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z})(\mathbf{x} + \Delta\mathbf{x})$$

The structured minimization problem can be transformed into:

$$\min_{\Delta\mathbf{z}, \Delta\mathbf{x}} \left\| \begin{array}{c} w\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) \\ \mathbf{z} + \Delta\mathbf{z} \end{array} \right\|,$$

where w is a large penalty value between 10^8 and 10^{10} .

Penalty Method

Let the residue

$$\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) = b(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z}) - A(\mathbf{c} + \mathbf{z} + \Delta\mathbf{z})(\mathbf{x} + \Delta\mathbf{x})$$

The structured minimization problem can be transformed into:

$$\min_{\Delta\mathbf{z}, \Delta\mathbf{x}} \left\| \begin{array}{c} w\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) \\ \mathbf{z} + \Delta\mathbf{z} \end{array} \right\|,$$

where w is a large penalty value between 10^8 and 10^{10} .

Reference: Rosen et al. '96, Park et al. '99

Linear Approximation to the Residue

$$\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) \approx \mathbf{r}(\mathbf{z}, \mathbf{x}) + b(\Delta\mathbf{z}) - A(\mathbf{c} + \mathbf{z})\Delta\mathbf{x} - A(\Delta\mathbf{z})\mathbf{x}.$$

We introduce a constant matrix P with

$$b(\Delta\mathbf{z}) = P\Delta\mathbf{z},$$

and a Sylvester-like matrix $Y(\xi)$ such that

$$Y(\xi)\zeta = A(\zeta)\xi, \quad \text{in particular} \quad A(\Delta\mathbf{z})\mathbf{x} = Y(\mathbf{x})\Delta\mathbf{z}.$$

Hence the first order approximation of the residue is

$$\mathbf{r}(\mathbf{z} + \Delta\mathbf{z}, \mathbf{x} + \Delta\mathbf{x}) \approx \mathbf{r}(\mathbf{z}, \mathbf{x}) + P\Delta\mathbf{z} - A(\mathbf{c} + \mathbf{z})\Delta\mathbf{x} - Y(\mathbf{x})\Delta\mathbf{z}.$$

First Order Iterative Update

$$\min_{\Delta \mathbf{x}, \Delta \mathbf{z}} \left\| \begin{bmatrix} w(Y(\mathbf{x}) - P) & wA(\mathbf{c} + \mathbf{z}) \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z} \\ \Delta \mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r}(\mathbf{z}, \mathbf{x}) \\ \mathbf{z} \end{bmatrix} \right\|$$

The iterative update

$$\mathbf{x} = \mathbf{x} + \Delta \mathbf{x}, \quad \mathbf{z} = \mathbf{z} + \Delta \mathbf{z}$$

is stopped when

$$\|\Delta \mathbf{x}\| \leq tol \quad \text{and/or} \quad \|\Delta \mathbf{z}\| \leq tol$$

Random Complex Perturbation in the Initialization

We initialize \mathbf{z} as:

$$\mathbf{z} = -H(\mathbf{v})^{Tr} (H(\mathbf{v})H(\mathbf{v})^{Tr})^{-1} S(\mathbf{c} + i\Delta\mathbf{c}_{rand})\mathbf{v},$$

where \mathbf{v} is the first sing. vec. of $S(\mathbf{c} + i\Delta\mathbf{c}_{rand})$, $\Delta\mathbf{c}_{rand}$ is a random real vector of small noise.

We initialize \mathbf{x} by normalizing \mathbf{v} , as before then

$$A(\mathbf{c} + i\Delta\mathbf{c}_{rand} + \mathbf{z})\mathbf{x} = b(\mathbf{c} + i\Delta\mathbf{c}_{rand} + \mathbf{z}).$$

Multiple Local Minima

Consider the polynomials

$$f = 1000y^{10} + y^3 - 1 \quad \text{and} \quad g = y^2 - \frac{1}{100}.$$

We seek to compute the nearest pair of polynomials \tilde{f} and \tilde{g} that have a non-trivial GCD.

The algorithm converges after about ten iterations in average to the local minima:

$$0.0421579, 0.0463113, 0.0474087, 0.0493292, \dots$$

for different initializations.

Among solutions, the polynomials

$$\tilde{f} = 1000.0y^{10} + 0.0000147908y^9 + \dots - 0.991601,$$

$$\tilde{g} = 0.956139y^2 - 0.0887590y - 0.189618,$$

have a common divisor

$$y - 0.4941547,$$

and the backward error is

$$\|f - \tilde{f}\|_2^2 + \|g - \tilde{g}\|_2^2 = 0.0421579.$$

It is the non-monic global minimum found by the global methods

Karmarkar and Lakshman Y.N. 1998; Hitz and Kaltofen 1998.

Structured Condition Number

The nearest unstructured singular matrix \hat{A} has

$$\|S - \hat{A}\|_F^2 = 0.000000098975.$$

Structured Condition Number

The nearest unstructured singular matrix \hat{A} has

$$\|S - \hat{A}\|_F^2 = 0.000000098975.$$

Let \hat{S} be the nearest singular Sylvester matrix to S ,

$$\begin{aligned}\|S - \hat{S}\|_F^2 &= 2\|f - \hat{f}\|_2^2 + 10\|g - \hat{g}\|_2^2 \\ &\geq 2\|f - \hat{f}\|_2^2 + 2\|g - \hat{g}\|_2^2 \\ &\geq 2\|f - \tilde{f}\|_2^2 + 2\|g - \tilde{g}\|_2^2 \geq 0.084315.\end{aligned}$$

Structured Condition Number

The nearest unstructured singular matrix \hat{A} has

$$\|S - \hat{A}\|_F^2 = 0.000000098975.$$

Let \hat{S} be the nearest singular Sylvester matrix to S ,

$$\begin{aligned}\|S - \hat{S}\|_F^2 &= 2\|f - \hat{f}\|_2^2 + 10\|g - \hat{g}\|_2^2 \\ &\geq 2\|f - \hat{f}\|_2^2 + 2\|g - \hat{g}\|_2^2 \\ &\geq 2\|f - \tilde{f}\|_2^2 + 2\|g - \tilde{g}\|_2^2 \geq 0.084315.\end{aligned}$$

Sylvester \neq Toeplitz à la Rump [2003]

App. GCD of Multivariate Polynomials

<i>Ex.</i>	<i>d_i</i>	<i>k</i>	<i>e</i>	<i>it.</i>	<i>error</i> (Zeng)	<i>error</i> (GKMYZ)	<i>error</i> (STLN)
1	7,7	4	3	2	2.44360e-4	2.59476e-4	6.50358e-5
2	7,7	4	5	1	2.44404e-8	2.59194e-8	6.50357e-9
3	7,7	4	7	1	2.44405e-12	2.59191e-12	6.50357e-13
4	7,7	4	9	1	2.44396e-16	2.59187e-16	6.50361e-17
5	6,6	3	2	3	2.26617	1.49524	4.80154e-1
6	10,10	5	4	2		2.74672e-3	1.84914e-3
7	8,8	4	5	2	7.09371e-5	2.38059e-5	2.01393e-5
8	40,40	30	5	2	1.39858e-3	4.83931e-4	4.39489e-4
9	10,9,8	5	3	2			6.21772e-2
10	8,7,8,6	4	5	2			4.04458e-6

Linearly Constrained Input Coefficients

- Monicity, sparsity:

set $z_j = \text{constant}$.

Linearly Constrained Input Coefficients

- Monicity, sparsity:

$$\text{set } z_j = \text{constant.}$$

- Nearest singular polynomials, approximate squarefree factorization, and symmetry:

$$\text{set } z_j = \lambda z_i$$

where λ is an integer.

Linearly Constrained Input Coefficients

- Monicity, sparsity:

$$\text{set } z_j = \text{constant.}$$

- Nearest singular polynomials, approximate squarefree factorization, and symmetry:

$$\text{set } z_j = \lambda z_i$$

where λ is an integer.

- In general, the linear constraints on the goal coefficients are

$$\Gamma \zeta = \gamma.$$

Intuitive Approach

We add the constraints $\Gamma(\mathbf{c} + \mathbf{z}) = \gamma$ directly, and iterate on

$$\min_{\Delta \mathbf{x}, \Delta \mathbf{z}} \left\| \begin{bmatrix} w(Y(\mathbf{x}) - P) & wA(\mathbf{c} + \mathbf{z}) \\ \mathbf{I} & \mathbf{0} \\ w\Gamma & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z} \\ \Delta \mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r}(\mathbf{z}, \mathbf{x}) \\ \mathbf{z} \\ w(\Gamma(\mathbf{c} + \mathbf{z}) - \gamma) \end{bmatrix} \right\|.$$

However, that least squares problem has a **larger** dimension.

Our Approach

Let the linear constraints on the goal coefficients be

$$\Gamma\zeta = \gamma.$$

We construct for the linear system (\mathbf{c} is the input coeff. vec.)

$$\Gamma\zeta = \gamma - \Gamma\mathbf{c}$$

a matrix \mathbf{C} , a vector \mathbf{d} and a sub-vector of free parameters $[\zeta_{i_1}, \dots, \zeta_{i_\mu}]^{Tr}$ such that

$$\zeta = \mathbf{C}\zeta^- + \mathbf{d}, \quad \zeta^- = \begin{bmatrix} \zeta_{i_1} \\ \vdots \\ \zeta_{i_\mu} \end{bmatrix}; \quad \Gamma\mathbf{C} = 0, \quad \Gamma\mathbf{d} = \gamma - \Gamma\mathbf{c}.$$

Initialization of \mathbf{x} and \mathbf{z}

Let

$$\mathbf{z} = \mathbf{d} - C (H(\mathbf{v})C)^{Tr} (H(\mathbf{v})C (H(\mathbf{v})C)^{Tr})^{-1} S(\mathbf{c} + \mathbf{d})\mathbf{v},$$

\mathbf{v} is the first singular vector of the matrix $S(\mathbf{c} + \mathbf{d})$. We have

$$S(\mathbf{c} + \mathbf{z})\mathbf{v} = 0.$$

We initialize \mathbf{x} as

$$\mathbf{x} = \left[-\frac{\mathbf{v}[1]}{\mathbf{v}[t]}, \dots, -\frac{\mathbf{v}[t-1]}{\mathbf{v}[t]}, -\frac{\mathbf{v}[t+1]}{\mathbf{v}[t]}, \dots \right]^{Tr},$$

where $\mathbf{v}[t]$ is the absolutely largest component. We have

$$A(\mathbf{c} + \mathbf{z})\mathbf{x} = b(\mathbf{c} + \mathbf{z}).$$

First Order Iterative Update

The iterative update can be written as

$$\min_{\Delta \mathbf{x}, \Delta \mathbf{z}^-} \left\| \begin{bmatrix} w(Y(\mathbf{x}) - P)C & wA(\mathbf{c} + \mathbf{z}) \\ C & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z}^- \\ \Delta \mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r}(\mathbf{z}, \mathbf{x}) \\ \mathbf{z} \end{bmatrix} \right\|.$$

First Order Iterative Update

The iterative update can be written as

$$\min_{\Delta \mathbf{x}, \Delta \mathbf{z}^-} \left\| \begin{bmatrix} w(Y(\mathbf{x}) - P)C & wA(\mathbf{c} + \mathbf{z}) \\ C & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z}^- \\ \Delta \mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r}(\mathbf{z}, \mathbf{x}) \\ \mathbf{z} \end{bmatrix} \right\|.$$

The new coefficient values are

$$\mathbf{c} + \mathbf{z} + \Delta \mathbf{z} = \mathbf{c} + \mathbf{z} + C\Delta \mathbf{z}^-,$$

and satisfies

$$\Gamma(\mathbf{c} + \mathbf{z} + \Delta \mathbf{z}) = \Gamma(\mathbf{c} + \mathbf{z}) = \Gamma(\mathbf{c} + \mathbf{d}) = \gamma$$

throughout the iteration.

Nearest Singular Polynomials

Lemma 3. Let $f(y) \in F[y]$ with $\deg(f) = n$ over a field F of characteristic 0 , and let k be a multiplicity with $2 \leq k \leq n$.

Denote by $f^{[i]} = d^i f / dy^i$, then $\deg(\gcd(f^{[0]}, \dots, f^{[k-1]})) \geq 1$ iff the matrix $S_k^{sing}(f) =$

$$\begin{bmatrix} C^{[n-1]}(f^{[k-1]}) & \mathbf{0} & \dots & \mathbf{0} & C^{[n-k]}(f^{[0]}) \\ \mathbf{0} & C^{[n-2]}(f^{[k-1]}) & & \mathbf{0} & C^{[n-k]}(f^{[1]}) \\ \vdots & & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & C^{[n-k+1]}(f^{[k-1]}) & C^{[n-k]}(f^{[k-2]}) \end{bmatrix}$$

has rank deficiency at least one.

Weighted Minimization Problem

For the nearest k -fold singular polynomial one optimizes

$$\|\Delta f\|,$$

while the GCD problem with the corresponding constraints on the coefficients optimizes

$$\sum \|d^i \Delta f / dy^i\|,$$

which has a different minimum.

We introduce a weight matrix D to the minimization problem,

$$\min_{\Delta \mathbf{c} \in \mathbb{R}^v} \|D\Delta \mathbf{c}\| \text{ or } \min_{\Delta \mathbf{c} \in \mathbb{C}^v} \|D\Delta \mathbf{c}\| \quad \text{with} \quad A(\mathbf{c} + \Delta \mathbf{c})\mathbf{x} = b(\mathbf{c} + \Delta \mathbf{c}).$$

Then the first order iterative update becomes

$$\min_{\Delta \mathbf{x}, \Delta \mathbf{z}^-} \left\| \begin{bmatrix} w(Y(\mathbf{x}) - P)C & wA(\mathbf{c} + \mathbf{z}) \\ DC & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{z}^- \\ \Delta \mathbf{x} \end{bmatrix} + \begin{bmatrix} -w\mathbf{r}(\mathbf{z}, \mathbf{x}) \\ D\mathbf{z} \end{bmatrix} \right\|.$$

Reference: H. Park et al. 99

<i>Ex.</i>	<i>d</i>	<i>k</i>	<i>it.</i>	<i>error(ZNKW)</i>	<i>error(STLN)</i>
1	4	2	12	.1763296120	.1763296118
		3	34	.6261127476	.6261127498
2	4	2	4	.1552760123e-12	.1552723415e-12
		3	11	.8834609009e-9	.9814886696e-9
		4	4	.2021848972e-4	.1958553174e-4
3	4	2	4	.1645037985e-10	.16450617515e-10
		3	4	.4144531274e-6	.4144531274e-6
		4	12	.1049993144	.1049993152
4	5	2	1	.2460987981e-8	.2461467456e-8
	5	3	20	.3681785214	.3681784856
5	6	2	2	.3231668276e-5	.3231668277e-5
6	6	2	3	.3009788845e-11	.3009789157e-11
		3	3	.7453849284e-6	.7453849284e-6
		4	24	.4449023547	.4449023547
7	5	2	8	.8565349347	.8565327605
8	21	2	2	.190477e-8	.1893347157e-8
		3	6	.963776e-4	.9637591989e-4

Univariate Singular Polynomials

Current Investigations

- Compute nearest factorization with a given degree pattern, e.g., all linear factors.

Current Investigations

- Compute nearest factorization with a given degree pattern, e.g., all linear factors.
- Compute nearest polynomial system with a given structure of zeros.

Current Investigations

- Compute nearest factorization with a given degree pattern, e.g., all linear factors.
- Compute nearest polynomial system with a given structure of zeros.
- Nonlinear structure low rank approximation for Bezout matrix.

Current Investigations

- Compute nearest factorization with a given degree pattern, e.g., all linear factors.
- Compute nearest polynomial system with a given structure of zeros.
- Nonlinear structure low rank approximation for Bezout matrix.
- Investigation of displacement operators for generalized Sylvester matrix or Bezout matrix or Ruppert matrix.

Current Investigations

- Compute nearest factorization with a given degree pattern, e.g., all linear factors.
- Compute nearest polynomial system with a given structure of zeros.
- Nonlinear structure low rank approximation for Bezout matrix.
- Investigation of displacement operators for generalized Sylvester matrix or Bezout matrix or Ruppert matrix.
- How to certify a local minimum is a global minimum

Code + Benchmarks at:

<http://mmrc.iss.ac.cn/~lzhi/Research/hybrid/manystln.html>

or

[google->kaltofen](#) (click on “Software”)

Grazie mille!